Investigating genetic, microbial-viral, and other environmental factors in neuroinflammatory and neurodegenerative diseases of CNS through ‘-omics’ or ‘meta-research’ approaches

By

Alexios-Fotios A. Mentis, M.D., M.P.H., Pg.Cert.

Physician-public health scientist & Genomics Specialist

Alexander S. Onassis Foundation Scholar

Larissa & Athens, 2020

A thesis submitted

in partial fulfillment of the requirements for the degree Doctor of Philosophy (Ph.D.)
The acceptance of the doctoral thesis by the Faculty of Medicine of School of Health Sciences of the University of Thessaly does not necessarily indicate acceptance of the opinions of the author (according the provision of article 202, paragraph 2 of the law 5343/1932).
"We are troubled on every side, yet not distressed; we are perplexed, but not in despair. Persecuted, but not forsaken; cast down, but not destroyed".

Saint Paul, Letter to Corinthians 2; 4:8-18

"Τοις αγαπώσι τον Θεόν, πάντα συνεργεί εις αγαθόν"

Απόστολος Παύλος, Προς Ρωμαίους: Η, 28-34
Accepted by the members of the Seven-Member Evaluation Committee (13th/22-07-2020):

1st Member (Supervisor)  
Eftimios Dardiotis, M.D., M.Sc., Ph.D.  
Associate Professor of Neurology, School of Health Sciences, Faculty of Medicine, University of Thessaly, Larissa, Greece

2nd Member  
Efthimia Petinaki, M.D., Ph.D.  
Professor of Medical Biopathology and Clinical Microbiology, School of Health Sciences, Faculty of Medicine, University of Thessaly, Larissa, Greece

3rd Member  
Georgia Xiromerisiou, M.D., Ph.D.  
Assistant Professor of Neurology, School of Health Sciences, Faculty of Medicine, University of Thessaly, Larissa, Greece

4th Member  
George P. Chrousos, M.D., Sc.D.  
Research Director, University Research Institute of Maternal and Child Health and Precision Medicine, Medical School, National and Kapodistrian University of Athens, Athens, Greece. Distinguished Investigator Emeritus, NICHD, National Institutes of Health, Bethesda, Maryland, United States. Holder, UNESCO Chair on Adolescent Health Care

5th Member  
Flora Bacopoulou, M.D., Ph.D.  
Assistant Professor, Center for Adolescent Medicine and UNESCO Chair on Adolescent Health Care, First Department of Pediatrics, School of Medicine, National and Kapodistrian University of Athens, Aghia Sophia Children's Hospital

6th Member  
Konstantinos Bonotis, M.D., Ph.D.  
Assistant Professor of Psychiatry, School of Health Sciences, Faculty of Medicine, University of Thessaly, Larissa, Greece

7th Member  
Evangelia Tsironi, M.D., Ph.D.  
Professor of Ophthalmology, School of Health Sciences, Faculty of Medicine, University of Thessaly, Larissa, Greece
ACKNOWLEDGMENTS

To Professors Efthimios Dardiotis, Efthimia Petinaki, Georgios Hadjiigeorgiou, and Georgia Xiromerisiou for their continuous support throughout my years of studies

... To my mentor, Prof. George P. Chrousos, M.D., Sc.D., for being a continuous source of inspiration and admiration

... To Elder Gavriel of Agion Oros (“Agion Oros” – Mount Athos), to Elder Thimothei of Panagia Pantanassa (Crete) and nun Theodosia, to Father Kyriakos of “Saint Eirini Palaias Fokaias”, and Father Dimitrios of “Saint Isidoroi of Lycabettus” for being spiritual lighthouses in these difficult years

... To my colleagues at the University of Thessaly and the Hellenic Pasteur Institute for their solidarity

... To the sacred memory of my grandparents, Eugenia (Triseugeni) and Stylianos, Fotios and Maria

... To my parents, Annie and Andrea, for their moral and financial support, and for having tolerated me

ALEXIOS-FOTIOS A. MENTIS
PREFACE

My native place, Greece, has long suffered a tough and unrelenting period of austerity. Why return to my native place beaten by the financial crisis to pursue my biomedical PhD studies, after being trained abroad with the pulse of biomedicine there? I wrestled with this decision, against the flow of half a million scientists (including more than 17,000 physicians) having quitte[d] the country. I was aware beforehand of the drastic curtailments in research. But I decided to undergo this hard landing. Why? Because it was difficult to resist the nostalgia for my homeland and to overcome my strong family ties. Because I somehow felt a zest for a high risk, high reward experience, where I had much to lose, but many things to gain and prove to myself and for my country.

Some years ago, I enrolled as a PhD student in my home country, and became, now under the mentorship of a highly-cited physician-scientist and with the scholarship and financial aid from the Alexander Onassis Foundation, a Homo doctoratus, a term used by Eric J Dolin to describe a graduate student. A PhD studentship has inherent difficulties everywhere, but I found that I suffered more than I expected. To give some examples, pursuing academic excellence and a productive career pathway have been damned by public policy leaders as rattan and cholera, respectively (sic) (i.e., in the sense of stigma and obnoxiousness); even the reference style applied to PhD theses needs to be explicitly stated in governmental gazettes; and, those achieving the most with the least effort are societally praised nowadays.

To survive and stay motivated in this resource-poor environment but, perhaps more importantly, mentally frustrating environment, I developed mental tools based on the concept of happicles. This is a term coined by the philosopher Michael Epstein to reflect the small or short-lived reasons to feel a little happy. These mental and cultural happicles have kept me motivated and mentally healthy in my academic journey. Some of these — such as elements of my religious background, civilization’s roots, and interaction with special populations — I encounter on my daily trip on the underground through my city.

I am lucky to have been born and to start my day off in the neighborhood where the great philosopher Socrates is believed to have lived. Actually, when I brainstorm for new projects with peers, I apply the Socratic method. This is an exchange of arguments that becomes a dialogue hallmarked by formulating and replying to questions, and it emphasizes critical thinking and absence of prejudgments as the true values of science.

My embarkation point on my daily journey is the Saint John’s Church Metro Station. In an otherwise typical station, I am vividly reminded of the priest Saint Nikolaos Planas. For fifty years, he tirelessly performed divine liturgies in the church next to this station’s exit for long hours each day. I started to be comforted by the notion that the life of a graduate student
was, in some ways, similar to that of this committed and inspirational man as manifested, for example, in the frequent long hours of solitary work. More broadly, I feel that the ascetic, quasi-unconventional lifestyle of priests, nuns, and medieval scholars everywhere resembles the isolated, often financially poor lifestyle of a PhD student. Although, as a PhD student, I was initially seeking to focus only on science, I ultimately found that spiritual places, like churches and monasteries, also provided something that I deeply needed. By the way, I feel very blessed that during the years of my PhD studies in an otherwise harsh environment, Elders Paisios (January 2015), Iakovos Tsalikis (November 2017), as well as, Amphilohios Makris of Patmos (August 2018), Ieronymos Simonopetritis (November 2019), Daniel Katounakiotis (March 2020), Ioseph Hsyhastis Spilaïotis (March 2020), Efraïm Katounakiotis (March 2020) and Metropolitan Kallinikos of Edessa, Pella, and Almopia (June 2020), were included among the recognized saints (canonization) of the Eastern Orthodox Christian Church.

When I visit the above sacred places, the spiritual life of the monks’ and nuns’ who inhabit those walls amazes me. It is a living paradigm of patience, mental stability, humbleness, and endurance in the sense of resilience. They keep defining themselves to me as troubled on every side, yet not distressed and perplexed, but not in despair. Adopting these virtues can only add value to my life as physician-scientist.

I soon realized that, during my visit to monasteries, the nuns there embrace me with compassion and unconditional love because I am human, and because of external determinants such as academic success. When interacting with the nuns and monks, as a scientist, I am flooded with oxytocin, the hormone of empathy and of compassion. As a human being, I feel a sense of relief and self-worth that is much needed in these difficult times.

Interestingly, the ascetic lifestyle that these nuns and monks follow appears, at first glance, highly demanding for obeying moral rules (whose infringement could be connotated with the notion of punishment). However, the secret of these unexpected PhD mentors of mine is the idea of forgiveness instead of punishment. I cannot think of a more suitable example to understand that moral behavior, like PhD experiments or even biological systems, can be imperfect and fail; but it is the resilience to move forward that counts in every case.

During every single visit, I also pay ultimate respect to these places because historically monasteries were the leading medieval places for treating vulnerable groups and patients. They were also painstakingly copying the biomedical literature of that era by preserving the manuscripts of Classics (including Hippocrates and Galen) in papyrus scrolls. Thus, I consider their work similar to that of modern biomedicine scholars and librarians. This analogy is a consolation in the current era of the deep mental conflict between the scientific rigor required
for publications where work and writings are elegantly scrutinized for their merit on the one hand, and, on the other, the recent trend where anyone can publish online content of an unparalleled triviality and often lacking of truth.

When I ultimately complete my visit to these sacred places, I appreciate the almost unchanged, thousand-year-old rituals of these centers. This undiluted tradition has offered me an unparalleled mental relaxation from the constant need to stay tuned into the progress in science.

Continuing my journey through the underground, my civilization’s roots, so omnipresent in my city, keep providing inspiration and solace. I pass through the Acropolis Metro Station, which is near the birthplace of classic works of literature and art that encapsulate the essence of scientific rigor rooted in antiquity and remind me that a PhD is, historically, a doctorate in Philosophy. Concepts born near Acropolis include Aristotle’s seminal undertaking of the now-trendy field of secondary research (meta-cognition). They also refer to Plato’s concept of Logos, combining logic and wording, which has helped me in drafting manuscripts by applying logical flow and language.

When the metro bell finally rings at the General Hospital Station, I am reminded of another crucial source of inspiration: the vulnerable groups who I hope to help through my scientific and medical work. The educational activities at our Research Institute educate children of migrants to the realm of science by exposing them to simple experiments; these interactions have instilled in me a better sense of the true nature of human problems. I also am constantly reminded of Pasteur’s phrase: When I approach a child, he inspires me in two sentiments; tenderness for what he is, and respect for what he may become. By educating children to appreciate science, I am reminded of my purpose to educate.

On the other side, by keeping in touch with some pediatric and adult patients, I am reminded of my chosen purpose to heal, and that biological systems are imperfect and may fail. I perceive this patient-centered mindset as especially powerful for students in research resources-poor environments; in the latter, basic research may be financially out of reach. I also feel that interacting with patient groups can also be an antidote to the detachment that some scientists may have to the outcome of their work when the fruits of their labors are a decade away from clinical application.

Will I reach my final destination: become a productive physician-scientist who is improving the lives of patients through scientific research? Only the future will tell. But for now, I have derived happiness and solace in the happicles available in the humble life of a Homo doctoratus and physician-scientist, a species recently described as endangered, as well
as, in the phrase expressed in the book *Biotheology* regarding DNA’s wisdom *How miniaturized are your works, Lord! In wisdom you made them all* (by paraphrasing a psalm of Kind David). This is a life which is full of effort to maintain mental health in otherwise deeply inhospitable conditions, and on the other, it is enriched by the historical and spiritual legacy surrounding me on an everyday journey through the ancient town of my city.

ALEXIOS-FOTIOS A. MENTIS
LARISSA & ATHENS, 11 APRIL 2020
COVID-19 QUARANTINE PERIOD &
FEAST DAY OF SAINT LAZARUS
BRIEF RESUME

1. EDUCATION

November 2014 − Until now
Ph.D. candidate
Hellenic Pasteur Institute, Athens, in collaboration with Faculty of Medicine, University of Thessaly, Larissa, Greece

Graduate diploma in Sequence Analysis & Genomics (GPA: A+)
Advanced Academic Programs, Zanvyl Krieger School of Arts and Sciences
The Johns Hopkins University, Baltimore, MD, USA

August 2017
Master’s degree in Public Health (Concentration: Infectious Diseases) (Grade: Excellent)
National School of Public Health, Athens, Greece

November 2014
Master’s Thesis: Assessment of neutrophil-to-lymphocyte ratio in blood and cerebrospinal fluid as index of differential diagnosis between bacterial and viral meningitis
Mentor: Prof. G. Tzanakaki, Ph.D.

July 2011
Doctor of Medicine degree (Grade: Highest honors)
Faculty of Medicine, School of Health Sciences,
National and Kapodistrian University of Athens, Greece

June 2005
High School Baccalaureate (Major in Sciences) (Valedictorian)
“GEITONAS” Private High School, Athens, Greece

2. DISTINCTIONS – AWARDS

November 2019
Certificate of merit in recognition of excellence in the field of Medicine, Hellenic Medical Society of New York, New York, US

November 2019
Selected by Scientific Committee as participant (ranked 4th among >100 applicants) for the Singularity University Summit Greece 2019

June 2019
Selected by Scientific Committee as participant for the Johns Hopkins University & Stavros Niarchos Foundation Bioethics Academy 1st summer course, June 20-22, 2019, Athens, Greece
<table>
<thead>
<tr>
<th>Year</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>July 2017</td>
<td>Travel grant to attend the <em>2017 Lectures on ‘BIG DATA and Applications’ under the aegis of the President of Hellenic Republic, Alexander Onassis Public Benefit Foundation</em>, Heraklion, Crete</td>
</tr>
<tr>
<td>September 2016</td>
<td>Grant for Ph.D. studentship <em>ad honorem</em> as ex-Valedictorian and Alumnus of “GEITONAS” School, Athens, Greece</td>
</tr>
<tr>
<td>2008 – 2009</td>
<td>Award for excellence in studies of Medicine during the academic year 2008-9 Year’s score: 9.33/10, ranking: top 1%, National Scholarship Foundation (IKY), Athens, Greece</td>
</tr>
</tbody>
</table>
LIST OF PUBLICATIONS

The present thesis has been based, in part, on the following publications:


h. Mentis A. Epigenomic engineering for Down syndrome. *Neuroscience and Biobehavioral Reviews* 2016. 71, 323-327

Investigating genetic, microbial-viral, and other environmental factors in neuroinflammatory and neurodegenerative diseases of CNS through ‘-omics’ or ‘meta-research’ approaches

ALEXIOS-FOTIOS MENTIS
UNIVERSITY OF THESSALY, FACULTY OF MEDICINE, 2020

THREE-MEMBER SUPERVISORY COMMITTEE

1. Efthimios Dardiotis, M.D., M.Sc., Ph.D., Assistant Professor of Neurology & Director, Laboratory of Neurogenetics, Department of Neurology, Faculty of Medicine, University of Thessaly (Supervisor)
2. Efthimia Petinaki, M.D., Ph.D., Professor of Microbiology & Clinical Biochemistry & Director, Department of Microbiology, Faculty of Medicine, University of Thessaly
3. Georgia Xiromerisiou, M.D., Ph.D., Assistant Professor of Neurology, Department of Neurology, Faculty of Medicine, University of Thessaly
ABSTRACT

Neurological disorders represent a major component of the global disease burden. Nonetheless, in contrast to cancer and heart disorders, two broad disease categories which have been extensively studied in the last three decades, neurological disorders have been understudied. As a result, their underlying causes, and the potential contribution of genetic and non-genetic (also described as environmental) factors remain underexplored. To address this hovering research gap, and in light of an increasing body of studies in the field, I aimed to study the contribution of several genetic and non-genetic risk factors in neurodegenerative and neuroinflammatory disorders, harnessing Alzheimer’s and Parkinson’s diseases, and Multiple Sclerosis as exemplars, respectively. By applying both -omics and meta-research (i.e., secondary research) approaches, my analyses suggest that: a) the approach of meta-umbrella systematic review can provide valuable insights into non-genetic protective and risk factors, and represents a third-generation meta-research approach, preceded by umbrella and systematic reviews; b) APOE4/APOE4 (the chief genetic risk factor of Alzheimer’s disease) may be implicated in regulating the recently discovered meningeal lymphatic vessels by causing lymphosclerosis, the lymphatic equivalent of atherosclerosis, and in turn, meningeal and brain lymphedema, thus collectively impeding the flux of CSF toxic metabolites from the brain to the cervix and potentially leading to accumulation of amyloid-beta and tau in the brain; c) On the single-microbe level, the gut microbe *Helicobacter pylori* may be associated with increased risk for Parkinson’s disease, whose clinical scores such as the Unified Parkinson's Disease Rating Scale (UPDRS) may be reduced following *Helicobacter pylori*’s eradication; and, d) On the microbiome level, the CSF may not be devoid of the presence of viruses, for which our advanced bioinformatic analysis indicated separate clustering of CSF virome in Multiple Sclerosis patients versus controls. In parallel to the above data-intensive approaches, and to justify the historic(al) roots of a Ph.D. thesis as Doctor Philosophicus, I also present a critical analysis of the so-called ELSI (Ethical, Legal, and Societal Implications) of Neurological Research. In doing so, I argue that: a) the research niches of countries with limited resources for research (yet often classified as wealthy) can be regarded as hidden pearls that can boost local research pipelines in times of austerity, and to this end, I provide specific examples from patient population and biotechnological resources in Greece; b) the post-CRISPR/Cas9 genome editing era represents a tremendous opportunity as potential therapeutic armamentarium even for complex disorders with neurological manifestations, such as Down’s syndrome (Trisomy 21); however, rigorous bioethics regulations should be set prior to clinical trials, especially if the latter are applied during the embryological stage of patients with Down’s syndrome; and c) stronger public advocacy is needed for syndromes with neuropsychiatric manifestations, such as the Imposter syndrome, whose high prevalence and psychological burden has had a huge
impact in societally underrepresented groups. Collectively, my Ph.D. thesis aims to shed some light into the causality of certain neurological disorders. I anticipate that the results and conclusions presented here will open an avenue to future research questions; prior to these future steps, however, addressing the ELSI is pivotal to ensure bioethics compliance and societal acceptance of neurological research.
# TABLE OF CONTENTS

ACKNOWLEDGMENTS ................................................................................. 5  
PREFACE .................................................................................................. 6  
BRIEF RESUME .................................................................................... 10  
LIST OF PUBLICATIONS ....................................................................... 12  
ABSTRACT .............................................................................................. 14  
LIST OF FIGURES .................................................................................. 23  

CHAPTER 1 — INTRODUCTION: RISK AND PROTECTIVE FACTORS FOR NEUROLOGICAL DISORDERS ................................................................................. 27  
1. Introduction ......................................................................................... 27  
2. Methods ............................................................................................. 30  
   2.1 Structure of Umbrella Reviews ..................................................... 30  
   2.2 Search Strategy and Eligibility Criteria ....................................... 30  
   2.3 Data Extraction ............................................................................ 31  
   2.4 Data Analyses ............................................................................... 32  
   2.5 Assessment of Epidemiological Credibility .................................. 33  
   2.6 Data Availability Statement ......................................................... 34  
3. Results ................................................................................................ 34  
   3.1 General Results ........................................................................... 34  
   3.2 Commonly observed Protective and Risk Factors of Neurological Conditions ........................................................................ 35  
   3.3 Specific Risk and Protective Factors of Neurological Conditions ........ 36  
4. Discussion .......................................................................................... 37  
   4.1 Principal Findings ......................................................................... 38  
   4.2 Smoking as an Exemplar of studying Risk and Protective Factors for Neurological Disorders ................................................................. 40  
   4.3 Additional features with class I evidence ...................................... 42  
   4.4 Implications for Target Groups ..................................................... 44  
   4.5 Strengths of our meta-umbrella approach ..................................... 46
CHAPTER 2 ─ TOWARDS A MULTIFACETED ROLE OF CHIEF GENETIC RISK FACTORS: THE CASE OF APOE4 IN ALZHEIMER DISEASE ................................. 62

1. Risk Factors for Alzheimer Disease (AD) ........................................ 62
2. The role of APOE4 as genetic risk factor for AD .............................. 63
3. APOE4 in AD-associated neurovascular and cerebrovascular function .......................... 66
4. APOE4 and the meningeal vs. traditional peripheral lymphatic system .......... 67
   4.1 The traditional peripheral vs. meningeal lymphatic system ................. 67
   4.2 Crosstalk of meningeal and cervical lymphatic vessels ...................... 69
   4.3 Investigating the potential links of APOE4 and the lymphatic system ...... 70
5. APOE4, aquaporin 4, and the glymphatic vs. the meningeal lymphatic systems 72
6. Expression of lymphatic-vessel genes in APOE4-expressing cell types of the brain 73
   6.1 Cell-specific effects of APOE4 in AD pathology .............................. 73
   6.2 The common mesodermal origin of lymphatic vessel cells and microglia ... 74
   6.3 Re-analysis of data from previous studies ..................................... 74
   6.4 Assessing the relevance of re-analyzed data on CNS lymphatic biology ..... 76
   6.5 The notions of attenuated lymphaticness, meningeal lymphedema, and lymphosclerosis in APOE4-related AD ........................................ 77
   6.6 Potential alternative explanations for our proposed conceptual framework . 78
   6.7 Framing of our conceptual framework into the broader evidence regarding AD pathogenesis ................................................................. 79
7. Major challenges and limitations ..................................................... 80
8. Conclusions and future perspectives ............................................... 82
9. Methodology applied for Re-Analysis: Identification of Studies, and Differential Expression (Re)analyses ........................................................... 83
10. FIGURES .................................................................................... 86
CHAPTER 3 ─ THE SINGLE MICROBE CONTRIBUTION TO NEUROLOGICAL DISEASES: THE CASE OF HELICOBACTER PYLORI INFECTION

1. Helicobacter Pylori Infection and Gastric Cancer Biology: Tempering a Double-Edged Sword

1.1 Introduction

1.2 Gastric cancer dynamics

1.2.1 Cancer dynamics and evolution in gastric tumors

1.2.2 Heterogeneity of cancer predisposition in gastric tissue

1.2.3 Stem cells and their biomarkers in gastric cancer

1.3 Mechanobiology

1.3.1 Three-dimensional (3-D) organoid cultures for analyzing H. pylori interactions with gastric tissue during tumorigenesis

