Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____	_____
Martha Konar	Date

Blood Microbiota and Type 2 Diabetes: The Role of LPS and 16S rRNA in Altered Glucose Metabolism

By

Martha Konar
MPH

The Hubert Department of Global Health

_____

Professor Mohammed Ali
Committee Chair

Blood Microbiota and Type 2 Diabetes: The Role of LPS and 16S rRNA in Altered Glucose Metabolism

By

Martha Konar

M.A., Columbia University, 2006
B.S., Massachusetts Institute of Technology, 2003

Thesis Committee Chair: Professor Mohammed Ali, M.D., M.Sc., M.Sc., M.B.A.

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in The Hubert Department of Global Health
2021

Abstract

Blood Microbiota and Type 2 Diabetes: The Role of LPS
and 16S rRNA in Altered Glucose Metabolism
By Martha Konar

Evidence suggests bacterial translocation from the gastrointestinal tract into the systemic circulation may contribute to the inflammatory processes associated with the development of type 2 diabetes. To evaluate this relationship, a systematic literature review was conducted using observational studies published within the past 15 years that measured LPS or 16S rRNA levels in the blood of human adults and assessed the association between this exposure and an outcome of type 2 diabetes. A search of the published data identified 11 studies that met the established criteria from 1,224 records. Analysis of the data revealed that while many studies suggest a relationship may be present, aspects of study design and methodology limited the validity of the results and the relationship between blood microbiota and type 2 diabetes remains unclear. To determine whether bacteria may in fact have a role in this process, additional research will be needed that builds upon the lessons learned. In particular, combining large, longitudinal studies with random sampling, full reporting of and adjustment for confounding variables, newer detection methods using molecular-based techniques and type 2 diabetes diagnostic criteria that are testing-based are the key recommendations that will allow us to better ascertain the relevance of this unique exposure and its affect on glucose metabolism.

Blood Microbiota and Type 2 Diabetes: The Role of LPS and 16S rRNA in Altered Glucose Metabolism

By

Martha Konar

M.A., Columbia University, 2006
B.S., Massachusetts Institute of Technology, 2003

Thesis Committee Chair: Professor Mohammed Ali, M.D., M.Sc., M.Sc., M.B.A.

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in The Hubert Department of Global Health
2021

I would like to acknowledge my parents, Michael and Grace Konar, for their tireless support throughout this process and my sister, Michelle Konar, for her help in adapting some of the images presented. To Professor Ali, thank you for your expertise and guidance but most importantly, for always leaving my name on your whiteboard. I am also tremendously grateful to Theresa Nash, Dr. Juan Leon, Dr. Dabney Evans and Dr. Kimberly Jacob Arriola for the opportunity to continue this work. Your understanding has made a big difference in my life. Lastly, I would like to recognize Dr. Bobak Mozayeni for the conversation that inspired this idea. I would not have made it here without all of you.

Table of Contents

List of Tables

List of Figures

I. Introduction

A. Introduction and Rationale

Diabetes is a chronic condition with significant morbidity and mortality that, according to the World Health Organization (WHO, 2016) and International Diabetes Federation (IDF, 2019a), affects an estimated 422 million adults (aged over 18 years) to 463 million adults (aged 20 to 79 years) based on data from 2014 and 2019, respectively, equivalent to approximately 9.3% of adults worldwide (IDF, 2019a). This condition occurs when the body does not produce appropriate levels of insulin, a hormone needed for glucose metabolism, or when the body does produce insulin but is unable to use it effectively. The most common types of diabetes include type 1, type 2, and gestational diabetes mellitus.

In type 1 diabetes, an autoimmune reaction targets the insulin-producing β-cells in the pancreas (IDF, 2019a). Without insulin, glucose in the bloodstream cannot enter the body's cells to be used for energy and instead accumulates to high levels in the blood, also known as hyperglycemia. This occurs in type 2 diabetes as well but is due to insulin resistance, where the failure of the body cells to respond to the hormone leads to increasing levels of insulin in an attempt to get them to do so. Their failure to respond, however, leaves blood glucose levels elevated resulting in many of the same symptoms and complications.

These effects can range from so mild that individuals are unaware they have diabetes (American Diabetes Association [ADA], 2019c) to life-threatening (ADA, 2019a). In fact, an estimated 50.1% of adults with diabetes, or 231.9 million individuals, do not know they have this condition (IDF, 2019a). When symptoms are present, those with diabetes can experience increased urination, thirst, hunger, and fatigue (ADA, 2019c). Impaired wound healing and

blurry vision can also develop. Many of the symptoms of type 1 and type 2 diabetes overlap, but weight loss is associated with type 1 while numbness, tingling, or pain in the extremities is common in type 2. This latter set of symptoms typifies peripheral neuropathy, one of the forms of nerve damage possible with this condition and one of its many complications (ADA, 2019a). Other complications include kidney damage and eye problems such as cataracts, glaucoma, and retinopathy. An increased susceptibility to infections and poor circulation can lead to skin and foot manifestations like ulcers and amputations.

Often these symptoms and complications can be successfully managed, but some like diabetic ketoacidosis, which more commonly affects individuals with type 1 diabetes, can lead to coma and even death (ADA, 2019a). And because of its chronic nature, some of the features described above can progress to kidney failure, blindness, or life-threatening infections. Additionally, those with diabetes are at an increased risk for high blood pressure, stroke (ADA, 2019a), and cardiovascular disease (IDF, 2019a). Other co-morbidities, principal among them cancer, further put individuals with type 2 diabetes and high body mass index (BMI) at risk, with the likelihood of them developing certain cancers elevated two-fold over what they would otherwise be. For these reasons, type 1 and type 2 diabetes are serious conditions with profound long-term effects on the individuals who develop them.

In contrast, gestational diabetes affects women unable to produce enough insulin to overcome the insulin resistance triggered by hormones produced by the placenta, but this is typically transient and does not continue beyond pregnancy (IDF, 2019a). It can, however, lead to adverse outcomes as these mothers are more likely to have large babies for their gestational age making a normal birth more difficult. Furthermore, women who develop gestational diabetes are at a higher risk for developing type 2 diabetes, and the risk is higher for their babies as well.

With hyperglycemia affecting approximately 15.8% of live births, this represents an estimated 20.4 million pregnancies impacted by this condition in 2019 alone.

All told, the societal and economic impacts of diabetes are significant, with the WHO (2018) listing it seventh among the top 10 causes of death worldwide. This equates to approximately 1.6 million deaths directly caused by diabetes according to their most recent estimates from 2016. Yet, based on data from 2012, high blood glucose causes approximately 2.2 million additional deaths due to its role in increasing the risk of cardiovascular disease and other complications for a yearly total estimated burden of 3.7 million deaths (WHO, 2016). Very similar are the 2019 projections by the IDF (2019a) where they estimate 4.2 million deaths to result from diabetes and its complications.

Harder to estimate, however, is the toll this places on individuals and their families as one's health declines and healthcare expenditures take a larger share of earnings. One way to approximate this is by calculating disability-adjusted life years (DALYs) by taking into account years of life lost (YLL) from premature mortality and years lost due to disability (YLD) from the health consequences of living with this condition (WHO, 2019b). Using this measure, one DALY is equivalent to one lost year of healthy life. Based on data from the Global Burden of Diseases, Injuries, and Risk Factors Study 2017, a comprehensive assessment of all epidemiological data for 359 diseases and injuries in 73 age and sex groups for 195 countries and territories over 28 years, diabetes mellitus accounts for approximately 67.9 million all-age DALYs on an annual basis worldwide (GBD 2017 DALYs and HALE Collaborators, 2018).

As for the economic costs, the most recent estimate from a report by the WHO (2016) states that diabetes directly costs the world more than US$ 827 billion annually while a similar estimate by the IDF (2019a) puts the cost at US$ 760 billion, a figure that represents

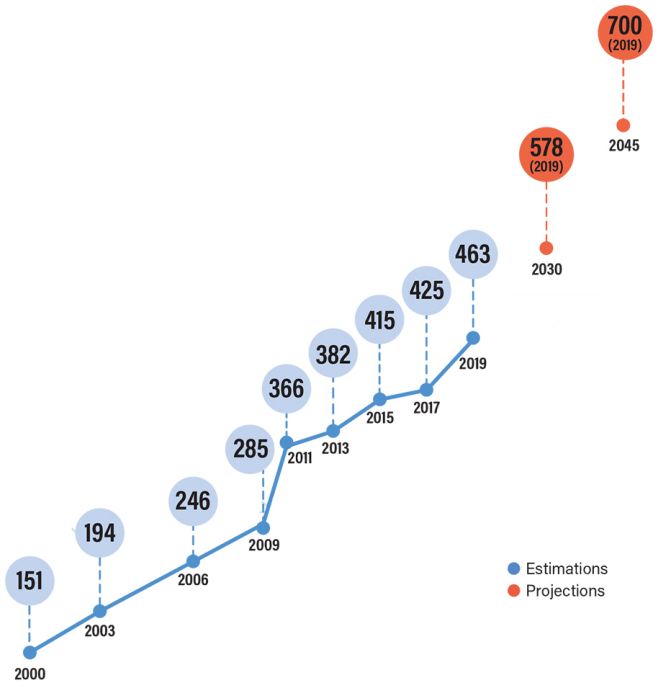
approximately 10% of global health expenditure. In addition to the direct costs are significant indirect costs, such as the fact that this condition primarily affects working age individuals with a subsequent loss in work and wages when health deteriorates. In the United States alone, an estimated US$ 90 billion is lost annually due to the indirect costs of diabetes with global estimates at US$ 455 billion in 2015. Taken together, this latter figure comprises 34.7% of the global total estimated cost of diabetes annually, which includes both direct and indirect costs—a staggering US$ 1.31 trillion in 2015.

Of the estimated 463 million adults affected by this condition, approximately 90% have type 2 diabetes (IDF, 2019a). This represents a tremendous opportunity for those in public health because these numbers are projected to increase to 578 million by 2030, yet many of these cases can be prevented (see Figure 1). This is no longer solely the burden of well-developed countries; currently, three in four people with diabetes, approximately 79%, live in middle- and low-income countries making diabetes a significant global health concern. While the prevalence of diabetes is highest in high-incomes countries at approximately 10.4% versus 9.5% and 4.0% in middle- and low-income countries, respectively, the burden can often be felt most acutely in poorer regions as these frequently have healthcare infrastructures unable to provide the same level of care.

Take for instance, the rates of those with diabetes who remain undiagnosed at approximately 66.8% in low-income countries compared to 52.6% and 38.3% in middle- and high-income countries, respectively (IDF, 2019a). Furthermore, for those with a diagnosis of type 2 diabetes, estimates suggest that one in two do not have access to insulin when prescribed, with this figure higher in low- and middle-income countries. For example, approximately 86% of those in Africa cannot access the insulin they need due to its limited availability and affordability. Sometimes these barriers to access result from a complete lack of availability of

Figure 1

Estimated and Projected Number of Adults (Aged 20-79 Years) with Diabetes (in Millions)



Note. Adapted from *Worldwide Toll of Diabetes*, by International Diabetes Federation, 2019b (https://www.diabetesatlas.org/en/sections/worldwide-toll-of-diabetes.html). Copyright 2020 by the International Diabetes Federation. Adapted with permission.

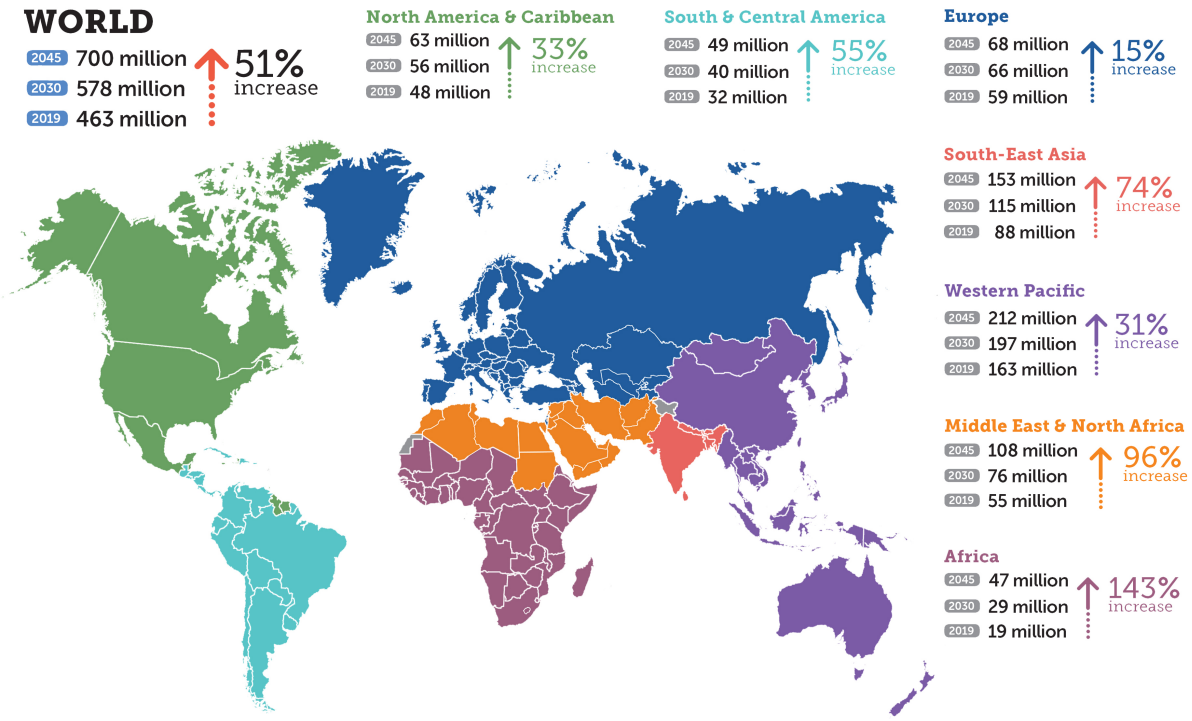
insulin in the health system as a whole or disruptions of the supply in certain areas or within levels of a health system. Other times, the high cost of medications prevents individuals and families from receiving appropriate care. In low-incomes countries, approximately 26.9% of households cannot afford the medication metformin versus 0.7% of households in high-income countries. When it comes to insulin, that figure is significantly higher with approximately 63% of households in low-income countries unable to afford this medication versus 2.8% of households in high-income countries.

These delays in diagnosis and the inadequate treatment individuals receive because they cannot afford to do otherwise are a significant cause for concern. Worldwide, no region remains unaffected by the rise in diabetes, and current projections suggest that many of the poorest regions will see the sharpest increases in number of cases (IDF, 2019a). So while on a global scale, diabetes cases will rise approximately 51% from 463 million in 2019 to 700 million by 2045, cases in South-East Asia, the Middle East and North Africa, and Africa will experience a 74%, 96%, and 143% increase in number of cases, respectively (see Figure 2). With annual global health expenditure on diabetes expected to rise alongside these trends from an estimated US$ 760 billion in 2019 to US$ 845 billion by 2045 (IDF, 2019a) and estimated losses in worldwide gross domestic product (GDP) from both the direct and indirect costs of diabetes at US$ 1.7 trillion just from 2011 to 2030 (WHO, 2016), it is undeniable the economic toll diabetes has globally. Add to this the individual cost of worsening health and lost years of life—all due to a condition that is largely preventable with the appropriate interventions, diagnosis, and treatment—and the true price of diabetes worldwide is immeasurable.

Key to prevention, therefore, is an understanding of who is at risk and why so that these measures, for type 2 diabetes in particular, can be targeted to those in the population most likely

Figure 2

Estimated and Projected Number of Adults (Aged 20-79 Years) with Diabetes Worldwide and by Region



Note. Adapted from *IDF Diabetes Atlas, 9th edition* (p. 4-5), by International Diabetes Federation, 2019a, International Diabetes Federation (https://www.diabetesatlas.org/upload/resources/2019/IDF_Atlas_9th_Edition_2019.pdf). Copyright 2019 by the International Diabetes Federation. Adapted with permission.

to benefit. According to the ADA (2019b), risk factors for developing prediabetes, a condition that often precedes type 2 diabetes, include being overweight; physically inactive; age 45 and older; or being Black, Hispanic/Latino, American Indian, Asian American, or Pacific Islander. Those with a parent or sibling with diabetes are at an increased risk as are those with high blood pressure, high triglycerides, or low high-density lipoprotein (HDL) cholesterol. Experiencing diabetes during a pregnancy or having a diagnosis of Polycystic Ovarian Syndrome are risk factors as well. While some of these traits are beyond an individual's control, many are positively influenced by maintaining a healthy lifestyle focused on improving nutrition (ADA, 2019e) and increasing fitness level (ADA, 2019d). Other recommendations include smoking cessation (ADA, 2019f), stress management (Hacket & Steptoe, 2017; Surwit et al., 2002), and sleep improvement (Walker, 2017).

In the progression towards type 2 diabetes, individuals transition from normal glucose regulation to an intermediate state of impaired glucose regulation before ultimately developing diabetes (WHO, 2019a). This intermediate state, described as prediabetes by the ADA (2020), is identified by the presence of impaired fasting glucose, impaired glucose tolerance, a hemoglobin A1C (HbA1C) value of 5.7-6.4% (39-47 mmol/mol), or a combination of these. To determine whether an individual has entered into one of these states, clinicians rely on four diagnostic tests. Random plasma glucose and fasting plasma glucose measure the level of glucose in plasma, but the former is performed at any particular point in time while the latter is after a period of no caloric intake for eight hours or more. Two-hour plasma glucose is similar but is measured after two hours during an oral glucose tolerance test where the individual first consumes a 75 g oral glucose load. In contrast, HbA1C represents blood glucose levels over the previous three months by measuring glycated hemoglobin, or glucose bound to hemoglobin, in red blood cells (National

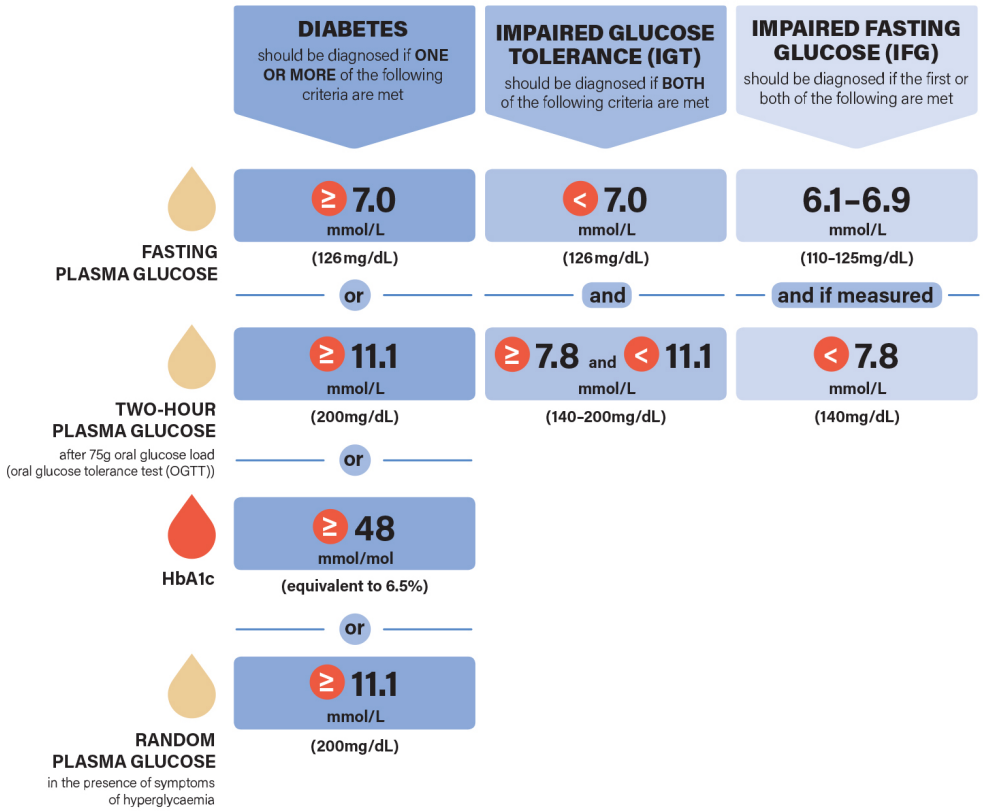
Institute of Diabetes and Digestive and Kidney Diseases [NIDDK], 2018). Higher levels of glucose in the blood over this period of time lead to higher levels of glycated hemoglobin. According to the WHO (2019a), IDF (2019a), and ADA (2020), the values of these various tests correspond to the state of glucose regulation, whether impaired fasting glucose, impaired glucose tolerance, or diabetes (see Figure 3). One exception, however, is the ADA (2020) defines impaired fasting glucose as having a fasting plasma glucose of 5.6-6.9 mmol/L (100-125 mg/dL). Another is their use of the term prediabetes as described above.

B. Problem Statement

Understanding why an individual with these risk factors develops type 2 diabetes is challenging as this condition results from a complex, multifactorial process involving both a genetic component and numerous environmental conditions (Tuomi et al., 2014). The distinction between type 1 and type 2 diabetes delineates a boundary that is, in actuality, sometimes unclear as both are polygenic in origin with over 60 associated genes identified in genome-wide association studies; furthermore, individuals have genetic predispositions to both. It is therefore likely that multiple overlapping mechanisms exist, and the interplay of this dynamic with environmental influences is not fully understood. There is growing consensus, however, that many of these pathways involve activation of systemic inflammation as more studies have demonstrated a link between inflammatory biomarkers and type 2 diabetes (Dandona et al., 2004; Lontchi-Yimagou et al., 2013; Pickup, 2004; Shoelson et al., 2006; Wellen & Hotamisligil, 2005). Interestingly, a major site for the production of these molecules appears to be adipose tissue (Lontchi-Yimagou et al., 2013), but even after adjusting for BMI, that association remains (Pickup, 2004).

Figure 3

Diagnostic Criteria for Diabetes, Impaired Glucose Tolerance and Impaired Fasting Glucose



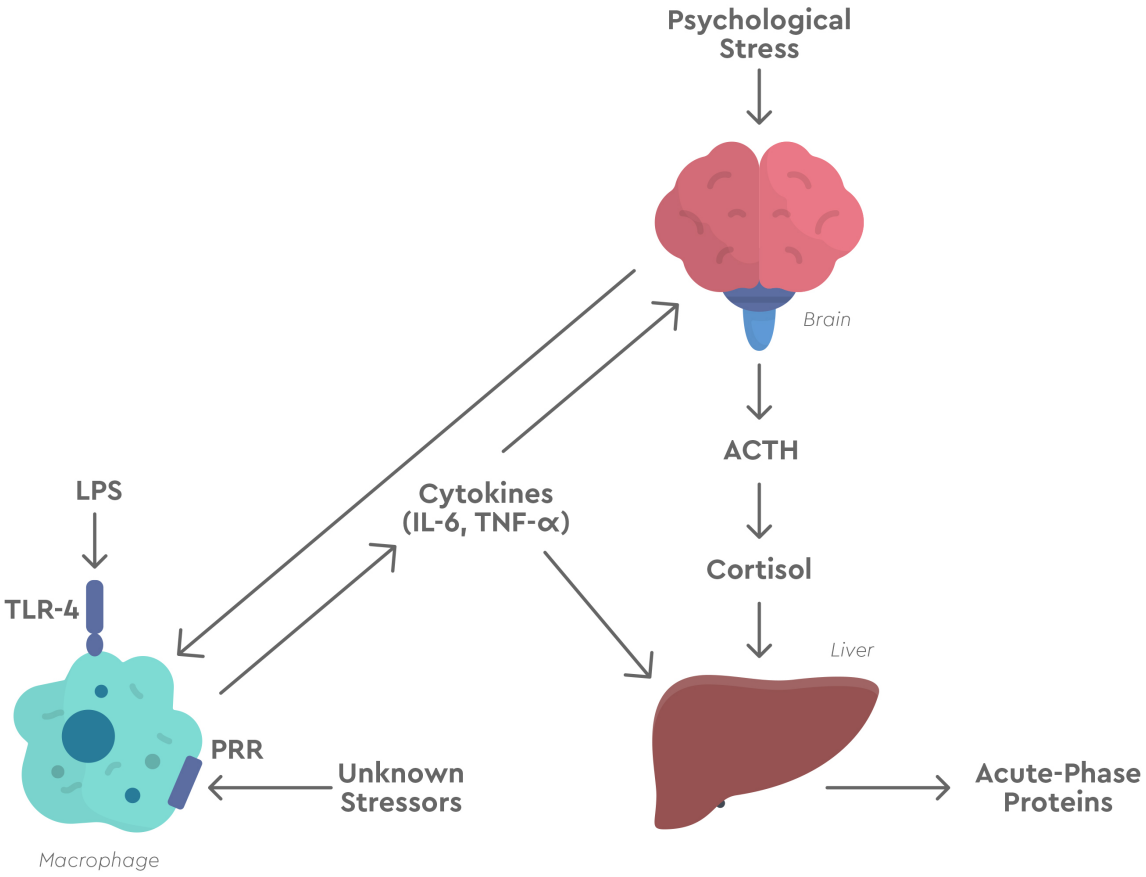
Note. The American Diabetes Association (2020) defines impaired fasting glucose as having a fasting plasma glucose of 5.6-6.9 mmol/L (100-125 mg/dL). Adapted from *IDF Diabetes Atlas, 9th edition* (p. 12), by International Diabetes Federation, 2019a, International Diabetes Federation (https://www.diabetesatlas.org/upload/resources/2019/IDF_Atlas_9th_Edition_2019.pdf). Copyright 2019 by the International Diabetes Federation. Adapted with permission.

A mechanism that has been suggested to explain this process involves interactions between the innate immune system, the acute-phase response, and the hypothalamic-pituitary-adrenal (HPA) axis (Pickup, 2004). The innate immune system defends against infection and injury via a set of reactions that are initiated to neutralize infectious agents, prevent tissue damage, and promote a return to homeostasis. Certain cells, both immune cells like macrophages but also other cells like adipocytes, contain pattern recognition receptors (PRRs) that recognize molecular patterns typical of certain pathogens (Pickup, 2004) as well as endogenous molecules from damaged cells (Takeuchi & Akira, 2010). Activating PRRs on or within a cell initiates signaling cascades that lead to the production of cytokines like interleukin-6 (IL-6) and tumor necrosis factor-α (TNF-α; Pickup, 2004). These inflammatory molecules trigger an acute-phase response and stimulate production of acute-phase proteins, mostly by the liver. These proteins have various roles but generally serve to prevent further injury and promote healing. Cytokine release also causes the brain to release corticotropin-releasing factor from the hypothalamus resulting in adrenocorticotropic hormone (ACTH) release from the pituitary, which culminates with cortisol release by the adrenal glands. This further stimulates production of acute-phase proteins by the liver. Lastly, the brain can also respond to psychological stress by triggering cytokine release, so it is capable of producing and mediating inflammation as well (see Figure 4).

Numerous authors refer to Hotamisligil et al. (1993) as the first study to demonstrate a relationship between an inflammatory marker, TNF-α, and insulin resistance in rodent models (Dandona et al., 2004; Lontchi-Yimagou et al., 2013; Wellen & Hotamisligil, 2005). Since then, many studies have shown an association between different inflammatory markers and type 2 diabetes in both animal models and humans (Dandona et al., 2004; Lontchi-Yimagou et al., 2013; Pickup, 2004; Shoelson et al., 2006; Wellen & Hotamisligil, 2005). Furthermore, clinical

Figure 4

Signaling Pathways of the Innate Immune System, Acute-Phase Response and Hypothalamic-Pituitary-Adrenal (HPA) Axis



Note. Adapted from "Inflammation and Activated Innate Immunity in the Pathogenesis of Type 2 Diabetes," by J. C. Pickup, 2004, *Diabetes Care*, *27*(3), p. 814 (https://doi.org/10.2337/diacare.27.3.813). American Diabetes Association Diabetes Care, American Diabetes Association, 2004. Copyright and all rights reserved. Material from this publication has been used with the permission of American Diabetes Association. Macrophage and brain icons made by Freepik from www.flaticon.com. Liver icon made by Nikita Golubev from www.flaticon.com. Macrophage and liver icons adapted by Michelle Konar.

studies have shown that different anti-inflammatory drugs have the ability to lower blood sugar levels (Donath, 2016), but treatment of the observed effect between inflammation and type 2 diabetes, as some are suggesting, does not get to the root cause of this association and has the downside of potentially suppressing essential immune pathways (Goldfine & Shoelson, 2017). A better long-term approach may be to focus instead on identifying the factors that trigger this inflammation so as to prevent its presence and the associated risk of type 2 diabetes. Given current trends, a better understanding of this mechanism is a public health imperative as this knowledge has the potential to save millions of lives as well as reduce the strain on health systems worldwide.

Various reasons for this relationship have been proposed. Excluding those for which individuals have no control, such as genetics and age, other possible factors include obesity (Dandona et al., 2004; Lontchi-Yimagou et al., 2013), nutrition (Dandona et al., 2004; Lontchi-Yimagou et al., 2013; Pickup, 2004), fitness level, smoking, and stress (Pickup, 2004). These areas of further exploration are in line with the risk factors and recommendations provided earlier by the ADA (2019b, 2019d, 2019e, 2019f) and other researchers (Hacket & Steptoe, 2017; Surwit et al., 2002). Additional areas of interest include the role of oxidative stress (Dandona et al., 2004; Wellen & Hotamisligil, 2005), gut microbiota, air pollutants, vitamin D deficiency, and epigenetic changes, which occur when differences in deoxyribonucleic acid (DNA) methylation patterns result in different levels of gene expression (Lontchi-Yimagou et al., 2013). Yet many of these suggestions fail to consider what the innate immune system is uniquely designed for, namely responding to microbial threats while the adaptive immune system prepares to respond.

This is a tremendous gap in our understanding of this public health problem as our

attitudes towards bacteria and other microorganisms have led us to see them as primarily acute exposures that result in certain symptoms or disease outcomes separate from those we classify as non-communicable. But bacteria are within us, on us, and exist throughout our environment (Yong, 2016). Estimates suggest there are roughly 39 trillion microbial cells interacting with the 30 trillion human cells that make up our bodies. Furthermore, these organisms persist from the bottom of ocean trenches all the way into the clouds. So a shift in perspective is necessary where, in addition to the traditional environmental exposures and lifestyle factors, we begin to consider bacteria and other microorganisms as chronic exposures as well and open up to the possibilities that presents for us. One area of interest is the increasing evidence that suggests these organisms can be found in the blood, a site typically thought to be sterile (Potgieter et al., 2015). They translocate, or move from one body site to another, without causing acute infection. The predominant source of this blood microbiota is the gut with studies supporting translocation from the oral cavity as well. So taken together, we have a highly significant exposure in a highly significant location, yet current thinking has largely overlooked this possibility and the number of studies examining it remain limited. This is a mistake and a missed opportunity.

As discussed earlier, various cells throughout the body contain PRRs that trigger the inflammatory pathways attributed to the development of type 2 diabetes (Pickup, 2004; Takeuchi & Akira, 2010). One group of PRRs are toll-like receptors (TLRs), which includes TLR-4. This surface receptor is known for binding lipopolysaccharide (LPS), a molecule specific to gram-negative bacteria, which then initiates the same processes that lead to increased levels of cytokines and acute-phase proteins (see Figure 4). Intracellular PRRs also exist that recognize viral components while others are secreted and circulate throughout the body, for example C-reactive protein (CRP), that flag microbial antigens for clearance by the immune system.

Therefore, a microbial component could in fact be the trigger that initiates this cascade of events. If so, this presents unique opportunities in terms of prevention, diagnosis, and treatment.

Rather than watching the number of diabetes cases continue to rise with strategies that have thus far been ineffective to curb this trend, we could instead explore an entirely new avenue of research and if this relationship does in fact exist, use it to develop novel diagnostic techniques that better predict who is at risk of developing type 2 diabetes. This would allow us to target therapies where they would have the greatest impact as well as begin to consider whether other treatment modalities typically reserved for infectious processes, such as antibiotics and vaccines, could have beneficial effects. Additionally, if we better understand the mechanisms underlying this association, we could identify relevant upstream factors, perhaps gut microbiota balance and intestinal permeability, that play into this process and target new diagnostic and treatment strategies towards these factors as well. Ultimately, information is one of our greatest tools as public health professionals in affecting change, and understanding the relationship between microorganisms and the development of type 2 diabetes places us one step closer towards finding real solutions.

C. Purpose Statement

In order to determine what is currently known about this relationship and to begin to piece together the possible role of bacteria in the development of type 2 diabetes, this systematic literature review examined current research focusing on the presence of microbiota in human blood and its association with this condition. Studies that detected LPS, also known as endotoxin, were taken into consideration as were studies that detected 16S ribosomal ribonucleic acid (16S rRNA), a method of identifying bacteria using molecular techniques. Data from the

selected studies were then evaluated to determine the strength of this association as well as the conclusions that could be drawn from this information.

Prior to this, one systematic literature review was published examining studies from 1984 to 2014 that looked at the presence of LPS in human blood and its role in numerous health conditions, including type 2 diabetes (Gnauck et al., 2016). However, as this gained more attention in recent years and molecular techniques became both increasingly sophisticated and more commonly utilized in research, this review ceased to contain the most current or comprehensive information on the topic. Therefore, an updated look into the subject that takes into account newer publications with stronger study designs and a shifting focus towards molecular methods was required to truly capture what is known about this relationship.

D. Research Questions

- Question 1: Is the presence of LPS in human blood associated with the presence of or development of type 2 diabetes?
- Question 2: Is the presence of 16S rRNA in human blood associated with the presence of or development of type 2 diabetes?

E. Significance Statement

It is easy to lose sight of the significance of something that cannot be seen. We have, since the very beginning, always lived in a microbial context, evolving alongside microbes yet unaware of their presence for most of human history (Yong, 2016). With their discovery, humanity could, for the first time, understand and address some of the most feared scourges of mankind with the unintended consequence of unfortunately locking us into the mindset of

microbe as pathogen.

What we often fail to recognize is the vastness of the domain of bacteria. They may be small and unseen to us but taken together, potentially the greatest form of life in biomass (Johnson, 2006). For humans, less than 100 species cause infectious disease while an untold number make up the remaining bacterial diversity of the biosphere (Yong, 2016). It is only more recently that scientists have begun to see these other microbes as beneficial and for the technology necessary to really study them and their relationships with us to become available.

Therefore, the research examining the link between blood microbiota and the development of type 2 diabetes marks an important shift. We have for the longest time either ignored bacteria or tried to get rid of them. Without an overt cause of type 2 diabetes tied to an infectious process, we assumed bacteria had no role to play and maintained this perspective for years. But our relationship with bacteria is more complex than choosing between pathogen or insignificant, and current data suggests a more nuanced view. It is possible for bacteria to have a myriad of different roles and for their presence in our bodies to exert an influence we have yet to fully understand.

At its very simplest, this idea helps to position us to identify individuals at risk of developing type 2 diabetes far earlier than would otherwise be possible should the presence of LPS or 16S rRNA prove useful as reliable biomarkers. From this, one could then explore other bacterial components or markers of bacterial presence to discover which are the strongest indicators and find interventions that work long before altered glucose metabolism becomes apparent. These interventions may be aimed at the bacteria themselves by adapting treatments already in use or towards various upstream factors that likely play a role.

This possibility has been overlooked in favor of focusing on lifestyle factors such as diet

and exercise. But why do these changes work as well as they do? Diet affects gut microbiota balance and intestinal permeability (Gentile & Weir, 2018; Tremaroli & Bäckhed, 2012), which appears to be the main source of bacteria present in the blood (Potgieter et al., 2015). Once present, it triggers a systemic inflammatory response amplified by adipose tissue (Creely et al., 2007), which decreases as individuals exercise more and improve their diet. But remove bacteria from the blood and you stop the inflammatory response from being amplified by adipose tissue. Alter one's gut microbiota balance and intestinal permeability and you prevent them from entering into the circulation in the first place. In all of this, bacteria play a key role. Seeing them as such opens up tremendous possibility in terms of prevention and further research.

This is why beyond the more immediate gains described above, the most significant may be an idea, some of which are the hardest to see when we think we already know. But to take the stance that our understanding of any given phenomenon, even its core principles, is complete because of where we stand currently and how far we have come to gather and test that information would be in error as new details, and sometimes complete revisions of previously held knowledge, occur as science progresses. For this is not a static process, but one of continued exploration.

Had it been assumed that there was no worth in looking at a drop of water through a microscope because our understanding of single-celled organisms did not yet exist, we would not have seen them (Yong, 2016). Nor would we have challenged the prevailing theory of "miasma," or bad air, to uncover the true cause of cholera in 1854 (Johnson, 2006) leaving John Snow thereafter regarded as the father of epidemiology (Centers for Disease Control and Prevention [CDC], 2017). Other conditions considered to have no infectious cause have turned out to have one after all, such as the relationships between human papilloma virus and cervical cancer

(Skloot, 2010), Epstein Barr virus and Burkitt's lymphoma (Shannon-Lowe & Rickinson, 2019), as well as *Helicobacter pylori* and gastric ulcers (Graham, 2014).

That is not to say that all scientific exploration should be pursued, but there is evidence to suggest bacteria can be present in the blood, a site where we did not previously believe them to be (Potgieter et al., 2015), and this contributes to a systemic inflammatory response that may play a role in the development of type 2 diabetes (Creely et al., 2007; Dandona et al., 2004; Lontchi-Yimagou et al., 2013; Pickup, 2004; Shoelson et al., 2006; Wellen & Hotamisligil, 2005). Thus, in summary, not only could this aid in prevention by identifying new biomarkers and new intervention strategies but perhaps the greatest benefit may in fact be how the concept of bacteria as a chronic exposure changes our thinking about and our approach to the treatment of diabetes in ways yet unseen.

F. Definition of Terms

Adult = 18 years and older

Significant = p value < 0.05

II. Literature Review

What information is currently known about this relationship was gathered primarily through searches of Google Scholar and PubMed although no systematic process was used. Topics are organized to first provide an introduction to the microbiome then a closer look at the relationship between diet and gut microbiota followed by the relationship between gut microbiota and type 2 diabetes. Next is an examination of LPS, proposed mechanisms of entry into the systemic circulation, and methods of detection with a similar look at 16S rRNA. Many of these interactions are summarized in the directed acyclic graph shown in Figure 5.

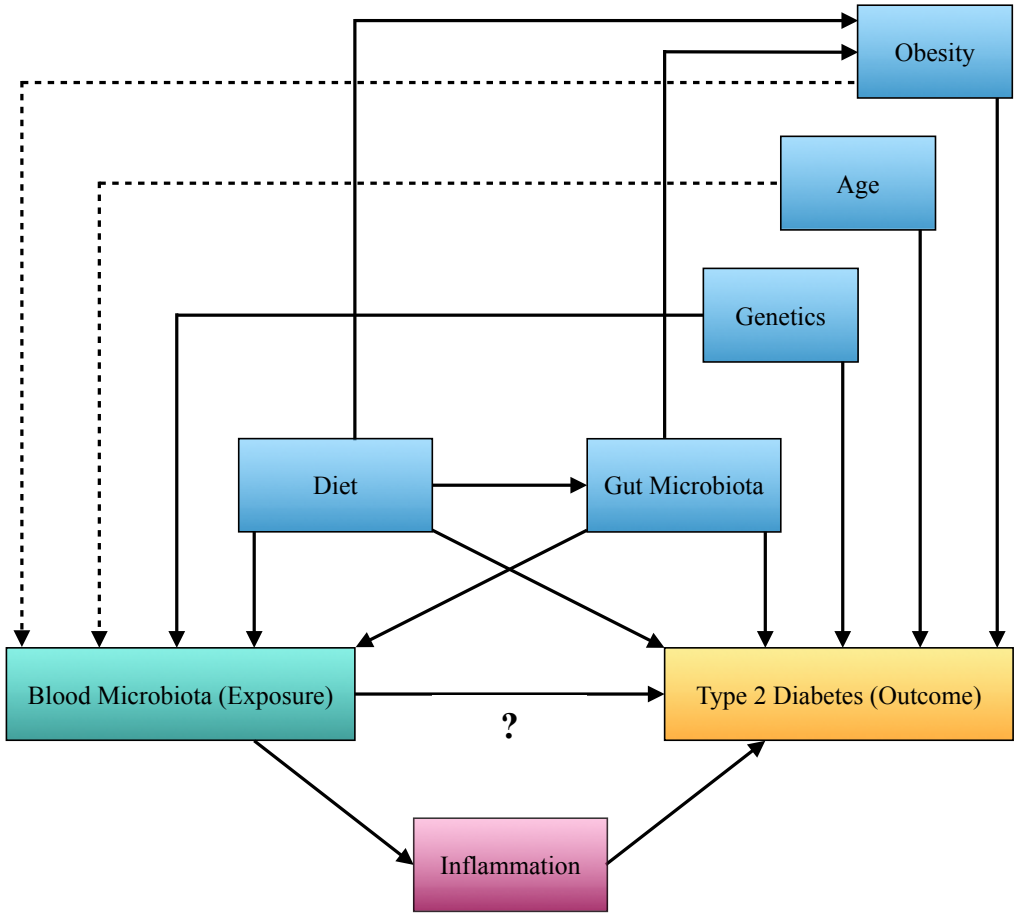
A. Exploring Host-Microbiome Interactions

We have entered into a new age of scientific inquiry with the development of molecular-based techniques, allowing us to first map our own human genome to subsequently mapping the microbial communities throughout our bodies (Proctor et al., 2019). One of the key efforts in this has been the National Institutes of Health (NIH) Human Microbiome Project (HMP), initiated in 2007 and conducted in two phases. The initial phase, or HMP1, sought to characterize microbial communities in the nose, mouth, gut, vagina, and skin of healthy adults with a subset of demonstration projects examining certain disease states. This research identified that on their own, the composition of these communities, often referred to as the microbiome or microbiota, do not consistently correspond with a state of health versus disease; this is likely because there is a great deal of variability in microbiota, even among healthy individuals (Huttenhower et al., 2012). Rather, what did appear consistent were the sets of metabolic pathways that these groups of organisms represent (see Figure 6).

The following phase, or HMP2, built upon these findings by exploring host-microbiome

Figure 5

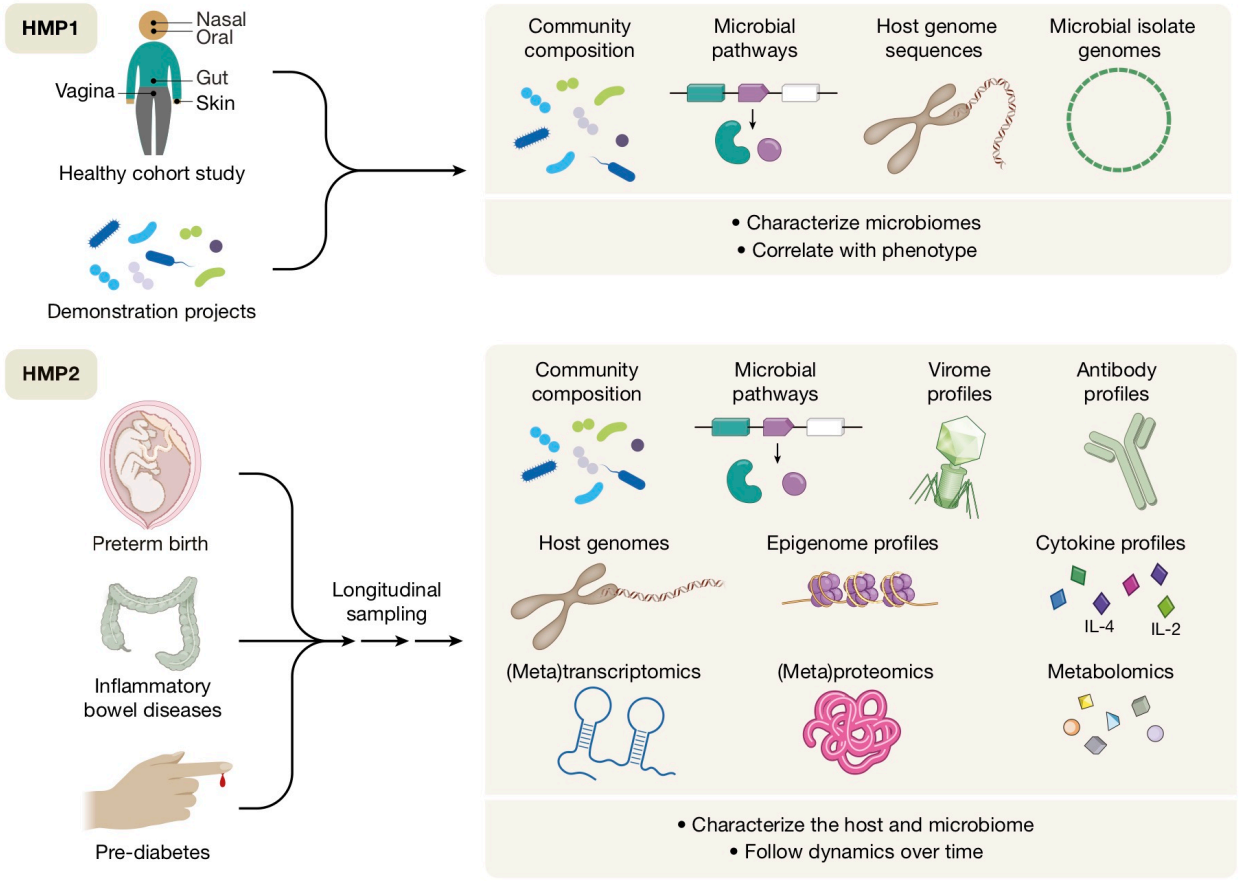
Directed Acyclic Graph of the Relationships Between Blood Microbiota, Type 2 Diabetes and Associated Variables



Note. Green = exposure, yellow = outcome, blue = confounder or possible confounder (indicated by presence of a dashed arrow), pink = mediator.

Figure 6

The First Phase (HMP1) and Second Phase (HMP2) of the National Institutes of Health Human Microbiome Project



Note. Adapted from "The Integrative Human Microbiome Project," by L. M. Proctor, H. H. Creasy, J. M. Fettweis, J. Lloyd-Price, A. Mahurkar, W. Zhou, G. A. Buck, M. P. Synder, J. F. Strauss, G. M. Weinstock, O. White, and C. Huttenhower, 2019, *Nature*, *569*, p. 642 (https://doi.org/10.1038/s41586-019-1238-8). CC BY 4.0.

interactions as it became clear that the bacteria present from a taxonomic standpoint is not as relevant to health outcomes as what those bacteria are doing and how the host responds (Proctor et al., 2019). Therefore, developing a functional understanding of the microbiome commenced including projects related to preterm birth, inflammatory bowel diseases, and prediabetes as these have known microbiome associations. Looking at dynamic molecular activity, these studies considered not only the microbial communities but also viruses present in both host and microbiome, known as the virome, and how all of these influence gene expression through epigenetic changes that affect RNA transcription, protein levels, and metabolism, also referred to as metatranscriptomics, metaproteomics, and metabolomics, respectively. Furthermore, host immunity was measured with antibody and cytokine profiles, and all of these multi-omic analyses occurred at multiple points to examine host-microbiome interactions over time (see Figure 6).

B. Relationship Between Diet and Gut Microbiota

As research into our microbiota began to identify what a healthy microbiome looks like and how it functions, we also became better able to identify the features of an imbalance or dysbiosis (Petersen & Round, 2014). These variations in our commensal communities can involve a decrease in beneficial organisms, an increase in potentially harmful organisms, reduction of overall diversity, or a combination of these. Interestingly, higher levels of gut microbial diversity exist in hunter-gatherers and those with rural agrarian diets compared to Westerners (Sonnenburg et al., 2016) with the highest levels seen in previously uncontacted Yanomami Amerindians (Clemente et al., 2015). One proposed explanation is the shift towards a Western diet high in fat and simple sugars and low in dietary fiber that contains an abundance of

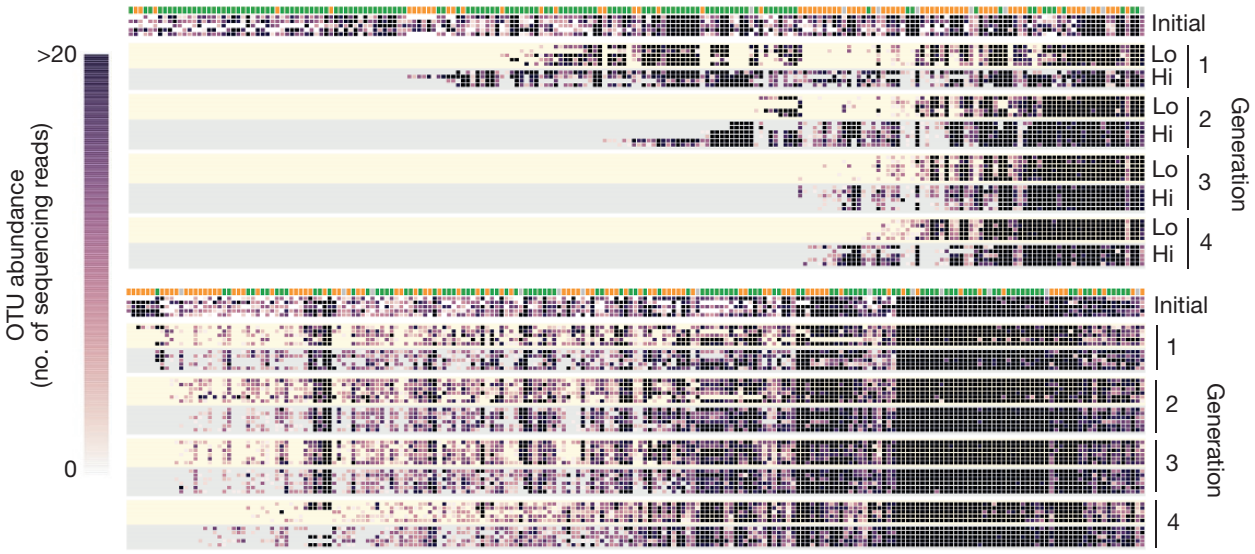
microbiota-accessible carbohydrates (MACs), the primary source of energy for gut microbiota (Sonnenburg et al., 2016).

In a study in mice with a human microbiota, switching from a high-MAC diet to a low-MAC diet resulted in a loss of microbial diversity, only some of which was recoverable after switching back to a high-MAC diet. Subsequent generations inherited this loss, which was further amplified when placed on a low-MAC diet themselves. By generation four, most taxonomic groups did not return even after switching back to a high-MAC diet with an approximately 68% overall loss of high-confidence operational taxonomic units (OTUs). These groups of sequences represent a group of organisms due to the high degree of genetic similarity they share with a reference sequence (Nguyen et al., 2016) and were considered high-confidence in this study as they were chosen for their prevalence and abundance to ensure reliable and consistent detection (see Figure 7). It was shown that, in addition to a loss of organisms, these mice also experienced a loss of function whenever all the microbiota with a particular enzyme were lost and not recovered. Therefore, this process results not just in changes to microbial diversity but also in how the microbiota works and what it can do..

Diets with reduced levels of MACs may not be the sole factor driving this shift in the gut microbiome since when we consume less of one substance, we invariably consume more of another (Sonnenburg & Bäckhed, 2016). It may be the increase in fat content of diets that plays a role, and although the dynamics are not fully understood, the connection to gut microbiota and obesity is clear. Firstly, germ-free mice that have no gut microbiota were shown to be resistant to weight gain and glucose intolerance on a Western diet until given gut microbiota from obese mice, which transferred the obese phenotype to them as well. Furthermore, gut microbiota from obese humans had the same effect when transferred to mice (Ridaura et al., 2013); thus, it

Figure 7

Operational Taxonomic Unit (OTU) Abundance in Generations of Mice Switching Between Diets Low and High in Microbiota-Accessible Carbohydrates (MACs)



Note. Taxonomic category indicated above each column by color with green for Bacteroidetes, orange for Firmicutes and gray for other. Diet-switching group above and controls below with microbiota of each mouse represented by each row.

appears these commensal organisms may be a necessary link between diet and its effects on metabolism. Secondly, germ-free mice became less resistant to obesity when on a high-fat diet with lower levels of sucrose than that described above, and this protective effect was lost when sucrose was not present at all, indicating that increasing the proportion of fat may be the reason for the observed effect (Sonnenburg & Bäckhed, 2016). Lastly, different types of fat have different influences with unsaturated fat resulting in a gut microbiota that prevents weight gain from saturated fat.

Others have shown consumption of simple sugars including glucose, sucrose, and fructose alters gut microbiota composition in mice (Khan et al., 2020). Alpha diversity is a measure of how varied and abundant organisms in a community are while beta diversity is a measure of how divergent the composition of one community is from another (Wagner et al., 2018). Both can be affected by a diet containing simple sugars as it was shown that glucose decreased species richness, a change in alpha diversity, whereas sucrose and fructose did not (Khan et al., 2020). However, all three sugars caused a change in beta diversity with a significant shift in microbial populations from that of controls. Similarly to that described above, the downstream effects of these changes were transferred to germ-free mice after colonization with microbiota from glucose-fed mice as well as prevented by antibiotic administration, indicating once again the role the gut microbiota likely play in modulating the effects of diet.

It is challenging, however, to say exactly what dietary factors are responsible, especially in humans. Our relationship with diet is complex and difficult to study as dietary interventions can be hard to maintain and self-reporting may be inaccurate (Sonnenburg & Bäckhed, 2016). Furthermore, different types of fiber, fat, and sugar as well as different proportions in every meal make it hard to discern which variations illicit which effects on the gut microbiota. Studies in

mice have allowed us to see what happens in more controlled environments, and this is beginning to make the relationships involved more clear.