1.3.2 Matrix metalloproteases and mechanobiology

1.4 H. pylori and its interaction with stomach microbiome and the host tissue

1.4.1 H. pylori in the microbiome and biofilm studies

1.4.2 H. pylori and host tissue response

1.5 Therapeutic perspectives: from bench to bedside

1.6 Conclusions

1.7 FIGURES

2. Helicobacter pylori Infection and Metachronous Gastric Cancer: A Critical Analysis

2.1 Introduction

2.2 Eradication of Helicobacter pylori infection and risk of metachronous gastric cancer

2.3 Methodological remarks in the recent clinical trial

2.4 Additional patient and pathogen factors that should have been considered

2.5 Conclusion

3. Helicobacter pylori infection in patients with Parkinson’s disease: a case-control study and meta-analysis

Institutional Repository - Library & Information Centre - University of Thessaly
01/11/2023 00:13:51 EET - 35.160.27.221
3.1 Introduction ............................................................................................................. 108

3.1.1 Parkinson’s disease ............................................................................................ 108

3.1.2 Helicobacter pylori ............................................................................................ 108

3.1.3 H. pylori and Parkinson’s disease ...................................................................... 108

3.2 Methods .................................................................................................................. 109

3.2.1 Case-control association study .......................................................................... 109

3.2.2 Meta-analysis ..................................................................................................... 110

3.2.3 Statistical analysis .............................................................................................. 111

3.2.4 Results ................................................................................................................ 111

3.2.5 Discussion .......................................................................................................... 113

3.3 Conclusions ............................................................................................................ 117

3.4 TABLES .................................................................................................................. 118

3.5 FIGURES ................................................................................................................. 120

CHAPTER 4 ─ THE CONTRIBUTION OF MICROBIOME TO NEUROLOGICAL
DISEASES: THE CASE OF CSF MICROBIOME AND VIROME IN MULTIPLE
SCLEROSIS .................................................................................................................. 125

1. Viruses and Multiple Sclerosis: From Mechanisms and Pathways to
Translational Research Opportunities ........................................................................... 125

1.1 Multiple Sclerosis: the clinical and historical context ........................................... 125

1.2 Virus-induced MS: from evolution to mechanistic insights from animal
models ............................................................................................................................ 126

1.3 Humans as models for multiple sclerosis ............................................................. 128

1.4 Stem cells-based modeling of multiple sclerosis: the role of viruses ................. 129

1.5 The viral pathophysiology of Multiple Sclerosis: a cornucopia of theories 129

1.5.1 Molecular mimicry .............................................................................................. 129

1.5.2 Bystander activation ........................................................................................... 130

1.5.3 Epitope spreading ............................................................................................... 131

1.5.4 Fertile field theory ............................................................................................. 131

1.5.5 Viral déjà vu ....................................................................................................... 131

1.6 Peripheral vs. central progression in multiple sclerosis ....................................... 132
New theories on viral etiology: the enteric virome and multiple sclerosis.. 133

“-Omics” and “systems medicine”: future perspectives in the MS-virus arena 134

Susceptibility biomarkers and next-generation sequencing ...................... 134

The search for direct evidence of viruses in MS ........................................ 135

Deciphering host-virus interactions ......................................................... 136

Beyond the brain: the contribution of extra-CNS immunity to MS ............ 136

Viruses in the treatment of multiple sclerosis .......................................... 137

Concluding remarks ............................................................................. 137

2. Viruses and Endogenous Retroviruses in Multiple Sclerosis: From Correlation to Causation ........................................................................ 143

Introduction ............................................................................................ 143

Search methods and selection criteria .................................................... 144

Epstein-Barr virus (EBV) ....................................................................... 144

Human herpesvirus 6 (HHV–6) ............................................................... 147

Varicella zoster virus (VZV) ................................................................... 148

Human endogenous retroviruses (HERVs) ............................................. 149

Human immunodeficiency virus (HIV) ................................................... 150

Cytomegalovirus (CMV) ......................................................................... 151

Measles and other morbilliviruses .......................................................... 152

Lymphocytic choriomeningitis virus (LCMV) ........................................... 152

Coronavirus ............................................................................................ 153

Saffold virus ............................................................................................ 153

Conclusions and suggestions for future research ..................................... 154

TABLES ................................................................................................. 156

FIGURES ............................................................................................... 157

3. Experimental analysis of CSF Microbiome and Virome and associated transcriptomics profiles ................................................................ 158

20
CHAPTER 5 — THE ELSI (ETHICAL, LEGAL, AND SOCIAL IMPLICATIONS) OF NEUROLOGICAL AND BIOMEDICAL RESEARCH: A CRITICAL VIEWPOINT

1. Neurological and Biomedical Research in times of social and financial austerity
   1.1 Biomedical Research: Lessons from the Last Decade’s Crisis and Austerity-stricken Small Countries to the Current COVID-19-related Crisis

2. The ethical aspects of neurological research
   2.1 The case of Epigenomic Engineering for Down Syndrome
      2.1.1 Introduction
      2.1.2 Epigenetics and the neurobiology of Down syndrome
      2.1.3 Epigenetic targets in Down syndrome: indicators from pre-clinical studies
      2.1.4 CRISPR-Cas9 for epigenomic editing in Down syndrome: opportunities and limitations
      2.1.5 Conclusions

3. Neurological and Biomedical Research and its psychological implications: The Example of Imposter syndrome as threat to diversity

APPENDICES, AND SUPPLEMENTARY FILES AND TABLES

1. Chapter 1
   1.1 Supplementary File – Update on Search Results
   1.2 Appendix 1

2. Chapter 2
2.1 Supplementary File 1.................................................................................... 223
2.3 Supplementary File 3.................................................................................... 233
2.4 Supplementary Figure 1................................................................................ 234
2.5 Supplementary Figure 2................................................................................ 235
2.6 Supplementary Figure 3................................................................................ 236
2.7 Supplementary Figure 4................................................................................ 237
2.8 Supplementary Figure 5................................................................................ 238

3. Chapter 3........................................................................................................... 239
   3.1 Supplementary file 1 ..................................................................................... 239
   3.2 Supplementary File 2.................................................................................. 239

REFERENCES ....................................................................................................... 240
LIST OF FIGURES

CHAPTER 2 - TOWARDS A MULTIFACETED ROLE OF CHIEF GENETIC RISK FACTORS: THE CASE OF APOE4 IN ALZHEIMER DISEASE

Figure 1. Hypothesized role of apolipoprotein e (apoe) and lymphatic vessels in alzheimer disease (AD). ................................................................. 86

Figure 2. Box and whisker plots showing meningeal lymphatic marker expression levels in ipsc-derived APOE4 and APOE3 (control) knock-in cells ............... 87

Figure 3. Box and whisker plots showing meningeal lymphatic marker expression levels in ipses derived from individuals with sporadic AD APOE4 (e4/e4) genotype ................................................................. 88

CHAPTER 3 - THE SINGLE MICROBE CONTRIBUTION TO NEUROLOGICAL DISEASES: THE CASE OF HELICOBACTER PYLORI INFECTION

CHAPTER 3.1 - HELICOBACTER PYLORI INFECTION AND GASTRIC CANCER BIOLOGY: TEMPERING A DOUBLE-EDGED SWORD

Figure 1. The development of gastric cancer as induced by H. Pylori and other factors. ................................................................. 102

CHAPTER 3.3 - HELICOBACTER PYLORI INFECTION IN PATIENTS WITH PARKINSON’S DISEASE: A CASE-CONTROL STUDY AND META-ANALYSIS

Figure 1. The selection process of eligible studies. ........................................ 120

Figure 2. Forest plot: meta-analysis of the prevalence of H. Pylori infection in the pd patient group compared with the healthy control group. .................. 121

Figure 3. Funnel plot to detect publication bias in the studies reporting H. Pylori infection in pd patients and healthy controls. ........................................ 122

Figure 4. Forest plot: standardized mean difference of mean updrs scores between H. Pylori infected (hp+) and non-infected (hp-) PD patients. ....................... 123

Figure 5. Funnel plot to detect publication bias in the studies reporting mean UPDRS scores in H. Pylori infected and non-infected PD patients. ....................... 124

CHAPTER 4 - THE CONTRIBUTION OF MICROBIOME TO NEUROLOGICAL DISEASES: THE CASE OF CSF MICROBIOME AND VIROME IN MULTIPLE SCLEROSIS

23
CHAPTER 4.1 - VIRUSES AND MULTIPLE SCLEROSIS: FROM MECHANISMS AND PATHWAYS TO TRANSLATIONAL RESEARCH OPPORTUNITIES

Figure 1. The viral pathophysiology of Multiple Sclerosis........................................ 139

Box 1 ......................................................................................................................... 140

CHAPTER 4.2 - VIRUSES AND ENDOGENOUS RETROVIRUSES IN MULTIPLE SCLEROSIS: FROM CORRELATION TO CAUSATION

Figure 1. Putative mechanisms of human endogenous retrovirus (HERV)-related autoimmunity in Multiple Sclerosis. ................................................................. 157
LIST OF TABLES

CHAPTER 1 - INTRODUCTION: RISK AND PROTECTIVE FACTORS FOR NEUROLOGICAL DISORDERS

Table 1. Common protective factors of non-communicable neurological disorders.. 56
Table 2. Common risk factors of non-communicable neurological disorders.......... 57
Table 3. Variables that are both risk and protective factors of non-communicable neurological disorders. ............................................................... 58
Table 4. Specific risk and protective factors of non-communicable neurological disorders ............................................................................. 59

CHAPTER 2 - TOWARDS A MULTIFACETED ROLE OF CHIEF GENETIC RISK FACTORS: THE CASE OF APOE4 IN ALZHEIMER DISEASE

Table 1. Genetic, environmental, and lifestyle risk factors for Alzheimer Disease (AD) appearance or progression.................................................. 89
Table 2. Markers of human lymphatic vessels. ...................................................... 90

CHAPTER 3 - THE SINGLE MICROBE CONTRIBUTION TO NEUROLOGICAL DISEASES: THE CASE OF HELICOBACTER PYLORI INFECTION

Table 1. Demographic data of pd patients and healthy controls included in case-control study ........................................................................ 118
Table 2. Characteristics of the studies included in the meta-analysis of h. Pylori infection in pd patients and healthy controls. .................................. 119

CHAPTER 4 - THE CONTRIBUTION OF MICROBIOME TO NEUROLOGICAL DISEASES: THE CASE OF CSF MICROBIOME AND VIROME IN MULTIPLE SCLEROSIS

CHAPTER 4.1 — VIRUSES AND MULTIPLE SCLEROSIS: FROM MECHANISMS AND PATHWAYS TO TRANSLATIONAL RESEARCH OPPORTUNITIES

Table 1. Important animal models used in multiple sclerosis research.............. 141

CHAPTER 4.2 - VIRUSES AND ENDOGENOUS RETROVIRUSES IN MULTIPLE SCLEROSIS: FROM CORRELATION TO CAUSATION

Table 1. The association between environmental viruses and herv elements and the downstream effects ........................................................................ 156

APPENDIX
Supplementary Table S1. Characteristics and quantitative synthesis of the 76 eligible meta-analyses of non-genetic (environmental) risk and protective factors for non-communicable diseases. .................................................................................................................. 205

Supplementary Table S2. Methodological quality of included umbrella reviews based on the amstar criteria and score. ...................................................................................................................... 222
Chapter 1 — Introduction: Risk and Protective Factors for Neurological Disorders

Non-genetic Risk and Protective Factors and Biomarkers for Neurological—Neuropsychiatric Disorders: A Meta-Umbrella Systematic Review of Umbrella Reviews

1. Introduction

Neurological disorders are the leading and second-leading causes of, respectively, disability and mortality worldwide (1, 2). Nowadays, because of improvements in quality of life, population growth, and longevity, a higher proportion of people are reaching ages with the highest neurological disorder prevalence (3). Furthermore, despite their high contribution of neurological disorders to the Global Burden of Disease, there has been only partial elucidation of the etiologies of non-communicable neurological disorders (for a discussion, which extends this study’ aims on the potential microbial (communicable) etiology of such disorders, (see (4), such as Alzheimer’s disease (AD), movement disorders, and neuro-inflammatory diseases (e.g., Multiple Sclerosis (MS)), among others. This important gap lies in contrast to other diseases, such as cancer and cardiovascular disorders, where research efforts have been far more prolific. This said, most published findings suggest an interplay of genetic predisposition risk and protective factors for neurological disorders (5-7). In parallel, major health and public policy reports provide annual updates assessing how much major risk factors contribute to the chronic burden of neurological diseases and have addressed urgent calls for action on such disorders, including mitigation of risky lifestyle factors (8-10).

The contribution of several genetic factors to neurological disorders has been examined. Single-nucleotide polymorphisms (SNPs), as well as genome (GWAS) - and transcriptome-wide association studies have revealed numerous possible related SNPs and mechanistic clues (11-13). Field-synopses, as well as, meta-analyses of GWAS have been reported (14-19). Nonetheless, as shown by updated study designs (phenome/ exposome/ environment-wide association studies) and relevant statistical tools (mediation and multivariable Mendelian randomization), this interplay has become even more complex because of many confounders (20-26). For instance, aging appears to be a major risk factor for neurodegenerative disorders; however, the aging process encompasses the (patho)physiological unfolding of life, as well as the contribution of genetic and lifestyle determinants (27). On the other hand, environmental factors in many cases are modifiable (further discussed in (28)). These factors contribute significantly to chronic noncommunicable disease progression, notably around 25% of global deaths may be due to threatening changes in our environment (29). Similarly, around 60% of cardiovascular mortality, a principal contributor to total

Institutional Repository - Library & Information Centre - University of Thessaly
mortality, can be attributed to eight major preventable risk factors (30). Thus, obtaining solid evidence on these modifiable factors is crucial for evidence-based clinical neurological counselling, health promotion strategies, and patient risk awareness, addressed either at high-risk individuals or at the population at large (31). Interestingly, recent attempts using health insurance datasets have been made to co-examine the contributions of genetic and environmental factors on the same individual’s clinical phenotype (32).

Umbrella reviews, whose number has been blossoming since the first endorsement of this review type by Cochrane in 2009, are structured through the systematic retrieval, collection, and assessment of information and tested for consistency of evidence of previously published systematic reviews and meta-analyses (31) (33), as initially discussed in (34). The end result is to collate compelling evidence into a single, informative review offering a broad view of a field to the medical community, hopefully covering knowledge gaps (35). In particular, an umbrella review facilitates the comparison between different meta-analyses by repeating the analyses of the latter in a so-called uniform approach for all factors, considering the expected variability in their quality, focus-of-interest, and the degree of reliability of evidence (31) (36). This methodology appears to have increased statistical power and is frequently employed to help synthesize the available literature to guide both clinical care and public health policies. Thus, umbrella reviews lie at the top of the hierarchy in the evaluation of evidence (2).

Examples of umbrella reviews have included the analysis of risk and protective factors for a certain disease or condition, or the effects of certain such factors on multiple health outcomes, based on meta-analyses or Mendelian Randomization studies (37), diagnostic criteria, and screening tools (38), diagnostic accuracy studies (39), therapeutic interventions (40), and/or interactions between genetic and environmental factors (41). With regards to brain health, several umbrella reviews have analyzed systematic reviews and meta-analyses reporting an association between environmental factors and a single non-communicable neurological disorder (e.g., the risk factors for Multiple Sclerosis) (42) (43) (44), while others have studied the role of a single factor (e.g., vitamin D) into multiple health outcomes, including neurological disorders (45).

As it is increasingly recognized that the factors in question may exert distinct (even opposite) effects in different neurological disorders, and that the evidence or its epidemiological credibility may be different across distinct neurological disorders (46). Thus, there is an urgent need to identify, compare, and contrast—common versus specific, frequent versus rare, similar versus opposite—risk as well as protective factors of neurological disorders. For example,
obesity may be one of the greatest risk factors for a plethora of disorders and their natural history (47); however, it may also be linked to increased survival in heart failure and coronary artery disease (obesity paradox) (48), while it is inversely associated with the onset and prognosis of amyotrophic lateral sclerosis (ALS) (49, 50), or with the disease activity of patients with systemic lupus erythematosus (51).

Such an *overarching* or encompassing study will be clinically important, as it will provide the opportunity to assess neurological disorders with shared *versus* specific risk and protective factors, which an umbrella review of a single risk factor is, by its design, not capable to address. Indeed, addressing common factors for comorbid disorders or comorbid patterns is an area of intense research (52, 53), having demonstrated so far *aggregates* of mental health, musculoskeletal, cardiovascular and metabolic problems, (54); this is especially important in light of the increasing aging of the population. Moreover, it is prudent for the busy clinician and public health scientist, to present all existing evidence on reliable and consistent factors in a single systematic review, in order to compare and contrast the contributing factors among different diseases, and the levels of their epidemiological credibility, and ultimately, to advocate for guidelines based on these findings. With that in mind, performing a *systematic review of umbrella reviews*—an approach we wish to call *meta-umbrella*—may save an enormous amount of time compared with obtaining and reading the large number of individual umbrella reviews.

Hence, the aim of this *meta-umbrella* review was five-fold: *a*) To critically review, in a systematic manner, the data presented in previous umbrella reviews regarding risk and protective factors for the sum of chronic non-communicable neurological disorders analyzed in these umbrella reviews, in order to offer an *overarching* field-wide overview; *b*) To assess the *cream-of-the cream* evidence spanning the last decades and to highlight factors that have displayed the most persuasive evidence of an association; *c*) To introduce an additional type of study design in the blossoming *meta-research* field; this approach could be applied to other disease categories (e.g., cardiovascular or neoplastic diseases) in the future; *d*) To equip clinicians, preventive medicine specialists, and policymakers with solid evidence for performing their health care-related tasks, and for creating policy-formulating guidelines to address neurological disorders with shared risk and protective factors; and, *e*) To provide a thorough discussion on the mechanisms underpinning the association of these risk and protective factors with neurological disorders, in order to address research gaps, at both translational and clinical levels, regarding how these factors interact with the pathogenesis of neurological diseases.
2. METHODS

2.1 Structure of Umbrella Reviews

We conducted a systematic review of umbrella reviews, which we call a *meta-umbrella review*; no advance registration of the review’s goals or protocol took place. Our systematic search of the literature demonstrated that published umbrella reviews follow two approaches—a review of known risk factors of a single clinical outcome and a review of the relation between a single risk factor and multiple clinical outcomes (for example, (44), (45)). For evidence from observational associations between chronic non-communicable neurological disorders and known genetic risk and/or protective factors or a review of the relations between a single risk and protective factor with multiple neurological disorders, we aimed to retrieve data from published systematic (i.e., not narrative) umbrella reviews of systematic reviews and meta-analyses. Such umbrella reviews (e.g., in (42)) were conducted using standardized methods (reviewed in (31)). Following guidelines for conducting umbrella reviews, we comprehensively analyzed the quantitative data of the meta-analyses conducted in published umbrella reviews (35) (31) (55) (56), while the qualitative results of systematic reviews discussed in these umbrella reviews were not further considered.

2.2 Search Strategy and Eligibility Criteria

Using a standardized search strategy (Appendix 1), we systematically explored the following databases: MEDLINE, SCOPUS, Web of Science, Cochrane Database of Systematic Reviews, Cumulative Index to Nursing and Allied Health Literature (CINAHL), and ProQuest Dissertations & Theses (in order to take into account grey literature), as well as JBI Database of Systematic Reviews and Implementation Reports, DARE, and the PROSPERO registered up to September 20, 2018, to identify umbrella reviews analyzing associations of genetic risk and protective factors with multiple chronic neurological disorders or umbrella reviews of single such factors with multiple clinical outcomes (including neurological disorders).

We used broad search terms (umbrella review$ OR umbrella review$.ti,ab.); other relevant keywords (stroke*, Alzheimer disease or dementia*, multiple sclerosis*, headache*, amyotrophic lateral sclerosis*, Parkinson disease, neurolog$) OR (multiple outcomes). Furthermore, harnessing a snowball procedure, citations and reference lists of the umbrella reviews were, also, systematically screened, following the example of other studies (57).

Based on predefined exclusion and inclusion criteria, four reviewers independently conducted a three-step evaluation of the title, abstract, and full text of the papers (Figure 1—
Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) Flowchart of the current meta-umbrella systematic review), and any discrepancy between the three investigators was resolved by consensus. We only opted to retain umbrella reviews that investigated the association of environmental and genetic risk or protective factors with all types of chronic neurological disorders, or umbrella reviews of a single risk factor with multiple clinical outcomes. There was no selective inclusion of umbrella reviews reporting only on systematic reviews of observational (in general) or prospective studies (in particular) or those reporting on clinical (randomized) trials. No language restrictions were applied in the search strings or in eligible study selection. Due to previous concerns that completely distinguished genetic from environmental risk and protective factors could be deceptive, we followed these studies’ pragmatic approach to refer to non-genetic factors with regards to risk and protective factors. This approach is defined as one applying the definitions of the papers that are included in the review process, instead of creating new definitions (28).

Studies were excluded if any of the following was relevant: a) referring to a protocol for umbrella reviews, and not an umbrella review study per se; b) the examined factor(s) was deemed as pure genetic factor or genetic biomarker (as these factors are non-modifiable; c) the factor(s) or neurological conditions referred to mostly neurosurgical and/or brain traumatic disorders (e.g., brain injury or neuro-oncological diseases), or neurological conditions with a subjective component (e.g., pain); d) the studies consisted of umbrella reviews assessing clinical outcomes (e.g., decline, impairment, relapse or remission) of neurological disorders, the severity of their clinical presentation factors, and the effects of a treatment or of an intervention for a neurological disease; and, e) umbrella reviews referring exclusively to studies in animals. Nonetheless, following the methodology of previous studies (58), we did not exclude umbrella reviews that evaluated cross-sectional studies; the latter may not allow causality inference, but they can certainly offer valuable associations.

When multiple meta-analyses on the same research question were qualified, the analyses with the highest number of factors were included. Considering that the umbrella reviews’ structure is following that of standard systematic reviews, its quality and integrity were validated with PRISMA (59).

2.3 Data Extraction

Four investigators, as set of groups of two, independently performed the data extraction from the studies. The first author, journal, and the year of publication of each article that qualified were documented. Additionally, we recorded the risk and protective factors, the biomarkers, as
well as chronic, non-communicable neurological conditions analyzed, number of studies reviewed, study-specific risk estimates [i.e., odds ratio (OR), hazard ratio, risk ratio (RR), or other pertinent effect size] alongside with their corresponding confidence interval (CI), and the number of participants in each study. We, also, reviewed whether the studies included performed any quality control over the reviews and meta-analyses used.

For data extracted from studies where one non-genetic risk factor was reviewed en face of multiple health outcomes, we extracted only those data that were relevant to neurological disorders.

As no standard criteria exist to assess the quality of umbrella reviews, we used the current expert recommendations and A Measurement Tool to Assess Systematic Reviews (AMSTAR) method to assess the quality of reviews (31, 60). Notwithstanding its limitations (which include, among others, a heavy dependence from the so-called reporting quality, and not the methodological one, as well as the lack of focus on the sample size and the design of single studies, as discusses in (61, 62)), AMSTAR is a tool which applies dichotomous scoring (i.e., 0 or 1) for eleven items (e.g., publication bias assessment) to evaluate systematic reviews, notably to assess: a) The quality of their methodology; b) If the search strategy is a thorough one; c) How much prone to bias is every systematic review; and, d) How appropriate are the statistical tools applied for meta-analysis (58). The AMSTAR method was applied by completing a checklist with specific questions on these systematic reviews. If graded between 8–11, 4–7, or 0–3, then AMSTAR scores were deemed of high, medium, and low quality, respectively.

2.4 Data Analyses

We performed a descriptive analysis of umbrella reviews. We specifically reviewed and recorded the summary effect size and its 95% CI using the random-effects model of the meta-analyses presented in these umbrella reviews. This model was opted instead of the fixed-effects model, because a) the random-effects model considers the potential heterogeneity of results between studies; and, b) because the classes of evidence (as below), which were relevant to our selection criteria, were based on values in random-effects model (63, 64). We also recorded the 95% prediction interval for each estimate; this a feature, which by considering the between-study heterogeneity, helps determining the uncertainty with regards to an effect anticipated in a new study in which the same association was assessed (65). In addition, we recorded the $I^2$ metric, which had been used to analyze any inter-study heterogeneity (66). The $I^2$ ranges, which estimate the proportion of inter-study variance over the sum of the intra- and inter-study variances.
variances, were between 0% and 100% (66). Values >50% or >75% represent large or very large heterogeneity, respectively (as in: (42-44)). We also reported whether small-study effects were described, i.e., whether smaller studies exaggerated a reported effect as opposed to larger studies using Egger’s regression asymmetry test (67), when applicable. The underlying rationale is, that both large heterogeneity and plausible reporting of excess significance bias reduce the accuracy of evidence of a contributing factor, regardless of the p-value and effect size.

2.5 Assessment of Epidemiological Credibility

Regarding the association between a risk or protective factor and a neurological disorder, we recorded the conclusions of each umbrella review according to the sub-categories of analyses that were based on the meta-analyses reviewed. Then, we assessed the strength of the association between risk and protective factors and neurological disorders. This strength was assessed using the criteria for the assessment of the epidemiological credibility based on previous umbrella reviews (42-44), (68), (69), (70), (71), (72), (73), (74).

In our meta-umbrella approach, we followed the criteria for assessing credibility of epidemiological evidence based on the above umbrella reviews. This was performed in alignment with recent expert recommendations that these criteria contribute to classify the accuracy level of evidence in a standardized and objective manner in umbrella reviews (46), and as a corollary, in the systematic reviews of umbrella reviews (i.e., meta-umbrella). In so doing, we categorized the observed associations into classes of evidence (i.e., class I–IV), based on whether each association was convincing by using a combination of the following criteria (which take into account both the p-value and the magnitude of the association): levels of significance of the random-effects model (P ≤ 0.05, P ≤ 0.001, and P < 10^-6); level of significance of the largest component study (P < 0.05); inclusion of >1000 participants (or number of participants greater than 20,000 with regards to continuous outcomes); absence of considerable heterogeneity (I^2 < 50%); as well as lack of evidence of either small study effect (P>0.10) or excess significance (P > 0.10); and, 95% prediction interval excluded the null value, as reported in previous studies (42-44), (68), (69), (70), (71), (72), (73), (37). It should also be kept in mind that the variables applied in the above criteria are continuous, whereas the chosen cut-off points are arbitrarily selected (31). More specifically, the rationale for the definitions of class I–IV lies on previous expert recommendations and relevant umbrella review studies, which were already conducted.
Following a pragmatic approach, we applied the above studies’ criteria and definitions on the following classes of evidence, which appear to take into account both p-value and the magnitude of the association (as expressed by prediction intervals): 

a) **Class I (convincing):** statistical significance based on the random-effects model with $P < 10^{-6}$, with $>1000$ cases or deaths (or number of participants greater than 20,000 with regards to continuous outcomes), with the largest component study reporting statistically significant effect ($P < 0.05$), with $95\%$ prediction interval excluding the null value, without large inter-study heterogeneity ($I^2 < 50\%$), with no evidence of excess of significance ($P > 0.10$), and with absence of evidence of small study effect ($P > 0.10$); 

b) **Class II (highly suggestive):** statistical significance with $P < 10^{-6}$, with $>1000$ cases or deaths (or number of participants greater than 20,000 with regards to continuous outcomes), and with the largest component of the study reported statistically significant effect ($P < 0.05$); 

c) **Class III (suggestive):** statistical significance with $P < 10^{-3}$, with $>1000$ cases or deaths (or number of participants greater than 20,000 with regards to continuous outcomes); and, 

d) **Class IV (weak):** the remaining statistically significant associations with $P < 0.05$; 

whereas, 

e) The rest were considered non-significant associations when $P \geq 0.05$ values were observed (reviewed also in (56)). Finally, the statistical analyses were retrieved from the umbrella reviews (when available).

### 2.6 Data Availability Statement

Data sharing is not pertinent or applicable to this article given that no datasets were produced or analyzed during the current study.

### 3. RESULTS

#### 3.1 General Results

The electronic search of the relevant databases yielded 2,797 potentially relevant reviews; of these, 14 umbrella reviews fulfilled the eligibility criteria and were included in the study (**Figure 1**) (42-45), (37, 69-73), (65, 75), (76), (77). On the contrary, no umbrella review on genetic risk or protective factors for neurological conditions was located. Also, no additional studies were identified through the snowball procedure.

Regarding how many studies addressed one disease as an outcome or multiple disease processes, the following separation can be made: 

a) For studies addressing one disease and multiple risk factors, four umbrella reviews identified; 

b) For studies addressing one risk factor for multiple health outcomes, ten umbrella reviews were identified. Also, in the single umbrella review assessing the similarities and differences between meta-analyses of observational...
studies and randomized controlled trials, there was no discordant direction observed for neurological disorders (45).

The 14 umbrella reviews corresponded to 106 unique meta-analyses of factors with marked association with eight neurological conditions (AD, PD, dementia, cognitive impairment, MS, ALS, neuromyelitis optica, and stroke). Among these umbrella reviews, four were identified as studies addressing multiple risk factors for a single neurological condition (42-44, 73), whereas ten umbrella reviews were identified as addressing one risk factor for multiple health outcomes, including neurological conditions (37, 45, 65, 69-72, 75-77).

The minimum and maximum number of studies included in the systematic reviews and meta-analyses of any of the risk factors was 3 and 67, respectively. **Supplementary Table 1** describes the factors, neurological conditions, and overall results of the selected studies, including the detailed data extraction. The median number of primary studies per meta-analysis was 9 (Interquartile range, IQR: 3–67), and the median number of participants was 6,673 (IQR: 121–9,902,859). The selected 106 meta-analyses investigated multiple non-genetic factors, such as diet, drugs, medical history, and comorbid disease, as well as habits and exposure to toxic environments. When multiple reviews reported the same factor, data were selected from the most recent review or the review of the highest quality. Overall, we summarized 72 factors associated with these neurological conditions.

**Supplementary Table 2** shows that the methodological quality of the selected umbrella reviews was assessed using the AMSTAR criteria (60). The total AMSTAR score of the reviews ranged from 7 to 9 points (range: 0–11 points; mean score, 8.0 points; standard deviation (SD), 0.39). Questions most frequently satisfied were questions 2–4 (related to duplicate study selection and data extraction, search comprehensiveness, and inclusion criteria). Other aspects of the AMSTAR score commonly satisfied by the reviews included questions 6–9 (related to the characteristics and scientific quality of the included studies, along with appropriateness of the methods used to combine the studies). Questions 1, 11 (related to a-priori study design and conflict of interest), and 5 (pertaining to reporting and provision of included and excluded list of studies) were the least frequently satisfied.

### 3.2 Commonly observed Protective and Risk Factors of Neurological Conditions

**Table 1** summarizes the protective factors of neurological conditions. Mediterranean diet was a common protective factor for dementia, AD, cognitive impairment, and stroke. Additionally, elevated plasma uric acid levels were associated with a lower risk of PD based on a class II
Serum uric acid levels were markedly associated with a lower risk of AD, dementia, multiple sclerosis, neuromyelitis optica, and ALS (class IV evidence). Physical activity, alcohol intake, and serum vitamin C were associated with a reduced risk of developing PD and dementia (Table 1). High serum vitamin B-12 levels were associated with a lower risk of PD and multiple sclerosis. High serum vitamin D levels were associated with a lower risk of AD based on class II evidence; it was also associated with a lower risk of developing multiple sclerosis, dementia/cognitive impairment, and PD (Table 1). Furthermore, high coffee consumption exerted a protective effect against PD and AD, based on class II and III evidence, respectively.

**Table 2** presents the common risk/protective factors of neurological disorders. Based on class III evidence, both exposures to farming and pesticides were risk factors of ALS and PD (Table 2). Exposure to low-frequency electromagnetic fields (class III evidence) was a risk factor of ALS and dementia. Additionally, exposure to organic solvents (class IV evidence) was a risk factor of PD and multiple sclerosis. Based on class IV evidence, *Chlamydia pneumonia* infection and the occurrence of the organism in CSF were risk factors of dementia and multiple sclerosis, respectively.

Two factors—tobacco smoking and hypertension—exerted a mixed (protective and risk) effect on neurological disorders (Table 3). On the one hand, tobacco smoking (based on class I and IV evidence, respectively) contributed to the development of MS and dementia. On the other hand, tobacco smoking was associated with a reduced risk for the development of PD according to class II evidence. Individuals with hypertension exhibited lower risk of developing PD but had higher risk of dementia.

### 3.3 Specific Risk and Protective Factors of Neurological Conditions

**Table 4** presents the specific risk and protective factors of neurological conditions; the findings based on class I–III evidence are summarized here. High β-carotene and n-3 fatty acid intake was significantly associated with a lower ALS risk based on class II and III evidence, respectively. In contrast, exposure to lead and other heavy metals was significantly associated with a higher risk of developing ALS based on class I and II evidence, respectively. A high level of exposure to welding was associated with a lower risk of developing PD than rural living, dairy product intake, constipation, head injury, and having anxiety or depression—all associated with a higher risk of developing the disease (Table 4).

Exposure to non-steroidal anti-inflammatory agents and a higher frequency of social contacts was significantly associated with a lower risk of developing dementia or cognitive
impairment. Contrarily, having a previous stroke, type 2 diabetes mellitus, exposure to aluminum, or lower education were significantly associated with a higher risk of dementia or cognitive impairment. The occurrence of depression at any age, early-life depression, and late-life depression were significantly associated with higher risk of developing dementia or cognitive impairment.

Diphtheria and tetanus vaccination, as well as a higher bone-mineral density in the femoral neck, lumbar spine, and hip, were significantly associated with lower risk of developing MS. Conversely, anti-Epstein Barr Nuclear Antigen (anti-EBNA) IgG seropositivity, infectious mononucleosis, appendectomy at age ≤20 years, Epstein Barr virus (EBV) DNA in mononuclear cells and serum, tonsillectomy at age ≤20 years, traumatic injury, anti-Viral Capsid Antigen (anti-VCA) IgG seropositivity, and chronic cerebrospinal venous insufficiency were significantly associated with a higher risk of developing MS.

Dietary factors, such as a high dietary whole grain intake, high dietary fiber intake, and a high chocolate intake were significantly associated with lower risk of stroke (Table 4).

4. DISCUSSION
Increasing accretion of data led to recent calls for comprehensive, field-wise analyses of risk and protective factors for many human disorders (78). Our study provided a meta-umbrella systematic review of genetic and environmental risk and protective factors associated with chronic neurological disorders published in earlier umbrella reviews and corresponding systematic reviews and meta-analyses. In so doing, our study provides an encompassing and, in parallel, systematic (overarching) perspective for the entire field of risk and protective factors for all diseases affecting a specific system (herein, the nervous system), albeit not with quantitative approaches. In contrary, field-wide meta-analyses using quantitative approaches have assessed the entire field of putative risk and protective factors but for a specific disease (78).

Notably, following previous characterizations of umbrella reviews as next-generation systematic review (35), our approach can be conceived as a third-generation systematic review. It is an approach that aims to offer a new perspective of secondary research (meta-research); this is a field hallmarked by the need to provide the most integrated evidence possible, and one in which several novel study designs have appeared during the last years, e.g., series of systematic reviews and meta-analyses in a single publication, where the analytical unit is the umbrella review study design (79). Likewise, other attempts refer to the field-wide meta-
analyses, in which a meta-analysis of observational studies is conducted on the sum of risk factors under consideration (78), or to synthesis of systematic reviews (e.g., in Neurology (80)), or to a systematic review of systematic reviews (81), as well as overviews of systematic reviews (82). Therefore, our meta-umbrella review could be seen as another study design added to the armamentarium of meta-research.

Although our primary aim was to study the higher possible number of neurological conditions (which we expected to have been analyzed in umbrella reviews), it then turned out that the umbrella reviews had studied only these eight neurological conditions in question. For instance, we could not find umbrella reviews of risk or protective factors for some common or major chronic neurological disorders such as migraine, headache, brain cancer, motor neuron disease, and epilepsy. Therefore, future umbrella reviews should be considered regarding the non-genetic risk factors of these conditions. Interestingly, almost all studies seem to have focused on neurological disorders of resource-rich countries, which could be indicative of the disproportionately lower number of publications regarding meta-research for global health neurology, namely neurological diseases of resource-poor countries (e.g., meningitis, neurocysticercosis).

4.1 Principal Findings

We studied 72 risk/protective factors with a marked association with chronic non-communicable neurological disorders, including genetic and nongenetic biomarkers, dietary factors, drugs, exposure to toxic environmental agents, habits, and medical history or comorbid diseases. Eight factors exhibited a decreased risk for an extensive number of the non-communicable neurological disorders, with those factors ranging in strength from class I to IV.

Notably, the following associations appeared with Class I evidence: a) In dementia, Mediterranean diet and frequency of social contacts were protective factors, while late-life depression and type 2 diabetes mellitus were risk factors; b) In MS, smoking, anti-EBNA IgG seropositivity, and infectious mononucleosis were risk factors; c) In ALS, lead was a risk factor; and d) In PD, physical activity as protective factor and constipation as risk factor (although serious concerns have been previously raised (43)).

While five risk and protective factors had class III and IV evidence of being significant in the occurrence of these neurological conditions, two such factors—tobacco smoking and hypertension—exerted a mixed risk and beneficial effect. Besides, 4, 12, 22, 16, and 3 factors exhibited a specific association with ALS, PD, dementia/cognitive impairment, multiple
sclerosis, and stroke, respectively. Based on the $I^2$ metric, heterogeneity was present in published reports, and few studies were consistent with non-heterogeneous evidence, when data had a prediction interval excluding the null.

With regards to dietary factors, we found substantial evidence highlighting the potential role of the Mediterranean diet in lowering the risk of dementia, AD, cognitive impairment, and stroke. Until now, several meta-analyses have reported quite solid evidence of the beneficial effect of the Mediterranean diet with neurodegenerative diseases, such as AD and other dementias (for an example, see the umbrella review (69), and (83)). However, discrepancies have reportedly occurred in the cardiovascular benefits related to the Mediterranean diet across socioeconomic groups (84). Because of different reporting methods across studies in the field, development of standardized tools is imperative for the assessment of the effectiveness of the Mediterranean diet to prevent cognitive impairment and neurodegenerative diseases. In a similar context, a recent study, which consisted of a series of meta-analyses and which included more than 130 million person-years of data from more than 240 original studies, presented quite robust evidence of low glycaemic index food intake and stroke reduction (79).

In parallel, negative associations between coffee consumption and PD and AD (85) have been reported; these findings have been consistent across study designs and geographical settings. The biological mechanism(s) underlying this protective effect remain(s) unclear. For example, regular coffee intake enhances insulin sensitivity, and hence, reduces the risk of diabetes mellitus type 2, which itself is a strong risk factor for cognitive decline (86). Also, recent meta-analyses, having considered the plausible roles of numerous modifiers, suggest that a 3.5 cups/day intake is inversely associated with all-cause mortality; this association remained undiluted even after adjusting for major modifiers, such as ageing, smoking, and alcohol (87).

This systematic review of umbrella reviews revealed counterintuitively significant associations of low serum uric acid levels with an increased risk of several neurological diseases (i.e., AD, PD, dementia, MS, neuromyelitis optica, and ALS). Our credibility assessment revealed that with the exception of PD (with class II evidence), these significant associations were within class IV evidence (72). Hence, no definitive conclusion could be made in favour or not of intensive lowering of serum uric acid levels in light of a putative higher risk for neurological diseases (88, 89). Further mechanistic studies are needed in this field, using appropriate animal models for each distinct disease entity. Also, clinical trials of increasing serum uric acid in neurological disorders have been conducted (90, 91).
According to class I and III evidence, physical activity exerts a beneficial effect against PD and dementia, respectively. Physical exercise can increase serum uric acid levels, which has been associated with a lower risk of developing PD and dementia (72, 89). However, patients with PD may be unable to exercise much, owing to neurological dysfunction, which might indicate reverse causation (92).

Serum vitamins B12, C, and D levels were associated with lower risk of different neurological conditions, such as MS (as also reported recently (93)), AD, dementia/cognitive impairment, and PD. Around 80% of these meta-analyses represented heterogeneous evidence ($I^2 > 50\%$), which cautioned against false interpretations. The observed heterogeneity most likely arose from different comparison groups in prospective, retrospective, and case-control studies, causing some of the meta-analyses to be derived from studies with diverse, contrasted categories of serum vitamin B12, C, and D levels (42-44). Furthermore, strong evidence linked presence of anti-EBV antibodies to MS (for further discussion, see (94)).

Our meta-umbrella review provides some evidence for a positive association between exposures to farming, pesticides, low-frequency electromagnetic fields, organic solvents, sand, C. pneumonia infection, and the occurrence of several neurological conditions (MS, PD and dementia); however, most of these associations were based on class III and IV evidence that could have resulted from the substantial heterogeneity of the primary studies. Hence, these associations warrant cautious interpretation. We also suggest that the findings on chronic cerebrospinal venous insufficiency should be interpreted with caution, considering both the wide range of the corresponding confidence intervals, and previous reports in the field (95, 96).

Framing our meta-umbrella review into the broader context of studies reviewing risk and protective factors for neurological disorders, we noticed that in another comprehensive review of systematic reviews, exposure to pesticides was identified as the commonest risk factor for AD, ALS, and PD, followed by smoking; the latter was associated with AD and MS (80).

4.2 Smoking as an Exemplar of studying Risk and Protective Factors for Neurological Disorders

Below, we discuss the findings on the effects of tobacco smoking in a separate section. This is because we consider that with all the body of evidence surrounding this field, smoking should represent an exemplar for studying risk and protective factors for neurological disorders, or, as other authors have previously claimed, representing the poster child of causal relations (97).
To begin with, we observed that tobacco smoking was associated with increased risk of MS (class I evidence) and dementia (class IV evidence), but with a decreased risk of PD (class II evidence). We observed a mildly significant association between hypertension and an elevated risk of dementia and a decreased risk of PD (both class IV evidence). A positive association exists between tobacco smoking and MS, with convincing (i.e., class I) evidence of, at least, a modest effect (44); however, confounding effects cannot be totally denied. More broadly, tobacco smoking has been included in the five principal risk factors that could explain around two out of three initial manifestations of demyelination (further reviewed in (98), (99)). Mechanistically, adverse immuno-modulatory effects, demyelination, and the disruption of the blood brain barrier could be accountable for the positive association between smoking and MS; however, this remains to be proven (100). Of note, the effects of smoking are now well-established regarding lung inflammation, the latter have also been linked with a high risk for MS (99). Of particular interest is also the role of oral tobacco (snuff), which was considered to be associated with a lower risk of MS, potentially through nicotine-mediated effects on subunits of immune cells expressing acetylcholine receptor (99).

Another possibility could be that people suffering from a certain neurological disorder, such as MS, prefer to smoke, whereas those unaffected choose to stop smoking more easily, as previously observed in schizophrenia) (101). Therefore, there is concern that if retrospective studies were included in the initial meta-analyses could influence the results of this meta-umbrella approach. Perhaps, in this specific field, it would have been probably wiser to select from the umbrella reviews only the meta-analyses of prospective studies. Similarly, another possibility could be to consider only those umbrella reviews that had examined credibility ceilings to assess effect estimates in combination with other sensitivity analyses (i.e., to include only prospective studies to assess temporality and reverse causation, or to perform the so-called credibility ceilings, which take into consideration limitations regarding the methodology of the studies) (31) (61) (102). Nonetheless, this option would have been a rather laborious process in the context of this, already extensive, meta-umbrella approach. Besides, it is commonly known that extensively performing sub-analyses into many subgroups could be linked to artificially increasing events of statistical significance. However, the teaching example of cross-sectional studies on lung cancer and smoking (in which case, lung cancer patients tend to quit smoking) for causing inverse causation should always be kept in mind (31).