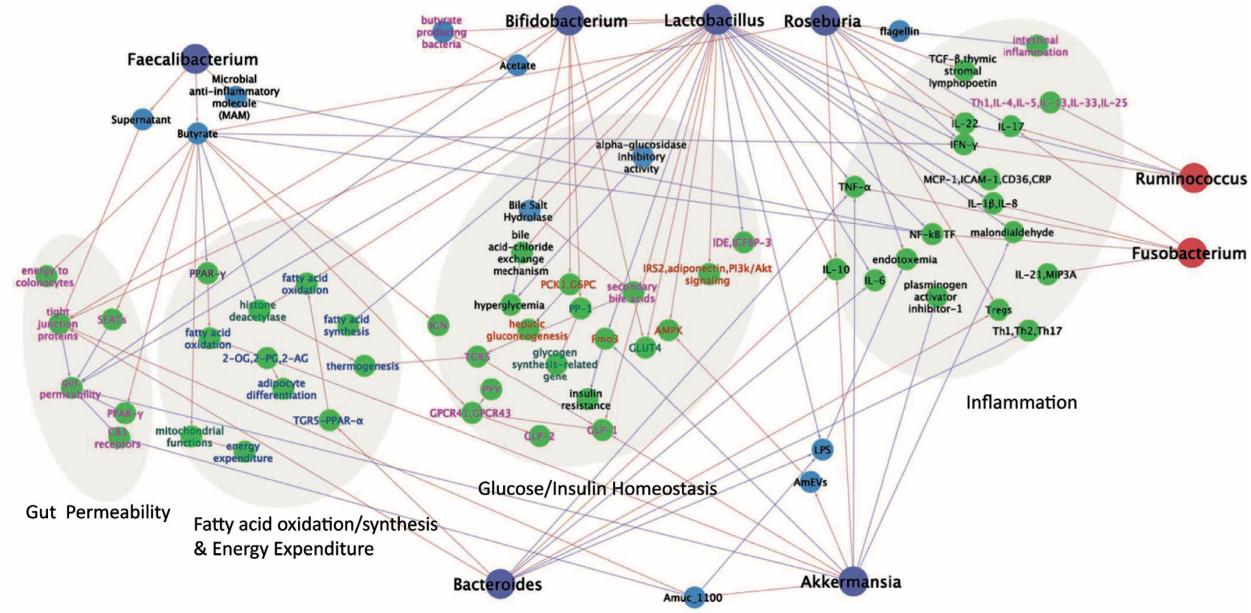
C. Relationship Between Gut Microbiota and Type 2 Diabetes

Continued exploration into the connections between the gut microbiota and obesity and type 2 diabetes will enable us to say with greater certainty which features contribute to health and which to disease as diet alone does not determine outcomes. This has been demonstrated in genetically identical mice housed together and fed the same high-fat diet (Burcelin et al., 2002). While approximately 47% developed obesity and diabetes, others developed only one condition or none at all. The past few years have seen a rapid increase in interest and research in an attempt to answer why. Numerous studies have examined various mechanisms to explain the role of gut microbiota in metabolic disease with data supporting its influence in glucose metabolism and insulin sensitivity, lipid metabolism and energy expenditure, gut permeability, as well as inflammation (Gurung et al., 2020). Additionally, from a taxonomic perspective, common findings included *Bifidobacterium*, *Bacteroides*, *Faecalibacterium*, *Akkermansia*, and *Roseburia* genera to be negatively associated with the development of type 2 diabetes and *Ruminococcus*, *Fusobacterium*, and *Blautia* genera to be positively associated with the development of type 2 diabetes (see Figure 8). The *Lactobacillus* genus had varied results (Gurung et al., 2020), which may be due to the fact that different species within genera have different roles within the gut microbiota and therefore different effects (Moreno-Indias et al., 2014).

Interestingly, while broader measures of gut microbiota health such as alpha and beta diversity have demonstrated positive associations with type 2 diabetes, other studies have found no association (Gurung et al., 2020). This was true as well for the ratio of the phyla Bacteroides

Figure 8

Association Map Between Gut Microbiota Genera, Microbial Products and Host Response



Color of labels in the node represents different organs where microbiota potentially elicit its effect: adipose tissue, gut, liver, liver&adipose tissue, muscle, systemic
Color of nodes: host features/products; microbial products; microbe positively associated with disease, microbe negatively associated with disease
Color of edge: blue, negative association between nodes, red, positive association between nodes

Note. From "Role of Gut Microbiota in Type 2 Diabetes Pathophysiology," by M. Gurung, Z. Li, H. You, R. Rodrigues, D. B. Jump, A. Morgun, and N. Shulzhenko, 2020, *EBioMedicine*, *51*, p. 4 (https://doi.org/10.1016/j.ebiom.2019.11.051). CC BY-NC-ND.

to Firmicutes, which has also been explored as a possible marker. These two phyla make up over 90% of the gut microbiota and display plasticity in that they quickly return to their original composition after experiencing a change (Magne et al., 2020). Modern stressors, however, exert effects that could lead to chronic modifications of relevance to obesity and type 2 diabetes. Thus, it may be more so that great variability exists in study methods or participant characteristics that leads to these contradictory results (Magne et al., 2020) in addition to the fact, as mentioned previously, that different members of the same taxonomic category may have different roles (Moreno-Indias et al., 2014). Furthermore, the relationships between gut microbiota, obesity, and type 2 diabetes are complex. We may as yet not fully know which factors are important to control for and every person is different. Interventions that work in some individuals have no effect in others and possibly depend on existing microbial diversity and functional capacity (Sonnenburg & Bäckhed, 2016). What is becoming apparent is the need for further research to clarify the mechanisms involved and to identify intervention points as the gut microbiota is a potentially powerful factor not only in the development of obesity and type 2 diabetes but also in its reversal.

D. LPS, Intestinal Absorption, and the Limulus Amoebocyte Lysate (LAL) Assay

Although diet influences the development of type 2 diabetes, perhaps through its effect on gut microbiota, additional research suggests there could be more at play to explain why. A shift towards a Western diet, with its abundance of fats and carbohydrates, often results in higher calories and nutritional density from the foods consumed. This results in weight gain and as the body repeatedly contends with elevated glucose levels after a meal, insulin resistance develops when the body's cells become less responsive to the elevated insulin levels the pancreatic β-cells

produce to maintain homeostasis. Ultimately reaching a state of fatigue, they no longer secrete adequate amounts of insulin and type 2 diabetes develops. However, variations in diet also affect absorption of intestinal contents as well as gut permeability, an interesting thing to note as the gut has an abundance of bacteria that can and do enter into the systemic circulation along with food. LPS, found in the outer cell membrane of gram-negative bacteria, contains an inner lipid A portion responsible for its toxic activity as well as a core polysaccharide and an antigenic O-specific polysaccharide (Mahon et al., 2015). The gut is known for being the site in the body where numerous members of the family Enterobacteriaceae reside, all gram negative and all harboring LPS. Especially after a high-fat meal, studies have shown LPS entering into the systemic circulation likely via the absorption of fatty acids following their incorporation into chylomicrons (Gnauck et al., 2016). This allows for transcellular passage into the lymphatic system thereby bypassing the hepatic-portal vein and clearance by the liver. Normally, bacteria and bacterial components that cross into the blood are removed in this manner, but absorption along with fats allows for entry into the systemic circulation through the lymphatic system. The extent to which this process contributes to the development of type 2 diabetes is unknown, but it may be one pathway that induces inflammation and puts an individual at increased risk.

This rise in LPS after a high-fat meal has been demonstrated in mice; furthermore, maintaining such a diet not only caused LPS levels to remain elevated throughout the day, it also shifted gut microbiota towards a higher ratio of LPS-containing bacteria in the first place (Cani et al., 2007). The role of the gut microbiota in this process is significant since treatment with antibiotics was shown to reduce LPS levels in the blood of mice on a high-fat diet similar to that seen in controls as well as improve numerous parameters related to obesity and diabetes such as weight gain and plasma glucose levels (Cani et al., 2008). In humans, similar effects have been

observed after consuming equicaloric drinks of water, glucose, and cream; of these, only cream was associated with elevated levels of LPS in the blood (Deopurkar et al., 2010). Consuming a mixed meal, however, reduces this effect as those who ate an equicaloric meal high in fruit and fiber with fat making up only 27% of calories experienced no rise in LPS whereas those who ate a high-fat, high-carbohydrate meal with fat accounting for 42% of calories did experience a significant rise (Ghanim et al., 2009). Determining the role of LPS, especially after a meal when it may be elevated, will be important since mice infused with LPS for 4 weeks demonstrated changes in glucose and insulin homeostasis as well as weight gain (Cani et al., 2007), which indicates a similar process may be occurring in humans.

To detect LPS in the blood, researchers have utilized the LAL assay, which uses a protein found in the blood of the North American horseshoe crab, or *Limulus polyphemus*, to initiate a clotting cascade (Gnauck et al., 2016). This was discovered when it was noted that *L. polyphemus* blood clots upon exposure to gram-negative bacteria. Thus, a lysate of amoebocytes, a cellular component of that blood, began to be used to identify the presence of LPS. Originally, the gel-clot LAL assay relied on the formation of a rigid gel, but the improved ability to isolate and purify the enzymes and proteins involved in the reaction led to the creation of turbidimetric and chromogenic LAL assays. The turbidimetric LAL assay measures the turbidity within the sample when exposed to LPS, and the chromogenic LAL assay measures a synthetic chromogenic substrate that replaces one of the proteins in the clotting cascade and produces a yellow color when activated. Both were approved by the Food and Drug Administration (FDA) to quantify LPS in pharmaceuticals and other medical products and have since been adapted by researchers for use in animal and human studies, although the assays were not developed nor approved for this purpose.

E. 16S rRNA, Bacterial Translocation, and Molecular-Based Techniques

With newer molecular-based techniques, researchers began to move away from a reliance on LPS to gauge bacterial presence in the blood and instead turned towards identifying bacterial 16S rRNA to provide that information. One of the advantages to this approach is this molecule, a subunit of ribosomes, is found in all bacteria and not just those that are gram negative (Matsuda et al., 2007). Additionally, multiple copies within each cell make its detection more sensitive than assays relying on detection of genomic sequences. But how do these bacteria enter the systemic circulation? Intestinal permeability depends upon the regulation of tight junctions between epithelial cells, which contain certain transmembrane and peripheral membrane proteins including claudins, occludin, and zonula occludens-1 (ZO-1; Turner, 2009). Studies in mice fed a high-fat diet have shown a significant reduction in the expression of ZO-1 and increased intestinal permeability (Cani et al., 2008). Similar to that described above, this effect resolved with antibiotic treatment suggesting that gut microbiota are involved in the regulation of this process. However, this may not be sufficient to explain the presence of 16S rRNA in the blood, an indicator of intact bacteria.

While some have suggested altered expression and distribution of tight junction proteins as a mechanism driving gut permeability in obese mice (Everard & Cani, 2013), other research supports a different hypothesis. Using fluorescently-labeled *Escherichia coli* (GFP-*E. coli*) administered to mice on a high-fat diet for one week, it was shown that these organisms adhered to gut mucosal surfaces at increased rates compared to controls as well as translocated into the blood and mesenteric adipose tissue (Amar et al., 2011a). This adherence continued into the fourth week on a high-fat diet at which point the mice had developed diabetes and interestingly, co-localization was observed between GFP-*E. coli* and dendritic cells. These and other

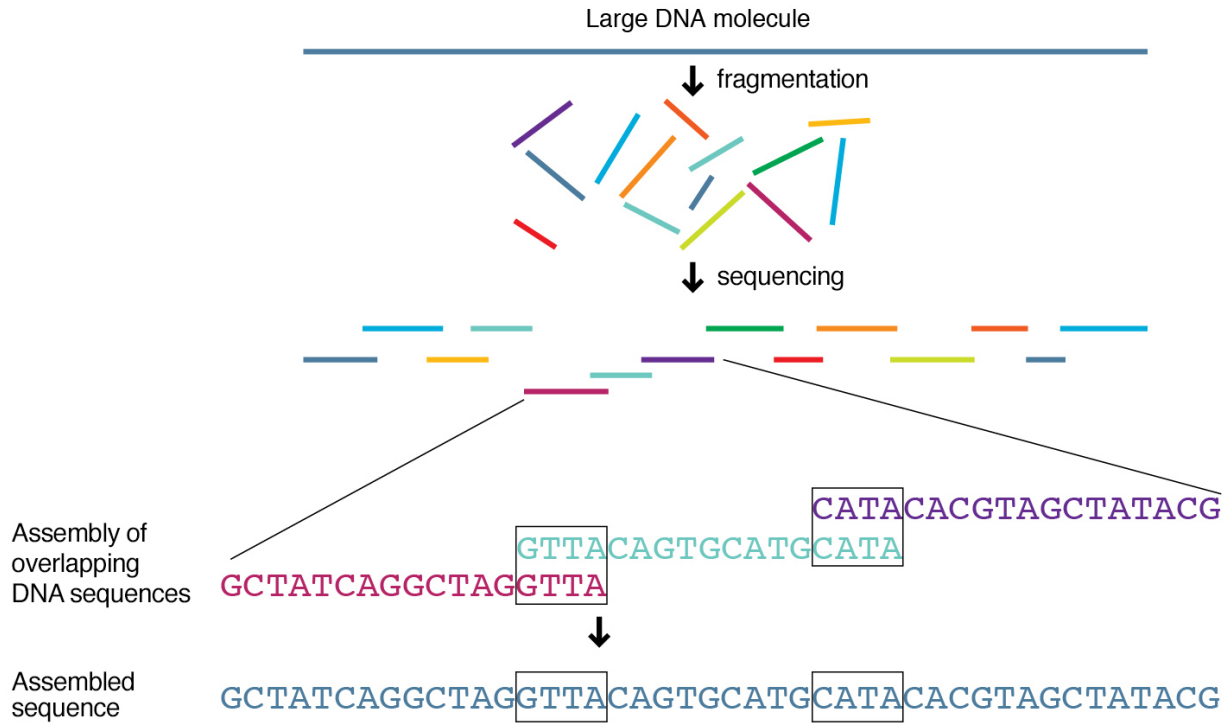
phagocytic cells such as macrophages ingest bacteria, and because it was previously shown that macrophage infiltration into adipose tissue increases with increasing BMI in humans (Weisberg et al., 2003), the study also looked for the presence of GFP-*E. coli* in mesenteric lymph nodes (Amar et al., 2011a). It was found that continued duration on a high-fat diet led not only to increased numbers of GFP-*E. coli* detected in blood and mesenteric adipose tissue but to their presence within mesenteric lymph nodes after four weeks as well. Therefore, a connection may exist between bacteria and innate immune cells in the gut that is driven by diet and leads to changes in adipose tissue.

These advances in our understanding of this process would not have be possible without developments in molecular-based techniques, which have led us to the point where we can not only sequence the full genome of an organism, but can also rapidly sequence the genomes of an entire community of microbes. This has led to the creation of databases where researchers can compare sequences identified in samples with reference sequences as well as to the growth of bioinformatics to analyze the vast quantities of information generated from multi-omic analyses that are now exploring metagenomic, metatranscriptomic, metaproteomic, and metabolomic relationships between these organisms with health and disease. In the past, it could take years and hundreds of thousands of dollars to undertake sequencing projects, but this has changed as we moved through the big revolutions that came with shotgun sequencing, next-generation sequencing, and single-molecule long-read sequencing (Loman & Pallen, 2015).

In shotgun sequencing, genomic material is broken into small fragments, which are sequenced individually (National Human Genome Research Institute [NHGRI], n.d.). Software then finds where these sequences overlap and uses these data to place fragments into the correct order thereby recreating the full genomic sequence (see Figure 9). Next-generation sequencing

Figure 9

Steps Involved in Shotgun Sequencing



Note. Adapted from *Shotgun Sequencing*, by National Human Genome Research Institute, n.d. (https://www.genome.gov/genetics-glossary/Shotgun-Sequencing). In the public domain.

allows for higher throughput because numerous small fragments can be sequenced simultaneously but sacrifices read length and therefore, the ability to reconstruct entire genomes as accurate assembly is not possible across long repeats (Loman & Pallen, 2015). This led to the development of single-molecule long-read sequencing that allows for sequencing of longer fragments (Goodwin et al., 2016). Using this technique, genomic material either moves along a bound polymerase that separates fluorophores from dNTPs as they are incorporated into a matching DNA strand with the color of light emitted used to generate sequences or moves through a nanopore with current passing through to generate shifts in voltage that can be interpreted.

In preparation for sequencing, nucleic acids may need to be isolated and amplified using either polymerase chain reaction (PCR) in the case of DNA or reverse transcription PCR (RT-PCR) in the case of RNA (Buckingham, 2012). A related technique known as quantitative PCR (qPCR), or real-time PCR, allows for measurement of the PCR product during the reaction and enables quantification of the starting material (Buckingham, 2012), which can complement or compete with next-generation sequencing (Goodwin et al., 2016). To amplify microbial content in a sample, these methods require the use of primers specific to either bacteria in general or to individual species of interest that will bind to the template (Amar et al., 2011b; Sato et al., 2014). This makes 16S rRNA particularly well suited as it contains conserved sequences that remain the same in all bacteria interspersed with hypervariable regions that are more unique and can be used to identify a sequence as belonging to a single genus or species (Jenkins et al., 2012). This allows one to use primers targeting the conserved regions to amplify nucleic acids within the hypervariable regions and compare the sequences generated to those within a database (Jenkins et al., 2012) or to use more specific primers to look for just the microbiota a researcher may be

interested in (Matsuki et al., 2004).

This strategy, as compared to using LPS, enables both quantification as well as identification, which is significant as LPS molecules are not all the same (Gnauck et al., 2016). Variations in the configuration of the inner lipid A portion, which is dependent upon species, result in differing levels of endotoxicity as measured by the LAL assay. Therefore, being able to not only say bacteria is present in the blood but also which bacteria are present has tremendous value. All of this data is suggestive of a possible relationship between bacteria and changes in health that result in type 2 diabetes. Using both measures, LPS and 16S rRNA, enables a review of what is known from studies before and after this shift towards molecular-based techniques occurred. As research continues into the taxonomic and functional changes that happen in response to external exposures such as diet, we will begin to better understand the downstream factors that result and how the interplay between all these relationships can lead to disease within individuals.

III. Methods

A. Inclusion Criteria

Prior to commencing this systematic literature review, parameters for included studies were set forth; these criteria identified cross sectional, cohort, and case-control studies published within the past 15 years that examined LPS or 16S rRNA in the blood of human adults, 18 years and older, and whether this exposure was associated with an outcome of type 2 diabetes (see Figure 10).

B. Exclusion Criteria

During this process, exclusion criteria were also implemented to remove records that did not meet the parameters set forth for included studies. If these were not clearly present, records remained included until the next round of review.

1. Population

As the intention of this systematic literature review was to assess studies in human adults, any records focused on models of type 2 diabetes either in animals or using in vitro methods such as culturing of human cell lines were excluded from consideration. Furthermore, records that referred to children or adolescents were excluded as were those in which the participants were deceased.

2. Exposure

With the goal of identifying the role of bacteria and bacterial components as an

Figure 10

Inclusion Criteria for Studies

Population	Human adults 18 years and older
Intervention or Exposure	Presence of LPS or 16S rRNA in blood
Comparison	Preferred but not required
Outcome	Development of type 2 diabetes
Timing	Published within past 15 years
Setting	Any setting
Study Design	Cross sectional, cohort, case-control
Language	Any language

exposure through the presence of LPS or 16S rRNA in blood, records focused on other infectious agents including viruses and helminths were excluded. In addition, although originating from bacteria themselves, records examining bacterial metabolites were excluded as these are downstream byproducts and do not indicate the presence of intact bacteria or parts thereof. Likewise, indirect markers of bacterial presence such as antibodies, procalcitonin, LPS-binding protein, and CRP were excluded as these are made by the body in response to bacteria but are not bacteria or bacterial components themselves. Furthermore, records discussing the cellular proteins which make up the multi-receptor complex that binds LPS, namely TLR-4, CD14, and MD-2 (Zanoni & Granucci, 2013), were excluded as these are involved in the recognition and signaling processes that occur downstream following this exposure. Another category of records that were excluded involved treatments that impact bacteria rather than the presence or absence of bacteria or bacterial components. These include dietary interventions as well as drugs and supplements such as antibiotics, prebiotics, plants, plant extracts, and antioxidants in addition to fecal transplantation, probiotics, and probiotic foods where the treatment is bacteria itself.

A key component in regards to exposure was location, and records examining the effect of microbiota in areas other than blood or tissue, such as the gut and mouth, were excluded from consideration. The term tissue was included because blood is regarded as a connective tissue (MedlinePlus, 2019). In the case of specific bacteria or bacterial components, when no mention of location was made or implied using terms such as translocation or biomarker, records remained included if the entity was known to enter the systemic circulation whereas those that remain primarily localized to a particular site

were excluded. Examples include *Helicobacter pylori*, *Mycobacterium* spp., and *Chlamydophila pneumoniae*.

3. Outcome

Records focused on outcomes such as inflammation that did not also specifically mention type 2 diabetes or conditions like insulin resistance that can lead to it were excluded from consideration. In addition, records where infection was the outcome rather than the exposure were excluded as well.

4. Study Design

To identify peer-reviewed cross sectional, cohort, and case-control studies, records whose method differed were excluded such as editorials, list of papers of note by editors, letters, articles in response to another article, case reports, meeting proceedings, conference abstracts, and review protocols. Reviews remained included if no other exclusion criteria were present.

C. Search Methods

Search terms were generated for use in PubMed and EMBASE (see Figure 11). As microbial influence in the pathogenesis of type 2 diabetes has been largely unexplored, words that describe bacteria both more broadly (i.e. microbiota, microbiome, dysbiosis, bacteria, metagenome, metagenomic, metatranscriptome, metatranscriptomic) and more narrowly (i.e. endotoxin, lipopolysaccharide, 16S ribosomal RNA) were selected in addition to adjective, plural, and abbreviated forms of the above. This was done to ensure a sufficiently wide pool of

Figure 11

PubMed and EMBASE Search Strategy

PubMed Search Strategy		
Search	Terms Used	Entries Recovered
1	(“microbiota”[MeSH] OR “microbiota”[TIAB] OR “microbiome”[TIAB] OR “dysbiosis”[MeSH] OR “dysbiosis”[TIAB] OR “bacteria”[MeSH] OR “bacteria"[TIAB] OR “bacterial”[TIAB] OR “bacteriologic”[TIAB] OR “bacteriological”[TIAB] OR “endotoxins”[MeSH] OR “endotoxin”[TIAB] OR “endotoxins”[TIAB] OR “lipopolysaccharides”[MeSH] OR “lipopolysaccharide"[TIAB] OR “lipopolysaccharides”[TIAB] OR “LPS”[TIAB] OR “metagenome”[MeSH] OR “metagenomics”[MeSH] OR “metagenome”[TIAB] OR “metagenomic”[TIAB] OR “metagenomics”[TIAB] OR “metatranscriptome”[TIAB] OR “metatranscriptomic”[TIAB] OR “metatranscriptomics”[TIAB] OR “RNA, ribosomal, 16S”[MeSH] OR “16S ribosomal RNA”[TIAB] OR “16S rRNA”[TIAB] OR “shotgun sequencing”[TIAB]) **AND** (”blood”[MeSH] OR “blood”[TIAB]) **AND** (“diabetes mellitus, type 2”[Majr] OR (“type 2”[TI] AND “diabetes”[TI]) OR “type 2 diabetes”[TI])	403
2	(“microbiota”[MeSH] OR “microbiota”[TIAB] OR “microbiome”[TIAB] OR “dysbiosis”[MeSH] OR “dysbiosis”[TIAB] OR “bacteria”[MeSH] OR “bacteria"[TIAB] OR “bacterial”[TIAB] OR “bacteriologic”[TIAB] OR “bacteriological”[TIAB] OR “endotoxins”[MeSH] OR “endotoxin”[TIAB] OR “endotoxins”[TIAB] OR “lipopolysaccharides”[MeSH] OR “lipopolysaccharide"[TIAB] OR “lipopolysaccharides”[TIAB] OR “LPS”[TIAB] OR “metagenome”[MeSH] OR “metagenomics”[MeSH] OR “metagenome”[TIAB] OR “metagenomic”[TIAB] OR “metagenomics”[TIAB] OR “metatranscriptome”[TIAB] OR “metatranscriptomic”[TIAB] OR “metatranscriptomics”[TIAB] OR “RNA, ribosomal, 16S”[MeSH] OR “16S ribosomal RNA”[TIAB] OR “16S rRNA”[TIAB] OR “shotgun sequencing”[TIAB]) **AND** (”blood”[Majr] OR “blood”[TI]) **AND** (“diabetes mellitus, type 2”[MeSH] OR (“type 2”[TIAB] AND “diabetes”[TIAB]) OR “type 2 diabetes”[TIAB])	159
3	(“bacteremia”[MeSH] OR “bacteremia”[TIAB] OR “endotoxemia”[MeSH] OR “endotoxemia”[TIAB]) **AND** (“diabetes mellitus, type 2”[Majr] OR (“type 2”[TI] AND “diabetes”[TI) OR “type 2 diabetes”[TI])	91
EMBASE Search Strategy		
Search	Terms Used	Entries Recovered
1	‘microflora’/exp OR ‘dysbiosis’/exp OR ‘bacterium’/exp OR ‘endotoxin’/exp OR ‘lipopolysaccharide’/exp OR ‘metagenome’/exp OR ‘metagenomics’/exp OR ‘metatranscriptome’/exp OR ‘metatranscriptomics’/exp OR ‘RNA 16S’/exp OR ‘shotgun sequencing’/exp OR (microbiota OR microbiome OR dysbiosis OR bacteria OR bacterial OR bacteriologic OR bacteriological OR endotoxin OR endotoxins OR lipopolysaccharide OR lipopolysaccharides OR LPS OR metagenome OR metagenomic OR metagenomics OR metatranscriptome OR metatranscriptomic OR metatranscriptomics OR ‘16S ribosomal RNA’ OR ‘16S rRNA’ OR ‘shotgun sequencing’):ti,ab **AND** (‘blood’/exp OR blood:ti,ab) **AND** (‘non insulin dependent diabetes mellitus’/exp/mj OR (‘type 2’:ti AND diabetes:ti) OR ‘type 2 diabetes’:ti)	603
2	‘microflora’/exp OR ‘dysbiosis’/exp OR ‘bacterium’/exp OR ‘endotoxin’/exp OR ‘lipopolysaccharide’/exp OR ‘metagenome’/exp OR ‘metagenomics’/exp OR ‘metatranscriptome’/exp OR ‘metatranscriptomics’/exp OR ‘RNA 16S’/exp OR ‘shotgun sequencing’/exp OR (microbiota OR microbiome OR dysbiosis OR bacteria OR bacterial OR bacteriologic OR bacteriological OR endotoxin OR endotoxins OR lipopolysaccharide OR lipopolysaccharides OR LPS OR metagenome OR metagenomic OR metagenomics OR metatranscriptome OR metatranscriptomic OR metatranscriptomics OR ‘16S ribosomal RNA’ OR ‘16S rRNA’ OR ‘shotgun sequencing’):ti,ab **AND** (‘blood’/exp/mj OR blood:ti) **AND** (‘non insulin dependent diabetes mellitus’/exp OR (‘type 2’:ti,ab AND diabetes:ti,ab) OR ‘type 2 diabetes’:ti,ab)	234
3	‘bacteremia’/exp OR ‘endotoxemia’/exp OR (bacteremia OR endotoxemia):ti,ab **AND** (‘non insulin dependent diabetes mellitus’/exp/mj OR (‘type 2’:ti AND diabetes:ti) OR ‘type 2 diabetes’:ti)	129

studies to capture those focused on this exposure of interest and either paired with the term blood to specify body site in search 1 and 2 or replaced with terms already indicating this location (i.e. bacteremia, endotoxemia) in search 3. In all searches, these terms for the exposure of interest were paired with terms for the outcome of interest, namely type 2 diabetes, and included both intact phrasing as well as alternative phrasing where type 2 and diabetes may be written separately. If MeSH or Emtree equivalents for any of these terms were available in PubMed or EMBASE, respectively, these were included as well.

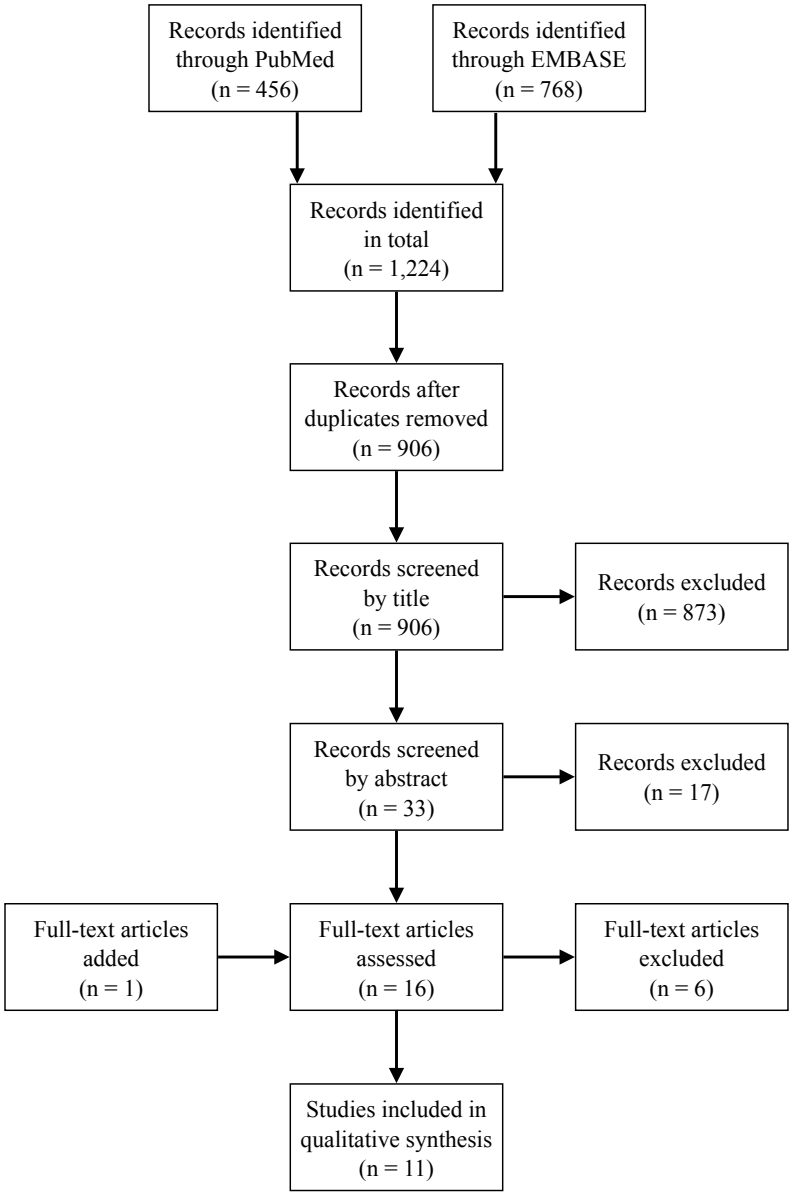
D. Search Process

This systematic literature review had one reviewer and all searches were performed on July 3rd, 2019. After saving each individually, searches 1, 2, and 3 were combined within PubMed and within EMBASE to remove duplicates resulting in 526 and 834 records retrieved, respectively. After applying a filter to limit results to those published within the past 15 years, that number changed to 456 and 768 records retrieved, respectively (see Figure 12). These were transferred into EndNote version X9.2 with a total of 1,224 records. During manual deduplication, identical results were removed if the first author, year of publication, and title were an exact match. Otherwise, such as in the case of small variations in spelling, a decision was made based on the digital object identifier or if not available, the international standard serial number. After this process, the final number of records remaining was 906.

All records went through title review and 33 remained after exclusion criteria were applied. These then entered abstract review after which, 16 articles were identified for full text review. Five reviews were identified (Gnauck et al., 2016; Klekotka et al., 2018; Piya et al., 2013; Pomytkin et al., 2015; Sato et al., 2017), and any studies cited related to the exposure of

Figure 12

Study Flow Diagram



Note. Adapted from "Preferred Reporting Items for Systematic Reviews and Meta-analyses: The PRISMA Statement," by D. Moher, A. Liberati, J. Tetzlaff, D. G. Altman, and The PRISMA Group, 2009, *PLoS Medicine*, *6*(7), e1000097 (https://doi.org/10.1371/journal.pmed.1000097). CC BY 4.0.

interest entered title review followed by abstract review. One additional study was identified in this manner (Hawkesworth et al., 2013). Of the original 16 articles that entered full text review, one study was excluded because no English translation was available (Liu et al., 2013). Therefore, the total number of articles requiring full text review was 17 with 11 studies ultimately included in the results of this systematic literature review.

E. Extraction

Various data were extracted including study design and sample size along with descriptive statistics, measurements of LPS or 16S rRNA, type 2 diabetes-related indices, and associations between these variables. Descriptive statistics included details and measurements such as type 2 diabetes diagnosis method, setting, sex, age, family history of diabetes, BMI, dietary macronutrients, physical activity, smoking, blood pressure, and lipid profile. Measurements of LPS and 16S rRNA included information regarding fasting status, sample type, and assay method in addition to quantitative values. Type 2 diabetes-related indices included HbA1C, fasting plasma glucose, two hour plasma glucose, impaired fasting glucose, type 2 diabetes prevalence, fasting plasma insulin, homeostasis model assessment of insulin resistance (HOMA-IR), insulin sensitivity index (ISI), quantitative insulin sensitivity check index (QUICKI), disposition index (DI), homeostasis model assessment of β-cell function (HOMA-B), and insulinogenic index (IGI). In instances where there were discrepancies between values for a particular measure, the value that was reported the same way numerous times was chosen over a single value that differed from the others. Otherwise, when all were used equally, the most conservative was selected.

F. Analysis

The data were analyzed and the results of the systemic literature review presented as a qualitative synthesis.

G. Reporting

While aspects of PRISMA were utilized, such as in the creation of the study flow diagram in Figure 12 (Moher et al., 2009), it was not used in its entirety to guide reporting.

H. Limitations

One of the greatest challenges was the diversity of language used to describe similar observations as the terminology regarding these microbial communities became more consistent and widely accepted over time. Some terms have been used interchangeably although subtle differences exist between them (Ursell et al., 2012) thereby necessitating a comprehensive strategy to identify records most likely to meet the search criteria. However, even with this in mind, the process of a systematic literature review may not capture relevant studies in the results if the vocabulary used differs from that of the search terms. The use of MeSH and Emtree equivalents reduces this likelihood, but in the case of newly published literature, these may not yet be assigned (Mao & Lu, 2017). In addition, as an understudied topic, there may be challenges researchers face in making their results more widely available. Thus, publication bias could be present due to a lack of interest for such manuscripts, especially those with nonsignificant findings.

Furthermore, in order to narrow in on the relationship between a particular exposure and outcome, the selection of individual search terms often comes at the exclusion of others. This can

be a difficult balance but is necessary so that specific interactions can be evaluated. To accomplish this goal in regards to exposure, the scope of the project remained focused on bacteria in general and LPS or 16S rRNA in particular. No other microorganisms or bacterial components were included in the search terms. Similarly, in regards to outcome, the scope of the project remained focused on type 2 diabetes. No other conditions, even those that may subsequently lead to it, were included in the search terms. There may exist studies outside of this scope with relevance to the topic; however, this approach was chosen to identify those that would most clearly demonstrate a relationship if present.

I. Ethical Considerations

This project was not considered human subjects research and therefore, did not require approval by the Emory University Institutional Review Board.

IV. Results

A. Cross Sectional Studies Examining the Relationship Between LPS and Type 2 Diabetes

Of the 11 studies included in this systematic literature review, seven provided data that were cross sectional in nature (Al-Obaide et al., 2017; Creely et al., 2007; de Waal et al., 2018; Hawkesworth et al., 2013; Huang et al., 2019; Jayashree et al., 2014; Pussinen et al., 2011). One study involved random sampling where both LPS and type 2 diabetes status were unknown at the time of recruitment (Huang et al., 2019) while five studies involved cross sectional data collection where type 2 diabetes status was known and part of the recruitment criteria (Al-Obaide et al., 2017; Creely et al., 2007; de Waal et al., 2018; Hawkesworth et al., 2013; Jayashree et al., 2014). One study provided baseline measurements of LPS and type 2 diabetes status within a cohort before those with prevalent diabetes were excluded from further analysis (Pussinen et al., 2011).

1. Descriptive Statistics

Descriptive statistics of the participants in these studies are summarized in Table 1. Only the sample sizes were available for all studies and ranged from 21 to 7,169 participants. Six studies provided details about age and five provided details about setting, sex, and BMI; these were the next most commonly reported variables. Four studies provided complete information regarding type 2 diabetes diagnosis method while an additional study provided partial information. These methods were used either on their own or in combination to delineate comparison groups with testing being the strongest as it relied on measurement of type 2 diabetes-related indices within the study whereas

Table 1

Descriptive Statistics From Cross Sectional Studies Examining the Relationship Between LPS and T2DM

Study	Groups	T2DM diagnosis method [a]	Setting [b]	*n*	Sex, *n* (%)		Age (years)	Family history, *n* (%)
					Male	Female		
Al-Obaide et al., 2017 [c]	Control	Physician-diagnosed	Hospital	20			54.3 ± 3.2 [d]	
	T2DM	Physician-diagnosed or physician-diagnosed and testing [f]	Hospital	20			64.4 ± 2.3 [d]	
							NS	
Creely et al., 2007 [c]	Control			25	20 (80) [g]	5 (20) [g]	48.1 ± 19.2 [g]	
	T2DM			25	20 (80) [g]	5 (20) [g]	52.2 ± 11.7 [g]	
					NS		NS	
de Waal et al., 2018 [c]	Control			11	5 (45.5)	6 (54.5)	48 ± 3.240 [d]	
	T2DM			10	5 (50)	5 (50)	57.5 (51 - 67.5)	
Hawkesworth et al., 2013 [c]	Non-obese control		Community	31	0 (0)	31 (100)	41.5 ± 6.2	
	Obese control		Community	33	0 (0)	33 (100)	43.4 ± 5.4	
	T2DM	Physician-diagnosed	Hospital	29	0 (0)	29 (100)	45.1 ± 5.2	
							$p = 0.06$	
Huang et al., 2019 [i]	T1 [j]	Self-reported or testing	Research	851	325 (38.2)	526 (61.8)	43.93 ± 15.05	
	T2 [j]	Self-reported or testing	Research	851	372 (43.7)	479 (56.3)	44.62 ± 15.63	
	T3 [j]	Self-reported or testing	Research	851	391 (45.9)	460 (54.1)	48.38 ± 15.76	
						$p = 0.004$	$p < 0.001$	
Jayashree et al., 2014 [c]	Control	Testing	Research	45	25 (55.5)	20 (44.5)	46 ± 9	
	T2DM	Testing	Research	45	23 (51.5)	22 (48.5)	51 ± 6	
					$p = 0.402$		$p < 0.05$	
Pussinen et al., 2011 [n]	Control	Self-reported and physician-diagnosed	Community	6,632				
	T2DM	Self-reported or physician-diagnosed	Community	537				

Note. Data presented as $M \pm SD$ or *Mdn* (IQ) unless otherwise indicated. BMI = body mass index; BP = blood pressure; eMERGE = electronic medical records and genomics; EU = endotoxin unit; HDL = high-density lipoprotein; IQ = interquartile range; LDL = low-density lipoprotein; LPS = lipopolysaccharide; NS = non-significant at $p < 0.05$; *SEM* = standard error of the mean; T1 = tertile 1 (LPS < 0.29 EU/ml); T2 = tertile 2 (LPS = 0.29 - 0.52 EU/ml); T2DM = type 2 diabetes mellitus; T3 = tertile 3 (LPS > 0.52 EU/ml).

[a] Physician-diagnosed = physician or medical record indicated T2DM diagnosis, medication or both; testing = measurement of T2DM-related indices; self-reported = T2DM status provided by participants. [b] Community = general population; hospital = outpatient clinic; research = participants from existing studies. [c] Cross sectional data collection. [d] $M \pm SEM$. [e] mg/dl. [f] eMERGE phenotype definition for T2DM (Richesson et al., 2013; Pacheco & Thompson, 2011). [g] Sex-, age-, and BMI-matched. [h] Hypertension defined as systolic BP > 140 mm Hg. [i] Cross sectional. [j] LPS *Mdn* (IQ) = 0.36 (0.27 - 0.62) EU/ml. [k] Estimated by the International Physical Activity Questionnaire. [l] Unclear whether former smokers were included. [m] Mean of two measurements. [n] Baseline measurement within a cohort.

Study	Groups	BMI (kg/m^2)	Dietary macronutrients (%)			Weekly physical activity, *n* (%)			Smoking, *n* (%)
			Protein	Fat	Carbohydrate	< 0.5 hr	0.5 - 1 hr	≥ 1 hr	
Al-Obaide et al., 2017 [c]	Control	28.17 ± 1.1 [d]	17	32.7	52				
	T2DM	33.2 ± 2.9 [d]	12	50.0	47				
		NS	$p < 0.05$	$p < 0.03$	NS				
Creely et al., 2007 [c]	Control	29.5 ± 4.3 [g]							
	T2DM	31.8 ± 4.5 [g]							
		NS							
de Waal et al., 2018 [c]	Control								
	T2DM								
Hawkesworth et al., 2013 [c]	Non-obese control	20.8 ± 1.8							
	Obese control	34.3 ± 4.5							
	T2DM	33.3 ± 5.7							
Huang et al., 2019 [i]	T1 [j]	22.83 ± 3.38				502 (59.0) [k]	215 (25.3) [k]	134 (15.7) [k]	170 (20.0) [l]
	T2 [j]	23.36 ± 3.55				503 (59.2) [k]	202 (23.8) [k]	145 (17.1) [k]	178 (20.9) [l]
	T3 [j]	24.28 ± 3.90				509 (59.9) [k]	193 (22.7) [k]	148 (17.4) [k]	177 (20.8) [l]
		$p < 0.001$					$p = 0.732$		$p = 0.879$
Jayashree et al., 2014 [c]	Control	26.9 ± 3.9							
	T2DM	27.2 ± 6.0							
		$p = 0.766$							
Pussinen et al., 2011 [n]	Control								
	T2DM								

Note. Data presented as *M* ± *SD* or *Mdn* (IQ) unless otherwise indicated. BMI = body mass index; BP = blood pressure; eMERGE = electronic medical records and genomics; EU = endotoxin unit; HDL = high-density lipoprotein; IQ = interquartile range; LDL = low-density lipoprotein; LPS = lipopolysaccharide; NS = non-significant at $p < 0.05$; *SEM* = standard error of the mean; T1 = tertile 1 (LPS < 0.29 EU/ml); T2 = tertile 2 (LPS = 0.29 - 0.52 EU/ml); T2DM = type 2 diabetes mellitus; T3 = tertile 3 (LPS > 0.52 EU/ml).

[a] Physician-diagnosed = physician or medical record indicated T2DM diagnosis, medication or both; testing = measurement of T2DM-related indices; self-reported = T2DM status provided by participants. [b] Community = general population; hospital = outpatient clinic; research = participants from existing studies. [c] Cross sectional data collection. [d] *M* ± *SEM*. [e] mg/dl. [f] eMERGE phenotype definition for T2DM (Richesson et al., 2013; Pacheco & Thompson, 2011). [g] Sex-, age-, and BMI-matched. [h] Hypertension defined as systolic BP > 140 mm Hg. [i] Cross sectional. [j] LPS *Mdn* (IQ) = 0.36 (0.27 - 0.62) EU/ml. [k] Estimated by the International Physical Activity Questionnaire. [l] Unclear whether former smokers were included. [m] Mean of two measurements. [n] Baseline measurement within a cohort.

Study	Groups	Blood pressure (mm Hg)		Hypertension (%)	Cholesterol (mmol/L)			Triglycerides (mmol/L)
		Systolic	Diastolic		Total	HDL	LDL	
Al-Obaide et	Control				195 ± 11 [d,e]	58 ± 3.3 [d,e]	72 ± 3.2 [d,e]	85 ± 4.1 [d,e]
	T2DM				175.7 ± 13 [d,e]	37.7 ± 3.0 [d,e]	99 ± 10 [d,e]	206.9 ± 21 [d,e]
					$p < 0.05$	$p < 0.05$	$p < 0.03$	$p < 0.001$
Creely et al.,	Control							
	T2DM							
de Waal et al.,	Control							
	T2DM							
Hawkesworth	Non-obese control	114.0 ± 14.1	73.3 ± 8.8	6.7 [h]				
	Obese control	126.6 ± 19.7	84.0 ± 12.3	26.7 [h]				
	T2DM	125.3 ± 17.1	84.2 ± 9.3	13.8 [h]				
Huang et al.,	T1 [j]	126.32 ± 19.22 [m]					3.09 ± 0.99	1.28 ± 1.12
	T2 [j]	129.62 ± 18.46 [m]					3.21 ± 0.96	1.39 ± 1.10
	T3 [j]	134.03 ± 20.16 [m]					3.39 ± 1.01	1.72 ± 1.68
		$p < 0.001$					$p < 0.001$	$p < 0.001$
Jayashree et	Control	127 ± 16	80 ± 11		171 ± 37 [e]	43 ± 5 [e]	99 ± 28 [e]	111 ± 40 [e]
	T2DM	132 ± 19	82 ± 11		189 ± 36 [e]	36 ± 6 [e]	120 ± 38 [e]	156 ± 66 [e]
		$p = 0.187$	$p = 0.466$		$p < 0.01$	$p < 0.001$	$p < 0.01$	$p < 0.001$
Pussinen et	Control							
	T2DM							

Note. Data presented as $M \pm SD$ or *Mdn* (IQ) unless otherwise indicated. BMI = body mass index; BP = blood pressure; eMERGE = electronic medical records and genomics; EU = endotoxin unit; HDL = high-density lipoprotein; IQ = interquartile range; LDL = low-density lipoprotein; LPS = lipopolysaccharide; NS = non-significant at $p < 0.05$; *SEM* = standard error of the mean; T1 = tertile 1 (LPS < 0.29 EU/ml); T2 = tertile 2 (LPS = 0.29 - 0.52 EU/ml); T2DM = type 2 diabetes mellitus; T3 = tertile 3 (LPS > 0.52 EU/ml).

[a] Physician-diagnosed = physician or medical record indicated T2DM diagnosis, medication or both; testing = measurement of T2DM-related indices; self-reported = T2DM status provided by participants. [b] Community = general population; hospital = outpatient clinic; research = participants from existing studies. [c] Cross sectional data collection. [d] $M \pm SEM$. [e] mg/dl. [f] eMERGE phenotype definition for T2DM (Richesson et al., 2013; Pacheco & Thompson, 2011). [g] Sex-, age-, and BMI-matched. [h] Hypertension defined as systolic BP > 140 mm Hg. [i] Cross sectional. [j] LPS *Mdn* (IQ) = 0.36 (0.27 - 0.62) EU/ml. [k] Estimated by the International Physical Activity Questionnaire. [l] Unclear whether former smokers were included. [m] Mean of two measurements. [n] Baseline measurement within a cohort.

physician-diagnosed utilized healthcare providers or medical records to establish a diagnosis. Otherwise, type 2 diabetes status was self-reported by participants. Information about blood pressure and lipid profile was less consistent. Three studies reported values for systolic blood pressure, two for diastolic blood pressure, and one for hypertension. Three studies provided details regarding low-density lipoprotein (LDL) cholesterol and triglycerides while two studies provided details regarding total and HDL cholesterol. Lastly, information about dietary macronutrients, weekly physical activity, and smoking each was available from a single study, and family history of type 2 diabetes was not reported by any study.

2. Measurement of LPS

At the time of LPS sampling, participants were either fasting or semifasting; however, details about fasting status were unknown for two studies (de Waal et al., 2018; Jayashree et al., 2014). LPS levels were determined using either serum or platelet poor plasma with a range of assays that included enzyme-linked immunosorbent assay (ELISA), fluorescent antibody staining, endpoint chromogenic LAL, and kinetic chromogenic LAL. These details are presented in Table 2 along with measurements of LPS and type 2 diabetes-related indices. In all studies, LPS levels were higher in the type 2 diabetes group as compared to controls with the exception of one study whose comparison groups were based on tertiles of LPS rather than type 2 diabetes status in which case, higher levels of LPS corresponded to a higher prevalence of type 2 diabetes (Huang et al., 2019). These differences were significant in six of the seven studies (Creely et al., 2007; de Waal et al., 2018; Hawkesworth et al., 2013; Huang et al., 2019;

Table 2

Measurements of LPS and T2DM-Related Indices From Cross Sectional Studies Examining the Relationship Between LPS and T2DM

Study	Groups	Fasting	Sample type	Assay method	LPS (EU/ml)	LPS activity [a]
Al-Obaide et al., 2017 [e]	Control T2DM	Fasting	Serum	ELISA	21.04 (14.12 - 32.9) [f] 36.46 (17.96 - 48.52) [f] $p = 0.094$	
Creely et al., 2007 [e]	Control T2DM	Fasting	Serum	Endpoint chromogenic LAL	3.1 ± 1.7 [g,h] 5.5 ± 1.6 [g,h] $p = 0.0031$	
de Waal et al., 2018 [e]	Control T2DM		Platelet poor plasma	Fluorescent antibody staining	0.5612 (0.2898 - 0.6275) [k] 1.847 (1.387 - 2.012) [k] $p < 0.0001$	
Hawkesworth et al., 2013 [e]	Non-obese control Obese control T2DM	Fasting	Serum	Endpoint chromogenic LAL	3.89 (3.20 - 4.73) [m] 3.86 (3.30 - 4.52) [m] 5.19 (3.43 - 7.87) [m] $p = 0.02$	
Huang et al., 2019 [p]	T1 [q] T2 [q] T3 [q]	Fasting	Serum	Endpoint chromogenic LAL		
Jayashree et al., 2014 [e]	Control T2DM		Serum	Endpoint chromogenic LAL	0.47 ± 0.02 [l,v,w] 0.57 ± 0.028 [l,v,w] $p < 0.05$	— [l,w] — [l,w] $p < 0.05$
Pussinen et al., 2011 [y]	Control T2DM	Semifasting [z]	Serum	Kinetic chromogenic LAL	62.18 ± 36.77 [u] 70.73 ± 42.62 [u] $p < 0.001$	

Note. Data presented as *M* ± *SD* or *Mdn* (IQ) unless otherwise indicated. 2 hr PG = 2 hr plasma glucose; BMI = body mass index; CI = confidence interval; DI = disposition index; ELISA = enzyme-linked immunosorbent assay; EU = endotoxin unit; FPG = fasting plasma glucose; FPI = fasting plasma insulin; HbA1C = hemoglobin A1C; HDL = high-density lipoprotein; HOMA-B = homeostasis model assessment of beta cell function; HOMA-IR = homeostasis model assessment of insulin resistance; IFG = impaired fasting glucose; IGI = insulinogenic index; IQ = interquartile range; ISI = insulin sensitivity index; LAL = limulus amoebocyte lysate; LPS = lipopolysaccharide; QUICKI = quantitative insulin sensitivity check index; *SEM* = standard error of the mean; T1 = tertile 1 (LPS < 0.29 EU/ml); T2 = tertile 2 (LPS = 0.29 - 0.52 EU/ml); T2DM = type 2 diabetes mellitus; T3 = tertile 3 (LPS > 0.52 EU/ml).

[a] LPS activity = LPS/HDL cholesterol. [b] Testing performed at baseline and participants met at least one T2DM diagnostic criteria from Figure 3. [c] HOMA-IR = fasting serum insulin (mU/L) x FPG (mmol/L)/22.5. [d] QUICKI = 1/(log(insulin) + log(glucose)). [e] Cross sectional data collection. [f] ng/ml. [g] Geometric mean. [h] Sex-, age-, and BMI-matched. [i] Serum. [j] IU/ml. [k] Mean fluorescence intensity normalized by secondary antibody control to account for non-specific binding. [l] *M* ± *SEM*. [m] Geometric mean (95% CI). [n] Plasma or serum unspecified. [o] pg/ml. [p] Cross sectional. [q] LPS *Mdn* (IQ) = 0.36 (0.27 - 0.62) EU/ml. [r] *n* = 2,494. [s] *N* (%) = 247 (9.7). [t] mU/L. [u] Log transformed. [v] Unclear whether EU/ml or µg/ml. [w] Age-adjusted. [x] mg/dl. [y] Baseline measurement within a cohort. [z] Participants asked to fast for 4 hr and avoid heavy meals; *Mdn* (IQ) = 5 (3 - 7) hr; no correlation between fasting time and log LPS; additional details about meals unknown.

Study	Groups	HbA1C (%)	FPG (mmol/L)	2 hr PG (mmol/L)	IFG (%)	T2DM (%) [b]
Al-Obaide et al., 2017 [e]	Control					
	T2DM					
Creely et al., 2007 [e]	Control		5.6 ± 0.9 [h,i]			
	T2DM		8.6 ± 2.5 [h,i]			
			$p < 0.0001$			
de Waal et al., 2018 [e]	Control					
	T2DM	7.4 ± 0.456 [l]				
Hawkesworth et al., 2013 [e]	Non-obese control		5.00 (4.84 - 5.17) [m]		3.5	0.0
	Obese control		5.51 (5.23 - 5.80) [m]		12.5	3.1
	T2DM		9.19 (7.75 - 10.89) [m]		82.1	67.9
Huang et al., 2019 [p]	T1 [q]	5.52 ± 0.70	5.10 ± 0.88	6.30 ± 2.17 [r]		4.1 [s]
	T2 [q]	5.59 ± 0.82	5.28 ± 1.05	6.68 ± 2.54 [r]		8.5 [s]
	T3 [q]	5.81 ± 1.20	5.62 ± 1.76	7.62 ± 3.79 [r]		16.5 [s]
		$p < 0.001$	$p < 0.001$	$p < 0.001$		$p < 0.001$
Jayashree et al., 2014 [e]	Control	5.6 ± 0.4	91 ± 7 [x]	115 ± 26 [x]		0.0
	T2DM	8.0 ± 2.2	144 ± 62 [x]	241 ± 92 [x]		100.0
		$p < 0.01$	$p < 0.01$	$p < 0.01$		
Pussinen et al., 2011 [y]	Control					
	T2DM					

Note. Data presented as $M \pm SD$ or *Mdn* (IQ) unless otherwise indicated. 2 hr PG = 2 hr plasma glucose; BMI = body mass index; CI = confidence interval; DI = disposition index; ELISA = enzyme-linked immunosorbent assay; EU = endotoxin unit; FPG = fasting plasma glucose; FPI = fasting plasma insulin; HbA1C = hemoglobin A1C; HDL = high-density lipoprotein; HOMA-B = homeostasis model assessment of beta cell function; HOMA-IR = homeostasis model assessment of insulin resistance; IFG = impaired fasting glucose; IGI = insulinogenic index; IQ = interquartile range; ISI = insulin sensitivity index; LAL = limulus amoebocyte lysate; LPS = lipopolysaccharide; QUICKI = quantitative insulin sensitivity check index; *SEM* = standard error of the mean; T1 = tertile 1 (LPS < 0.29 EU/ml); T2 = tertile 2 (LPS = 0.29 - 0.52 EU/ml); T2DM = type 2 diabetes mellitus; T3 = tertile 3 (LPS > 0.52 EU/ml).