With regards to PD, the potential underlying genetic and non-genetic roots (or/and bias) of the association between tobacco smoking and PD are reviewed elsewhere (103); however, caution is needed in distinguishing epidemiological terminology (e.g., suggesting that longer
duration of smoking is needed for a risk reduction, as cited in the above study), from core public health messages.

In every case, we feel that the core message of promoting tobacco smoking cessation as an effective public health intervention should remain undiluted, because of its multiple well-established benefits (104, 105), regardless of the extent to which tobacco cessation might also decrease the incidence and/or severity of MS (106), regardless of genetic susceptibility to smoking habits (104, 107). The case of smoking acting both as a risk factor for certain diseases and a protective factor for others should not serve as an opportunity of potentially dilute a key public health message, or even counselling MS-affected patients or their family member who are at higher-than-normal risk in favour of smoking (108).

4.3 Additional features with class I evidence

Another feature with class I evidence was chronic occupational exposure to lead, which presented higher risk for ALS. Arguably, one of the underlying mechanisms in ALS pathogenesis is based on lead toxicity (109, 110). In humans, lead toxicity manifests as ALS-like clinical symptoms, such as weakness originating in the wrist and finger extensors and ultimately spreading to other muscles; the blood of patients with ALS revealed higher levels of lead exposure-biomarkers than healthy controls (110). In potential future research, lead toxicity should not be considered in vacuum but rather in association with other heavy metals and welding, even though the latter two are classified in lower levels of evidence (class II, and class III, respectively). In the same vein, when lead levels are taken into account, the association of copper with ALS risk appears attenuated (111). Certain isotopic compositions of copper have been detected at higher levels in the CSF of ALS patients than that of AD patients or healthy controls (112). Moreover, occupational exposure to silica has also been implicated in ALS risk (113). It would be worth exploring whether the fact that silica belongs to the same family of the periodic table with lead) could explain these traits.

Constipation was positively associated with PD (class 1 evidence). A prospective cohort study reported a significant association with a similar effect size in the meta-analyses (reviewed in (43)). Another study reported that constipation could be a symptom of PD, but also a premorbid symptom preceding motor dysfunction symptoms of PD by at least 10–20 years (114). Nowadays, constipation is regarded as a manifestation of PD via the peripheral nervous system, a condition in which the threshold for the appearance of symptoms may be decreased. This is perhaps because of the larger functional reserve of the midbrain dopamine and integrated basal ganglia motor systems to control movement (115). In any case, the
connection between PD and gut dysfunction seems quite robust. In this context, laboratory studies have demonstrated abnormal deposition of α-synuclein within the submucosal and myenteric plexuses of the enteric nervous system, while, recently, the spread of α-synuclein from the gut to brain (known as the Braak hypothesis) through the vagus nerve was demonstrated in mouse models (43, 116, 117). Moreover, a recent study of a huge cohort of 1.6 million subjects reported that the healthy human appendix contains intraneuronal α-synuclein and misfolded aggregates, and that the early removal of the appendix reduces the risk of developing PD (118). Last, any causal association between beta-2-adrenoreceptor antagonist (beta-blocker) and higher risk for PD appears weak in terms of its evidence (119).

Our meta-umbrella review assessed the specific risk factors related to dementia and AD. While only late-life depression and type 2 diabetes mellitus were positively associated with AD, depression at any stage in life was linked to all types of dementia. In fact, late-life depression was markedly associated with both dementia and AD (120). It remains unclear whether depression is a risk factor for developing dementia or just a prodrome of dementia manifested by progressive cognitive decline (121). The class II evidence of the association of type 2 diabetes mellitus with all types of dementia might reflect type 2 diabetes mellitus-driven susceptibility to different types of dementia, with a modest increase in the risk for AD (42).

Low levels of social interaction markedly affected the occurrence of dementia. Thus, social networking, along with educational and leisure activities, are modifiable protective factors, which might aid in the maintenance of cognitive function with increasing age (for systematic reviews of modifiable factors in dementia, see: (122); (123)). The above describe the concept of brain reserve, which refers to the ability of the individual to tolerate age- and disease-related pathology changes in the brain, without developing clear clinical symptoms or signs (42) (124).

Last, while serum 25-hydroxy-vitamin D has been studied in umbrella reviews on neurological disorders, the same does not hold true for 1,25-hydroxy-vitamin D; the latter was only assessed for cancer (45).

Overall, despite this extensive body of evidence, we wish to stress that the majority of epidemiologically identified protective and risk factors do not lie at the bottom of the health impact pyramid, in which the main social and economic determinants of health, such as education, race, income, and housing, are included (125) (for an umbrella review on how these determinants affect health, see (126)). Interestingly, modification of the epidemiologically
identified protective and risk factors is expected to have the most pronounced impact at the population level, even though they have received significantly less research attention than the socioeconomic determinants—an issue of health equity we attempted to address elsewhere (107). Thus, core public health actions should be undertaken not only top-down, but also bottom-up (i.e., tackling not only the disease-specific, but also the fundamental determinants of health) (127-129).

In addition, there are potentially less appreciated or less easily quantifiable protective and risk factors, such as:

a) the family environment, now-studied through Family-Wide Association Studies (130);
b) the accumulation of physical and emotional stress along the human lifespan (131);
c) living in urban versus rural environments, and in slum versus non-slum urban environments (132-134); and
d) specific nutritional habits, such as milk consumption (135). These factors may be worth exploring in the future, regarding their association with specific and integrated neurological conditions, thus combining epidemiological and environmental neuroscience (136).

4.4 Implications for Target Groups

Major implications for several target groups, namely patients and their caregivers, healthy subjects, clinicians, researchers, environmental health specialists, policy makers, and educational institutions, could be anticipated from this meta-umbrella review. In a way similar to umbrella reviews in other fields (28) (137), this meta-umbrella study provides the opportunity to:

a) Stimulate more comprehensive, patient-centered approaches allowing truly informed decisions during genetic counselling or/and coaching for lifestyle changes (137-139);
b) Enhance the accuracy of predicted onset and natural history of neurological conditions at high-risk populations, especially if coupled with polygenic risk scores (140). In so doing, our study can help advocating disease prognostication based on the protective and risk protective factors identified here-in; and,
c) offer guidance on future prevention interventions to promote protective factors and mitigate amenable risk factors in the general population, especially in young and middle-aged subjects, in whom the so-called window of opportunity still exists (141).

Thus, our approach could assist in promoting campaigns on brain health aimed towards the general public, could increase the level of awareness of neurological conditions, following the successful examples of campaigns regarding cancer and cardiovascular conditions; d) assist policy makers at the local, national, regional, international, and global level to draft new guidelines or updating to existing ones, and to explore how modifying protective and risk factors should be incorporated into national health plans; e) stimulate additional mechanistic, translational, and clinical research on the etiology of neurological conditions, and the many
unanswered questions. Ultimately, this meta-umbrella study could contribute to physicians understanding the contribution of the environment as a risk factor for neurological disorders, as well as, to environmental health specialists appreciating the ties of the environment with the nervous system.

Although several authors have argued that deciphering how the mechanistic effects of certain protective and risk factors differ between distinct neurological disorders (e.g., AD and ALS) is potentially important (142), we feel that maintaining a public health lens approach is always crucial; and, equally importantly; f) compare the magnitude of the associations between several protective and risk factors and specific neurological conditions, a gap in the literature. In this case, the commonality of some protective and risk factors could present an opportunity for holistic policy-making (e.g., promoting Mediterranean diet to prevent a wide spectrum of neurological disorders). It could also serve as an impetus to develop transdiagnostic approaches in neurology as is currently done in psychiatry (28) (for further discussion on the transdiagnostic theory, see: (143) (144)).

Our approach could advocate on behalf of further studies to estimate the mean percentage population attributable fractions of protective and risk factors in all relevant diseases, and in developing the appropriate statistical tools to account for these fractions, in alignment with (145); g) our approach could generate a broader discussion on the role of umbrella and meta-umbrella review approaches as the highest level of evidence in the meta-research field. In the latter context, this study may lead to developing criteria and tools necessary to assess the quality of umbrella reviews and to allow inter-comparisons between such analyses. In the same direction lies the need for consistent a priori publication of protocol of umbrella reviews, in alignment with previous calls (146). Further adherence to common, standardized methodologies will improve upon the earlier suggestions from a decade ago (commented in (46)); h) our meta-umbrella review could hopefully serve as teaching material for courses on preventive neurology offered by the relevant medical education institutions.

Many of the class I evidence results of our study seem to re-affirm previous opinions on implementation science (for a discussion on geopolitical factors affecting implementation of policies on chronic diseases, see: (9)). As previously supported (147), scientists should stop advocating the need for yet another clinical trial on the cognitive benefits of healthy lifestyles and lobby decision-makers to implement societal polices to actively promote propitious lifestyles. This approach will substantially produce benefit not only the brain, but society overall. Complicated situations, in which a factor has both a beneficial and a risk effect (e.g.,
hypertension in PD and dementia), provide an opportunity to highlight the broader potential discrepancies between public health and precision medicine (107); interestingly, this gap in the research literature also calls for implementation science research to guide health policy.

4.5 Strengths of our meta-umbrella approach

This study has several strengths, as it is public health policy−, clinical−, and meta-research−oriented, representing, as similar calls-for-action in other diseases (148), the need to focus, first and foremost, on public health and preventive policies. The first strength includes the use of a methodical and systematic approach for gathering and evaluating all published appropriate quality umbrella literature regarding protective and risk factors for the chronic non-communicable neurological disorders. This may be quite useful to the busy clinician who may not possess the adequate time to perform reviews on his/her own (149), and who, in turn, is offered an overarching, and up-to-date knowledge on a wide array of contributing epidemiological factors. In this context, our approach attempts to address the challenge of evaluating evidence provided by a number of high quality meta-analyses (150).

Our systematic, overarching approach allows study of the top evidence. In general, umbrella designs are of special value when applied to provide an overall picture that can inform guidelines; for instance, rather than examining one risk factor or disease, a meta-umbrella review can consider multiple risk factor for multiple disease processes. For this reason, we ensured that four researchers participated in the search and quality assessment of published studies, enhancing the validity of our analyses. Notably, whereas earlier non-umbrella reviews examined fourteen neurological disorders (and the associated protective and risk factors) as part of the National Population Health Studies-related Guidelines (58), our study attempted to review all neurological disorders in the sum of relevant umbrella reviews. This will become more important if we consider that umbrella reviews themselves are already regarded as the highest level in the hierarchical pyramid of evidence.

In the same context, by performing an umbrella review, a neurology-wide notion is created based on the extent to which systematic reviews produce similar results and conclusions, and, in turn, unravel the consistency or contradiction of evidence in this field (as commented in (2)). In this sense, by co-examining several factors in parallel, umbrella reviews, and, as a corollary, our meta-umbrella approach, can detect irregular patterns in the associations observed (e.g., as the case for farming and pesticides in the study by (150)). To enhance the encompassing character of our study, we have included biomarkers acting as surrogate measurements of epidemiological or clinical features. Although these surrogates might
lower the strength of the evidence, they can still be indicative of the above features, and, thus, offering potentially valuable observations.

Our meta-umbrella approach should be viewed as a third-generation systematic review study design, allowing, among others, to discuss research gaps and to determine the potential mechanistic underpinnings of such associations. In other words, because of the advent in the wealth of umbrella reviews conducted, this meta-umbrella review should be seen as a logical next step, where available umbrella reviews will serve as the analytical unit of the review, in which the meta-umbrella approach will allow selection and inclusion of the highest level of evidence, notably other umbrella reviews (i.e., being the meta equivalent of umbrella reviews, following their description in (2)).

This study obviates the need to perform further research on many different factors, which demonstrated no significant association, as shown in previous studies (108). Generally, it is important to keep in mind that cautious interpretation of p-values should be made. Namely, when a p-value is found on the non-significant range, this means that either no difference exists between the groups or, alternatively, that the number of participants is too low to show if such a difference exists. Similarly, when a confidence interval lies in both sides of the line of no effect, this means that either no difference exists between the groups, or, alternatively, that the number of participants (sample size) is too low to show if such a difference exists. A confidence interval not crossing the line of no effect can provide reassurance on the strength or weakness of the evidence, and importantly, whether or not the study’s results are definitive, with no further need to perform additional similar studies.

4.6 Limitations of our meta-umbrella approach

Some limitations and considerations should be acknowledged.

First, the population and interventions discussed could be perceived as too broad; however, our study is in alignment with its overarching goal. Besides, the wide presentation of protective and risk factors could be perceived as attenuated if not followed by investigations on their biological plausibility.

Second, our analyses from the existing umbrella reviews have heavily relied on the quality of information, meta-analytical methods, and systematic reviews included; this has some major caveats. Indeed, all issues from the original studies, which may harbor distinct study designs, classification criteria, and sample sizes that should preferably not be combined
in meta-analyses), can be transferred up to the meta-analyses, umbrella reviews, and here-in, our meta-umbrella review (discussed further in (73). For example, it has been suggested that in the case that only few studies (i.e., with less than one-thousand patients) have examined a specific protective or risk factor, which may otherwise have a very strong effect, this factor will still be classified as in class IV evidence (31). The underlying reason is that, by the very nature of the study design, the quality of evidence offered by and the conclusions derived from this meta-umbrella approach cannot surpass those that of the umbrella reviews (108). Likewise, the biases and limitations of the meta-analyses are transferred to the umbrella and meta-umbrella reviews, including those relevant to confounding factors (31) (46). Collectively, our meta-umbrella review depends on the choices of the umbrella systematic reviewers, putting confidence on their choices (137).

A similar caveat is that a protective or risk factor, or even systematic reviews not previously analyzed in umbrella reviews may be neglected in a meta-umbrella study; for instance, the role of physical and emotional stress is not discussed in (or even captured by) umbrella reviews. The same holds true for pregnancy-related maternal health, epilepsy, myasthenia, Tourette syndrome and Huntington disease, to name a few examples (in contrast to other kind of reviews (80)). Likewise, certain factors, such as air pollution and other environmental risks, appear under-investigated by the umbrella reviews examined, despite their major contribution to the pathophysiology of chronic diseases (151). Similarly, intermediate risk factors (as defined in (152)), such as birth weight, or other increasingly recognized factors, such as social networks, whose impact on health outcomes is being steadily better understood (153), could have been neglected by umbrella reviews. Another major example is that studies on early life risk factors, including in utero exposure, seem absent in umbrella reviews, even though the early life-related (and even those evoked in a transgenerational manner by progenies, e.g., through epigenetic regulations (154) (155, 156), or those described in the capacity-load model (157)) play crucial roles in the development of adult diseases (158) (159) (160). Interpreting these studies with caution would is essential as, in theory, several early life factors may influence the intercept of adult disease manifestations, whereas other factors may exercise an influence later in a disease course. The above limitation could be studied in detail in the future.

That said, with regards to the majority of protective and risk factors which cannot be examined using standard epidemiological approaches or are not frequently assessed, the level of evidence will be most likely low given the scarcity of the existing relevant data (as further discussed in (28)). Similarly, in light of the massive amount of primary and secondary research
studies, the possibility of a medical field being totally unexplored at the systematic review/meta-analysis level is not high (31).

Umbrella reviews and, as corollary, our meta-umbrella approach are inherently biased in favor of commonly assessed factors, or towards more common neurological disorders (perhaps with the exception of neuromyelitis optica) or those of adult but not pediatric populations, because, so far, all umbrella reviews have been conducted on these disorders. This is because of the lack of umbrella reviews examining factors that are less common in adult or rare neurological disorders, besides the conditions included. Assessing the input of genetic and environmental factors in rare diseases by comparing evidence from common diseases could be hampered by the heterogeneity of studies (further discussed in (73)). Perhaps, key events may lie in the interactions of genetic and non-genetic factors; notably epigenetic modifications (e.g., DNA methylation, chromatin modifications, etc.), the gut microbiome, and others, summing up to the so-called exposome—a field that is still in its infancy (161). Likewise, our study may be biased towards more easily quantifiable factors, neglecting other crucial factors, such as stress—a factor established nowadays for its multiple negative health effects, including neurological disorders (162-166).

Despite our efforts to be as thorough as possible in employing a sensitive literature search strategy, no umbrella reviews except in the English language were found in our systematic literature search; however, we cannot exclude the likelihood that some relevant papers in other languages might have been overlooked. That said, Moher et al. have shown that omitting studies in non-English languages may not introduce considerable bias (167). Similarly, while we assessed the main results of the reviews, we may have overlooked the unreported protective and risk factors that could have been significant. Likewise, our meta-umbrella review might have missed some available evidence, e.g., recently published studies that had not been included in the prior meta-analyses, or because of the blossoming field of umbrella reviews, such as those with publication data from September 21st, 2018, and onwards. To this end, and following the example of previous studies (80), we have provided a Supplementary file with all relevant publications (umbrella reviews) that have appeared in PubMed until December 31st, 2019, so that the reader remains updated. Moreover, despite our search strategy to include grey literature and policy documents, following previous urging (2), no relevant documents were identified.

It is possible that conclusions drawn in some studies may have been confounded by biases other than sample size; these include reverse causality, which might operate in the
associations assessed that might have affected the findings on the nature of linkage. This consideration highlights the need for prospective studies to demonstrate the direction of causality, and the nature of the linkage; this issue becomes of particular importance when a broad range of disorders are considered (e.g., association of smoking with stroke vs. that with lung cancer metastases to the brain). Similar considerations could also include other sources of bias, either unknown or known, e.g., the so-called survivor bias (as in ALS, or stroke). This latter bias refers to factors that are highly prevalent following the disease onset; these factors can turn out to be protective. The phenomenon is worse in studies assessing the prevalence of a disease, because factors contributing to both the disease’s onset and its progression may be underestimated (as reviewed in (142)). Moreover, several additional types of bias remain unaddressed; for example, the gender bias; how controls are selected; how the measures of exposure are defined; and, what is the composition of the population examined in terms of demographic and geographical variation (for further discussion on these issues, see: (108)).

More broadly, causal inference has nowadays been requiring sophisticated methods (e.g., multivariable Mendelian randomization, or Bayesian approaches), whose description and usage exceed the present study (for a discussion, see: (97) (168)).

Another source of bias could refer to the heterogeneity of the data in combining protective and risk factors for multiple different disease processes. Multiple risk factors, which may be ill-defined or based on studies that are observational in design, may be associated with very distinct disease processes, with differing pathophysologies. For instance, PD is particularly prone to recall bias as it tends to have a longer prodromal phase with symptoms that might not be diagnosed as PD until manifestation of the later characteristic motor symptoms. Likewise, the comparator groups for different neurological disorders included in this meta-umbrella review are expected to be different. Importantly, the so-called vibration of effects in observational studies, which is linked to how the selection of adjusting variables leads to results’ variability, should be taken into account during interpreting our results (169).

The overarching character of presenting more-than-one protective and risk factors for the sum of neurological conditions should be regarded through the public health and clinical counselling lenses rather than from a mere epidemiological perspective. One of our chief goals was to reduce reporting bias—this is, indeed, one of the main reasons for which systematic reviews (in our case a systematic review of umbrella reviews) are performed. In so doing, we wish to highlight the challenges preventive medicine specialists can face (and the balanced, informed decisions outweighing benefits vs. harms in both societal and personalized approaches), while performing counseling services as it regards how well they can
communicate their message—e.g., when the same factor is protective for one neurological disease and risk factor for the other—and the potential inherent difficulties of addressing public health guidelines.

In the same vein, other limitations could include: a) the potentially different, or, in some cases, potentially inaccurate definitions regarding healthy control groups among meta-analyses or/and individual studies leading to inaccuracies, thus introducing another source of bias (commented in (28); and, b) in some cases, the protective and risk factors examined may be due to the interaction of other, more primordial factors (e.g., educational level as result of parental education and family wealth status), therefore, representing risk markers, or there is an apparent heterogeneity in their definition (46). That said, and in alignment with previous studies (28), we pursued a pragmatic approach by showing confidence to the choice of the primary systematic reviewers—and, this should be recognized (31).

The issue of overlapping systematic reviews cannot be excluded; however, we have undertaken every effort to retain the systematic review and meta-analyses that included the highest number of primary studies. Therefore, as discussed elsewhere (150), similar direction (i.e., positive or negative association) and order of magnitude may be due to similar biases existing in these overlapping studies.

Of note, our approach presents additional difficulties in comparing the effect sizes spanning the sum of factors investigated, due to the issue of directly comparing the effect sizes between different meta-analyses (e.g., HR, RR, MD, and SMD), let alone when these effect sizes stemmed from different study designs (e.g., HR presents difficulties if used and interpreted in cross-sectional studies), or when distinct effect size metrics are not converted to a single metric of reference. Thus, our study’s results should be interpreted more through the qualitative rather than the quantitative lens (31). Additionally, pooled effects cannot be estimated; the same is pertinent to umbrella reviews of meta-analyses. In this context, data could have been presented in forest plots without depicting pooled effects (31). Furthermore, another effect size, the attributable risk (AR), has the potential of assessing public health impact, because it allows capturing how much of an outcome could have been prevented from occurring if there was no presence of exposure to a certain factor (170); however, AR has not been frequently used in systematic reviews (80). Thus, future studies should be encouraged to focus on calculating the attributable risk (AR) for neurological disorders based on the exposure to all the protective and risk factors, to ultimately guide health and public policy interventions. In addition, previous studies have suggested replacing the term risk factor with predictor and
explanatory factor for risk stratification and causal studies, respectively [70]. Similarly, others have proposed the notion of risk marker to reflect that many risk factors, such as immigration and ethnicity, may stem from the interaction of other risk factors (28). Collectively, future umbrella and meta-umbrella studies should agree on the effect size to be used for comparisons, and on the preferred terminology regarding protective and risk factors (for the lack of standardization in medical terminology, see (171)).