[a] LPS activity = LPS/HDL cholesterol. [b] Testing performed at baseline and participants met at least one T2DM diagnostic criteria from Figure 3. [c] HOMA-IR = fasting serum insulin (mU/L) x FPG (mmol/L)/22.5. [d] QUICKI = 1/(log(insulin) + log(glucose)). [e] Cross sectional data collection. [f] ng/ml. [g] Geometric mean. [h] Sex-, age-, and BMI-matched. [i] Serum. [j] IU/ml. [k] Mean fluorescence intensity normalized by secondary antibody control to account for non-specific binding. [l] $M \pm SEM$. [m] Geometric mean (95% CI). [n] Plasma or serum unspecified. [o] pg/ml. [p] Cross sectional. [q] LPS *Mdn* (IQ) = 0.36 (0.27 - 0.62) EU/ml. [r] $n = 2{,}494$. [s] N (%) = 247 (9.7). [t] mU/L. [u] Log transformed. [v] Unclear whether EU/ml or µg/ml. [w] Age-adjusted. [x] mg/dl. [y] Baseline measurement within a cohort. [z] Participants asked to fast for 4 hr and avoid heavy meals; *Mdn* (IQ) = 5 (3 - 7) hr; no correlation between fasting time and log LPS; additional details about meals unknown.

Study	Groups	FPI (pmol/L)	Measures of insulin resistance			Measures of insulin secretion		
			HOMA-IR [c]	ISI	QUICKI [d]	DI	HOMA-B	IGI
Al-Obaide et al., 2017 [e]	Control							
	T2DM							
Creely et al., 2007 [e]	Control	8.9 ± 1.9 [g,h,i,j]						
	T2DM	12.0 ± 1.8 [g,h,i,j]						
		$p = 0.08$						
de Waal et al., 2018 [e]	Control							
	T2DM							
Hawkesworth et al., 2013 [e]	Non-obese control	4.31 (3.28 - 5.65) [m,n,o]			0.35 ± 0.08			
	Obese control	11.08 (9.08 - 13.52) [m,n,o]			0.25 ± 0.04			
	T2DM	12.11 (9.71 - 15.09) [m,n,o]			0.22 ± 0.03			
Huang et al., 2019 [p]	T1 [q]	5.45 (4.18 - 7.28) [i,t]	— [u]					
	T2 [q]	6.91 (5.21 - 8.53) [i,t]	— [u]					
	T3 [q]	9.97 (7.63 - 12.82) [i,t]	— [u]					
		$p < 0.001$	$p < 0.001$					
Jayashree et al., 2014 [e]	Control							
	T2DM							
Pussinen et al., 2011 [y]	Control							
	T2DM							

Note. Data presented as *M* ± *SD* or *Mdn* (IQ) unless otherwise indicated. 2 hr PG = 2 hr plasma glucose; BMI = body mass index; CI = confidence interval; DI = disposition index; ELISA = enzyme-linked immunosorbent assay; EU = endotoxin unit; FPG = fasting plasma glucose; FPI = fasting plasma insulin; HbA1C = hemoglobin A1C; HDL = high-density lipoprotein; HOMA-B = homeostasis model assessment of beta cell function; HOMA-IR = homeostasis model assessment of insulin resistance; IFG = impaired fasting glucose; IGI = insulinogenic index; IQ = interquartile range; ISI = insulin sensitivity index; LAL = limulus amoebocyte lysate; LPS = lipopolysaccharide; QUICKI = quantitative insulin sensitivity check index; *SEM* = standard error of the mean; T1 = tertile 1 (LPS < 0.29 EU/ml); T2 = tertile 2 (LPS = 0.29 - 0.52 EU/ml); T2DM = type 2 diabetes mellitus; T3 = tertile 3 (LPS > 0.52 EU/ml).

[a] LPS activity = LPS/HDL cholesterol. [b] Testing performed at baseline and participants met at least one T2DM diagnostic criteria from Figure 3. [c] HOMA-IR = fasting serum insulin (mU/L) x FPG (mmol/L)/22.5. [d] QUICKI = 1/(log(insulin) + log(glucose)). [e] Cross sectional data collection. [f] ng/ml. [g] Geometric mean. [h] Sex-, age-, and BMI-matched. [i] Serum. [j] IU/ml. [k] Mean fluorescence intensity normalized by secondary antibody control to account for non-specific binding. [l] *M* ± *SEM*. [m] Geometric mean (95% CI). [n] Plasma or serum unspecified. [o] pg/ml. [p] Cross sectional. [q] LPS *Mdn* (IQ) = 0.36 (0.27 - 0.62) EU/ml. [r] *n* = 2,494. [s] *N* (%) = 247 (9.7). [t] mU/L. [u] Log transformed. [v] Unclear whether EU/ml or µg/ml. [w] Age-adjusted. [x] mg/dl. [y] Baseline measurement within a cohort. [z] Participants asked to fast for 4 hr and avoid heavy meals; *Mdn* (IQ) = 5 (3 - 7) hr; no correlation between fasting time and log LPS; additional details about meals unknown.

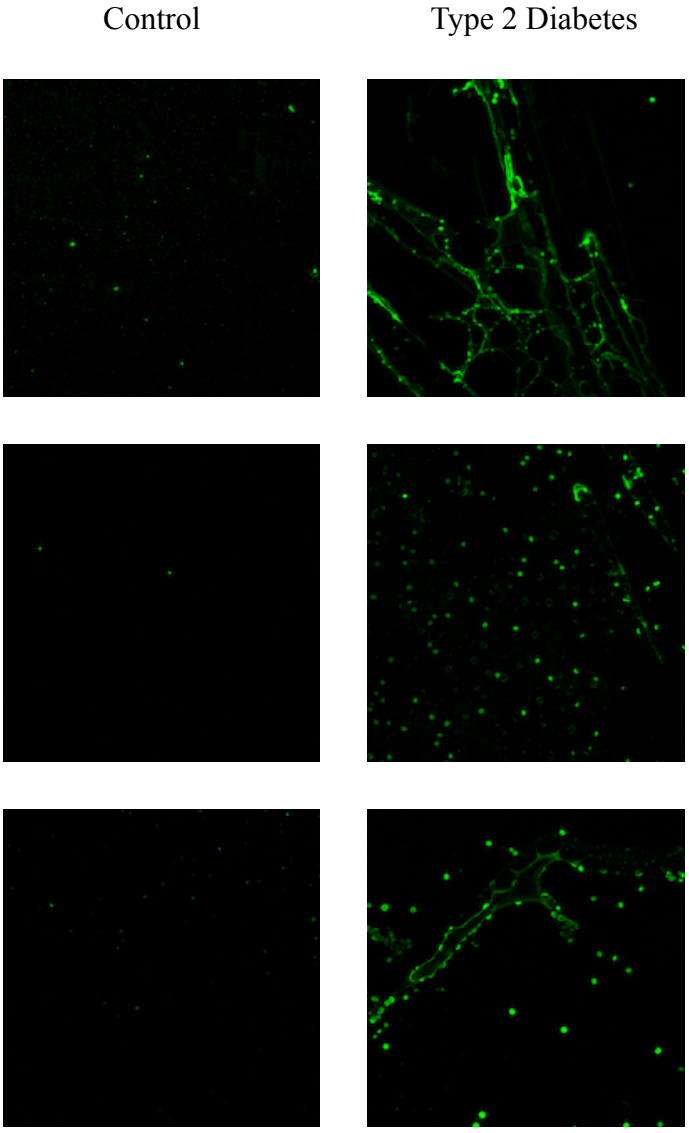
Jayashree et al., 2014; Pussinen et al., 2011), yet there was high variability among the values. Of the five studies whose assay method included LAL and thus measured LPS in EU/ml, or endotoxin units per milliliter, means and geometric means ranged from 0.47 to 70.73 within groups (Creely et al., 2007; Hawkesworth et al., 2013; Huang et al., 2019; Jayashree et al., 2014; Pussinen et al., 2011). LPS measurements from two studies were obtained by assay methods that differed with one using ELISA to quantify LPS in nanograms per milliliter (Al-Obaide et al., 2017) and the other using fluorescent antibody staining to quantify LPS in mean fluorescence intensity as shown in Figure 13 (de Waal et al., 2018). Conversion to EU/mL was not available since an EU is a measure of LPS activity, not quantity, as determined by the LAL assay after comparison to a standard (Gnauck et al., 2016). A final measure of LPS was LPS activity defined as LPS/HDL cholesterol, a separate assessment to that described above that takes into account HDL cholesterol levels since they may affect the clearance of LPS from the bloodstream (Jayashree et al., 2014). Only one study provided this information and found significantly higher levels in those with type 2 diabetes as compared to controls.

3. Measurement of Type 2 Diabetes-Related Indices

Four studies included information about fasting plasma glucose while three provided details regarding type 2 diabetes prevalence and fasting plasma insulin. HbA1C values were available for two studies with an additional study providing values for the type 2 diabetes group but not for controls. Two-hour plasma glucose was reported by two studies as well. Lastly, impaired fasting glucose, HOMA-IR, and QUICKI were each provided by one study while ISI, DI, HOMA-B, and IGI were not reported by any study.

Figure 13

Fluorescent Antibody Staining of Platelet Poor Plasma With Anti-E. coli LPS Antibody



Note. Adapted from "Correlative Light-Electron Microscopy Detects Lipopolysaccharide and its association with fibrin fibres in Parkinson's Disease, Alzheimer's Disease and Type 2 Diabetes Mellitus," by G. M. de Waal, L. Engelbrecht, T. Davis, W. J. S. de Villiers, D. B. Kell, and E. Pretorius, 2018, *Scientific Reports*, *8*, Article 16798 p. 7 (https://doi.org/10.1038/s41598-018-35009-y). CC BY 4.0.

4. Associations Between LPS and Type 2 Diabetes-Related Indices

Statistical analysis of the associations between LPS and various type 2 diabetes-related indices, as summarized in Table 3, revealed no relationships between fasting glucose (Creely et al., 2007), fasting insulin (Hawkesworth et al., 2013), or QUICKI index and LPS in two studies. Of the values that indicated a positive association, many were weak, age-adjusted correlations between HbA1C, fasting glucose, or two-hour plasma glucose with LPS activity (Jayashree et al., 2014). However, there was also a significant difference in total cholesterol, LDL cholesterol, and triglycerides between those with type 2 diabetes and controls. Further analysis found no correlation between total and LDL cholesterol and LPS activity (total: $r = 0.137$, $p = 0.184$; LDL: $r = 0.179$, $p = 0.198$), but there was a weak, age-adjusted, positive correlation between triglycerides and LPS activity ($r = 0.353$, $p < 0.001$) that was unadjusted for. An additional value showed an unadjusted, positive association between fasting glucose and LPS with the difference in age between groups analyzed and found to be nonsignificant but no other comparisons were made (Hawkesworth et al., 2013). Only one study showed a moderate to strong relationship between any type 2 diabetes-related indices and LPS with a moderate, unadjusted, positive association between HOMA-IR and LPS ($r = 0.692$, $p < 0.001$) as well as a strong; sex-, age-, and BMI-adjusted; positive association between fasting insulin and LPS ($r = 0.731$, $p < 0.001$; Creely et al., 2007). It is important to note that this relationship was examined in the control group only, which included 25 participants, and no other information about possible confounding variables was reported.

Four studies provided data exploring the relationship between LPS and type 2 diabetes prevalence (Hawkesworth et al., 2013; Huang et al., 2019; Jayashree et al., 2014;

Table 3

Cross Sectional Associations Between LPS and T2DM-Related Indices

Variable	Study	Positive association	No association	*p*	Adjusted for [a]							
					S	A	B	L	BP	LI	SE	BI
HbA1C [b,c]	Jayashree et al., 2014 [d]	r = 0.334 [e]		< 0.001								
FPG [c,f]	Creely et al., 2007 [d]		— [g,h]	NS								
	Hawkesworth et al., 2013 [d]	0.24 [0.09 - 0.38] [g,i,j,k]		0.001								
	Jayashree et al., 2014 [d]	r = 0.229 [e,l]		0.026								
2 hr PG [c,f]	Jayashree et al., 2014 [d]	r = 0.341 [e,l]		< 0.001								
FPI [c,m]	Creely et al, 2007 [d]	r = 0.623 [g,h,n,o,p]		< 0.001								
		r = 0.731 [g,h,n,o,p]		< 0.001								
	Hawkesworth et al., 2013 [d]		0.20 [-0.12 - 0.51] [g,j,k,n,q]	0.22								
HOMA-IR [c,r]	Creely et al., 2007 [d]	r = 0.692 [g,p]		< 0.001								
QUICKI [c,s]	Hawkesworth et al., 2013 [d]		—	NS								
T2DM [t]	Hawkesworth et al., 2013 [d]	0.43 [0.16 - 0.70] [g,u,v]		0.002								
		— [g,u,v]		< 0.05								
	Huang et al., 2019 [w]	2.16 [1.42 - 3.27] [x]		< 0.001 [y]								
		4.59 [3.13 - 6.74] [z]										
		2.05 [1.33 - 3.17] [x]		< 0.001 [y]								
		3.28 [2.18 - 4.93] [z]										
	Jayashree et al., 2014 [d]	13.43 [1.998 - 18.9]		0.003								
			5.45 [0.216 - 6.25]	0.303								
	Pussinen et al., 2011 [aa,bb]		1.043 [0.759 - 1.433] [g,cc]	0.79								
		1.717 [1.290 - 2.284] [g,dd]		< 0.001								
		2.232 [1.689 - 2.949] [g,ee]		< 0.001								
		1.005 [1.003 - 1.008] [g,ff]		< 0.001								

Note. LPS measured in EU/ml. Data presented as OR [95% CI] unless otherwise indicated. 2 hr PG = 2 hr plasma glucose; BMI = body mass index; CI = confidence interval; EU = endotoxin unit; FPG = fasting plasma glucose; FPI = fasting plasma insulin; HbA1C = hemoglobin A1C; HDL = high-density lipoprotein; HOMA-IR = homeostasis model assessment of insulin resistance; LDL = low-density lipoprotein; LPS = lipopolysaccharide; NS = non-significant at $p < 0.05$; OR = odds ratio; Q1 = quartile 1; Q2 = quartile 2; Q3 = quartile 3; Q4 = quartile 4; QUICKI = quantitative insulin sensitivity check index; T1 = tertile 1 (LPS < 0.29 EU/ml); T2 = tertile 2 (LPS = 0.29 - 0.52 EU/ml); T2DM = type 2 diabetes mellitus; T3 = tertile 3 (LPS > 0.52 EU/ml).

[a] Blue = variable adjusted; S = sex; A = age; B = BMI or obesity; L = lifestyle (physical activity, smoking, or drinking); BP = blood pressure (systolic); LI = lipid profile (HDL cholesterol, LDL cholesterol, or triglycerides); SE = socioeconomic factors (education, assets owned, water within the home, flush toilet, or urban region); BI = biomarkers (zonula occludens-1). [b] %. [c] Independent variable. [d] Cross sectional data collection. [e] LPS activity = LPS/HDL cholesterol. [f] mmol/L. [g] LPS log transformed. [h] Serum. [i] Glucose log transformed. [j] Plasma or serum unspecified. [k] β [95% CI]. [l] mg/dl. [m] pmol/L. [n] Insulin log transformed. [o] IU/ml. [p] Measured in control group only. [q] pg/ml. [r] HOMA-IR = FPI (mU/L) x FPG (mmol/L)/22.5 (Matthews et al., 1985; Wallace et al., 2004). [s] QUICKI = 1/(log(insulin) + log(glucose)). [t] Dependent variable. [u] T2DM defined as FPG ≥ 7.0 mmol/L (WHO & IDF, 2006), which differs from original group criteria. [v] Mean difference [95% CI]. [w] Cross sectional. [x] T2 versus T1. [y] *p* for trend. [z] T3 versus T1. [aa] Baseline measurement within a cohort. [bb] LPS quartile ranges unspecified. [cc] Q2 versus Q1. [dd] Q3 versus Q1. [ee] Q4 versus Q1. [ff] Per unit increase.

Pussinen et al., 2011). One showed no association between these variables once lipid profile and biomarkers, specifically ZO-1, were added to the model (Jayashree et al., 2014). Those that demonstrated a positive relationship included one that examined the mean difference between groups and found an association between LPS and type 2 diabetes after adjusting for age, obesity, and socioeconomic factors (Hawkesworth et al., 2013). Two showed a strong, positive association with the highest LPS levels corresponding to a 2.2- to 3.3-fold greater prevalence of type 2 diabetes, respectively, after adjusting for sex and age (quartile 4 (Q4) to quartile 1 (Q1): odds ratio (OR) [95% confidence interval (CI)] = 2.232 [1.689-2.949], $p < 0.001$; Pussinen et al., 2011) or for sex, age, BMI, lifestyle, blood pressure, lipid profile, and socioeconomic factors (tertile 3 (T3) to tertile 1 (T1): OR [95% CI] = 3.28 [2.18-4.93], p for trend < 0.001; Huang et al., 2019). However, the former used baseline measurements of LPS and type 2 diabetes within a cohort. Those with the condition were excluded from further analysis making it difficult to determine if any confounding variables were present at a significant level whereas the latter compared and adjusted for a wider range of factors.

B. Prospective Cohort Studies Examining the Relationship Between LPS and Type 2 Diabetes

Two studies included in this systematic literature review provided information on LPS and the development of type 2 diabetes from prospective cohorts (Camargo et al., 2019; Pussinen et al., 2011). One was discussed previously, although with limited data, since it also provided cross sectional measures of LPS and type 2 diabetes prevalence at the outset of the study (Pussinen et al., 2011). This cohort had a median follow-up period of 10.8 years (Pussinen et al.,

2011) while the other had a median follow-up period of five years (Camargo et al., 2019).

1. Descriptive Statistics

Descriptive statistics of the participants in these studies are summarized in Table 4. Details regarding type 2 diabetes diagnosis method, setting, sample size, sex, age, BMI, total, HDL and LDL cholesterol, as well as triglycerides were available from both whereas details regarding family history of type 2 diabetes, dietary macronutrients, physical inactivity, smoking, systolic and diastolic blood pressure, as well as hypertension were provided by only one. Sample sizes for these two studies were 462 and 6,632 participants, and neither used self-reported methods to establish a type 2 diabetes diagnosis. However, one study chose testing to identify individuals with incident diabetes (Camargo et al., 2019) whereas the other did not (Pussinen et al., 2011).

2. Measurement of LPS

Participants provided both fasting and semifasting samples in one study (Camargo et al., 2019) whereas only semifasting samples were provided in the other (Pussinen et al., 2011). These were taken four hours after consuming a standardized mixed meal with fat content adjusted by body weight in the former and at a range of times in the latter (median = 5; interquartile range = 3-7 hours) with the only instructions being to avoid heavy meals. LPS levels were determined using either plasma or serum by endpoint or kinetic chromogenic LAL assay. These details are presented in Table 5 along with measurements of LPS and type 2 diabetes-related indices.

In regard to LPS, the two studies provided contradictory results. One found no

Table 4

Descriptive Statistics From Prospective Cohort Studies Examining the Relationship Between LPS and T2DM

Study	Groups	T2DM diagnosis method [a]	Setting [b]	*n*	Sex, *n* (%)		Age (years)	Family history, *n* (%)
					Male	Female		
Camargo et al., 2019 [c]	Control	Testing [d]	Research	355	302 (85.1)	53 (14.9)	57.33 ± 0.50 [e]	
	T2DM	Testing [d]	Research	107	87 (81.3)	20 (18.7)	58.75 ± 0.87 [e]	
							$p = 0.171$	
Pussinen et al., 2011 [h]	Control	Physician-diagnosed	Community	6,170	2,962 (50.5)	2,900 (49.5)	53.2 ± 11.0	1,278 (20.7) [i]
	T2DM	Physician-diagnosed	Community	462	261 (60.1)	173 (39.9)	57.3 ± 9.4	129 (27.9) [i]
					$p < 0.001$		$p < 0.001$	$p < 0.001$

Note. Data presented as *M* ± *SD* or *Mdn* (IQ) unless otherwise indicated. BMI = body mass index; HDL = high-density lipoprotein; IQ = interquartile range; LDL = low-density lipoprotein; LF = low-fat diet intervention; LPS = lipopolysaccharide; MED = mediterranean diet intervention; *SEM* = standard error of the mean; T2DM = type 2 diabetes mellitus.

[a] Physician-diagnosed = physician or medical record indicated T2DM diagnosis, medication or both; testing = measurement of T2DM-related indices; self-reported = T2DM status provided by participants. [b] Community = general population; hospital = outpatient clinic; research = participants from existing studies. [c] Median follow-up = 5 years. [d] (ADA, 2011). [e] *M* ± *SEM*. [f] Nonsignificant difference in LPS between LF and MED. [g] mg/dl. [h] Median follow-up = 10.8 years. [i] Mother or father. [j] Level 1 on a scale of 1 to 4 from a self-administered questionnaire. [k] Never = never smoked regularly; former = smoked regularly but quit ≥ 1 month ago; current = smoking regularly for ≥ 1 year and within the previous month (Vartiainen et al., 2010). [l] Mean of two measurements. [m] Hypertension defined as systolic or diastolic BP > 140 or 90 mm Hg, respectively, or antihypertensive medication. [n] Semifasting. [o] Log transformed. [p] Calculated fasting triglycerides (mmol/L) = 1.34 ± 0.85 and 2.01 ± 1.18 for control and T2DM, respectively ($p < 0.001$); log transformed.

Study	Groups	BMI (kg/m²)	Dietary macronutrients (%)			Physical inactivity, *n* (%)	Smoking, *n* (%)		
			Protein	Fat	Carbohydrate		Never	Former	Current
Camargo et al., 2019 [c]	Control	29.88 ± 0.22 [e]	15	LF < 30 MED ≥ 35 [f]	LF ≥ 55 MED ≤ 50 [f]				
	T2DM	31.39 ± 0.47 [e]							
		$p = 0.002$							
Pussinen et al., 2011 [h]	Control	26.7 ± 4.1				1,271 (20.6) [j]	3,321 (53.8) [k]	1,507 (24.4) [k]	1,342 (21.8) [k]
	T2DM	31.6 ± 5.2				169 (36.6) [j]	234 (50.6) [k]	117 (25.3) [k]	111 (24.0) [k]
		$p < 0.001$				$p < 0.001$		$p = 0.374$	

Note. Data presented as *M* ± *SD* or *Mdn* (IQ) unless otherwise indicated. BMI = body mass index; HDL = high-density lipoprotein; IQ = interquartile range; LDL = low-density lipoprotein; LF = low-fat diet intervention; LPS = lipopolysaccharide; MED = mediterranean diet intervention; *SEM* = standard error of the mean; T2DM = type 2 diabetes mellitus.

[a] Physician-diagnosed = physician or medical record indicated T2DM diagnosis, medication or both; testing = measurement of T2DM-related indices; self-reported = T2DM status provided by participants. [b] Community = general population; hospital = outpatient clinic; research = participants from existing studies. [c] Median follow-up = 5 years. [d] (ADA, 2011). [e] *M* ± *SEM*. [f] Nonsignificant difference in LPS between LF and MED. [g] mg/dl. [h] Median follow-up = 10.8 years. [i] Mother or father. [j] Level 1 on a scale of 1 to 4 from a self-administered questionnaire. [k] Never = never smoked regularly; former = smoked regularly but quit ≥ 1 month ago; current = smoking regularly for ≥ 1 year and within the previous month (Vartiainen et al., 2010). [l] Mean of two measurements. [m] Hypertension defined as systolic or diastolic BP > 140 or 90 mm Hg, respectively, or antihypertensive medication. [n] Semifasting. [o] Log transformed. [p] Calculated fasting triglycerides (mmol/L) = 1.34 ± 0.85 and 2.01 ± 1.18 for control and T2DM, respectively ($p < 0.001$); log transformed.

Study	Groups	Blood pressure (mm Hg)		Hypertension, *n* (%)	Cholesterol (mmol/L)			Triglycerides (mmol/L)
		Systolic	Diastolic		Total	HDL	LDL	
Camargo et al., 2019 [c]	Control				160.65 ± 1.62 [e,g]	44.58 ± 0.53 [e,g]	91.10 ± 1.33 [e,g]	119.45 ± 3.24 [e,g]
	T2DM				164.97 ± 3.41 [e,g]	43.52 ± 1.04 [e,g]	93.40 ± 2.66 [e,g]	132.60 ± 6.60 [e,g]
					$p = 0.217$	$p = 0.355$	$p = 0.421$	$p = 0.059$
Pussinen et al., 2011 [h]	Control	138.3 ± 20.4 [l]	83.8 ± 10.8 [l]	3,080 (50.0) [m]	5.7 ± 1.0	1.40 ± 0.36	3.61 ± 0.91	1.48 ± 0.95 [n,o,p]
	T2DM	148.3 ± 20.0 [l]	88.7 ±11.5 [l]	341 (73.8) [m]	5.9 ± 1.1	1.21 ± 0.33	3.81 ± 0.94	2.19 ± 1.30 [n,o,p]
		$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$

Note. Data presented as *M* ± *SD* or *Mdn* (IQ) unless otherwise indicated. BMI = body mass index; HDL = high-density lipoprotein; IQ = interquartile range; LDL = low-density lipoprotein; LF = low-fat diet intervention; LPS = lipopolysaccharide; MED = mediterranean diet intervention; *SEM* = standard error of the mean; T2DM = type 2 diabetes mellitus.

[a] Physician-diagnosed = physician or medical record indicated T2DM diagnosis, medication or both; testing = measurement of T2DM-related indices; self-reported = T2DM status provided by participants. [b] Community = general population; hospital = outpatient clinic; research = participants from existing studies. [c] Median follow-up = 5 years. [d] (ADA, 2011). [e] *M* ± *SEM*. [f] Nonsignificant difference in LPS between LF and MED. [g] mg/dl. [h] Median follow-up = 10.8 years. [i] Mother or father. [j] Level 1 on a scale of 1 to 4 from a self-administered questionnaire. [k] Never = never smoked regularly; former = smoked regularly but quit ≥ 1 month ago; current = smoking regularly for ≥ 1 year and within the previous month (Vartiainen et al., 2010). [l] Mean of two measurements. [m] Hypertension defined as systolic or diastolic BP > 140 or 90 mm Hg, respectively, or antihypertensive medication. [n] Semifasting. [o] Log transformed. [p] Calculated fasting triglycerides (mmol/L) = 1.34 ± 0.85 and 2.01 ± 1.18 for control and T2DM, respectively ($p < 0.001$); log transformed.

Table 5

Measurements of LPS and T2DM-Related Indices From Prospective Cohort Studies Examining the Relationship Between LPS and T2DM

Study	Groups	Fasting	Sample type	Assay method	LPS (EU/ml)	Baseline ΔLPS	3 year ΔLPS [a]	HbA1C (%)	FPG (mmol/L)
Camargo et al., 2019 [g]	Control	Fasting and semifasting [h]	Plasma	Endpoint chromogenic LAL [i]	— [j,k]	— [k,l,m,n]	— [k,l,m,n]	5.86 ± 0.02 [o]	92.59 ± 0.53 [o,p]
	T2DM				— [j,k]	— [k,m,s,t]	— [k,m,n,u]	6.03 ± 0.03 [o]	96.18 ± 1.04 [o,p]
					NS	$p = 0.925$	$p = 0.252$	$p < 0.001$	$p = 0.002$
Pussinen et al., 2011 [v]	Control	Semifasting [w]	Serum	Kinetic chromogenic LAL	61.06 ± 36.11 [k]				5.06 ± 0.63 [q,x,y]
	T2DM				77.03 ± 42.03				5.84 ± 1.92 [q,x,y]
					$p < 0.001$				$p < 0.001$

Note. Data presented as $M \pm SD$ or *Mdn* (IQ) unless otherwise indicated. 2 hr PG = 2 hr plasma glucose; BMI = body mass index; DI = disposition index; EU = endotoxin unit; FPG = fasting plasma glucose; FPI = fasting plasma insulin; HbA1C = hemoglobin A1C; HOMA-B = homeostasis model assessment of beta cell function; HOMA-IR = homeostasis model assessment of insulin resistance; IFG = impaired fasting glucose; IGI = insulinogenic index; IQ = interquartile range; ISI = insulin sensitivity index; LAL = limulus amoebocyte lysate; LPS = lipopolysaccharide; NS = non-significant at $p < 0.05$; OGTT = oral glucose tolerance test; QUICKI = quantitative insulin sensitivity check index; *SEM* = standard error of the mean; T2DM = type 2 diabetes mellitus.

[a] ΔLPS = change in LPS between fasting and semifasting. [b] Testing performed at baseline and participants met at least one T2DM diagnostic criteria from Figure 3. [c] HOMA-IR = fasting insulin (μIU/ml) x fasting glucose (mmol/ml)/22.5 (Blanco-Rojo et al., 2016; Song et al., 2007). [d] ISI = 10,000/√(fasting insulin (pmol/L) x fasting glucose (mmol/L) x mean OGTT insulin (pmol/L) x mean OGTT glucose (mmol/L)) (Blanco-Rojo et al., 2016). [e] DI = ISI x ($AUC_{30\ min\ insulin}/AUC_{30\ min\ glucose}$), where $AUC_{30\ min}$ is the area under the curve between baseline and 30 min of the OGTT for insulin (pmol/L) and glucose (mmol/L), respectively, calculated by the trapezoidal method (Blanco-Rojo et al., 2016). [f] IGI = (30 min insulin (pmol/L) - fasting insulin (pmol/L))/(30 min glucose (mmol/L) - fasting glucose (mmol/L)) (Blanco-Rojo et al., 2016). [g] Median follow-up = 5 years. [h] Semifasting samples taken 4 hours after consuming a standardized mixed meal with fat content adjusted by body weight. [i] (Lonza, n.d.-b). [j] Fasting. [k] Log transformed. [l] Within control group at 5 years. [m] Sex-, age-, and BMI-adjusted. [n] p within group = NS. [o] $M \pm SEM$. [p] mg/dl. [q] Serum. [r] mU/L. [s] Within T2DM group at 5 years. [t] p within group < 0.05. [u] Within T2DM group at 3 years ($n = 78$ out of 107). [v] Median follow-up = 10.8 years. [w] Participants asked to fast for 4 hr and avoid heavy meals; *Mdn* (IQ) = 5 (3 - 7) hr; no correlation between fasting time and log LPS; additional details about meals unknown. [x] Semifasting. [y] $n = 5,691$.

Study	Groups	2 hr PG (mmol/L)	IFG (%)	T2DM (%) [b]	FPI (pmol/L)	Measures of insulin resistance			Measures of insulin secretion		
						HOMA-IR [c]	ISI [d]	QUICKI	DI [e]	HOMA-B	IGI [f]
Camargo et al., 2019 [g]	Control			0.0	8.34 ± 0.31 [o,q,r]	2.58 ± 0.09 [o]	4.32 ± 0.14 [o]		1.03 ± 0.03 [o]		1.08 ± 0.06 [o]
	T2DM			0.0	10.51 ± 0.66 [o,q,r]	3.37 ± 0.30 [o]	3.35 ± 0.20 [o]		0.83 ± 0.05 [o]		0.64 ± 0.30 [o]
					$p = 0.001$	$p = 0.001$	$p = 0.001$		$p = 0.003$		$p = 0.025$
Pussinen et al., 2011 [v]	Control										
	T2DM										

Note. Data presented as $M \pm SD$ or *Mdn* (IQ) unless otherwise indicated. 2 hr PG = 2 hr plasma glucose; BMI = body mass index; DI = disposition index; EU = endotoxin unit; FPG = fasting plasma glucose; FPI = fasting plasma insulin; HbA1C = hemoglobin A1C; HOMA-B = homeostasis model assessment of beta cell function; HOMA-IR = homeostasis model assessment of insulin resistance; IFG = impaired fasting glucose; IGI = insulinogenic index; IQ = interquartile range; ISI = insulin sensitivity index; LAL = limulus amoebocyte lysate; LPS = lipopolysaccharide; NS = non-significant at $p < 0.05$; OGTT = oral glucose tolerance test; QUICKI = quantitative insulin sensitivity check index; *SEM* = standard error of the mean; T2DM = type 2 diabetes mellitus.

[a] ΔLPS = change in LPS between fasting and semifasting. [b] Testing performed at baseline and participants met at least one T2DM diagnostic criteria from Figure 3. [c] HOMA-IR = fasting insulin (μIU/ml) x fasting glucose (mmol/ml)/22.5 (Blanco-Rojo et al., 2016; Song et al., 2007). [d] ISI = 10,000/√(fasting insulin (pmol/L) x fasting glucose (mmol/L) x mean OGTT insulin (pmol/L) x mean OGTT glucose (mmol/L)) (Blanco-Rojo et al., 2016). [e] DI = ISI x ($AUC_{30\ min\ insulin}/AUC_{30\ min\ glucose}$), where $AUC_{30\ min}$ is the area under the curve between baseline and 30 min of the OGTT for insulin (pmol/L) and glucose (mmol/L), respectively, calculated by the trapezoidal method (Blanco-Rojo et al., 2016). [f] IGI = (30 min insulin (pmol/L) - fasting insulin (pmol/L))/(30 min glucose (mmol/L) - fasting glucose (mmol/L)) (Blanco-Rojo et al., 2016). [g] Median follow-up = 5 years. [h] Semifasting samples taken 4 hours after consuming a standardized mixed meal with fat content adjusted by body weight. [i] (Lonza, n.d.-b). [j] Fasting. [k] Log transformed. [l] Within control group at 5 years. [m] Sex-, age-, and BMI-adjusted. [n] *p* within group = NS. [o] $M \pm SEM$. [p] mg/dl. [q] Serum. [r] mU/L. [s] Within T2DM group at 5 years. [t] *p* within group < 0.05. [u] Within T2DM group at 3 years ($n = 78$ out of 107). [v] Median follow-up = 10.8 years. [w] Participants asked to fast for 4 hr and avoid heavy meals; *Mdn* (IQ) = 5 (3 - 7) hr; no correlation between fasting time and log LPS; additional details about meals unknown. [x] Semifasting. [y] $n = 5,691$.

significant differences between those with type 2 diabetes and controls, which included comparisons of fasting LPS levels at baseline as well as the change in LPS levels between fasting and semifasting at baseline and three years into the study (Camargo et al., 2019). However, when a comparison was made between baseline measurements of fasting and semifasting LPS levels within each group, there was a significant increase in LPS following the standardized mixed meal after adjusting for sex, age, and BMI in those who developed type 2 diabetes at 5 years that was not present in controls at 5 years. A similar assessment between three year measurements of fasting and semifasting LPS levels after the standardized mixed meal did not find a significant difference in those who developed type 2 diabetes at 3 years (n = 87 out of 107). In comparison, the other study showed significantly higher semifasting LPS levels in the type 2 diabetes group as compared to controls (Pussinen et al., 2011).

3. Measurement of Type 2 Diabetes-Related Indices

Details regarding fasting plasma glucose were provided by two studies while those regarding HbA1C, type 2 diabetes prevalence, fasting plasma insulin, HOMA-IR, ISI, DI, and IGI were each provided by one. No studies reported information about two-hour plasma glucose, impaired fasting glucose, QUICKI, and HOMA-B.

4. Associations Between LPS and Type 2 Diabetes-Related Indices

Statistical analysis of the associations between LPS and various type 2 diabetes-related indices, as summarized in Table 6, revealed both studies showed a moderate to strong, positive association between the highest LPS levels (Pussinen et al., 2011), or

Table 6

Prospective Cohort Associations Between LPS and T2DM-Related Indices

Variable	Study	Positive association	No association	*p*		Adjusted for [a]									
						S	A	F	B	L	BP	LI	SE	BI	T2
T2DM [b]	Camargo et al., 2019 [c]		1.752 [0.966 - 3.177] [d,e,f,g]	NS											
		2.074 [1.147 - 3.747] [d,e,g,h]		< 0.05											
	Pussinen et al., 2011 [i]		1.073 [0.768 - 1.498] [e,j]	0.681											
		1.703 [1.255 - 2.309] [e,l]		0.001	< 0.001 [k]										
		2.751 [2.071 - 3.654] [e,m]		< 0.001											
			0.939 [0.652 - 1.351] [e,j,n]	0.733											
			1.233 [0.885 - 1.718] [e,l,n]	0.216	0.012 [k]										
		1.518 [1.090 - 2.114] [e,m,n]		0.013											
		1.008 [1.006 - 1.010] [e,n,o]		< 0.001											
		1.004 [1.001 - 1.007] [e,.n,o]		0.019											

Note. LPS measured in EU/ml. Data presented as HR [95% CI] unless otherwise indicated. BMI = body mass index; CI = confidence interval; EU = endotoxin unit; HbA1C = hemoglobin A1C; HDL = high-density lipoprotein; HR = hazard ratio; ISI = insulin sensitivity index; LPS = lipopolysaccharide; NS = non-significant at $p < 0.05$; Q1 = quartile 1 (LPS = 2.40 - 38.10 EU/ml); Q2 = quartile 2 (LPS = 38.20 - 54.10 EU/ml); Q3 = quartile 3 (LPS = 54.20 - 77.0 EU/ml); Q4 = quartile 4 (LPS = 77.10 - 475.8 EU/ml); T1 = tertile 1; T2 = tertile 2; T2DM = type 2 diabetes mellitus; T3 = tertile 3. [a] Blue = variable adjusted; S = sex; A = age; F = family history; B = BMI; L = lifestyle (diet, physical activity, or smoking); BP = blood pressure (hypertension); LI = lipid profile (total cholesterol, HDL cholesterol, or triglycerides); SE = socioeconomic factors (education); BI = biomarkers (gamma-glutamyltransferase or C-reactive protein); T2 = T2DM-related indices (HbA1C or ISI). [b] Dependent variable. [c] Median follow-up = 5 years. [d] Association between baseline LPS fold change = semifasting LPS/fasting LPS and T2DM at 5 years. [e] LPS log transformed. [f] T2 versus T1. [g] LPS fold change tertile ranges unspecified. [h] T3 versus T1. [i] Median follow-up = 10.8 years. [j] Q2 versus Q1. [k] *p* for trend. [l] Q3 versus Q1. [m] Q4 versus Q1. [n] Adding semifasting serum glucose (n = 5,691) or replacing semifasting with calculated fasting triglycerides did not notably affect the results. [o] Per unit increase.

highest change in LPS levels (Camargo et al., 2019), corresponding to a 1.5- to 2.1-fold higher risk of incident type 2 diabetes, respectively, after adjusting for sex, age, family history, BMI, lifestyle, blood pressure, lipid profile, socioeconomic factors, and biomarkers, specifically gamma-glutamyltransferase and CRP, in the former (Q4 to Q1: hazard ratio (HR) [95% CI] = 1.518 [1.090-2.114], $p = 0.013$, p for trend = 0.012) or for sex, age, BMI, lifestyle, lipid profile, and type 2 diabetes-related indices, specifically HbA1C and ISI, in the latter (T3 to T1: HR [95% CI] = 2.074 [1.147-3.747], $p < 0.05$).

C. Studies Examining the Relationship Between 16S rRNA and Type 2 Diabetes

As the method of using 16S rRNA to identify bacteria was developed more recently, only three studies in this systematic literature review utilized this approach to examine the relationship between bacteria in the blood and the development of type 2 diabetes (Amar et al., 2011b; Qui et al., 2019; Sato et al., 2014). Each adopted a different study design: one looked at a prospective cohort with a follow-up period of 9 years (Amar et al., 2011b), another chose a nested case-control design with a follow-up period of 2 years (Qui et al., 2019), and the final approach analyzed data from a cross sectional study (Sato et al., 2014).

1. Descriptive Statistics

Descriptive statistics of the participants in these studies are summarized in Table 7. All provided data on type 2 diabetes diagnosis method, setting, sample size, sex, age, BMI, total cholesterol, and triglycerides. Information about smoking, hypertension, and HDL cholesterol were reported by two studies while family history of type 2 diabetes, dietary macronutrients, systolic and diastolic blood pressure, as well as LDL cholesterol

Table 7

Descriptive Statistics From Studies Examining the Relationship Between 16S rRNA and T2DM

Study	Groups	T2DM diagnosis method [a]	Setting [b]	*n*	Sex, *n* (%)		Age (years)	Family history, *n* (%)
					Male	Female		
Amar et al., 2011b [c,d]	Control	Physician-diagnosed or testing	Research	3,149	1,516 (48)	1,633 (52)	47 ± 10	574 (18)
	T2DM	Physician-diagnosed or testing	Research	131	98 (75)	33 (25)	51 ± 9	30 (23)
					$p = 0.0001$		$p = 0.0001$	$p = 0.2$
Qui et al., 2019 [i,j]	Control	Physician-diagnosed	Research	100	64 (64) [k]	36 (36) [k]	51.98 ± 8.05 [k]	
	T2DM	Physician-diagnosed	Research	50	35 (70) [k]	15 (30) [k]	51.64 ± 6.18 [k]	
					$p = 0.465$		$p = 0.775$	
Sato et al., 2014 [n]	Control	Testing	Hospital	50	26 (52)	24 (48)	60.2 ± 12.9	
	T2DM	Physician-diagnosed	Hospital	50	26 (52)	24 (48)	62.5 ± 10.8	
					NS		NS	

Note. Data presented as $M \pm SD$ or *Mdn* (IQ) unless otherwise indicated. 16S rRNA = 16S ribosomal ribonucleic acid; BDHQ = brief-type self-administered diet history questionnaire; BMI = body mass index; BP = blood pressure; HDL = high-density lipoprotein; IQ = interquartile range; LDL = low-density lipoprotein; NS = non-significant at $p < 0.05$; T2DM = type 2 diabetes mellitus. [a] Physician-diagnosed = physician or medical record indicated T2DM diagnosis, medication or both; testing = measurement of T2DM-related indices; self-reported = T2DM status provided by participants. [b] Community = general population; hospital = outpatient clinic; research = participants from existing studies. [c] Prospective cohort. [d] Follow-up = 9 years. [e] Former smokers not included. [f] Hypertension defined as systolic or diastolic BP ≥ 140 or 90 mm Hg, respectively, or antihypertensive medication. [g] $n = 3,242$. [h] Log transformed. [i] Nested case-control. [j] Follow-up = 2 years. [k] Sex- and age-matched. [l] Unclear whether former smokers were included. [m] Hypertension criteria unspecified. [n] Cross sectional data collection. [o] Dietary habits in the preceding month assessed by BDHQ. [p] mg/dl.

Study	Groups	BMI (kg/m²)	Dietary macronutrients (%)			Physical inactivity, *n* (%)	Smoking, *n* (%)
			Protein	Fat	Carbohydrate		
Amar et al., 2011b [c,d]	Control	23.8 ± 2.8					586 (19) [e]
	T2DM	25.8 ± 2.5					36 (27) [e]
		p = 0.0001					*p* = 0.01
Qui et al., 2019 [i,j]	Control	23.42 ± 2.93 [k]					45 (45) [k,l]
	T2DM	25.14 ± 2.88 [k]					21 (42) [k,l]
		p = 0.001					*p* = 0.727
Sato et al., 2014 [n]	Control	21.7 (20.9 - 23.5)	17.5 ± 3.0 [o]	28.6 ± 5.1 [o]	53.9 ± 6.6 [o]		
	T2DM	25.5 (23.5 - 30.8)	17.1 ± 3.5 [o]	26.8 ± 5.7 [o]	56.1 ± 7.9 [o]		
		p < 0.01	NS	NS	NS		

Note. Data presented as *M* ± *SD* or *Mdn* (IQ) unless otherwise indicated. 16S rRNA = 16S ribosomal ribonucleic acid; BDHQ = brief-type self-administered diet history questionnaire; BMI = body mass index; BP = blood pressure; HDL = high-density lipoprotein; IQ = interquartile range; LDL = low-density lipoprotein; NS = non-significant at $p < 0.05$; T2DM = type 2 diabetes mellitus. [a] Physician-diagnosed = physician or medical record indicated T2DM diagnosis, medication or both; testing = measurement of T2DM-related indices; self-reported = T2DM status provided by participants. [b] Community = general population; hospital = outpatient clinic; research = participants from existing studies. [c] Prospective cohort. [d] Follow-up = 9 years. [e] Former smokers not included. [f] Hypertension defined as systolic or diastolic BP ≥ 140 or 90 mm Hg, respectively, or antihypertensive medication. [g] *n* = 3,242. [h] Log transformed. [i] Nested case-control. [j] Follow-up = 2 years. [k] Sex- and age-matched. [l] Unclear whether former smokers were included. [m] Hypertension criteria unspecified. [n] Cross sectional data collection. [o] Dietary habits in the preceding month assessed by BDHQ. [p] mg/dl.

Study	Groups	Blood pressure (mm Hg)		Hypertension, *n* (%)	Cholesterol (mmol/L)			Triglycerides (mmol/L)
		Systolic	Diastolic		Total	HDL	LDL	
Amar et al., 2011b [c,d]	Control			984 (31) [f]	5.69 ± 0.97			0.90 (0.65 - 1.28) [g,h]
	T2DM			71 (54) [f]	6.01 ± 1.10			1.29 (0.97 - 1.78) [g,h]
				$p = 0.0001$	$p = 0.0002$			
Qui et al., 2019 [i,j]	Control	119.68 ± 14.50 [k]	77.96 ± 9.82 [k]	6 (6) [k,m]	4.54 ± 0.87 [k]	1.17 (0.95 - 1.48) [k]	2.80 (2.32 - 3.60) [k]	1.26 (0.93 - 1.79) [k]
	T2DM	128.64 ± 20.61 [k]	82.78 ± 12.94 [k]	12 (24) [k,m]	4.92 ± 0.97 [k]	0.99 (0.84 - 1.35) [k]	3.16 (2.49 - 3.60) [k]	2.04 (1.34 - 3.09) [k]
		$p = 0.007$	$p = 0.023$	$p = 0.001$	$p = 0.017$	$p = 0.013$	$p = 0.057$	$p < 0.001$
Sato et al., 2014 [n]	Control				212.9 ± 28.1 [p]	61.5 ± 16.3 [p]		109.5 ± 83.3 [p]
	T2DM				190.3 ± 45.5 [p]	46.8 ± 13.9 [p]		124.9 ± 59.1 [p]
					$p < 0.01$	$p < 0.01$		$p < 0.05$

Note. Data presented as $M \pm SD$ or *Mdn* (IQ) unless otherwise indicated. 16S rRNA = 16S ribosomal ribonucleic acid; BDHQ = brief-type self-administered diet history questionnaire; BMI = body mass index; BP = blood pressure; HDL = high-density lipoprotein; IQ = interquartile range; LDL = low-density lipoprotein; NS = non-significant at $p < 0.05$; T2DM = type 2 diabetes mellitus.
[a] Physician-diagnosed = physician or medical record indicated T2DM diagnosis, medication or both; testing = measurement of T2DM-related indices; self-reported = T2DM status provided by participants. [b] Community = general population; hospital = outpatient clinic; research = participants from existing studies. [c] Prospective cohort. [d] Follow-up = 9 years. [e] Former smokers not included. [f] Hypertension defined as systolic or diastolic BP ≥ 140 or 90 mm Hg, respectively, or antihypertensive medication. [g] $n = 3{,}242$. [h] Log transformed. [i] Nested case-control. [j] Follow-up = 2 years. [k] Sex- and age-matched. [l] Unclear whether former smokers were included. [m] Hypertension criteria unspecified. [n] Cross sectional data collection. [o] Dietary habits in the preceding month assessed by BDHQ. [p] mg/dl.

were reported by one. No studies provided details regarding physical inactivity. Sample sizes ranged from 100 to 3,280 participants, and none used self-reported methods to establish a type 2 diabetes diagnosis. These three studies instead used testing, physician-diagnosed, or a combination to determine comparison groups.

2. Measurement of 16S rRNA

Participants provided fasting samples and either plasma or peripheral blood leukocytes were used to assess 16S rRNA by targeting the gene encoding it (16S rDNA) or 16S rRNA directly. The assay methods varied and included qPCR, reverse transcription qPCR (RT-qPCR), and next-generation sequencing. These details are presented in Table 8 along with measurements of 16S rRNA and type 2 diabetes-related indices. Overall, two of the three studies showed significantly higher 16S rDNA or 16S rRNA in the type 2 diabetes group as compared to controls by either measuring the concentration (Amar et al., 2011b) or determining the rate of detection in participants (Sato et al., 2014). However, one used peripheral blood leukocytes to do so, and there was a significant difference in leukocyte count between the two groups (control: 6.3 ± 1.7 x 10^9/L, type 2 diabetes: 6.7 ± 2.0 x 10^9/L, $p = 0.002$, $n = 3{,}242$) as well as an unadjusted, positive association between 16S rDNA and leukocyte count ($r = 0.041$, $p = 0.02$, $n = 891$), which may have contributed to the results observed (Amar et al., 2011b). In addition, the second study used 21 group-, genus-, and species-specific primer sets to amplify bacterial sequences rather than universal primers (Sato et al., 2014). These would have identified many but not all bacteria. Still, by targeting 16S rRNA, these three studies provided more comprehensive information about bacterial presence than LPS-based

Table 8

Measurements of 16S rRNA and T2DM-Related Indices From Studies Examining the Relationship Between 16S rRNA and T2DM

Study	Groups	Fasting	Sample type	Assay method	16S rRNA, *n* (%)	16S rDNA (ng/µl)	HbA1C (%)	FPG (mmol/L)	2 hr PG (mmol/L)
Amar et al., 2011b [c,d]	Control	Fasting	Peripheral blood leukocytes [e,f]	qPCR		0.057 (0.032 - 0.120) [g,h]	5.39 ± 0.38	5.23 ± 0.49	
	T2DM					0.076 (0.042 - 0.141) [g,h]	5.83 ± 0.51	6.07 ± 0.55	
						$p = 0.04$	$p = 0.0001$	$p = 0.0001$	
Qui et al., 2019 [k,l]	Control	Fasting	Plasma	Next-generation sequencing		62,442 [m,n]		6.05 (5.23 - 6.28)	
	T2DM					60,007 [m,o]		6.12 (5.95 - 6.43)	
								$p = 0.453$	
Sato et al., 2014 [p]	Control	Fasting	Plasma	RT-qPCR	2 (4)		5.6 (5.4 - 5.8)	94.1 ± 12.3 [q,r]	
	T2DM				14 (28)		8.7 (8.0 - 9.5)	155.3 ± 44.7 [q,r]	
					$p < 0.01$		$p < 0.01$	$p < 0.01$	

Note. Data presented as *M* ± *SD* or *Mdn* (IQ) unless otherwise indicated. 2 hr PG = 2 hr plasma glucose; 16S rDNA = gene encoding 16S rRNA; 16S rRNA = 16S ribosomal ribonucleic acid; DI = disposition index; FPG = fasting plasma glucose; FPI = fasting plasma insulin; HbA1C = hemoglobin A1C; HOMA-B = homeostasis model assessment of beta cell function; HOMA-IR = homeostasis model assessment of insulin resistance; IFG = impaired fasting glucose; IGI = insulinogenic index; IQ = interquartile range; ISI = insulin sensitivity index; qPCR = quantitative polymerase chain reaction; QUICKI = quantitative insulin sensitivity check index; RT-qPCR = reverse transcription-qPCR; T2DM = type 2 diabetes mellitus.

[a] Testing performed at baseline and participants met at least one T2DM diagnostic criteria from Figure 3. [b] Computed using software downloaded from www.dtu.ox.ac.uk. [c] Prospective cohort. [d] Follow-up = 9 years. [e] Leukocyte count for control and T2DM was 6.3 ± 1.7 and 6.7 ± 2.0 x 10^9/L, respectively ($p = 0.002$); $n = 3,242$. [f] Unadjusted, positive association between log 16S rDNA and leukocyte count (g/L) ($r = 0.041$, $p = 0.02$); $n = 891$. [g] Glass microbeads were not used to improve extraction because isolation of bacterial DNA was not the intended use during sample preparation; validation of this method showed microbeads increased extraction 10-fold but the proportion between samples remained the same and allowed for comparison. [h] Log transformed. [i] Plasma or serum unspecified. [j] $n = 3,242$. [k] Nested case-control. [l] Follow-up = 2 years. [m] High-quality sequences per sample. [n] Total high-quality sequences = 6,244,227. [o] Total high-quality sequences = 3,000,391. [p] Cross sectional data collection. [q] Serum. [r] mg/dl.

Study	Groups	IFG (%)	T2DM (%) [a]	FPI (pmol/L)	Measures of insulin resistance			Measures of insulin secretion		
					HOMA-IR [b]	ISI	QUICKI	DI	HOMA-B [b]	IGI
Amar et al., 2011b [c,d]	Control			37.14 (27.46 - 50.62) [h,i,j]	0.97 ± 0.45				83.35 ± 25.73	
	T2DM			51.70 (36.28 - 79.52) [h,i,j]	1.41 ± 0.76				79.01 ± 32.59	
					$p = 0.0001$				$p = 0.0009$	
Qui et al., 2019 [k,l]	Control									
	T2DM									
Sato et al., 2014 [p]	Control		0.0							
	T2DM									

Note. Data presented as $M \pm SD$ or *Mdn* (IQ) unless otherwise indicated. 2 hr PG = 2 hr plasma glucose; 16S rDNA = gene encoding 16S rRNA; 16S rRNA = 16S ribosomal ribonucleic acid; DI = disposition index; FPG = fasting plasma glucose; FPI = fasting plasma insulin; HbA1C = hemoglobin A1C; HOMA-B = homeostasis model assessment of beta cell function; HOMA-IR = homeostasis model assessment of insulin resistance; IFG = impaired fasting glucose; IGI = insulinogenic index; IQ = interquartile range; ISI = insulin sensitivity index; qPCR = quantitative polymerase chain reaction; QUICKI = quantitative insulin sensitivity check index; RT-qPCR = reverse transcription-qPCR; T2DM = type 2 diabetes mellitus.