Another limitation is that our meta-umbrella review was not registered for its goal or protocol at a database like PROSPERO; however, a) our search criteria were predefined and typical of systematic reviews; b) our exclusion criteria were not intending to alter the number or nature of umbrella reviews identified; and, c) no re-synthesis or quantitative synthesis with regards to outcomes of interest or of any other data in any sort of meta-analysis (as discussed in: (2)) took place. Besides, there has been no established protocol to assess the quality of umbrella reviews, so AMSTAR application should be treated with caution. In addition, we performed a systematic review of umbrella reviews to the best of our ability, in the absence of established standard criteria exist for assessing the quality of umbrella reviews. Furthermore, pooled effects cannot be estimated, as the same is pertinent to umbrella reviews of meta-analyses. In this context, data could have been presented in forest plots without depicting pooled effects (31).

Other limitations could be that a) we merged the systematic reviews referring to observational studies to those of clinical trials, given that observational studies should be treated with high cautiousness, when referring to protective and risk factors (commented in (172)); b) we did not distinguish between the two major categories of observational studies, i.e., case-control vs. cohort studies; and, c) we did not perform a sensitivity analysis. Recently, interesting methodologies have been proposed to reconcile the distinct study designs of observational studies and randomized trials (173). Nonetheless, we feel merging, an undeniable risk in principle, has not impacted our results, because of: a) our pragmatic approach; b) the non-quantitative approach applied here; and, c) the lack of evident internal discrepancy among the findings reviewed. Besides, both study types offer real-word results (commented in (150)). Another similar concern could be why retrospective studies were included in this meta-umbrella review, especially for those influential factors that cannot be randomized (smoking). Nevertheless, selecting for retrospective studies to include can affect, in principle, the meta-analyses (e.g., there are meta-analyses of observational or cohort studies), and not a meta-umbrella review; in this sense, performing a sensitivity analysis may not be feasible for a meta-umbrella review. An additional concern could refer to the heterogeneity of the data when
combining risk factors for multiple different disease processes. Multiple risk factors, which may be ill-defined as they are observational in design, are associated with very distinct disease processes, with separate pathophysiology.

Last, our search strategy for umbrella reviews does not allow to identify studies titled *systematic reviews of systematic reviews, meta-reviews, systematic meta-review* (examples of this study types include (174-176)), *systematic reviews of meta-analyses, meta-epidemiological study, review of meta-analyses*, or combinations of the above (examples in (36, 177-179) (180)); however, omitting such studies does not contrast our primary goal to consider only the data reported in reviews clearly defined as umbrella reviews. What is more, an *a posteriori* search strategy using similar search strings still revealed only one study with potential applicability (i.e., (181)). Under all circumstances, a call for uniform terminology in the field of *meta-research* could be made, in alignment with previously expressed opinions (171), in particular those arguing that the used definitions of systematic reviews is surrounded by ambiguity and lack of clarity (182).

5. CONCLUSIONS AND FUTURE DIRECTIONS

Notwithstanding potential considerations that the topic under study may appear broad enough, this exposure-wide assessment of risk and protective factors of chronic neurological disorders, which is conducted using a systematic review of umbrella reviews (and the corresponding meta-analyses), offers an *overarching* outline of different aspects of the assumed risk and protective factors’ in the field. Such comprehensive analyses provide a way to systematically analyze the qualitative and quantitative characteristics of the existing literature, the validity of evidence, and how these studies could provide readers with information on knowledge gaps. In so doing, this *meta-umbrella* review allows to identify the sum, so far, of preventable or/and modifiable risk factors, which may be relevant—either as common to most or specific to certain—neurological disorders. To all these goals, standardizing how a study is designed and how exposure to certain factors is defined would be essential (108). Likewise, we should be prompted to develop and standardize the criteria for assessing the quality of umbrella reviews.

Additionally, our *meta-umbrella* approach provides a wealth of discussion on potential biases that could occur in reporting umbrella reviews, and it provides a rationale to develop reproducible robust methodologies while performing umbrella studies. To enhance extensive analyses of risk/protective factors, we recommend developing standardized and reproducible methods to identify these factors; performing a field-based systematic review of umbrella
reviews might simplify the process. Equally importantly, as almost all umbrella review studies have so far focused on neurological disorders of resource-rich countries, a call for intensifying meta-research on neurological disorders affecting resource-poor countries could be addressed. Last, this meta-umbrella review offers new perspectives in the meta-research field. The fact that not all risk factors or/and diseases were uniformly mentioned, that fact per se could serve as call-for-action to work on developing guidelines and a uniform framework for Systematic reviews of Umbrella reviews.

Being distant from any overly expectations on reaching final conclusions on causal versus casual associations, this study attempted to highlight the commonality of certain risk and protective factors, and the relevant level of evidence surrounding these factors (150). Emphasis should also be given on the challenges that preventive medicine specialists can face, while they perform counselling services, and on the need to communicate health promotion messages to the patients’ community in a consistent manner. Interestingly, regarding the commonality of diseases in terms of symptoms, there seems to exist an underlying basis on shared genetic traits, as well as the level of connectivity in protein interactions (disease network neighborhoods) (183, 184); however, in parallel to a public health approach, further research is needed to better understand the distinct pathophysiology of conditions affecting the nervous system. For example, a key question would be whether the cellular reactions following exposure to risk factors are genetically determined, and if so, whether this genetic background is similar between distinct neurological patients remains to be deciphered.

Under all circumstances, we feel that in order to be able to reap further epidemiological rewards, this sort of study (i.e., meta-umbrella review) should be jointly assessed with studies coupling mechanistic, biological insights with statistical methods on causal inference (97, 150). Future research on the field may require a more thorough applications of statistics focusing on causal effects, such as mediation (a statistical technique allowing to study the effects of a third mediator variable between an independent and dependent variable) and multivariable Mendelian randomization studies, which can render some risk factors as causal determinants (97); however, caution on appropriately interpreting such methodologies has been expressed (185). In parallel to this, novel robust evidence from novel umbrella reviews should steadily update this type of study (in a way similar to living systematic review (186)), which, in turn, could offer periodic updates for policy formulation.

Ultimately, neurological research and health policy trends could focus on creating a Neurological Diseases Atlas, in which every risk and protective factor will be both qualitatively
and quantitatively linked to all associated diseases (similarity to causal mapping (97)). This Atlas could be accompanied by a score-based approach we wish to call poly-non-genc risk scores; these scores could assess the exposure to common factors (in a similar pattern to polygenic scores. In so doing, this Atlas—if appropriately assessing mean percentage population attributable fractions of each risk factor for the sum of diseases (as in (145), for single disease)—can lead to significant clinical and public health impact.
### Table 1. Common protective factors of non-communicable neurological disorders.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Effect size metric</th>
<th>Effect size (95% CI)</th>
<th>Neurological condition</th>
<th>Level of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mediterranean diet</td>
<td>RR 0.58 (0.35–0.95)</td>
<td>Stroke</td>
<td>III</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RR 0.51 (0.31–0.84)</td>
<td>Cognitive impairment</td>
<td>II</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RR 0.60 (0.48–0.77)</td>
<td>Alzheimer’s disease</td>
<td>I</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RR 0.69 (0.57–0.84)</td>
<td>Dementia</td>
<td>I</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RR 0.67 (0.58–0.76)</td>
<td>Parkinson’s disease</td>
<td>II</td>
<td></td>
</tr>
<tr>
<td>Coffee consumption</td>
<td>RR 0.73 (0.55–0.97)</td>
<td>Alzheimer’s disease</td>
<td>III</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RR 0.29 (0.11–0.76)</td>
<td>Alzheimer’s disease</td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RR 0.39 (0.27–0.57)</td>
<td>Parkinson’s disease</td>
<td>II</td>
<td></td>
</tr>
<tr>
<td>Serum uric acid</td>
<td>RR 0.58 (0.41–0.83)</td>
<td>Dementia</td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RR 0.49 (0.27–0.87)</td>
<td>Multiple Sclerosis</td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RR 0.22 (0.10–0.45)</td>
<td>Neuromyelitis optica</td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RR 0.21 (0.14–0.32)</td>
<td>Amyotrophic lateral sclerosis</td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td>Alcoholic intake</td>
<td>RR 0.72 (0.61–0.86)</td>
<td>Dementia</td>
<td>III</td>
<td></td>
</tr>
<tr>
<td>Physical activity</td>
<td>HR 0.75 (0.66–0.85)</td>
<td>Parkinson’s disease</td>
<td>III</td>
<td></td>
</tr>
<tr>
<td>Serum Vitamin D</td>
<td>RR 0.76 (0.66–0.86)</td>
<td>Dementia</td>
<td>III</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OR 0.44 (0.24–0.78)</td>
<td>Multiple sclerosis</td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HR 0.66 (0.57–0.78)</td>
<td>Parkinson’s disease</td>
<td>I</td>
<td></td>
</tr>
<tr>
<td>Serum Vitamin B₁₂</td>
<td>OR 0.16 (0.05–0.50)</td>
<td>Parkinson’s disease</td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OR 0.50 (0.40–0.63)</td>
<td>Parkinson’s disease</td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OR 0.64 (0.44–0.93)</td>
<td>Multiple Sclerosis</td>
<td>II</td>
<td></td>
</tr>
<tr>
<td>Vitamin C intake</td>
<td>OR 0.81 (0.67–0.98)</td>
<td>Parkinson’s disease</td>
<td>IV</td>
<td></td>
</tr>
</tbody>
</table>

OR, odds ratio; RR, relative risk; HR, hazard ratio; 95% CI, 95% confidence interval.
Table 2. Common risk factors of non-communicable neurological disorders.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Metric</th>
<th>Effect size (95% CI)</th>
<th>Neurological condition</th>
<th>Level of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Farming</td>
<td>OR</td>
<td>1.42 (1.17–1.73)</td>
<td>Amyotrophic lateral sclerosis</td>
<td>III</td>
</tr>
<tr>
<td></td>
<td>OR</td>
<td>1.30 (1.16–1.46)</td>
<td>Parkinson’s disease</td>
<td>III</td>
</tr>
<tr>
<td></td>
<td>OR</td>
<td>1.42 (1.17–1.73)</td>
<td>Amyotrophic lateral sclerosis</td>
<td>III</td>
</tr>
<tr>
<td>Pesticides</td>
<td>OR</td>
<td>1.62 (1.40–1.88)</td>
<td>Parkinson’s disease</td>
<td>III</td>
</tr>
<tr>
<td>Low-frequency electromagnetic fields</td>
<td>OR</td>
<td>1.29 (1.03–1.62)</td>
<td>Amyotrophic lateral sclerosis</td>
<td>III</td>
</tr>
<tr>
<td></td>
<td>RR</td>
<td>1.74 (1.37–2.21)</td>
<td>Dementia</td>
<td>III</td>
</tr>
<tr>
<td></td>
<td>OR</td>
<td>1.22 (1.01–1.47)</td>
<td>Parkinson’s disease</td>
<td>IV</td>
</tr>
<tr>
<td>Organic solvents</td>
<td>OR</td>
<td>1.54 (1.03–2.29)</td>
<td>Multiple sclerosis</td>
<td>IV</td>
</tr>
<tr>
<td>Chlamydia pneumonia DNA in CSF</td>
<td>OR</td>
<td>3.22 (1.20–8.59)</td>
<td>Multiple sclerosis</td>
<td>IV</td>
</tr>
<tr>
<td>Chlamydia pneumonia infection</td>
<td>OR</td>
<td>6.00 (1.93–18.66)</td>
<td>Dementia</td>
<td>IV</td>
</tr>
</tbody>
</table>

OR, odds ratio; RR, relative risk; HR, hazard ratio; 95% CI, 95% confidence interval.
Table 3. Variables that are both risk and protective factors of non-communicable neurological disorders.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Effect size metric</th>
<th>Effect size (95% CI)</th>
<th>Neurological condition</th>
<th>Level of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smoking</td>
<td>RR</td>
<td>1.26 (1.05–1.50)</td>
<td>Dementia</td>
<td>IV</td>
</tr>
<tr>
<td></td>
<td>RR</td>
<td>1.52 (1.39–1.66)</td>
<td>Multiple Sclerosis</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td>RR</td>
<td>0.64 (0.60–0.69)</td>
<td>Parkinson’s disease</td>
<td>II</td>
</tr>
<tr>
<td></td>
<td>HR</td>
<td>1.59 (1.20–2.11)</td>
<td>Dementia</td>
<td>IV</td>
</tr>
<tr>
<td>Hypertension</td>
<td>RR</td>
<td>0.75 (0.61–0.90)</td>
<td>Parkinson’s disease</td>
<td>IV</td>
</tr>
</tbody>
</table>

RR, relative risk; HR, hazard ratio; 95% CI, 95% confidence interval.
Table 4. Specific risk and protective factors of non-communicable neurological disorders

<table>
<thead>
<tr>
<th>Neurological condition</th>
<th>Factors</th>
<th>Effect size metric</th>
<th>Effect size (95% CI)</th>
<th>Levels of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amyotrophic lateral sclerosis</td>
<td>β-Carotene intake</td>
<td>RR</td>
<td>0.92 (0.87–0.97)</td>
<td>II</td>
</tr>
<tr>
<td></td>
<td>n-3 fatty acids intake</td>
<td>RR</td>
<td>0.71 (0.59–0.85)</td>
<td>III</td>
</tr>
<tr>
<td></td>
<td>Other heavy metals</td>
<td>OR</td>
<td>2.13 (1.33–3.41)</td>
<td>II</td>
</tr>
<tr>
<td></td>
<td>Lead</td>
<td>OR</td>
<td>1.81 (1.39–2.35)</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td>Welding</td>
<td>RR</td>
<td>0.86 (0.80–0.92)</td>
<td>III</td>
</tr>
<tr>
<td></td>
<td>Nigral volume</td>
<td>OR</td>
<td>0.31 (0.17–0.55)</td>
<td>IV</td>
</tr>
<tr>
<td></td>
<td>Serum urate</td>
<td>RR</td>
<td>0.65 (0.43–0.97)</td>
<td>IV</td>
</tr>
<tr>
<td></td>
<td>Hydrocarbon</td>
<td>OR</td>
<td>1.36 (1.13–1.63)</td>
<td>IV</td>
</tr>
<tr>
<td>Parkinson’s disease</td>
<td>Rural living</td>
<td>OR</td>
<td>1.32 (1.18–1.48)</td>
<td>III</td>
</tr>
<tr>
<td></td>
<td>Well water intake</td>
<td>RR</td>
<td>1.21 (1.05–1.40)</td>
<td>IV</td>
</tr>
<tr>
<td></td>
<td>Dairy products intake</td>
<td>RR</td>
<td>1.40 (1.20–1.63)</td>
<td>III</td>
</tr>
<tr>
<td></td>
<td>Carbohydrate intake</td>
<td>RR</td>
<td>1.24 (1.05–1.48)</td>
<td>IV</td>
</tr>
<tr>
<td></td>
<td>Energy intake</td>
<td>RR</td>
<td>1.39 (1.01–1.92)</td>
<td>IV</td>
</tr>
<tr>
<td></td>
<td>Constipation</td>
<td>RR</td>
<td>2.30 (2.02–2.63)</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td>Head injury</td>
<td>OR</td>
<td>1.55 (1.33–1.81)</td>
<td>II</td>
</tr>
<tr>
<td></td>
<td>Anxiety or depression</td>
<td>RR</td>
<td>1.86 (1.64–2.10)</td>
<td>II</td>
</tr>
<tr>
<td></td>
<td>Melatonin</td>
<td>MD</td>
<td>–2.64 (–5.99 to 0.71)</td>
<td>IV</td>
</tr>
<tr>
<td></td>
<td>Vitamin E</td>
<td>RR</td>
<td>0.80 (0.67–0.95)</td>
<td>IV</td>
</tr>
<tr>
<td></td>
<td>Statins</td>
<td>RR</td>
<td>0.72 (0.59–0.89)</td>
<td>IV</td>
</tr>
<tr>
<td></td>
<td>Aspirin</td>
<td>RR</td>
<td>0.77 (0.63–0.95)</td>
<td>IV</td>
</tr>
<tr>
<td></td>
<td>Non-Aspirin NSAIDS</td>
<td>RR</td>
<td>0.65 (0.49–0.86)</td>
<td>IV</td>
</tr>
<tr>
<td></td>
<td>NSAIDS</td>
<td>RR</td>
<td>0.74 (0.64–0.86)</td>
<td>III</td>
</tr>
<tr>
<td></td>
<td>Fish intake</td>
<td>RR</td>
<td>0.88 (0.79–0.98)</td>
<td>IV</td>
</tr>
<tr>
<td></td>
<td>Antihypertensive drugs</td>
<td>HR</td>
<td>0.84 (0.75–0.94)</td>
<td>IV</td>
</tr>
<tr>
<td></td>
<td>RR (95% CI)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>------------</td>
<td>---</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social participation</td>
<td>0.71 (0.57–0.88)</td>
<td>IV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frequency of social contacts</td>
<td>0.64 (0.54–0.76)</td>
<td>I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stroke</td>
<td>1.59 (1.25–2.02)</td>
<td>II</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aluminum</td>
<td>1.72 (1.33–2.21)</td>
<td>III</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loneliness</td>
<td>1.58 (1.19–2.09)</td>
<td>IV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low education</td>
<td>1.81 (1.36–2.43)</td>
<td>III</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Midlife obesity</td>
<td>1.81 (1.22–2.69)</td>
<td>IV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Traumatic brain injury</td>
<td>1.40 (1.03–1.90)</td>
<td>IV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spirochaetal infection</td>
<td>10.65 (3.4–33.4)</td>
<td>IV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Herpesviridae infection</td>
<td>1.38 (1.14–1.65)</td>
<td>IV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depression at any age</td>
<td>1.99 (1.84–2.16)</td>
<td>I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early-life depression</td>
<td>1.63 (1.27–2.11)</td>
<td>III</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Late-life depression</td>
<td>1.85 (1.67–2.05)</td>
<td>I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type 2 diabetes mellitus</td>
<td>1.54 (1.39–1.72)</td>
<td>I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diphtheria vaccination</td>
<td>0.60 (0.40–0.91)</td>
<td>II</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tetanus vaccination</td>
<td>0.71 (0.57–0.88)</td>
<td>II</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anti-EBV IgG seronegativity</td>
<td>0.13 (0.05–0.32)</td>
<td>IV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMD in femoral neck</td>
<td>0.36 (0.21–0.61)</td>
<td>III</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMD in lumbar spine</td>
<td>0.34 (0.24–0.50)</td>
<td>III</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMD in hip</td>
<td>0.33 (0.18–0.60)</td>
<td>III</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anti-EBNA IgG seropositivity</td>
<td>4.46 (3.26–6.09)</td>
<td>I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infectious mononucleosis</td>
<td>2.17 (1.97–2.39)</td>
<td>I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Appendectomy at age ≤20 years</td>
<td>1.17 (1.02–1.34)</td>
<td>II</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EBV DNA in mononuclear cells and serum</td>
<td>1.84 (1.02–3.20)</td>
<td>II</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tonsillectomy at age ≤20 years</td>
<td>1.32 (1.09–1.61)</td>
<td>II</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Traumatic injury</td>
<td>1.41 (1.03–1.92)</td>
<td>II</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Dementia or cognitive impairment**

**Multiple Sclerosis**
<table>
<thead>
<tr>
<th>Condition</th>
<th>Measure</th>
<th>OR</th>
<th>95% CI</th>
<th>Evidence Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anti-VCA IgG seropositivity</td>
<td>OR</td>
<td>4.52</td>
<td>(2.85–7.15)</td>
<td>III</td>
</tr>
<tr>
<td>Chronic cerebrospinal venous insufficiency</td>
<td>OR</td>
<td>8.45</td>
<td>(3.47–20.56)</td>
<td>III</td>
</tr>
<tr>
<td>Intrathecal production of Ig for <em>Chlamydia pneumoniae</em></td>
<td>OR</td>
<td>3.84</td>
<td>(1.32–11.21)</td>
<td>IV</td>
</tr>
<tr>
<td>Serum homocysteine</td>
<td>OR</td>
<td>4.57</td>
<td>(1.40–14.89)</td>
<td>IV</td>
</tr>
<tr>
<td>Dietary whole grain intake</td>
<td>RR</td>
<td>0.86</td>
<td>(0.45–0.92)</td>
<td>III</td>
</tr>
<tr>
<td>Dietary fiber intake</td>
<td>RR</td>
<td>0.81</td>
<td>(0.73–0.90)</td>
<td>III</td>
</tr>
<tr>
<td>Chocolate intake</td>
<td>RR</td>
<td>0.86 (0.79–0.94)</td>
<td>IV</td>
<td></td>
</tr>
</tbody>
</table>

OR, odds ratio; RR, relative risk; HR, hazard ratio at follow-up; Mean difference (MD); 95% CI, 95% confidence interval. References for levels of evidence.
Chapter 2 — Towards a multifaceted role of chief genetic risk factors: The case of APOE4 in Alzheimer Disease

Apolipoprotein E4 and Meningeal Lymphatics in Alzheimer Disease: A Conceptual Framework

1. Risk Factors for Alzheimer Disease (AD)

Alzheimer disease (AD) is among the principal causes of death and a global public health priority (187, 188). AD is increasingly recognized as an heterogeneous disease (and, as corollary, an umbrella term), which encompasses several cognitive subtypes, and whose biological and clinical manifestations may not be always co-prevalent; thus, a call for more individualized approaches has been made (189-192). While most patients with AD manifest the late-onset, sporadic form of the disease, - which may also harbor a genetic component -, a small percentage (< 1%) of patients carry inherited gene mutations that lead to a much earlier onset of the symptomatology (188). Both genetic and environmental factors, reviewed in (42) and summarized in Table 1, contribute to the pathogenesis of AD. Regarding the environmental factors, we and others have shown that chronic stress might contribute to AD (reviewed also in (193-195) via a) sustained Corticotrophin-releasing Hormone and cortisol effects in brain cells (196), and b) chronic increases in systemic and brain levels of inflammatory cytokines (such as Tumor Necrosis Factor (TNF) (197); the latter, as components of innate immunity, are increasingly recognized for their neuro-inflammatory roles in AD (198, 199).

Novel insights into the etiology of AD highlight the potential contribution of the meningeal lymphatic system in the manifestations of the disease (200). Indeed, the aforementioned connection between chronic stress, inflammation, and AD may be facilitated by the meningeal lymphatic system, as is the case in other diseases (such as Multiple Sclerosis (MS) (165) and post-traumatic stress disorder (201)). This could reflect the known roles of the traditional peripheral lymphatic system in inflammation and cancer, as revealed by several animal models; for instance, chronic stress in mice leads to structural changes in the lymphatic vessels of tumors and increased dissemination of metastases via transport of tumor cells through the peripheral lymphatic system (202). Also, patients taking beta-blockers, which antagonize the endogenous stress-stimulated catecholamines at the level of the beta-adrenergic receptors, show significantly fewer proximal lymph node and distant metastases than those not receiving beta-blockers (202); this is likely due to a reduction in catecholamine-mediated signaling events by the receptor antagonists. These findings may have broader implications for AD, granted the
established link between chronic stress and neuroinflammation (reviewed in (203) and (163), as well as the known connection between neuroinflammation and AD (199, 204-207).

Here, we suggest that meningeal lymphatic vessel function is influenced by apolipoprotein E4 (APOE4), which has been established as the leading genetic risk factor for developing AD (208). We also present novel analyses of previously published RNA-Seq data that offer new insights into how meningeal lymphatic vessels, in association with APOE4, may contribute to the pathogenesis of AD. Our main goal is to propose a new conceptual framework on the role of reduced lymphatic function (or *meningeal attenuated lymphaticness*) and *lymphedema* in APOE4-related AD. Ultimately, in light of recent evidence, which strongly suggests that impaired brain capillary function contributes to cognitive dysfunction and AD manifestations (209), we wish to provide further evidence on the *neurovascular*-centered view of AD.

2. **The role of APOE4 as genetic risk factor for AD**

The human *APOE* gene expresses three isoforms: APOE2, APOE3, and APOE4, corresponding to the APOE2, APOE3, and APOE4 proteins (Figure 1). Although APOE has been extensively investigated for its role in liver lipoprotein metabolism (for a historical review and future perspectives, see (210), recent data suggested that APOE is also involved in the pathophysiology of the central nervous system (CNS). For instance, APOE’s genetic ablation in elderly mice caused a reduction in neuritic plaques and a decrease in the associated neuronal synaptic loss and glia activation, suggesting that APOE might be helpful against certain features of aging *per se* (211); however, these neuroprotective effects have not been observed in the context of the AD-associated *APP/PSEN1* mouse model, in which APOE may be neurotoxic. Interestingly, the role of APOE in the CNS is further illustrated by a recently described ultra-rare mutation in the *APOE3* gene, which confers resistance to developing familial autosomal dominant AD (212). From a cognitive standpoint, it was previously suggested that APOE4 might increase memory in young adults; however, a recent meta-analysis failed to verify such a functional association (213).

Carrying and expressing the *APOE4*-coding allele is the chief genetic risk factor for AD, with predictive values exceeding polygenic scores for cognitive ageing in elderly populations (214),(215). While the APOE4 isoform is present in approximately 13-15% of the population, it is carried by more than 50% of individuals with late-onset AD (216). Notably, the population attributable fraction of APOE4 for AD (i.e., the theoretical reduction in AD
incidence in the absence of the APOE4-coding allele) is around seven percent (217). These APOE4 effects enhance genetic anticipation, principally in late-onset AD (218). Of note, and according to some studies, APOE4 may be a promising therapeutic target for the disease (219).

APOE4 status is also linked to dementia with Lewy body disease, with decreased levels of APOE methylation having also been implicated in this disease (220) (221) (222). APOE4 status has also been linked to Parkinson disease-related dementia (223), MS, tauopathy, vascular dementia (whose partial genetic overlap with AD may be explained by APOE genetic status) (224), mixed vascular dementia and AD (225), chronic traumatic brain injury (226), cerebral amyloid angiopathy (227), and cerebrovascular disease (228), as well as positive status of Transactive Response DNA-Binding Protein 43 (a protein previously linked to AD, frontotemporal dementia, and amyotrophic lateral sclerosis) (229). While APOE2 status can postpone age-at-onset of AD (230), an individual possessing a single APOE4 allele has a 3-fold increased risk for AD late in life, and carrying two alleles has been associated with a higher than 10-fold risk (231). Of note, age, region, and ethnicity may partially modulate this association (232-235); for instance, Hispanic carriers of APOE4 seem to have a higher amyloid load than non-Hispanic carriers (236).

Mechanistically, diploidy for human APOE4 is thought to contribute to the aggregation of amyloid-beta in the brain (for comprehensive reviews, see (227, 237, 238), possibly through the neuronal receptor LRP1 (239). Further experimental evidence supporting a role of APOE4 in amyloid-beta aggregation came from studies using Pittsburgh Compound-B–positron emission tomography (PET) imaging (240). In one of these studies, APOE4 status was associated with accelerated cognitive dysfunction in individuals whose PET results were positive for amyloid-beta (241). From a biochemical point of view, amyloid-beta fibrils were more strongly associated with APOE4 than with APOE3 or APOE2 (242). Conversely, amyloid-beta protofibrils had higher stability in their association with APOE3 or APOE2 than with APOE4 (243).

Patients with sporadic AD may exhibit impairments in amyloid-beta clearance, without major changes in the de novo production of the peptide (244). In mice, APOE4 directly disrupts clearance of amyloid-beta across the Blood-Brain-Barrier (BBB), suggesting that impairment in the neurovascular function of the BBB may contribute to AD etiology. To determine the critical stage of amyloid fibril (seeding stage) and plaque (plaque stage) formation, in which APOE4 exerts its strongest effect, Liu and colleagues developed a cell-type specific, Cre-floxed-mediated inducible mouse model to control expression of astrocytic APOE4 during
amyloid fibril and plaque formation(245). Their data indicated that APOE4 had its greatest impact during the seeding stage of amyloid-beta formation, probably by impeding amyloid-beta clearance and promoting its aggregation(245). Interestingly, these associations may share similarities to those observed in brain pathologies in Down syndrome (trisomy 21) patients. These patients carry an extra copy of chromosome 21, in which the amyloid precursor protein (APP) is located. Trisomy 21 results in APP overexpression(246); however, others studies support that trisomy-21–related dementia may be caused by overexpression of non-APP genes leading to a decrease in the soluble amyloid-beta-38, and amyloid-beta-40(247). Over two thirds of older adults with trisomy 21 die from dementia; among them, the risk of premature death is increased by almost seven-fold in APOE4 isoform carriers(248). Lastly, the modulatory role of APOE4 on amyloid-beta formation has been partly attributed to decreased serum and brain concentrations of APOE in APOE4 carriers vs. non-APOE4 carriers, and not to the APOE4 allele(s) per se(227).

AD has been associated with tauopathy, another proteinopathy characterized by the pathological accumulation of tau protein in the brain (for a review on tau’s role in physiology, see(249). The link between tauopathy and APOE4 has been addressed in several studies. In tau-expressing transgenic mice, APOE4 exacerbated tau-mediated neurodegeneration, causing increases in brain atrophy and neuro-inflammation, and alterations in glial cell function; this effect was seemingly independent of amyloid-beta pathology (250). In addition, APOE4 increased levels of phosphorylated tau and its extracellular release by neurons, a process independent of glia cells (251).

APOE4 may also contribute to AD symptomatology via mechanisms that are not related to either amyloid-beta or tau. First, APOE4 causes a decrease in the levels of exosomes released in the brain interstitial space, together with a reduction in endosome/exosome pathway-related gene expression(252). These phenomena are linked to malfunction of the lysosomal system and the impaired degradation of cellular debris, which, in turn, may lead to accumulation of amyloid-beta, and hence, to neurodegeneration(252). Second, the presence of APOE4 hinders neuronal responsiveness to reelin, a glycoprotein that controls neuronal migration and synaptic transmission, while it promotes thrombosis and hemostasis(253). In doing so, APOE4 may reduce, in patients with AD, the protective effects of reelin against beta-amyloid-induced cognitive impairment(254). Third, APOE4 may lead to dysfunction and eventual death of gamma-aminobutyric acid (GABA)–expressing interneurons in the hippocampus, a brain region that is severely impaired in AD(255). Because these neurons are inhibitory, their absence causes hyperexcitability of the entire hippocampal network, leading to the impairments
observed in AD(256) (for this conceptual framework, see(257). Finally, APOE4’s role in AD may involve a dysfunctional immune system(258). Indeed, activation of the innate immune response is considered a disease-promoting factor in AD(259), and APOE4 regulates different aspects of the inflammatory reaction (e.g., microglia activation)(260).

In particular, one of APOE4’s effects on innate immunity appears to be, at least partly, via the Triggering Receptor Expressed on Myeloid cells 2 (TREM2), which is primarily expressed in the CNS microglia (250) and involved in microglial disorders (microgliopathies) (261). Indeed, TREM2 is involved in APOE4’s downstream activation of microglia (262, 263). TREM2 double knock–out mice exhibit increased amyloid plaque formation with less plaque-associated APOE (264). Importantly, genome-wide association studies (GWAS) have corroborated the crucial roles of TREM2 and other microglia-associated proteins, such as PLCG2 and ABI3, in AD pathophysiology (265-270). This evidence points to a delicate interaction between APOE (including the APOE4 isoform), TREM2, and amyloid-beta. Further research will determine whether this interaction plays a key role in the pathophysiology of human AD.

3. APOE4 in AD-associated neurovascular and cerebrovascular function

Recent advances in our understanding of AD highlight the role of neurovascular function in the pathogenesis of AD, including cognitive impairment, thus warranting further investigation of vascular markers of the disease (271). Studies linking AD and neurovascular function have focused on fibrinogen (272), imaging biomarkers of amyloid-beta and neurodegeneration (e.g., plasma neuronal-enriched extracellular vesicles), MRI, amyloid PET, Tau PET, and fluorodeoxyglucose PET (reviewed in (273-276). Of particular interest is the role of the neurovascular system in clearing amyloid-beta from the brain. For example, Nation and colleagues (277) detected hippocampal capillary damage and BBB breakdown in people with early cognitive dysfunction, independently of amyloid-beta and tau pathologies; these findings suggested that BBB dysfunction can serve as an early biomarker of human cognitive dysfunction, a precursor of AD (278).

Notably, APOE4 itself has been linked to impaired BBB function. In transgenic mice, expression of APOE4 (but not other APOE variants) leads to breakdown of the BBB; this phenomenon is mediated by the activation of a proinflammatory pathway in pericytes (i.e., perivascular cells) (279). BBB breakdown leads to passage of neurotoxic proteins and other substances in the brain parenchyma, as well as, decreased blood flow in the brain microvasculature. These events likely occur prior to neuronal dysfunction, and they may lead to neurodegenerative changes (279). In the clinic, similar observations on BBB breakdown have
been inferred from increases in cerebrospinal fluid (CSF)/plasma albumin quotients in APOE4 carriers (280).

Based on neuroimaging data, cerebrovascular diseases can increase the risk for AD and tau pathology (281, 282; 283). High levels of low-density lipoprotein (LDL), also associated with vascular diseases, may lead to amyloid accumulation in APOE4 carriers (284). Moreover, cardiovascular risk factors and type 2 diabetes mellitus also increase the risk of dementia, suggesting that impaired metabolism also contributes to AD (285). Experiments in cells and animal models have shed light on some putative mechanisms. For instance, APOE and circulating high-density lipoprotein (HDL)–mediated amyloid-beta transport in bioengineered human cerebral vessels, with APOE4 being less effective than APOE2 in this regard (286). Also, exogenously administered amyloid-beta was more effectively cleared through the BBB in mice overexpressing the low-density-lipoprotein receptor (LDLR) (287). Taking this further, the loss of mouse ApoE in pericytes was shown to inhibit the clearance of aggregated amyloid-beta-42 on multi-spot glass slides; this inhibition was rescued by human APOE3 but not APOE4 (288). Collectively, the above findings partly explained APOE4’s pathogenic role and provided further insights into the neurovascular-centered view of AD.

4. APOE4 and the meningeal vs. traditional peripheral lymphatic system

Considering the strong interactions between the vascular and peripheral lymphatic systems, as well as, the shared features between the meningeal and traditional peripheral lymphatic systems, it is worth exploring the potential roles of APOE4 in the meningeal lymphatic system, following insights from the traditional peripheral lymphatic system as well.

4.1 The traditional peripheral vs. meningeal lymphatic system

The peripheral lymphatic system plays a crucial role in the clearance of many substances, such as cellular debris, proteins, lipids, and other macromolecules from peripheral organs. Thus, impaired clearance of various interstitial space macromolecules has been often linked to abnormalities in peripheral lymphatic vessels, such as enlargement of these vessels and leakage of their contents into the surrounding space (for reviews on this field, see (289, 290)). According to the predominant theories, peripheral lymphatic vessels originate in early development from the venous endothelial system by forming lymphatic endothelial cell progenitors. However, according to recent lineage tracing studies, progenitors of peripheral lymphatic endothelial cells may also include dermal blood capillaries, lymphangioblasts, blood cell progenitors, hemogenic endothelial cells, and organ-specific (e.g., heart) endothelial cells. These progenitors, at least in animal models, may also include venous endothelial cells outside the heart, and lymphatic progenitors stemming from the blood-producing yolk sac endothelium.
Progenitor cells derived from the yolk sac may be particularly relevant in the cardiac lymphatic vasculature (for all above, see (291-293)).

Until recently, it was believed that the meninges lacked lymphatic vessels, principally because of the dogmatic view on the immuno-privileged status of the brain. The Italian scientist Paolo Mascagni first described the meningeal lymphatic system in the 1800’s, but his conclusions were disputed until recently (for works spanning all this period, see (294) (295) (296) (297) (298) (299) (300)). In addition, despite description of Mendelian disorders with peripheral lymphedema and cognitive impairment (e.g., in Hennekan syndrome (301, 302)) or non-specific neurological signs (e.g., generalized lymphedema associated with neurologic signs –GLANS– syndrome (303), the existence of meningeal lymphatics remained disputed. Therefore, until some years ago, it was thought that meningeal lymphatic vessels did not exist, let alone play any role in CNS physiology or pathophysiology.

Intriguingly, several studies have now demonstrated that lymphatic vessels exist within the meninges of the CNS (304-306) (reviewed in(307, 308). This is a re-discovery, one heralding a delayed history (309). Alongside with the confirmatory findings suggesting that Schlemm’s canal in the eyes had features resembling lymphatics that were evolutionarily conserved (310), this re-discovery suggests inter alia the notion that investing on mining of hidden pearls can be crucial in advancing biomedical research. Fascinatingly, even peripheral lymphatic vessels were also re-discovered after the initial Hippocrates and Erasistratus of Ceos analyses of white blood-filled vessels, and mesenteric milky arteries, respectively, till the 17th century’s Gasparo Aselli’s observation of a dog’s milky veins (311), mentioned in(312, 313).

These initial descriptions of meningeal lymphatic vessels were followed by further studies in humans. In particular, by coupling MRI findings with post-mortem analyses, researchers demonstrated that meningeal lymphatic vessels existed alongside blood vessels in humans and non-human primates (314). More recently, meningeal lymphatic vessels in the basal part of the mouse skull were shown to support the drainage of macromolecules (metabolites, debris, and immune cells) from the cerebrospinal fluid (CSF). This most likely follows exchange of solutes, debris and immune cells with the brain interstitial fluid (ISF), via the efferent paravascular glial lymphatic (glymphatic) system into the meningeal lymphatic vessels, ultimately draining into the conventional cervical lymphatic system, including local lymph nodes (200, 304, 306) (for a review, see(315).

Drainage of glymphatic fluid through the newly described basal meningeal lymphatic vessels in the mouse skull is quite efficient(316). This higher draining capacity of the basal
than the dorsal meningeal lymphatic vessels may be partly explained by their respective proximity to the subarachnoid space, and by their location, which is independent of nerve fibers. These meningeal lymphatic vessels are presumably distinct from the traditional peripheral lymphatic system(316). Notably, the function of the basal meningeal lymphatic vessels was compromised in ageing mice(316). Mechanistically, lymphatic wall hyperplasia may take place in response to increased capillary lymphatic pressure, as it occurs in peripheral lymphedema. The latter has been linked to aberrant type IV collagen distribution, a reduction in lymphatic valves, potentially via decreased Proxl and Foxc2 expression(316). Moreover, lymphatic endothelial cells in aged mice showed significantly altered intercellular junction types, along with reduced functionality for lymphatic drainage through the basal meningeal lymphatic vessels(316). Altogether, these findings suggest that aged basal meningeal lymphatic vessels are associated with reduced lymph flow; however, to our knowledge, whether similar pathology is present in human AD has not been investigated as yet.

Although not directly applicable to human pathophysiology, valuable lessons can be learned from observations in animal models, such as the zebrafish. In the latter, mural lymphatic endothelial cells produce vascular growth factors and promote accumulation of low-density lipoproteins from the bloodstream(305, 317). As noted above, meningeal lymphatic cells are derived from the blood vasculature and in zebrafish, this phenomenon is mediated by the Vegfc-Vegfd-Ccbe1-Vegfr3 pathway; this suggests that these cells are probably lymphatic capillary-type endothelial cells(317). Ablation of meningeal lymphatic vessels in mice using a photodynamic drug, Visudyne (verteporfin), resulted in a) reduced drainage of macromolecules, and b) impairments in cognitive performance. Altogether, these findings point to a major role of meningeal lymphatic vessels as a conduit in the clearance of macromolecules from the mouse brain. These macromolecules most likely reach the meningeal lymphatic system through the efferent lymphatic drainage system formed by astrocytes around brain blood vessels, as described in(318). The functions of both the glymphatic and meningeal lymphatic systems are compromised during ageing(200, 319, 320).

In summary, the studies discussed above suggest that the meningeal lymphatic vessels can be, at least in part, responsible for the dysfunctional clearance of macromolecules from the brain. Here, we theorize that this process could contribute to the decline of cognitive function associated with age and neurodegenerative diseases, such as AD (200).

4.2 Crosstalk of meningeal and cervical lymphatic vessels

The (re)discovery of meningeal lymphatic vessels is now an acknowledged paradigm shift. However, the presence of meningeal lymphatic vessels should not be considered in isolation
but rather in the context of the conventional lymphatic system, including the most proximal cervical lymph vessels and nodes (CLNs). CLNs are already known to play a major role in MS and other neuroinflammatory diseases (321), in which brain-peripheral immune crosstalk has been observed (322) (potentially because of the presence of auto-antigens or microbial risk factors (323, 324)). The role of the CLNs in AD, however, has been investigated far less. Indeed, although follicular dendritic cells (FDCs) within the CLNs develop from perivascular precursors (325), their potential links to meningeal lymphatic vessels or to AD are largely unknown.

To our knowledge, FDCs have been mostly linked to the prolonged retention of human immunodeficiency virus (HIV)-1/ simian immunodeficiency virus (SIV) infection in the context of CLNs (326), through a number of mechanisms, including cycling endosomes (327). Taken further, FDCs seem to indirectly contribute to transmitting the virus to CD4+ T follicular helper cells; this is performed in a B-cell–mediated manner(328). Elucidating how meningeal lymphatic vessels interact with CLN-associated FDCs could shed light on their crosstalk in MS and, notably, AD. Interestingly, a very recent single-cell transcriptomics study reported APOE among the genes that are differentially expressed in mouse CLNs(329), thus suggesting a role of this gene (and, as corollary, of its variants such as APOE4) in CLNs-related pathological processes.

In the next section, we summarize the evidence linking APOE4 to the meningeal lymphatic system, and how this may contribute to the pathophysiology of AD.

4.3 Investigating the potential links of APOE4 and the lymphatic system

Brain tissue samples from patients with AD compared to normal brains show several microvascular alterations, with morphological studies reporting fusiform dilations, tortuosities, and abnormal branching, an overall decrease in the density of capillaries, mitochondrial abnormalities in capillary endothelial cells, and degeneration of pericytes(330). These studies are aligned with previous data, showing focal constriction of many terminal arterioles and irregularly shaped smooth muscle cells, and capillaries with both abnormal constrictions and dilatations, in patients with AD(331).

In addition, data from both biopsies of patients with AD and mouse AD models have provided some insights into the mechanisms underlying capillary constriction and reduced blood flow; in this regard, amyloid-beta enhanced oxidative stress in pericytes via NADPH oxidase-4, leading to endothelin-1–mediated effects of endothelin A receptors on capillary-related pericytes(209). In line with these observations, clinical studies suggested that blood-
pressure-lowering drugs, such as calcium channel blockers (which lower arterial rather than venous resistance(332)) effectively increased cerebral blood flow in patients with AD(333). To our knowledge, the effects of selective endothelin receptor A antagonists in cerebral blood flow in AD patients have not been investigated.