[a] Testing performed at baseline and participants met at least one T2DM diagnostic criteria from Figure 3. [b] Computed using software downloaded from www.dtu.ox.ac.uk. [c] Prospective cohort. [d] Follow-up = 9 years. [e] Leukocyte count for control and T2DM was 6.3 ± 1.7 and 6.7 ± 2.0 x 10^9/L, respectively ($p = 0.002$); $n = 3{,}242$. [f] Unadjusted, positive association between log 16S rDNA and leukocyte count (g/L) ($r = 0.041$, $p = 0.02$); $n = 891$. [g] Glass microbeads were not used to improve extraction because isolation of bacterial DNA was not the intended use during sample preparation; validation of this method showed microbeads increased extraction 10-fold but the proportion between samples remained the same and allowed for comparison. [h] Log transformed. [i] Plasma or serum unspecified. [j] $n = 3{,}242$. [k] Nested case-control. [l] Follow-up = 2 years. [m] High-quality sequences per sample. [n] Total high-quality sequences = 6,244,227. [o] Total high-quality sequences = 3,000,391. [p] Cross sectional data collection. [q] Serum. [r] mg/dl.

methods since they could and did detect gram-positive as well as gram-negative bacteria present in participants.

3. Measurement of Type 2 Diabetes-Related Indices

All studies provided information about fasting plasma glucose while two provided details regarding HbA1C. Fasting plasma insulin, HOMA-IR, and HOMA-B were each reported by one, with partial information about type 2 diabetes prevalence available from one study as well, whereas details regarding two-hour plasma glucose, impaired fasting glucose, ISI, QUICKI, DI, and IGI were not reported by any study.

4. Associations Between 16S rRNA and Type 2 Diabetes-Related Indices

Statistical analysis of the associations between 16S rRNA and various type 2 diabetes-related indices, as summarized in Table 9, revealed no association between 16S rDNA and fasting insulin, HOMA-B, and HOMA-IR and a negative, unadjusted association between 16S rDNA and fasting plasma glucose (Amar et al., 2011b). Only one study provided this data, and the analysis was performed in 27% of participants (n = 891 out of 3,280). In contrast, the relationship between 16S rDNA and type 2 diabetes was assessed in all individuals and showed a weak, positive association with a 1.35-fold higher risk of incident type 2 diabetes after adjusting for sex, age, family history, BMI, waist circumference, lifestyle, blood pressure, and type 2 diabetes-related indices, specifically fasting plasma glucose (standardized OR [95% CI] of incident type 2 diabetes for 1 standard deviation of log(16S rDNA) = 1.35 [1.11-1.64], p = 0.002). However, total cholesterol was not included in this model although it was significantly

Table 9

Associations Between 16S rRNA and T2DM-Related Indices

Variable	Study	Positive association	Negative association	No association	*p*	Adjusted for [a]						
						S	A	F	B	L	BP	T2
FPG [b,c]	Amar et al., 2011b		r = -0.04 [d,e]		0.02							
FPI [b,f]	Amar et al., 2011b			r = -0.0013 [d,e,g,h]	0.94							
HOMA-B [b]	Amar et al., 2011b			r = -0.0047 [d,e,i]	0.80							
HOMA-IR [b]	Amar et al., 2011b			r = -0.033 [d,e,i]	0.07							
T2DM [b]	Amar et al., 2011b	1.36 [1.08 - 1.72] [j,k]			0.009							
				1.15 [0.84 - 1.58] [k,l]	0.38							
				1.10 [0.87 - 1.39] [m]	0.45							
		1.61 [1.23 - 2.10] [n]			0.0005							
		1.29 [1.08 - 1.55]			—							
		1.35 [1.11 - 1.64]			0.002							

Note. Data presented as standardized OR [95% CI] of incident type 2 diabetes for 1 *SD* of log(16S rDNA) unless otherwise indicated. 16S rDNA = gene encoding 16S rRNA; 16S rRNA = 16S ribosomal ribonucleic acid; BMI = body mass index; CI = confidence interval; FPG = fasting plasma glucose; FPI = fasting plasma insulin; HOMA-B = homeostasis model assessment of beta cell function; HOMA-IR = homeostasis model assessment of insulin resistance; NS = non-significant at $p < 0.05$; OR = odds ratio; T2DM = type 2 diabetes mellitus.
[a] Blue = variable adjusted; S = sex; A = age; F = family history; B = BMI or waist circumference; L = lifestyle (smoking); BP = blood pressure (hypertension); T2 = T2DM-related indices (FPG). [b] Dependent variable. [c] mmol/L. [d] 16S rDNA log transformed. [e] $n = 891$. [f] pmol/L. [g] Plasma or serum unspecified. [h] Insulin log transformed. [i] Computed using software downloaded from www.dtu.ox.ac.uk. [j] FPG < 6.1 mmol/L. [k] p between FPG strata = NS. [l] FPG ≥ 6.1 mmol/L. [m] Incident type 2 diabetes within 0 - 3 years. [n] Incident type 2 diabetes within 6 - 9 years.

different between comparison groups.

An additional finding was a weak, positive association between 16S rDNA and type 2 diabetes in those without elevated fasting plasma glucose (< 6.1 mmol/L) that was not present in those with higher levels (standardized OR [95% CI] of incident type 2 diabetes for 1 standard deviation of log(16S rDNA) = 1.36 [1.08-1.72], $p = 0.009$; Amar et al., 2011b). Furthermore, the study showed a moderate, positive association between these variables in those who developed type 2 diabetes within six to nine years that was not present in those who developed it within zero to three years (standardized OR [95% CI] of incident type 2 diabetes for 1 standard deviation of log(16S rDNA) = 1.61 [1.23-2.10], $p = 0.0005$). Both of these measures were adjusted for sex, age, family history, BMI, waist circumference, lifestyle, and blood pressure with the latter value adjusted for type 2 diabetes-related indices as well, specifically fasting plasma glucose. The study also reported a C statistic of 0.564 for 16S rDNA and type 2 diabetes, which demonstrated it had limited ability to provide predictive discrimination of this outcome.

V. Discussion

While in numerous instances the data in this systematic literature review seem to indicate a relationship between LPS or 16S rRNA and type 2 diabetes may be present, details in the study design, population, and assessment of the exposure and outcome variables ultimately limit what conclusions may be drawn.

A. Study Design

One of the greatest difficulties in determining what effect the presence of bacteria or bacterial components in the blood, as measured by LPS or 16S rRNA, may have on the development of type 2 diabetes is the cross sectional nature of many of the studies (Al-Obaide et al., 2017; Creely et al., 2007; de Waal et al., 2018; Hawkesworth et al., 2013; Huang et al., 2019; Jayashree et al., 2014; Pussinen et al., 2011; Sato et al., 2014). Although this information allows us to assess the prevalence of type 2 diabetes in those with elevated levels of LPS or 16S rRNA and vice versa, one cannot infer causality as these are both measured at the same timepoint. Thus, determining with greater certainty whether one influences the other is only possible with a stronger, longitudinal study design.

Overall, and without taking other issues into consideration, it appears seven of the eight studies that provided cross sectional data found at least some indication that elevated LPS or 16S rRNA and type 2 diabetes are associated with each other at significantly higher levels than that seen in controls (Creely et al., 2007; de Waal et al., 2018; Hawkesworth et al., 2013; Huang et al., 2019; Jayashree et al., 2014; Pussinen et al., 2011; Sato et al., 2014). Three prospective cohorts (Amar et al., 2011b; Camargo et al., 2019; Pussinen et al., 2011) and one nested case-

control (Qui et al., 2019) were also included in the review. The three prospective cohorts had similar findings to the cross-sectional studies. Yet the reliability of these results remains in question given various drawbacks in the methodology used to explore this relationship.

B. Population

An additional consideration in assessing the quality of the results includes both sample size and characteristics of the populations in which these variables were measured. Of the eight studies that provided cross sectional data, six had sample sizes of 100 participants or less (Al-Obaide et al., 2017; Creely et al., 2007; de Waal et al., 2018; Hawkesworth et al., 2013; Jayashree et al., 2014; Sato et al., 2014). Of the remaining, only one used random sampling of the population to recruit study participants (Pussinen et al., 2011). However, this same study also provided no details about possible confounding variables as it was a baseline measurement of LPS and type 2 diabetes status within a cohort after which those with prevalent diabetes were excluded. Thus, not only do these results come from studies with a weaker study design that already have certain limitations, but their usefulness is further limited by none truly taking a large, random sample of the population and accounting for, or at least reporting, information about other variables known to affect this outcome.

This remains an issue in some of the prospective cohort and nested case-control studies. While the sample sizes were all larger than 100 participants, only one used random sampling of the population to recruit study participants, the same study described above that provided both cross sectional measures and data from a prospective cohort (Pussinen et al., 2011). In the case of the prospective cohort, the study provided thorough information about possible confounding variables, with the exception of dietary macronutrients. This makes the moderate, adjusted,

positive association between LPS and type 2 diabetes with a 1.5-fold increased risk the most reliable when taking into account study design, sample size, sampling method, and treatment of confounding variables (Q4 to Q1: HR [95% CI] = 1.518 [1.090-2.114], $p = 0.013$, p for trend = 0.012). The remaining prospective cohorts and nested case-control drew their participants from existing studies (Amar et al., 2011b; Camargo et al., 2019; Qui et al., 2019), one of which, for instance, included only coronary heart disease patients (Camargo et al., 2019). This approach restricts the ability to then make broader statements about what these results may mean for a population in general since most individuals at risk for type 2 diabetes do not have such a serious medical condition. Furthermore, none of these studies provided as thorough information about confounding variables leaving it unclear whether the relationship we see is truly related to LPS or 16S rRNA (Amar et al., 2011b; Camargo et al., 2019; Qui et al., 2019). Even the strongest study in this regard, by omitting data regarding diet as well as taking semifasting measurements at variable lengths of time after a nonstandardized meal, had a possible confounding variable in its population that was never addressed in addition to the variables related to diet and its immediate effect on LPS introduced during sampling that both could have prevented the true relationship from being seen (Pussinen et al., 2011). Therefore, given that no study within this systematic literature review fully accounted for all possible confounding variables as well as the preference for primarily weak study designs, small sample sizes, and convenience sampling means that despite whatever relationship the results may suggest, a determination as to whether LPS or 16S rRNA has a role in the development of type 2 diabetes cannot be made at this time without additional research.

While producing a cost-effective study is desirable, many variables affect this outcome. Designing a well thought-out study that ensures the results are representative of the population

and identifies other factors that may be influencing them is important. Thus, of additional relevance are the inclusion and exclusion criteria used to determine comparison groups since how we define who is healthy affects the comparison made. For instance, when a condition such as impaired liver function influences the removal of LPS (Nesseler et al., 2012) and elevated bilirubin interferes with the LAL assay (Gnauck et al., 2016), having studies that exclude for this (Al-Obaide et al., 2017; Sato et al., 2014) whereas others do not may introduce variability into the results seen. And this is only one example of many. Of particular concern is the fact that while all the studies paid careful attention to type 2 diabetes status, other forms of altered glucose metabolism such as impaired glucose tolerance or an elevated HbA1C indicative of prediabetes were handled differently. In one study, these individuals were included within the cohort (Camargo et al., 2019) whereas in a different study, those with impaired glucose tolerance were counted in the prevalent type 2 diabetes group and excluded from the cohort (Pussinen et al., 2011). Still another used the criteria of a HbA1C less than 6.0% to qualify as healthy, which excludes some but not all of those with prediabetes (Sato et al., 2014). Most, however, did not take these measures into account. But the changes that lead towards type 2 diabetes happen in a progressive manner, and fasting LPS can be elevated in other forms of altered glucose metabolism besides just type 2 diabetes, such as in impaired glucose tolerance (Harte et al., 2012). Thus, how these individuals get included or excluded and what group they belong to will affect the results obtained. Ultimately, a consensus on what constitutes a healthy individual is needed unless random sampling of the population is used without any exclusion. While this latter approach would eliminate selection bias and is preferable as we do not as yet understand this process fully and all of the conditions that may or may not influence it, it is often difficult to undertake and does not on its own fully resolve the question of who should be considered a

control. Incorporating both in one's methodology would ensure a truly well-designed study.

A final issue concerns the use of medication among participants. Only four studies provided any kind of information about the type 2 diabetes medications participants were taking (Al-Obaide et al., 2017; Creely et al., 2007; Hawkesworth et al., 2013; Sato et al., 2014) with two of these studies also providing some details regarding other medications (Al-Obaide et al., 2017; Sato et al., 2014). This information is relevant as the initiation of treatment for type 2 diabetes may affect gut microbial balance and permeability (Sato et al., 2017) as well as LPS levels (Al-Attas et al., 2009) whereas treatment for other conditions may affect the development of type 2 diabetes. This has been demonstrated with antidepressants (Pomytkin et al., 2015) and antihypertensive medications (Taylor et al., 2006). Thus, details regarding treatment for type 2 diabetes and other conditions should be considered essential information as they may affect both the exposure and outcome variables.

C. Assessment of LPS and 16S rRNA Levels

A possible flaw in the determination of LPS and 16S rRNA levels may be the decision to take fasting samples as opposed to postprandial or semifasting samples. Given that one of the proposed mechanisms for entry of LPS into the systemic circulation is via chylomicrons (Gnauck et al., 2016) and studies in both mice and humans have demonstrated an increase in LPS after a high-fat meal (Cani et al., 2007; Deopurkar et al., 2010), measurement of LPS and 16S rRNA in a fasted state may not be the opportune moment to examine this relationship. Only two studies used semifasting samples (Camargo et al., 2019; Pussinen et al., 2011), one of which had the novel approach of not only taking both fasting and semifasting samples but also assessing whether the change in levels from fasting to semifasting was significant within and between

groups (Camargo et al., 2019). While the results were inconsistent with some measures indicating a significant difference and others indicating none, this method is ideal for assessing a relationship if one is present. Furthermore, by providing a standardized mixed meal with fat content adjusted by body weight, the study helped to ensure that whatever differences were measured had more to do with the processes occurring within the individual than with the composition of the meal itself, since fat composition of the meal prior to sampling influences whether a rise in LPS will occur (Ghanim et al., 2009).

Diet and factors related to intestinal permeability may not be the only variables influencing this mechanism. Components in the blood whose role it is to bind and clear LPS also reduce the time it remains in circulation. This sequestration and neutralization occurs through binding to lipoproteins such as apoE (Kell & Pretorius, 2015) as well as through the action of LPS-binding protein and soluble CD14, which transfer LPS to monocytes and macrophages (Munford, 2005). Antibodies also inactivate LPS molecules until they can be removed (Munford, 2005) and interestingly, showed greater reproducibility after two weeks within the same individuals ($r = 0.82$, $n = 48$) than assessing LPS directly ($r = 0.2$, $n = 48$; Hawkesworth et al., 2013). Thus, unlike measures of type 2 diabetes risk such as BMI and HbA1C that remain stable over periods of time and allow for easy assessment in a clinical setting, the variability of LPS and 16S rRNA complicates its use in research and as a possible clinical screening tool.

Furthermore, numerous issues inherent in the LAL assay limit the usefulness of LPS in particular as this test was developed to identify bacterial contamination in pharmaceuticals and medical products, but not to test human samples (Gnauck et al., 2016). By adapting it for this purpose, it becomes important to be aware of interference from EDTA or clot activator in tubes as well as from biological components in plasma and serum. While dilution and heating may

resolve the latter, it cannot release sequestered LPS although some additional treatment steps have been investigated and may be able to do so. Of additional concern are the reaction kinetics with untreated or diluted serum initiating the reaction more quickly and with an altered reaction curve than standards prepared in water. In diluted and heated serum, the reaction is delayed but the reaction curve is not altered. Use of an endpoint LAL assay where the reaction is stopped and measured after a period of time would not identify these variations and either overestimate or underestimate LPS levels depending on pretreatment method. In comparison, a kinetic LAL assay takes measurements at multiple points in time (Lonza, n.d.-a) and would be better able to detect such deviations. Only one study in this systematic literature review used such an approach (Pussinen et al., 2011) while the remaining five studies that chose a LAL assay used the endpoint method (Camargo et al., 2019; Creely et al., 2007; Hawkesworth et al., 2013; Huang et al., 2019; Jayashree et al., 2014). This same study was also the only one to report any sort of pretreatment while the others did not use any or did not report doing so. Communicating methods fully, even when pretreatment was not performed, is essential given the variability that exists in adapting this assay for research use. Adopting a standardized approach to the information that should be included in a study, such as incorporating the STROBE guidelines to ensure key elements get reported (von Elm et al., 2008), would help to resolve this issue. Unfortunately, in this instance, the lack of information regarding pretreatment methods coupled with the drawbacks of using untreated plasma and serum with an endpoint LAL assay raises additional questions about the reliability of these values and what significance they may have. Lastly, a limitation of all LAL assays is that by measuring the endotoxicity of LPS using a clotting cascade in North American horseshoe crabs, the results may not necessarily correlate with its inflammatory potential in humans (Gnauck et al., 2016). LPS molecules also differ in their structure depending on the

species of bacteria that produce them and thus have differing endotoxic potentials. In effect, the same quantity of LPS from two unrelated bacteria may not result in the same value if measured in endotoxin units. These dynamics are significant as the assay appears to have a higher sensitivity for bacteria that North American horseshoe crabs may encounter in their environment as well as the potential to produce false positives since fungi can trigger the same clotting cascade.

Many of these issues are resolved by adopting a molecular-based approach that focuses on the detection of 16S rRNA. While fasting state and the meal prior to sampling may still affect the entry of bacteria or bacterial components into the systemic circulation and features of the innate and adaptive immune response may influence their clearance, the measurement of 16S rRNA using universal primers does not vary based on bacterial species nor does it face the challenges seen with interference and reaction kinetics as in the LAL assay. It also has the advantages of detecting gram-positive bacteria and providing qualitative information about the bacterial species present in addition to quantitative data. This is significant as our relationship with the microbes that inhabit us is complex, and some may prove to have a greater role in this process than others. By focusing solely on gram-negative bacteria, LPS could only ever provide a partial representation of what was happening in the body. And evidence suggests lipoteichoic acid from the cell walls of gram-positive bacteria stimulates the production of cytokines through TLR-2 similarly to that seen with LPS and TLR-4 (Sato et al., 2017). Toxic shock syndrome toxin-1, one of the virulence factors of the gram-positive bacteria *Staphylococcus aureus*, has also been shown to induce systemic inflammation and impaired glucose tolerance in animal models (Vu et al., 2015). Rather than targeting individual bacterial components, this method of using 16S rRNA is both more broad in being able to detect all bacteria regardless of their

makeup and more nuanced in being able to provide species-specific information. This was an advantage of the three studies within this systematic literature review that chose this approach (Amar et al., 2011b; Qui et al., 2019; Sato et al., 2014). But as with any newer method, a lack of consensus in how information is reported can limit comparison between studies and each used a different approach to measure 16S rRNA levels. One study reported concentration (Amar et al., 2011b), another provided high-quality sequences per sample (Qui et al., 2019), and the last relied on the rate of detection in participants (Sato et al., 2014). Additionally, although analysis of the qualitative data from these studies was beyond the scope of this project, similar issues are likely to be encountered as we move this research forward and determine how best to interpret the vast quantities of data generated.

D. Assessment of Type 2 Diabetes

An additional factor that affected the results in this systematic literature review was how type 2 diabetes was defined differently or not at all. Two studies provided no details about how diagnosis was established (Creely et al., 2007; de Waal et al., 2018) and one provided partial information (Hawkesworth et al., 2013). Of the remaining eight studies, only two relied solely on testing to determine comparison groups (Camargo et al., 2019; Jayashree et al., 2014), while the rest used physician-diagnosed or some combination of self-reported, physician-diagnosed, and testing for diabetes. This is significant because testing identifies clinically mild cases of type 2 diabetes that other methods may miss resulting in an underreporting of this outcome. While self-reported and physician-diagnosed diabetes often relies on testing at some point to diagnose a participant as having type 2 diabetes, it is often not clear when this occurred and whether glucose and insulin homeostasis may have changed since that time, especially as many individuals begin

lifestyle modifications and medication. This issue was demonstrated clearly in the study that provided partial information with a non-obese control, obese control, and type 2 diabetes group, only the last of which was physician-diagnosed and the rest unknown (Hawkesworth et al., 2013). When testing was performed to determine type 2 diabetes prevalence within each group based on fasting plasma glucose, there were no cases in the non-obese controls but the obese controls and type 2 diabetes group had a prevalence of 3.1% and 67.9%, respectively. Thus, the study showed how when testing is not used to determine comparison groups, you may have misclassification whereby participants who are considered controls actually have type 2 diabetes as well as participants who are considered to have type 2 diabetes no longer have an elevated blood glucose. Variations such as this can ultimately affect the comparisons made and their significance.

With the method used to establish a type 2 diabetes diagnosis in many cases not dependent solely on testing, the measurement of type 2 diabetes-related indices helped to further support the delineation of comparison groups in these studies. While overall the values that were reported did provide additional confirmation that the criteria used were sufficient to identify meaningful differences, many gaps were evident including measures for which information was not available or only partially available as well as a lack of statistical analysis to determine significance. Even for a measure as common as HbA1C or fasting plasma glucose, only eight studies provided details for all groups and made a statistical comparison for at least one of these two variables (Amar et al., 2011b; Camargo et al., 2019; Creely et al., 2007; Huang et al., 2019; Jayashree et al., 2014; Pussinen et al., 2011; Qui et al., 2019; Sato et al., 2014). Furthermore, some of these studies measured glucose using serum samples (Creely et al., 2007; Pussinen et al., 2011; Sato et al., 2014); however, the values obtained with plasma and serum are not the same

(Kang et al., 2016). With serum, the delay in sample processing results in a decrease in glucose concentration that can lead to a misdiagnosis. For instance, a study found when comparing both methods that 263 out of 1,254 individuals who met the criteria for a type 2 diabetes diagnosis using fasting plasma would have been missed by applying the same criteria to fasting serum samples. While the difference may not have a large impact when used simply to compare groups within a study, as it has internal validity and would affect both similarly, it does not have external validity and complicates the ability to easily use this measure to make comparisons between studies, especially when researchers will look to this variable to get a quick sense of glucose status among controls and those with type 2 diabetes if testing was not the criteria used to establish these groups.

E. Final Recommendations

In conclusion, although this systematic literature review was unable to answer the question of whether LPS or 16S rRNA may have a role in the development of type 2 diabetes, the proposed mechanism for how this could occur and may be a contributing factor warrants further exploration before discounting it as a possibility. New research is needed that builds upon the lessons learned thus far and incorporates strong study designs, such as cohort and case-control studies, with large sample sizes, random sampling, and a greater consensus on who should be considered a control, especially in regards to those with altered glucose metabolism who do not as yet meet the criteria for type 2 diabetes. Confounding variables should be fully reported and adjusted for when significant differences between comparison groups are identified, and information about medication needs to be reported as well. Ideally, taking both fasting and semifasting samples would best allow us to determine the relationship present since we do not in

fact know the ideal timing for assessment and having participants consume a standardized mixed meal before a semifasting sample would eliminate variation introduced by diet. Given the numerous drawbacks of LPS and the LAL assay, future studies should opt for molecular-based methods that focus on 16S rRNA and use universal primers, which will provide the most comprehensive information about the bacterial species present within samples. Since we also do not know what level of exposure may be significant, providing a rate of detection among participants along with more detailed quantitative information such as concentration or number of sequences per sample is necessary as we work to determine the relevance of these values. Lastly, utilizing testing to determine type 2 diabetes status is essential to be able to truly evaluate the impact of bacteria on this outcome. These details must be a part of any future studies if we are to make this relationship clear.