The extent to which the meningeal lymphatic system of patients with AD exhibits the aforementioned abnormalities, and how APOE affects meningeal lymphatic vessels, are both worthy of further investigation. Characterizing the branching morphology of these vessels to detect any pathological remodeling in APOE4-related AD could provide novel insights into the pathophysiology of the disease. In this context, Lim and colleagues (2009) reported that ApoE-deficient \((ApoE^{-/-})\) mice exhibited distinct lymphatic phenotypes, including tissue swelling, leaky peripheral lymphatic vessels, a significant dilatation of capillary peripheral lymphatic vessels, and a reduction in the transport of lymphatic fluid and dendritic cells from peripheral tissues(334). Moreover, peripheral lymphatic vessels reduced their recruitment of smooth muscle cells and showed an altered distribution of the lymphatic endothelial hyaluronic acid receptor-1 (LYVE-1)(334).

It should be noted, however, that meningeal and peripheral lymphatic vessels have different characteristics. Of note, the meningeal lymphatic vessels are less complex, and show smaller lymphatic branching and fewer valves to prevent back-flow of lymph(306); nonetheless, basal meningeal lymphatic vessels have more clearly defined valves than dorsal meningeal lymphatic vessels(316). Interestingly, the metabolic pathways mediating cholesterol homeostasis in the brain also differ from those of peripheral tissues. The BBB prevents peripheral cholesterol from entering the brain, in which cholesterol is largely synthesized by astrocytes and oligodendrocytes(335).

Omnipresent endothelial cells and peripheral lymphatic vessels may share some ontogenetic features. Specifically, they may share common embryonic cellular origins; thus, peripheral lymphatic vessels may be reprogrammed to become blood vessels in case of blood flow-related events, such as shear stress(336). Given these close ontogenetic similarities between blood and peripheral lymphatic vessels, and APOE’s established role in the peripheral lymphatic system(334)(330), it would be intriguing to decipher whether \(a\) meningeal lymphatic vessels play a role in amyloid-beta clearance, and \(b\) amyloid-beta clearance is compromised by the \(APOE4\) isoform (analogous to the \(APOE4\) isoform’s effects on the BBB and meningeal lymphatic disruption(337)). In this regard, a recently published study showed that ablation of the meningeal lymphatic vessels in the 5xFAD mouse model of AD resulted in
a striking deposition of amyloid-beta in the meninges, macrophage recruitment to large amyloid-beta aggregates, and an increase in the amyloid-beta plaque load in the hippocampus(200). Moreover, as compared with healthy control mice, 5xFAD mice showed vascular amyloid-beta pathology in the cortical leptomeninges and beta-amyloid depositions in the dura mater adjacent to the superior sagittal sinus(200). Together, these findings underscore the crucial roles of CSF and ISF fluid drainage through meningeal lymphatic vessels for normal brain physiology and pathophysiology.

In addition, the worsening of amyloid-beta pathology upon disruption of the meningeal lymphatic system in mouse models of AD suggests that dysfunction of these vessels exacerbates AD pathology. Importantly, the AD transgenic mouse models (J20 and 5xFAD) used in(200) showed no differences from controls in their meningeal lymphatics (for review, see (315)). These mouse models, however, are probably less relevant to APOE4-induced AD, given that they are driven by overexpression of mutated human APP transgenes. Notably, the 5xFAD model presents a more aggressive phenotype due to the expression of mutated presenilin (PS1) in the same construct(338). Hence, these models may be better suited for early-onset AD, whereas the involvement of APOE4 may be more relevant to late-onset AD; thus, these results may not capture the effects of APOE4 as a major AD risk factor on meningeal lymphatic vessels.

5. APOE4, aquaporin 4, and the glymphatic vs. the meningeal lymphatic systems

We contend here-in that the efferent glymphatic system of paravascular astroglial channels draining into the meningeal lymphatic system could also be involved in AD pathogenesis. The glymphatic system supports CSF-ISF exchange and clearance of interstitial waste from the brain parenchyma(318), not only in mice, but also in humans(339). Moreover, the glymphatic and, as corollary, the meningeal lymphatic system (for a discussion on their interconnection and a schematic representation of their differences, see (340)) could potentially play a role in the clearance of metabolites, macromolecules, such as amyloid-beta, and inflammatory mediators, as well as immune cells. The glymphatic system’s physiology may be affected by the natural arterial pulsations of solute exchange in the CSF-ISF interphase, with major increases of this exchange taking place in the brain during deep sleep(341, 342).

In mice, disruption of the astroglial aquaporin 4 (AQP4) water channel (which regulates the glymphatic clearance of macromolecules(343)) resulted in the accumulation of amyloid-beta in the hippocampus after blockade of meningeal lymphatic drainage via ligation of deep cervical lymphatic nodes (LdcLNs)(344). Mice with deficits in both glymphatic and meningeal lymphatic clearance exhibited increased microglial activity and activation of the
microglial inflammasome, as well as enhanced hippocampal neural apoptosis and reduction of cognitive function(344). Moreover, tau levels were increased in the LdcLNs mice, but not in the *Aqp4*-null mice(344).

The lymphatic system contributes to the transport of lentiviral-delivered APOE3 to neurons(337). Thus, in addition to its clearance role, the lymphatic system may help distribute essential molecules throughout the brain (although the brain retention of APOE3 or APOE2 is lower than that of APOE4)(337). However, this notion is still in debate; for example, concerns have been raised on the role of AQP4 in the convective transport of solutes produced by the CSF-ISF exchange, calling for a reappraisal of the notion of a lymphatic system(345). In addition, there is evidence supporting a more diffusive mode of transport that does not involve AQP4(346). Nevertheless, a recent thorough investigation and meta-analysis suggested that AQP4 dysfunction could impair CSF influx (and, as a possible corollary, the exchange of solutes and cells with the ISF)(347). Therefore, further studies are needed to help us understand how the lymphatic and meningeal lymphatic systems are interconnected with regards to AD pathology, and how APOE4 influences this cross-talk.

6. Expression of lymphatic-vessel genes in *APOE4*-expressing cell types of the brain

6.1 Cell-specific effects of APOE4 in AD pathology

The cell-specific effects of APOE4 in AD pathology are poorly understood. Lin and colleagues attempted to determine which cell types in the human brain are affected by the expression of the *APOE4* isoform. To this end, they applied CRISPR/Cas9 gene editing to create *APOE4* knock-in human induced pluripotent stem cells (iPSCs) stemming from a single iPSC line of a human subject without AD; all these modifications were conducted in an otherwise isogenic background(348). Using a differentiation system for iPSCs, they generated various brain cell types and compared gene expression resulting from the *APOE4* knock-in to that of an analogous *APOE3* knock-in(97). They found that the expression of hundreds of genes was altered in iPSC-derived neurons, astrocytes, and microglia; many of these were also aberrantly expressed in postmortem samples of patients with AD(348). The observed cellular defects caused by *APOE4* expression included a) an increased number of synapses and elevated amyloid-beta-42 secretion in neurons, b) defects in amyloid-beta uptake and cholesterol accumulation in astrocytes, and, c) an aberrant morphology correlating with reduced amyloid-beta phagocytosis in microglia(348). Therefore, by harnessing gene editing approaches in an otherwise isogenic background, this study has the potential to offer interesting mechanistic insights concerning *APOE4*. 

73
To this end, we theorize below that specific cell types could be used as surrogates for lymphatic vessels to provide initial clues for our conceptual framework.

### 6.2 The common mesodermal origin of lymphatic vessel cells and microglia

To provide a rationale for the use of surrogates to understand meningeal lymphatic vessel functions, we attempt here to highlight the common mesodermal origin of lymphatic vessels and microglia. On the one hand, the ontogenetic origins of microglia cells, now shown to comprise at least nine transcriptionally-distinct subtypes (349), have remained controversial. Experimental evidence suggested that microglial homeostasis in the adult brain was not mediated by postnatal hematopoietic progenitors, whereas microglia development in mice required the colony stimulating factor-1 receptor (CSF-1R). Based on an in vivo lineage tracing study, the origins of adult microglia was attributed to primitive myeloid progenitors, which, in turn, originated from the yolk sac (i.e., an element which arises before mouse embryonic day 8, and which is associated with the splanchnic layer of the lateral plate mesoderm (350). On the other hand, the ontogenetic origin of lymphatic vessels has been extensively discussed. Some authors suggested that the lymphatic vessels emerge from the primary lymph sacs, which develop from primitive veins. Others suggested that they first arise from mesenchymal spaces, and then connect to the primary sacs, thereby, indirectly, to the venous system (312). However, in both sides of the argument, a common mesodermal origin has been proposed for microglia and lymphatic vessels. The above suggests that microglial cells might be used as a proxy to study lymphatic cells, particularly in the context of an iPSC line. Potente and Mäkinen (351) have summarized the evidence on peripheral lymphatic vessel origin, noting a) the undefined, in many cases, origin of non-venous lymphatic endothelial cells, b) the migration of lymphatic endothelial cells from the lymphatic vasculature, and c) the fact that blood-forming endothelial progenitors undergo vasculogenesis to form lymphatic vessels.

### 6.3 Re-analysis of data from previous studies

Considering the above information, we theorize that the cell-type specific RNA-seq dataset from Lin and colleagues could allow us to examine, at the level of gene expression, the hypothesis that meningeal lymphatic vessels are affected by APOE4 (348). Encouraged by both recent calls to make use of open public genomic data (352) and the notion that iPSCs can be an attractive way to model AD pathology (353), we re-analyzed the data from the study of Lin and colleagues deposited in GEO (GSE102956). Our goal was to identify genes that were differentially expressed in the APOE4 (pathological state) vs. the APOE3 (control) knock-in. We were interested in studying the expression of selected gene markers in parental cells in the starting iPSCs and iPSCs-derived neurons, astrocytes, and microglia (see the Appendix section.
for the methods applied). These selected gene markers referred to lymphatic vessel-related genes or/and genes in which well-characterized missense mutations have been linked to peripheral lymphedema *(Table 2).*

Using the available technical replicates derived from these iPSCs, our re-analysis did not reveal statistically significant differences in the expression of genes of interest in iPSCs, neurons, and astrocytes (data not shown); this was based on searches on all gene markers presented in Table 2. It should be noted, however, that some of these gene markers may not be uniquely expressed in lymphatic endothelial cells. Interestingly, in the microglial cells, we found statistically significant differences between the *APOE4* and *APOE3* knock-in cells in the expression of several genes related to features of lymphatic vessels (Figure 2). In the *APOE4* knock-in cells, we detected significantly decreased levels of expression of *CCBE1* (a marker linked to lymphedema) and *PDPN* (adjusted *p*-values < 0.05, in both cases), and significantly increased levels of *CD68* (a marker whose absence of detection during immunohistochemistry staining is linked to features of lymphatic vessels), as well as, *PECAM1*, compared to the *APOE3* parental cells (adjusted *p*-value < 0.05, in both cases). Collectively, except for *PECAM1*, these findings indicate a more pronounced presence of RNA transcripts corresponding to lymphatic or lymphatic-related markers in cells expressing *APOE3* than in cells expressing *APOE4* (Figure 2).

To avoid the possibility of selective reporting bias in our re-analysis, we performed a systematic query to identify other studies in the GEO database focusing on human-derived samples and APOE (Appendix). Twenty-two GEO datasets were identified in total, including the one described above (GSE102956). Fourteen datasets were excluded for various reasons (for the reasons of exclusion, see Suppl. File 1). A total of eight datasets were ultimately considered, including GSE102956 described above (Suppl. File 1). Of interest, we found an additional *in vitro* study (354) (GSE117588), in which iPSCs were derived from individuals with sporadic AD who had the *APOE4* (e4/e4) genotype. Parental iPSC lines were then edited to *APOE3* (e3/e3) genotype, and cerebral organoids (with neural progenitor cells) were generated from each of the lines; arguably, this is more relevant to AD pathology. Using this dataset, we identified significant differences in the differential expression of a range of genes related to lymphatic cells in the *APOE4* vs. the non-*APOE4* cells. Specifically, we found that *APOE4* cells of the cerebral organoids showed significant decreases in the expression levels of *EFNB2, VEGFR3 (FLT4), FN1, ITGA9, LYVE1, PDPN, PECAM1*, and *PTPRC* (all with adjusted *p*-values < 0.001), as well as, *GJA1* (adjusted *p*-value < 0.05) (Figure 3). We observed, though, a significant increase in the housekeeping gene *Glyceraldehyde 3-phosphate*
dehydrogenase (GAPDH) (data not shown) in the non-APOE4 cells, which might raise concerns on the degree of validity in the current analysis. However, some authors have expressed concerns about accounting for house-keeping genes in genome-wide RNA-Seq analyses, thus suggesting global normalizations rather than normalizations based on housekeeping genes (for a discussion on this issue, see (355)).

In parallel to this, we did not find significant differences in the differential expression of genes of interest in studies from human tissues, such as in (GSE48350) (356) (Supplementary Figure 1), in (GSE106241) (357) (Supplementary Figure 2) or in (GSE125050)(358) (Supplementary Figure 3) aside from the PDPN gene in (GSE29652)(359)(Supplementary Figure 4). This could indicate that, contrary to the isogenic conditions of in vitro studies on APOE4 vs. APOE3, other parameters (e.g., stage of disease, gender, and age) could have decreased the power of analysis by adding significant non-APOE-relevant variation. Finally, lack of replication in the SAGE-based study (GSE6677) (360) (Supplementary Figure 5) precluded a statistical analysis of the data (for further details on the reanalyzed studies, see Supplementary File 2).

Although the above studies are limited in their ability to predict true anatomical effects, a logical extension of our initial results is that the APOE3-expressing cells exhibit more prevalent characteristics of lymphatic cells than APOE4-expressing cells. Furthermore, in APOE4-expressing cells, we observed reduced expression of gene markers linked to peripheral lymphedema (e.g., VEGFR3 (FLT4) (361)) or lymphatic valve formation (e.g., FN1 (362), GIA1 (363, 364), and ITGA9 (365)). Together, these findings indicate a deficient effect of APOE4 on certain, iPSCs-derived, lymphatics-related functions.

6.4 Assessing the relevance of re-analyzed data on CNS lymphatic biology

Could these insights, which are based on transcriptomic data obtained from iPSCs, be relevant to CNS lymphatics biology? Extensive experimental research on the effects of APOE4 on peripheral or CNS lymphatic endothelial cells/vessels is required before drawing any certain conclusion. With regard to the first set of data (Figure 2), a major consideration was that microglia should not express AQP4 (classical marker of astrocyte endfeet in the brain) or PECAM1 (classical marker of brain vascular endothelial cell). This is because, at the single-cell level, the transcripts for PECAM1 in vascular endothelial cells (either of blood or lymphatic origin) and for AQP4 in astrocytes are a hundred, if not a thousand, times higher than in microglia. This consideration could hinder the interpretation of the results regarding iPSC-derived microglia.
Another issue, in the studies that we reanalyzed, was the ontogenetic stage at which the APOE4-knock-in iPSC-derived cells started to express PECAM1, PDPN, CCL21, LYVE-1, FLT4, and PROX1, at similar transcript levels to lymphatic endothelial cells. This is a relevant issue because the above APOE4-knock-in cells may not directly reflect the physiology of lymphatic endothelial cells.

6.5 The notions of attenuated lymphaticness, meningeal lymphedema, and lymphosclerosis in APOE4-related AD

Despite these considerations, the preliminary findings of our re-analyses, regardless of the origin of iPSCs, point towards higher expression of RNA transcripts corresponding to lymphatic or lymphatic-related markers in APOE3 vs. APOE4 cells. Thus, despite the distinct genetic signature and function of meningeal lymphatic vessels, our data suggest that APOE4-related AD may be linked to attenuated lymphatic features, such as shrinkage of meningeal lymphatic vessels, weakened function of meningeal lymphatic valves, and, in turn, reduced lymphatic flow. Taking this further, and in line with previous findings on ageing in rodents (316), we propose to label these APOE4-mediated cellular events with the term attenuated lymphaticness. Accordingly, if meningeal lymphatic vessels undergo alterations resembling lymphosclerosis (as previously defined in(366, 367)), the CSF after the exchange of solutes and cells with ISF coming through the efferent glymphatic system (a term we wish to describe as glymph) could be obstructed in terms of their flow, causing local stagnation (i.e., meningeal lymphedema). Whether or not this meningeal lymphedema could affect the removal of cellular debris, amyloid-beta, and tau, is an issue worth exploring in future studies. Early evidence suggests that the meningeal lymphatic system may be implicated in the clearance of such elements(368) (200, 369-372).

According to field experts, lymphosclerosis can be classified into 4 groups based on severity: a) thin (translucent) and expandable lymphatic vessel walls, and identifiable lumen (s0); b) thin (white) and expandable lymphatic vessel walls wall, and identifiable lumen (s1); c) thick (white) non-expandable lymphatic vessel walls, and identifiable lumen (s2); and, d) very thick (white), non-expandable lymphatic vessel walls with non-identifiable lumen (s3) (373).

Detailed anatomical and mechanistic studies are needed to further support our proposed model of lymphosclerosis (Figure 1), and to determine which of the above four categories it fits. However, partial support comes from a) recent evidence showing that amyloid-beta causes constriction of brain capillaries in a pericyte-mediated manner (209), and, b) from the close
association of blood capillaries with lymphatic vessels, at least in the peripheral lymphatic system (351). Additional support comes from studies on the basal meningeal lymphatic cells in ageing rodents, as outlined above (316) (based on findings from (374) (375)). Others have reported differences between dilated lymphatic vessels (ectatic type) and non-dilated lymphatic vessels with thickened walls (contraction/sclerotic type). Therefore, two different characteristics of lymphatic vessels –diameter and sclerosis– should be evaluated (366). Thus, in parallel to the potential presence of meningeal lymphatic hyperplasia in APOE4–mediated AD, additional evidence should be provided on the presence or lack of markers of meningeal lymphosclerosis.

To this end, it would be tempting to derive speculations from the observation that, in peripheral lymphedema, an arrest of lymphatic contraction occurs in late stages of the disease; notably, this arrest is characterized by a gradual decrease in the contraction amplitude (but unaltered frequency of contraction) of the collecting lymphatic vessels (375). Nevertheless, the negative effects of the ligation of deep cervical lymph nodes, or of lymphosclerosis (if it is perceived as the functional analogue of the mechanical ligation of lymph nodes) on AD may be disease-specific (376). Notably, in the Experimental Autoimmune Encephalitis (EAE) model of MS, in which meningeal lymphatic vessels are not submitted to expansion (lymphangiogenesis), the obstruction of these vessels is linked to a reduction in brain reactive T cell-mediated inflammation and amelioration of the disease phenotype (377) (378).

6.6 Potential alternative explanations for our proposed conceptual framework

Another possibility could be that the altered neuroinflammatory response (204, 207, 379) and proinflammatory mediator levels (such as chitinase-3–like protein 1, and several cytokines, such as TNF-alpha, Interleukin-1 receptor antagonist, and complement component C1q and 3) in the CSF of patients with AD (380-382) may be caused by abnormal microglial function (383). Interestingly, although certain microglial types confer resistance to neurodegeneration (384), and higher resistance to certain regions, such as cerebellum and white matter (385), the so-called morphologically activated microglia exerts an influence on neurodegeneration that could be similar to the one caused by APOE4 (386). Therefore, in line with recent findings on APOE4’s role in microglial activation (387), APOE4-mediated neuroinflammatory response and proinflammatory cytokines may ultimately influence CNS lymphatic vessel drainage capacity, and disturb the composition and exchange of both CSF and ISF macromolecules, as well as CNS immunity.
6.7 Framing of our conceptual framework into the broader evidence regarding AD pathogenesis

Given the lack of studies exploring the role of APOE4 in meningeal lymphatic function, additional evidence will be needed to explore the concepts presented here. Of note, our informed hypothesis on the APOE4-mediated meningeal lymphedema may not be entirely surprising, considering that a) APOE4 may exercise its effects in a pathway that involves the vascular endothelial growth factor (VEGF), and VEGF upregulation can reverse APOE4 pathology(388), and b) VEGF-C administration can restore meningeal lymphatic vessel pathology in aged mice(200). Given that APOE4 is a principal genetic factor for AD, elucidating its role in AD remains pivotal, similarly to the roles of chief risk factors in other diseases (e.g., mechanistic investigations on the FTO region, which has the strongest genetic links to obesity(389)).

More broadly, our conceptual framework also raises the questions: why APOE4 has not been associated with peripheral edema? Why has APOE4 been discussed almost exclusively in relation to meningeal lymphedema? Based on data (not shown) that we extracted from the Human Protein Atlas(390), APOE seems to be predominantly expressed at high levels in cerebral cortex, hippocampus, caudate nucleus, and adrenal medulla. All these tissues share common embryological origins. Therefore, the downstream effects of APOE would most likely be pertinent to these tissues.

As previously reviewed(244), it is also still unclear how meningeal lymphatic vessels respond to, and also control, the high levels of amyloid-beta contained in the brain fluids of patients with AD. Possibilities include poor drainage of the CSF, decreased paravascular clearance through the glymphatic system, and/or decreased clearance through the meningeal lymphatic system(244). These questions become now highly relevant in light of the fact that disruption of meningeal lymphatic vessels in an AD mouse model leads to amyloid-beta deposits in the meninges(200). Thus, we suggest that future studies should clarify the involvement and mechanisms through which APOE4 influences amyloid-beta deposition in the meninges.

Unfortunately, no study in humans has yet linked levels of the well-established and emerging serum/CSF biomarkers of AD (i.e., tau protein, phosphorylated tau-181, neurogranin, chitinate-3-like protein 1, neurofilament light, synaptosomal-associated protein-25, amyloid-beta-40 and amyloid-beta-42 isoforms, visinin-like protein 1, and blood alpha-2 macroglobulin(379, 391-397)) with alterations of the meningeal lymphatic system. This could be clarified by using advanced microscopy techniques (e.g., cryo-electron microscopy) of
APOE4–expressing cells to elucidate abnormalities in the morphology of blood vasculature and lymphatic systems. In addition, 7-Tesla MRI imaging (314) could enable a comparison of vessel diameter between patients with AD with or without the APOE4 isoform (314).