References

Al-Attas, O. S., Al-Daghri, N. M., Al-Rubeaan, K., da Silva, N. F., Sabico, S. L., Kumar, S., McTernan, P. G., & Harte, A. L. (2009). Changes in endotoxin levels in T2DM subjects on anti-diabetic therapies. *Cardiovascular Diabetology*, *8*, Article 20. https://doi.org/10.1186/1475-2840-8-20

Al-Obaide, M. A. I., Singh, R., Datta, P., Rewers-Felkins, K. A., Salguero, M. V., Al-Obaidi, I., Kottapalli, K. R., & Vasylyeva, T. L. (2017). Gut microbiota-dependent trimethylamine-*N*-oxide and serum biomarkers in patients with T2DM and advanced CKD. *Journal of Clinical Medicine*, *6*(9), 86. https://doi.org/10.3390/jcm6090086

Amar, J., Chabo, C., Waget, A., Klopp, P., Vachoux, C., Bermúdez-Humarán, L. G., Smirnova, N., Bergé, M., Sulpice, T., Lahtinen, S., Ouwehand, A., Langella, P., Rautonen, N., Sansonetti, P. J., & Burcelin, R. (2011a). Intestinal mucosal adherence and translocation of commensal bacteria at the early onset of type 2 diabetes: Molecular mechanisms and probiotic treatment. *EMBO Molecular Medicine*, *3*, 559-572. https://doi.org/10.1002/emmm.201100159

Amar, J., Serino, M., Lange, C., Chabo, C., Iacovoni, J., Mondot, S., Lepage, P., Klopp, C., Mariette, J., Bouchez, O., Perez, L., Courtney, M., Marre, M., Klopp, P., Lantieri, O., Doré, J., Charles, M. A., Balkau, B., & Burcelin, R. (2011b). Involvement of tissue bacteria in the onset of diabetes in humans: Evidence for a concept. *Diabetologia*, *54*, 3055-3061. https://doi.org/10.1007/s00125-011-2329-8

American Diabetes Association. (2011). Diagnosis and classification of diabetes mellitus. *Diabetes Care*, *34*(Suppl. 1), S62-S69. https://doi.org/10.2337/dc11-S062

American Diabetes Association. (2019a). *Complications*. Retrieved December 24, 2019, from

https://www.diabetes.org/diabetes/complications

American Diabetes Association. (2019b). *Diabetes risk*. Retrieved December 27, 2019, from https://www.diabetes.org/diabetes-risk

American Diabetes Association. (2019c). *Diabetes symptoms*. Retrieved December 24, 2019, from https://www.diabetes.org/diabetes/type-1/symptoms

American Diabetes Association. (2019d). *Fitness*. Retrieved December 27, 2019, from https://www.diabetes.org/fitness

American Diabetes Association. (2019e). *Nutrition*. Retrieved December 27, 2019, from https://www.diabetes.org/nutrition

American Diabetes Association. (2019f). *Prevention*. Retrieved December 27, 2019, from https://www.diabetes.org/diabetes-risk/prevention

American Diabetes Association. (2020). Classification and diagnosis of diabetes: Standards of medical care in diabetes—2020. *Diabetes Care*, *43*(Suppl. 1), S14-S31. https://doi.org/10.2337/dc20-S002

Blanco-Rojo, R., Alcala-Diaz, J. F., Wopereis, S., Perez-Martinez, P., Quintana-Navarro, G. M., Marin, C., Ordovas, J. M., van Ommen, B., Perez-Jimenez, F., Delgado-Lista, J., & Lopez-Miranda, J. (2016). The insulin resistance phenotype (muscle or liver) interacts with the type of diet to determine changes in disposition index after 2 years of intervention: the CORDIOPREV-DIAB randomised clinical trial. *Diabetologia*, *59*, 67-76. https://doi.org/10.1007/s00125-015-3776-4

Buckingham, L. (2012). *Molecular diagnostics: Fundamentals, methods and clinical applications*. F. A. Davis

Burcelin, R., Crivelli, V., Dacosta, A., Roy-Tirelli, A., & Thorens, B. (2002). Heterogeneous

metabolic adaptation of C57BL/6J mice to high-fat diet. *American Journal of Physiology-Endocrinology and Metabolism*, *282*, E834-E842. https://doi.org/10.1152/ajpendo.00332.2001

Camargo, A., Jimenez-Lucena, R., Alcala-Diaz, J. F., Rangal-Zuñiga, O. A., Garcia-Carpintero, S., Lopez-Moreno, J., Blanco-Rojo, R., Delgado-Lista, J., Perez-Martinez, P., van Ommen, B., Malagon, M. M., Ordovas, J. M., Perez-Jimenez, F., & Lopez-Miranda, J. (2019). Postprandial endotoxemia may influence the development of type 2 diabetes mellitus: From the CORDIOPREV study. *Clinical Nutrition*, *38*, 529-538. https://doi.org/10.1016/j.clnu.2018.03.016

Cani, P. D., Amar, J., Iglesias, M. A., Poggi, M., Knauf, C., Bastelica, D., Neyrinck, A. M., Fava, F., Tuohy, K. M., Chabo, C., Waget, A., Delmée, E., Cousin, B., Sulpice, T., Chamontin, B., Ferrières, J., Tanti, J., Gibson, G. R., Casteilla, L., . . . Burcelin, R. (2007). Metabolic endotoxemia initiates obesity and insulin resistance. *Diabetes*, *56*, 1761-1772. https://doi.org/10.2337/db06-1491

Cani, P. D., Bibiloni, R., Knauf, C., Waget, A., Neyrinck, A. M., Delzenne, N. M., & Burcelin, R. (2008). Changes in gut microbiota control metabolic endotoxemia-induced inflammation in high-fat diet-induced obesity and diabetes in mice. *Diabetes*, *57*, 1470-1481. https://doi.org/10.2337/db07-1403

Centers for Disease Control and Prevention. (2017, March 14). *John Snow: A legacy of disease detectives*. https:// blogs.cdc.gov/publichealthmatters/2017/03/a-legacy-of-disease-detectives/

Clemente, J. C., Pehrsson, E. C., Blaser, M. J., Sandhu, K., Gao, Z., Wang, B., Magris, M., Hidalgo, G., Contreras, M., Noya-Alarcón, Ó., Lander, O., McDonald, J., Cox, M.,

Walter, J., Lyn Oh, P., Ruiz, J. F., Rodriguez, S., Shen, N., Jin Song, S., . . . Dominguez-Bello, M. G. (2015). The microbiome of uncontacted Amerindians. *Science Advances*, *1*(3), e1500183. https://doi.org/10.1126/sciadv.1500183

Creely, S. J., McTernan, P. G., Kusminski, C. M., Fisher, F. M., Da Silva, N. F., Khanolkar, M., Evans, M., Harte, A. L., & Kumar, S. (2007). Lipopolysaccharide activates an innate immune system response in human adipose tissue in obesity and type 2 diabetes. *American Journal of Physiology-Endocrinology and Metabolism*, *292*, E740-E747. https://doi.org/10.1152/ajpendo.00302.2006

Dandona, P., Aljada, A., & Bandyopadhyay, A. (2004). Inflammation: The link between insulin resistance, obesity and diabetes. *Trends in Immunology*, *25*(1), 4-7. https://doi.org/10.1016/j.it.2003.10.013

Deopurkar, R., Ghanim, H., Friedman, J., Abuaysheh, S., Sia, C. L., Mohanty, P., Viswanathan, P., Chaudhuri, A., & Dandona, P. (2010). Differential effects of cream, glucose, and orange juice on inflammation, endotoxin, and the expression of toll-like receptor-4 and suppressor of cytokine signaling-3. *Diabetes Care*. *33*(5), 991-997. https://doi.org/10.2337/dc09-1630

de Waal, G. M., Engelbrecht, L., Davis, T., de Villiers, W. J. S., Kell, D. B., & Pretorius, E. (2018). Correlative light-electron microscopy detects lipopolysaccharide and its association with fibrin fibres in parkinson's disease, alzheimer's disease and type 2 diabetes mellitus. *Scientific Reports*, *8*, Article 16798. https://doi.org/10.1038/s41598-018-35009-y

Donath, M. Y. (2016). Multiple benefits of targeting inflammation in the treatment of type 2 diabetes. *Diabetologia*, *59*(4), 679-682. https://doi.org/10.1007/s00125-016-3873-z

Everard, A, & Cani, P. D. (2013). Diabetes, obesity and gut microbiota. *Best Practice & Research: Clinical Gastroenterology*, *27*, 73-83. https://doi.org/10.1016/j.bpg.2013.03.007

Freepik. (2020a). *Bacteria free icon* [Clip art]. Flaticon. https://www.flaticon.com/free-icon/bacteria_822183?term=cell&page=1&position=40

Freepik. (2020b). *Brain free icon* [Clip art]. Flaticon. https://www.flaticon.com/free-icon/brain_827908

GBD 2017 DALYs and HALE Collaborators. (2018). Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990-2017: A systematic analysis for the Global Burden of Disease Study 2017. *Lancet*, *392*, 1859-1922. https://doi.org/10.1016/S0140-6736(18)32335-3

Gentile, C. L, & Weir, T. L. (2018). The gut microbiota at the intersection of diet and human health. *Science*, *362*(6416), 776-780. https://doi.org/10.1126/science.aau5812

Ghanim, H., Abuaysheh, S., Sia, C. L., Korzeniewski, K., Chaudhuri, A., Fernandez-Real, J. M., & Dandona, P. (2009). Increase in plasma endotoxin concentrations and the expression of toll-like receptors and suppressor of cytokine signaling-3 in mononuclear cells after a high-fat, high-carbohydrate meal: Implications for insulin resistance. *Diabetes Care*, *32*(12), 2281-2287. https://doi.org/10.2337/dc09-0979

Gnauck, A., Lentle, R. G., & Kruger, M. C. (2016). Chasing a ghost? - Issues with the determination of circulating levels of endotoxin in human blood. *Critical Reviews in Clinical Laboratory Sciences*, *53*(3), 197-215. https://doi.org/10.3109/10408363.2015.1123215

Goldfine, A. B., & Shoelson, S. E. (2017). Therapeutic approaches targeting inflammation for diabetes and associated cardiovascular risk. *The Journal of Clinical Investigation*, *127*(1), 83-93. https://doi.org/10.1172/JCI88884

Golubev, N. (2020). *Liver free icon* [Clip art]. Flaticon. https://www.flaticon.com/free-icon/liver_361496?term=liver&page=1&position=12

Goodwin, S., McPherson, J. D., & McCombie, W. R. (2016). Coming of age: Ten years of next-generation sequencing technologies. *Nature Reviews Genetics*, *17*, 333-351. https://doi.org/10.1038/nrg.2016.49

Graham, D. Y. (2014). History of *Helicobacter pylori*, duodenal ulcer, gastric ulcer and gastric cancer. *World Journal of Gastroenterology*, *20*(18), 5191-5204. https://doi.org/10.3748/wjg.v20.i18.5191

Gurung, M., Li, Z., You, H., Rodrigues, R., Jump, D. B., Morgun, A., & Shulzhenko, N. (2020). Role of gut microbiota in type 2 diabetes pathophysiology. *EBioMedicine*, *51*, 102590. https://doi.org/10.1016/j.ebiom.2019.11.051

Hackett, R. A., & Steptoe, A. (2017). Type 2 diabetes mellitus and psychological stress—a modifiable risk factor. *Nature Reviews Endocrinology*, *13*(9), 547-560. https://doi.org/10.1038/nrendo.2017.64

Harte, A. L., Varma, M. C., Tripathi, G., McGee, K. C., Al-Daghri, N. M., Al-Attas, O. S., Sabico, S., O'Hare, J. P., Ceriello, A., Saravanan, P., Kumar, S., & McTernan, P. G. (2012). High fat intake leads to acute postprandial exposure to circulating endotoxin in type 2 diabetic subjects. *Diabetes Care*, *35*, 375-382. https://doi.org/10.2337/dc11-1593

Hawkesworth, S., Moore, S. E., Fulford, A. J. C., Barclay, G. R., Darboe, A. A., Mark, H., Nyan, O. A., & Prentice, A. M. (2013). Evidence for metabolic endotoxemia in obese and

diabetic Gambian women. *Nutrition & Diabetes*, *3*, e83. https://doi.org/10.1038/nutd.2013.24

Hotamisligil, G. S., Shargill, N. S., & Spiegelman, B. M. (1993). Adipose expression of tumor necrosis factor-alpha: Direct role in obesity-linked insulin resistance. *Science*, *259*(5091), 87-91. https://doi.org/10.1126/science.7678183

Huang, X., Yan, D., Xu, M., Li, F., Ren, M., Zhang, J., & Wu, M. (2019). Interactive association of lipopolysaccharide and free fatty acid with the prevalence of type 2 diabetes: A community-based cross-sectional study. *Journal of Diabetes Investigation*, *10*(6), 1438-1446. https://doi.org/10.1111/jdi.13056

Huttenhower, C., Gevers, D., Knight, R., Abubucker, S., Badger, J. H., Chinwalla, A. T., Creasy, H. H., Earl, A. M., FitzGerald, M. G., Fulton, R. S., Giglio, M. G., Hallsworth-Pepin, K., Lobos, E. A., Madupu, R., Magrini, V., Martin, J. C., Mitreva, M., Muzny, D. M., Sodergren, E. J., . . . White, O. (2012). Structure, function and diversity of the healthy human microbiome. *Nature*, *486*, 207-214. https://doi.org/10.1038/nature11234

International Diabetes Federation. (2019a). *IDF diabetes atlas, 9th edition*. https://www.diabetesatlas.org/upload/resources/2019/IDF_Atlas_9th_Edition_2019.pdf

International Diabetes Federation. (2019b). *Worldwide toll of diabetes*. Retrieved December 21, 2019, from https://www.diabetesatlas.org/en/sections/worldwide-toll-of-diabetes.html

Jayashree, B., Bibin, Y. S., Prabhu, D., Shanthirani, C. S., Gokulakrishnan, K., Lakshmi, B. S., Mohan, V., & Balasubramanyam, M. (2014). Increased circulatory levels of lipopolysaccharide (LPS) and zonulin signify novel biomarkers of proinflammation in patients with type 2 diabetes. *Molecular and Cellular Biochemistry*, *388*, 203-210. https://doi.org/10.1007/s11010-013-1911-4

Jenkins, C., Ling, C. L., Ciesielczuk, H. L., Lockwood, J., Hopkins, S., McHugh, T. D., Gillespie, S. H., & Kibbler, C. C. (2012). Detection and identification of bacteria in clinical samples by 16S rRNA gene sequencing: Comparison of two different approaches in clinical practice. *Journal of Medical Microbiology*, *61*, 483-488. https://doi.org/10.1099/jmm.0.030387-0

Johnson, S. (2006). *The ghost map: The story of London's most terrifying epidemic — and how it changed science, cities, and the modern world*. Penguin Group.

Kang, J. G., Park, C., Ihm, S., & Park, S. W. (2016). A potential issue with screening prediabetes or diabetes using serum glucose: a delay in diagnosis. *Diabetes & Metabolism Journal*, *40*, 414-417. https://doi.org/10.4093/dmj/2016.40.5.414

Kell, D. B, & Pretorius, E. (2015). On the translocation of bacteria and their lipopolysaccharides between blood and peripheral locations in chronic, inflammatory diseases: the central roles of LPS and LPS-induced cell death. *Integrative Biology*, 7, 1339-1377. https://doi.org/10.1039/c5ib00158g

Khan, S., Waliullah, S., Godfrey, V., Khan, M. A., Ramachandran, R. A., Cantarel, B. L., Behrendt, C., Peng, L., Hooper, L. V., & Zaki, H. (2020). Dietary simple sugars alter microbial ecology in the gut and promote colitis in mice. *Science Translational Medicine*, *12*, eaay6218. https://doi.org/10.1126/scitranslmed.aay6218

Klekotka, R. B., Mizgała-Izworska, E., Drzastwa, W., & Mazur, B. (2018). The role of *Staphylococcus aureus* in the clinical diagnosis of diabetic patients. *Postępy Mikrobiologii*, *57*(2), 166-178. https://doi.org/10.21307/PM-2018.57.2.166

Liu, Y., Zhao, T., & Hou, L. (2013). Change and correlated factors of fasting level of the plasma endotoxin in subjects with different glucose tolerances and body mass indices. *Sichuan*

Da Xue Xue Bao Yi Xue Ban, *44*(5), 769-773. https://doi.org/10.13464/j.scuxbyxb.2013.05.017

Loman, N. J., & Pallen, M. J. (2015). Twenty years of bacterial genome sequencing. *Nature Reviews Microbiology*, *13*, 787-794. https://doi.org/10.1038/nrmicro3565

Lontchi-Yimagou, E., Sobngwi, E., Matsha, T. E., & Kengne, A. P. (2013). Diabetes mellitus and inflammation. *Current Diabetes Reports*, *13*(3), 435-444. https://doi.org/10.1007/s11892-013-0375-y

Lonza. (n.d.-a). *Chromogenic LAL assay*. Retrieved April 16, 2021, from https://bioscience.lonza.com/lonza_bs/US/en/chromogenic-lal-assay#accept

Lonza. (n.d.-b). *QCL-1000 endpoint chromogenic LAL assay*. Retrieved March 24, 2021, from https://bioscience.lonza.com/lonza_bs/US/en/Endotoxin-Detection/p/000000000000182373/QCL-1000-Endpoint-Chromogenic-LAL-Assay#accept

Magne, F., Gotteland, M., Gauthier, L., Zazueta, A., Pesoa, S., Navarrete, P., & Balamurugan, R. (2020). The Firmicutes/Bacteroidetes ratio: A relevant marker of gut dysbiosis in obese patients? *Nutrients*, 12(5), 1474. https://doi.org/10.3390/nu12051474

Mahon, C. R., Lehman, D. C., & Manuselis, G. (2015). *Textbook of diagnostic microbiology*. Elsevier.

Mao, Y., & Lu, Z. (2017). MeSH Now: Automatic MeSH indexing at PubMed scale via learning to rank. *Journal of Biomedical Semantics*, *8*, Article 15. https://doi.org/10.1186/s13326-017-0123-3

Matsuda, K., Tsuji, H., Asahara, T., Kado, Y., & Nomoto, K. (2007). Sensitive quantitative detection of commensal bacteria by rRNA-targeted reverse transcription-PCR. *Applied*

and Envrionmental Microbiology, *73*(1), 32-39. https://doi.org/10.1128/AEM.01224-06

Matsuki, T., Watanabe, K., Fujimoto, J., Takada, T., & Tanaka, R. (2004). Use of 16S rRNA gene-targeted group-specific primers for real-time PCR analysis of predominant bacteria in human feces. *Applied and Environmental Microbiology*, *70*(12), 7220-7228. https://doi.org/10.1128/AEM.70.12.7220-7228.2004

Matthews, D. R., Hosker, J. P., Rudenski, A. S., Naylor, B. A., Treacher, D. F., & Turner, R. C. (1985). Homeostasis model assessment: Insulin resistance and *β*-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia*, *28*, 412-419. https://doi.org/10.1007/BF00280883

MedlinePlus. (2019, May 6). *Tissue types*. Retrieved December 12, 2020, from https://medlineplus.gov/ency/imagepages/8682.htm

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, *6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

Moreno-Indias, I., Cardona, F., Tinahones, F. J., & Queipo-Ortuño, M. I. (2014). Impact of the gut microbiota on the development of obesity and type 2 diabetes mellitus. *Frontiers in Microbiology*, *5*, Article 190. https://doi.org/10.3389/fmicb.2014.00190

Munford, R. S. (2005). Detoxifying endotoxin: time, place and person. *Journal of Endotoxin Research*, *11*(2), 69-84. https://doi.org/10.1179/096805105X35161

National Human Genome Research Institute. (n.d.). *Shotgun sequencing*. Retrieved November 8, 2020, from https://www.genome.gov/genetics-glossary/Shotgun-Sequencing

National Institute of Diabetes and Digestive and Kidney Diseases. (2018, April). *The A1C test & diabetes*. Retrieved January 5, 2020, from https://www.niddk.nih.gov/health-

information/diabetes/overview/tests-diagnosis/a1c-test

Nesseler, N., Launey, Y., Aninat, C., Morel, F., Mallédant, Y., & Seguin, P. (2012). Clinical review: The liver in sepsis. *Critical Care*, *16*, Article 235. https://doi.org/10.1186/cc11381

Nguyen, N., Warnow, T., Pop, M., & White, B. (2016). A perspective on 16S rRNA operational taxonomic unit clustering using sequence similarity. *npj Biofilms and Microbiomes*, *2*, Article 16004. https://www.doi.org/10.1038/npjbiofilms.2016.4

Pacheco, J. A., & Thompson, W. (2011). *Type 2 diabetes mellitus electronic medical record case and control selection algorithms*. Phenotype Knowledgebase. https://phekb.org/sites/phenotype/files/T2DM-algorithm.pdf

Petersen, C., & Round, J. L. (2014). Defining dysbiosis and its influence on host immunity and disease. *Cellular Microbiology*, *16*(7), 1024-1033. https://doi.org/10.1111/cmi.12308

Pickup, J. C. (2004). Inflammation and activated innate immunity in the pathogenesis of type 2 diabetes. *Diabetes Care*, *27*(3), 813-823. https://doi.org/10.2337/diacare.27.3.813

Piya, M. K., Harte, A. L., & McTernan, P. G. (2013). Metabolic endotoxaemia: Is it more than just a gut feeling? *Current Opinion in Lipidology*, *24*(1), 78-85. https://doi.org/10.1097/MOL.0b013e32835b4431

Pomytkin, I. A., Cline, B. H., Anthony, D. C., Steinbusch, H. W., Lesch, K.-P., & Strekalova, T. (2015). Endotoxaemia resulting from decreased serotonin transporter (5-HTT) function: A reciprocal risk factor for depression and insulin resistance? *Behavioral Brain Research*, *276*, 111-117. https://doi.org/10.1016/j.bbr.2014.04.049

Potgieter, M., Bester, J., Kell, D. B., & Pretorius, E. (2015). The dormant blood microbiome in chronic, inflammatory diseases. *FEMS Microbiology Reviews*, *39*(4), 567-591.

https://doi.org/10.1093/femsre/fuv013

Proctor, L. M., Creasy, H. H., Fettweis, J. M., Lloyd-Price, J., Mahurkar, A., Zhou, W., Buck, G. A., Synder, M. P., Strauss, J. F., Weinstock, G. M., White, O., & Huttenhower, C. (2019). The integrative human microbiome project. *Nature*, *569*, 641-648. https://doi.org/10.1038/s41586-019-1238-8

Pussinen, P. J., Havulinna, A. S., Lehto, M., Sundvall, J., & Salomaa, V. (2011). Endotoxemia is associated with an increased risk of incident diabetes. *Diabetes Care*, *34*, 392-397. https://doi.org/10.2337/dc10-1676

Qui, J., Zhou, H., Jing, Y., & Dong, C. (2019). Association between blood microbiome and type 2 diabetes mellitus: A nested case-control study. *Journal of Clinical Laboratory Analysis*, *33*, e22842. https://doi.org/10.1002/jcla.22842

Richesson, R. L., Rusincovitch, S. A., Wixted, D., Batch, B. C., Feinglos, M. N., Miranda, M. L., Hammond, W. E., Califf, R. M., Spratt, S. E. (2013). A comparison of phenotype definitions for diabetes mellitus. *Journal of the American Medical Informatics Association*, *20*, e319-e326. https://doi.org/10.1136/amiajnl-2013-001952

Ridaura, V. K., Faith, J. J., Rey, F. E., Cheng, J., Duncan, A. E., Kau, A. L., Griffin, N. W., Lombard, V., Henrissat, B., Bain, J. R., Muehlbauer, O. I., Semenkovich, C. F., Funai, K., Hayashi, D. K., Lyle, B. J., Martini, M. C., Ursell, L. K., Clemente, J. C., van Treuren, W., . . . Gordon, J. I. (2013). Gut microbiota from twins discordant for obesity modulate metabolism in mice. *Science*, *341*(6150), 1241214. https://doi.org/10.1126/science.1241214

Sato, J., Kanazawa, A., Ikeda, F., Yoshihara, T., Goto, H., Abe, H., Komiya, K., Kawaguchi, M., Shimizu, T., Ogihara, T., Tamura, Y., Sakurai, Y., Yamamoto, R., Mita, T., Fujitani, Y.,

Fukuda, H., Nomoto, K., Takahashi, T., Asahara, T., . . . Watada, H. (2014). Gut dysbiosis and detection of "live gut bacteria" in blood of Japanese patients with type 2 diabetes. *Diabetes Care*, *37*, 2343-2350. https://doi.org/10.2337/dc13-2817

Sato, J., Kanazawa, A., & Watada, H. (2017). Type 2 diabetes and bacteremia. *Annals of Nutrition and Metabolism*, *71*(Suppl. 1), 17-22. https://doi.org/10.1159/000479919

Shannon-Lowe, C., & Rickinson, A. (2019). The global landscape of EBV-associated tumors. *Frontiers in Oncology*, *9*, Article 713. https://doi.org/10.3389/fonc.2019.00713

Shoelson, S. E., Lee, J., & Goldfine, A. B. (2006). Inflammation and insulin resistance. *The Journal of Clinical Investigation*, *116*(7), 1793-1801. https://doi.org/10.1172/JCI29069

Skloot, R. (2010). *The immortal life of Henrietta Lacks*. Crown Publishers.

Song, Y., Manson, J. E., Tinker, L., Howard, B. V., Kuller, L. H., Nathan, L., Rifai, N., & Liu, S. (2007). Insulin sensitivity and insulin secretion determined by homeostasis model assessment and risk of diabetes in a multiethnic cohort of women. *Diabetes Care*, *30*, 1747-1752. https://doi.org/10.2337/dc07-0358

Sonnenburg, E. D., Smits, S. A., Tikhonov, M., Higginbottom, S. K., Wingreen, N. S., & Sonnenburg, J. L. (2016). Diet-induced extinctions in the gut microbiota compound over generations. *Nature*, *529*, 212-215. https://doi.org/10.1038/nature16504

Sonnenburg, J. L., & Bäckhed, F. (2016). Diet-microbiota interactions as moderators of human metabolism. *Nature*, *535*, 56-64. https://doi.org/10.1038/nature18846

Surwit, R. S., van Tilburg, M. A., Zucker, N., McCaskill, C. C., Parekh, P., Feinglos, M. N., Edwards, C. L., Williams, P., & Lane, J. D. (2002). Stress management improves long-term glycemic control in type 2 diabetes. *Diabetes Care*, *25*(1), 30-34. https://doi.org/10.2337/diacare.25.1.30

Takeuchi, O., & Akira, S. (2010). Pattern recognition receptors and inflammation. *Cell*, *140*, 805-820. https://doi.org/10.1016/j.cell.2010.01.022

Taylor, E. N., Hu, F. B., & Curhan, G. C. (2006). Antihypertensive medications and the risk of incident type 2 diabetes. *Diabetes Care*, *29*, 1065-1070. https://doi.org/10.2337/dc05-2366

Tremaroli, V., & Bäckhed, F. (2012). Functional interactions between the gut microbiota and host metabolism. *Nature*, *489*, 242-249. https://doi.org/10.1038.nature11552

Tuomi, T., Santoro, N., Caprio, S., Cai, M., Weng, J., & Groop, L. (2014). The many faces of diabetes: a disease with increasing heterogeneity. *Lancet*, *383*, 1084-1094. https://doi.org/10.1016/S0140-6736(13)62219-9

Turner, J. R. (2009). Intestinal mucosal barrier function in health and disease. *Nature Reviews Immunology*, *9*, 799-809. https://doi.org/10.1038/nri2653

Ursell, L. K., Metcalf, J. L., Wegener Parfrey, L., & Knight, R. (2012). Defining the human microbiome. *Nutrition Reviews*, *70*(Suppl. 1), S38-S44. https://doi.org/10.1111/j.1753-4887.2012.00493.x.

Vartiainen, E., Laatikainen, T., Peltonen, M., Juolevi, A., Männistö, S., Sundvall, J., Jousilahti, P., Salomaa, V., Valsta, L., & Puska, P. (2010). Thirty-five-year trends in cardiovascular risk factors in Finland. *International Journal of Epidemiology*, *39*, 504-518. https://doi.org/10.1093/ije/dyp330

von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., Vandenbroucke, J. P., & STROBE Initiative. (2008). The strengthening the reporting of observational studies in epidemiology (STROBE) statement: Guidelines for reporting observational studies. *Journal of Clinical Epidemiology*, *61*(4), 344-349.

https://doi.org/10.1016/j.clinepi.2007.11.008

Vu, B. G., Stach, C. S., Kulhankova, K., Salgado-Pabón, W., Klingelhutz, A. J., & Schlievert, P. M. (2015). Chronic superantigen exposure induces systemic inflammation, elevated bloodstream endotoxin, and abnormal glucose tolerance in rabbits: Possible rold in diabetes. *mBio*, *6*(2), Article e02554-14. https://doi.org/10.1128/mBio.02554-14

Wagner, B. D., Grunwald, G. K., Zerbe, G. O., Mikulich-Gilbertson, S. K., Robertson, C. E., Zemanick, E. T., & Harris, J. K. (2018). On the use of diversity measures in longitudinal sequencing studies of microbial communities. *Frontiers in Microbiology*, *9*, Article 1037. https://doi.org/10.3389/fmicb.2018.01037

Walker, M. (2017). *Why we sleep: Unlocking the power of sleep and dreams* [Audiobook]. Simon & Schuster Audio.

Wallace, T. M., Levy, J. C., & Matthews, D. R. (2004). Use and abuse of HOMA modeling. *Diabetes Care*, *27*, 1487-1495. https://doi.org/10.2337/diacare.27.6.1487

Weisberg, S. P., McCann, D., Desai, M., Rosenbaum, M., Leibel, R. L., & Ferrante, A. W. (2003). Obesity is associated with macrophage accumulation in adipose tissue. *The Journal of Clinical Investigation*, *112*(12), 1796-1808. https://doi.org/10.1172/JCI19246

Wellen, K. E., & Hotamisligil, G. S. (2005). Inflammation, stress, and diabetes. *The Journal of Clinical Investigation*, *115*(5), 1111-1119. https://doi.org/10.1172/JCI200525102

World Health Organization. (2016). *Global report on diabetes*. https://www.who.int/publications-detail/global-report-on-diabetes

World Health Organization. (2018, May 24). *The top 10 causes of death*. https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death

World Health Organization. (2019a). *Classification of diabetes mellitus 2019*.

https://www.who.int/publications-detail/classification-of-diabetes-mellitus

World Health Organization. (2019b). *Metrics: Disability-adjusted life year (DALY)*. Retrieved December 26, 2019, from https://www.who.int/healthinfo/global_burden_disease/metrics_daly/en/

World Health Organization & International Diabetes Federation. (2006). *Definition and diagnosis of diabetes mellitus and intermediate hyperglycemia*. https://apps.who.int/iris/bitstream/handle/10665/43588/9241594934_eng.pdf?sequence=1

Yong, E. (2016). *I contain multitudes: The microbes within us and a grander view of life*. HarperCollins Publishers.

Zanoni, I., & Granucci, F. (2013). Role of CD14 in host protection against infections and in metabolism regulation. *Frontiers in Cellular and Infection Microbiology*, *3*, Article 32. https://doi.org/10.3389/fcimb.2013.00032