Lastly, given the single nucleotide variation difference between APOE3 and APOE4 (Figure 1), genome editing by CRISPR/Cas9, similar to applications in other neurological diseases (reviewed in (246, 398)), could offer novel opportunities for the study of APOE4-mediated pathogenetic mechanisms (227).

7. Major challenges and limitations

The data and conceptual framework discussed here must be interpreted in light of the following considerations. First, regulation of the transcriptome in one cell type does not necessarily reflect the modulation of the same genes in neighboring cells. Therefore, caution should be applied on how change in the analyzed cells reflects those in actual meningeal lymphatic endothelial cells. Potentially, the APOE4–mediated effects on meningeal lymphatic cells could be the outcome of a cross-talk between intra- and inter-cellular interactions between lymphatic endothelial cells, as also observed in developmental stages of the fetal and postnatal period (399). Second, the technical difficulties associated with the assessment of the differences in postmortem brain specimens between APOE4 vs. non-APOE4 patients with AD pose a major challenge in the proposed conceptual framework; for instance, histological fixation may affect the diameter of lymphatic vessels, suggesting that rapid freezing of tissues may be critical in future experiments. Third, the in vitro studies analyzed here compared the APOE3- with the APOE4-expressing cells; however, the feasible addition of a knock-in cell (e.g., with APOE2 genotype) would have been a valuable control. Doing so would be essential to draw further conclusions concerning the differential effects of APOE3 and APOE4 on the expression of the analyzed genes.

An important clinical consideration is that some of the lymphatic vessel-associated genes examined here may not be specific to the meningeal lymphatic vessels, and that large changes in their expression levels in lymphatic endothelial cells would cause lymphedema. However, there are no available data on the higher incidence of peripheral lymphedema in patients with mutated APOE4, or in patients with AD. The tissue-specific expression of APOE implies that the effects of APOE will most likely be exerted in brain-related tissues. Moreover, it would be crucial to link experimental data derived from this conceptual framework with impaired functional connectivity in AD, and with cellular surrogates of AD-related cognitive decline.
Given that peripheral lymphatic vessels in the small intestine contribute to the transport of lipids to the circulatory system (400), it is likely that meningeal lymphatic vessels similarly remove fatty acids from brain parenchyma. Although the mechanism for the metabolism of fatty acids in hyperactive neurons has been recently elucidated (according to this mechanism, fatty acids coupled with ApoE-positive lipid droplets are discarded from neurons, endocytosed by neighboring astrocytes, and metabolized through oxidative phosphorylation, ultimately leading to activation of molecular pathways to overcome fatty acid toxicity, and, in turn, protect neuronal function (401)), its relevance for AD neurodegeneration requires further investigation.

The above is significant, granted that the majority of differences in the brain transcriptome of patients with AD can be attributed to alterations of gene expression in excitatory neurons (along with those in oligodendrocytes) (402). However, whether the potential role of meningeal lymphatic vessels in brain fatty acid metabolism is affected by APOE4 (both with regard to APOE4 lipid droplets and our concept on meningeal lymphatic vessel architecture in APOE4-related AD) might be hard to interpret.

We also wish to point out that conducting single-cell studies from the prefrontal cortex of patients with AD (as in(402) may not be directly relevant to meningeal lymphatic vessel physiology, given that meningeal lymphatic cells do not cross the brain parenchyma but, most likely, receive fluid from the glymphatic system which does communicate with the brain parenchyma (for a review on CNS vasculature, see (403). This lack of relevance could be attributed to a) the fact that only brain parenchyma-traversing vascular structures exist in the prefrontal cortex, b) technical challenges, because of the small, statistically under-powered number of each category of cells, and, c) the difficulty to control for common factors, such as age and gender. Of note, we conducted a preliminary reanalysis on the above single-cell transcriptomics study (402), where we mapped the APOE genotypes to individual cells. This reanalysis revealed a bias of APOE4 allele-carrying patients for higher Braak stages, a method to classify the severity of AD based on autopsy findings. Notably, no samples with Braak stage 1 or 2 have the APOE4 allele, and no non-APOE4 samples have Braak stage VI. Mitigation of this bias might be possible in more abundant cell types (e.g., microglia, neurons, astrocytes, and oligodendrocytes) by limiting the analysis to the middle of severity level (i.e., where both genotypes are present). However, the latter approach is not equally possible in the two major cell populations of interest, namely pericytes and endothelial cells, because of their much smaller numbers (Supplementary File 3). Collectively, the above single-cell transcriptomic study (402) might have been biased, given that a) all APOE4 samples were also high-severity samples (as assessed using Braak stages), and, b) only AD samples were examined, with no
healthy controls but mere separation based on no or little pathology group and mild to severe pathology group.

Moreover, with regards to recent findings on basal meningeal lymphatic vessels (316), distinguishing the potential effects of APOE isoforms on basal vs. dorsal meningeal lymphatic vessels would also be of importance. Finally, in light of the very recently deciphered molecular anatomical connections between meninges and the skull, deciphering the complexity of meningeal architecture and how the latter could affect processes of neuroinflammation and neurodegeneration (including AD) will be of crucial importance (404).

8. Conclusions and future perspectives

The study of meningeal lymphatics will hopefully improve our understanding of AD. Indeed, the evidence discussed here raises fascinating questions about the connection between APOE4 and the meningeal lymphatic system in AD. Deciphering the role of APOE4, the strongest known genetic link to AD, in the meningeal lymphatic system could reveal a missing link in our understanding of the etiology and pathology of AD. The suggested association between APOE4 and molecules of the VEGF pathway, if further validated, could provide further insights into the demonstrated link between VEGF ligand and receptor genes with the cognitive decline and neuropathology of AD, although the precise mechanisms involved remain to be investigated (405).

More broadly, by considering a) the recent observations on capillary vasoconstriction in AD (406), b) the neuronal/astrocytic metabolism of fatty acids coupled with the established role of lymphatic vessels (at least in the peripheral system) for fatty acid transportation (401), c) the inverse epidemiologic associations of AD with diets with low input of saturated fats (e.g., Mediterranean diet) (83), and d) the absence of AD-like alterations in non-human primates (407), we could conceptually describe late-onset AD as an evolutionarily human-specific disease, in which several epidemiologic factors (such as modern Western lifestyle) exert a profound impact on a human-specific variant APOE4. In addition, lymphatic vessels share common features with other types of vessels; therefore, this conceptual framework on meningeal lymphatic cells (especially its aspect on lymphosclerosis) could be aligned with previous calls for further studies addressing how environmental factors affect, in both common and distinct ways, arterial (atherosclerosis), and venous capillaries (capillary vasoconstriction) in the skull region (406, 408).

Our conceptual framework, if further verified by additional, intensive mechanistic studies, could serve as a prelude to the development of CSF-implicating therapeutics in AD and
other neurodegenerative disorders. Indeed, recent studies have demonstrated a broad, previously unexpected role of CSF in various brain pathologies (409), while intrathecal administration of medications was employed successfully in major neurogenetic disorders (410). Therefore, we speculate that intrathecal administration of VEGF or other agents might restore normal anatomy and function of meningeal lymphatic vessels in AD, allowing passage of lymphatic and lymphatic fluid into the cervical lymph nodes and proper processing of its content by the resident FDCs. Interestingly, successful therapies involving cervical and other lymph nodes, potentially including FDCs, have been applied in other neurological diseases (e.g., the use of fingolimod in MS(411).

We hope that this novel conceptual framework, coupled with previous and future findings, may help develop the notion of meningeal lymphedema and lymphosclerosis in APOE4-related AD. In a broader context, informed hypotheses such as this could assist in the integration of epidemiologic data (such as the role of low-lipid diets in AD) with molecular signalling data (APOE4–downstream events). Finally, our approach highlights the power of reanalyzing open data to produce new perspectives in the precision medicine era (352, 412).

9. Methodology applied for Re-Analyses: Identification of Studies, and Differential Expression (Re)analyses

To identify studies in which we could re-analyze samples with AD based on APOE4 vs. APOE3 phenotype, we searched for all studies present on the GEO (https://www.ncbi.nlm.nih.gov/geo/) until June, 1st 2019 focusing on APOE in humans. More specifically, the URL to the query used to perform the search is: https://www.ncbi.nlm.nih.gov/gds/?term=(apoe)%20AND%20%Homo%20sapiens%22[por gn]%20AND%20(%22gse%22[Filter]). Overall, 22 datasets were identified (See Supplementary File 1 for included and excluded studies).

In general, a systematic meta-analysis of multiple studies pulled into a single dataset can increase the power to detect DE genes due to a larger sample size. However, the studies identified from the above GEO search originated from critically different experimental systems ranging from biopsies and iPSC lines to brain organoids, making these studies not comparable. Hence, forcing such different studies into a meta-analysis might not be meaningful. Indeed, combining datasets merely aiming to create a bigger dataset could lead to a number of biases (413); as such, we considered all individual studies. For the description of specific studies identified through the query and exclusion criteria, see Supplementary File 1. For each of the
respective studies, expression values were downloaded from the GEO repository. Whenever possible, using the pre-processed values deposited by the authors was preferred in favor of the raw data.

With regards to microarray-based studies, distribution of expression values across all genes within each sample was visually examined to establish whether the data had been normalized prior to submission. If necessary, raw intensities were log2-transformed and normalized using the quantile normalization method from the affyPLM package (414). Probes without gene annotation were removed. Differential expression analysis was conducted using limma package (414). Gene expression was modelled jointly for all samples in a given study, for which the APOE genotype status was available. The test contrast was set between samples containing at least one APOE4 allele (e4 group) and those containing no such allele (no e4 group). Additional variables, such as Braak stage, gender, and age, were included in the model, whenever reported. To take into account the false discovery rate, adjusted p-value was computed by Benjamini-Hochberg method on the genes of interest present within the study (415).

With regards to RNA-seq studies, gene count matrices were downloaded from the GEO repository (based on their accession numbers from corresponding studies), and processed using edgeR package (416). Genes with low expression levels were removed, as a result of our requirement for, at least 5 reads, in at least 3 samples. In the study GSE125050, which contained several samples with low library sizes, we additionally removed samples that had less than 10,000 genes with at least 5 reads, in order to ensure a minimal library complexity. Normalization factors were then calculated using the Trimmed Mean of M values (TMM) method in edgeR package (416). These normalization factors were used to calculate effective library sizes, in turn used to calculate normalized CPM values for genes of interest and visualized as boxplots and a heatmap. Differential expression testing was carried out using edgeR’s quasi-likelihood framework (417). Setting of test contrasts and inclusion of additional variables were conducted, as described above for microarray-based studies. Raw p-values were adjusted for multiple testing by applying the Benjamini-Hochberg method (415) on the genes of interest.

Concerning the analysis of SAGE-based dataset (GSE6677), mappings between SAGE tags and genes were downloaded from SAGEmap database (ftp://ftp.ncbi.nlm.nih.gov/pub/sage/mappings/SAGEmap_Hs_NlaIII_10_best.gz). No statistical analysis of the data was performed due to the lack of replication. Expression was
visualized using scatter plots. For genes that had more than one associated tag, all counts for tags were plotted, and colour was used to distinguish between different tags within the same gene.

With regard to the single-cell RNA-Seq studies, single-nucleus RNA-seq profiling data from prefrontal cortex of patients with AD was obtained from the ROSMAP study (https://www.synapse.org/#!Synapse:syn18485175), after approval of request from the Synapse Access and Compliance Team, and in alignment with a previous study (418). Clinical information for the participants, including the APOE genotype, was obtained from https://www.synapse.org/#!Synapse:syn3191087. We assigned the APOE genotype to each cell by matching via patient ID and calculated number of cells with a given genotype within each cell type.
Figure 1. Hypothesized role of Apolipoprotein E (APOE) and lymphatic vessels in Alzheimer disease (AD). Schematic depiction of mouse ApoE and human APOE isoforms (top right), with yellow indicating the presence of cysteine [C] and blue indicating the presence of arginine [R]. Illustration of brain and meningeal anatomy in a non-AD brain, showing meningeal lymphatic vessels, brain cell types, and outflow of CSF (top left and magnified in bottom left). The labeling “APOE e4” with an arrow pointing to the dysfunctional meningeal lymphatic vessel (bottom right) illustrates the hypothesis that APOE4 is associated with the reduction in size of meningeal lymphatic vessels, and the consequent blockade of clearance of macromolecules such as amyloid-beta. (adapted and modified from Ref. (419) © 2018, and Ref. (304), © 2015, both with permission from Springer Nature).
Figure 2. Box and whisker plots showing meningeal lymphatic marker expression levels in iPSC-derived APOE4 and APOE3 (control) knock-in cells, with iPSCs being derived from an unaffected subject carrying APOE3 alleles, in which cells were gene edited using CRISPR/Cas9 to generate APOE4 iPSCs from parental APOE3 cells, and isogenic iPSCs were then differentiated into neurons, astrocytes and microglia-like cells. Normalized gene expression of meningeal lymphatic markers and related genes in cells expressing knock-in of either non-APOE4 (blue) or APOE4 (yellow). Differentially expressed genes are marked with an asterisk to indicate statistical significance: * for false discovery rate (FDR) < 0.05; ** for FDR < 0.01; and *** for FDR < 0.001. The perpendicular bars represent the standard deviation (SD). Based on re-analysis of data from the published source in (348). Please note that the scale of the graph may differ between the different, depicted genes.
Figure 3. Box and whisker plots showing meningeal lymphatic marker expression levels in iPSCs derived from individuals with sporadic AD APOE4 (e4/e4) genotype, with parental iPSCs lines edited to E3 (e3/e3) genotype, and cerebral organoids generated from each of the lines. Normalized gene expression of meningeal lymphatic markers and related genes in cells expressing knock-in of either non–APOE4 (blue) or APOE4 (yellow). Differentially expressed genes are marked with an asterisk to indicate statistical significance: * for false discovery rate (FDR) < 0.05; ** for FDR < 0.01; and *** for FDR < 0.001. The perpendicular bars represent the standard deviation (SD). Based on re-analysis of data from the published source in (397). Please note that the scale of the graph may differ between the different, depicted genes.
### Table 1. Genetic, environmental, and lifestyle risk factors for Alzheimer disease (AD) appearance or progression.

<table>
<thead>
<tr>
<th>Non (purely) genetic risk factors</th>
<th>Genetic—Chromosomal factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depression at any age and late-life depression&lt;sup&gt;a&lt;/sup&gt;</td>
<td>APOE4 and other gene loci, including some variants more prevalent in APOE4&lt;sup&gt;+&lt;/sup&gt; patients&lt;sup&gt;c, m&lt;/sup&gt;, and SORLI (neuronal apolipoprotein E receptor)&lt;sup&gt;e&lt;/sup&gt;, and rare coding variants in Apolipoprotein B&lt;sup&gt;p&lt;/sup&gt;</td>
</tr>
<tr>
<td>Type 2 diabetes mellitus&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Amyloid precursor protein (inherited AD form)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>Frequency of social contacts—loneliness&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Presenilin-1 gene (inherited AD form; main-cause in autosomal-dominant, early-onset AD)&lt;sup&gt;y, u&lt;/sup&gt;</td>
</tr>
<tr>
<td>Benzodiazepines use&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Presenilin-2 gene (inherited AD form)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>Low adherence to Mediterranean diet&lt;sup&gt;b, l&lt;/sup&gt;</td>
<td>Trisomy 21 (Down syndrome)&lt;sup&gt;e&lt;/sup&gt;</td>
</tr>
<tr>
<td>Aging&lt;sup&gt;g&lt;/sup&gt;</td>
<td>Immune (epigenetically regulated) response&lt;sup&gt;f&lt;/sup&gt; and other epigenetic events&lt;sup&gt;x&lt;/sup&gt;</td>
</tr>
<tr>
<td>Anemia/Very low hemoglobin levels&lt;sup&gt;b&lt;/sup&gt;, as well as very high hemoglobin levels (U-shaped relation)&lt;sup&gt;y&lt;/sup&gt;</td>
<td>Variants in loci in TREM2 and soluble TREM2 modulators, i.e., MS4A4A and MS4A6A, and loci in CD2AP, IQCK, ACE, ADAM10, ADAMTS1, and WWOX, and rare variants&lt;sup&gt;j, w&lt;/sup&gt;</td>
</tr>
<tr>
<td>High cortical iron levels&lt;sup&gt;l, k&lt;/sup&gt; and plasma ferritin&lt;sup&gt;l&lt;/sup&gt;</td>
<td>Variants in loci in IGHG3, ZNF655, GPA1, OR8G5, IGHV3-7, SLCA2A3, and IncRNA AC099552&lt;sup&gt;q, r&lt;/sup&gt;</td>
</tr>
<tr>
<td>Grand Multiparity&lt;sup&gt;k&lt;/sup&gt;</td>
<td>Somatic brain mutations in MAPK, AMPK, and PI3K-AKT pathway&lt;sup&gt;n&lt;/sup&gt;</td>
</tr>
<tr>
<td>Chronic stress (psychological and biological) and inflammation&lt;sup&gt;g&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Low-density Lipoprotein (LDL) Cholesterol&lt;sup&gt;p&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Neuropsychiatric manifestations: psychosis, aggression/agitation, affective symptoms&lt;sup&gt;z&lt;/sup&gt;</td>
<td></td>
</tr>
</tbody>
</table>

**Sources:** a. (42); b. (83); c. (420); d. (421); e. (246); f. (204); g. (193); h. (422); i. (423, 424); j. (425); k. (426); l. (427); m. (428); n. (429); o. (430); p. (431); q. (432); r. (432); s. (433); t. (434); u. (435); v. (436), (437, 438); x. (439); y. (440); z. (441, 442)

<sup>*</sup>Following postmortem assessment, and mostly referring to risk for cognitive decline in already diagnosed patients with AD.
Table 2. Markers of human lymphatic vessels.

<table>
<thead>
<tr>
<th>Lymphatic markers &amp; Lymphatic-related markers</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Prox1</td>
<td>Podoplanin (PDPN)</td>
</tr>
<tr>
<td>CD31 (PECAM1)</td>
<td>VEGFR3</td>
</tr>
<tr>
<td>LYVE-1</td>
<td>CCL21</td>
</tr>
<tr>
<td>CD68 negativity⁺</td>
<td>CD45 (PTPRC)⁺</td>
</tr>
</tbody>
</table>

**Initiation of valve formation**

<table>
<thead>
<tr>
<th>GATA-2</th>
<th>Ephrin type B2 (EFNB2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FOXC2</td>
<td>Prox1</td>
</tr>
</tbody>
</table>

**Glymphatic and Astrocytic Markers**

| Aquaporin-4                                   |  |

**Lymphatic Valve Maturation**

| Integrin alpha-9 (ITGA9)                      | Fibronectin 1 (FN 1) |
| Connexin 43 (GJA1)                             |                      |

**Lymphedema**

<table>
<thead>
<tr>
<th>VEGF-C⁰</th>
<th>VEGFR-3 (or FLT4)⁺</th>
</tr>
</thead>
<tbody>
<tr>
<td>FOXC2⁺</td>
<td>GATA2⁺</td>
</tr>
<tr>
<td>CCBE1⁺</td>
<td>GJC2⁺</td>
</tr>
</tbody>
</table>

⁺According to (304), the absence of CD68 positivity on Lyve-1 positive structures has been noted.
⁺⁺According to (443), a striking co-localization of CD45(+) leukocytes with the developing lymphatics has been noted in mice.
⁺⁺⁺According to (444), astrocytes lacking GJA1 showed reduced Apoe protein levels as well as impaired amyloid-beta phagocytosis.
⁺⁺⁺⁺VEGFC exposure increases the diameter of meningeal lymphatic vessels; the latter completely fail to develop upon VEGFC and VEGFD inhibition.
⁺⁺⁺⁺⁺Marker of initiation of valve formation, and mutant gene forms linked to lymphedema-distichiasis syndrome.
⁺⁺⁺⁺⁺⁺Mutant gene forms, with functional, missense mutations linked to lymphedema.

Source: Gene markers based on previous reports from: (304); (445); (244, 316); (446); (447); (443); (448); (449); (450); (451) and OMIM (Online Mendelian Inheritance for Man) Database (www.omim.org) (Assessed January 2019).
Chapter 3 — The Single Microbe Contribution to Neurological Diseases: The case of Helicobacter Pylori infection

1. Helicobacter Pylori Infection and Gastric Cancer Biology: Tempering a Double-Edged Sword

1.1 Introduction

As one of the most “successful” pathogens in terms of causing infection to humans, Helicobacter pylori affects an estimated 50-60% of the world’s population or an estimated 4.4 billion people globally in 2015 (453). It is a common cause of treatable conditions, such as dyspepsia and peptic ulcer disease. Since the turn of the century, the prevalence of H. pylori infection has declined in industrialized countries and plateaued in developing and newly industrialized countries. In the 1980’s, Warren and Marshall made the landmark discovery (with earlier indirect indications from Lykoudis) (454) that H. pylori causes gastric ulcers, which led to a “paradigm shift” challenging axiomatic beliefs that microbial populations cannot survive in the gastric pH-intensive environment. More recent discoveries of urease- and pH-sensing BabA adhesin have provided clues to the underlying mechanisms for H. pylori’s ability to withstand the gastric environment (455).

H. pylori infection is also a critical risk factor for gastric cancer, contributing to approximately 75% of all gastric cancer cases, as well as to low-grade mucosa-associated lymphoid tissue (MALT) lymphoma, a subcategory of primary gastric lymphoma, which accounts for approximately 30-40% of extranodal lymphomas in total (456). Gastric cancer has remained a significant commonly diagnosed forms of gastrointestinal malignancy (457). Although morbidity and mortality from gastric cancer have decreased in the past few decades, gastric cancer remains the third leading cause of cancer for both men and women worldwide (453), constituting a global health problem. Thus, considerable efforts are focused on H. pylori eradication in vulnerable patients, with the goal of eliminating gastric symptoms and preventing the development of gastric cancer. This strategy has shown signs of success: for instance, the decrease in H. pylori infections in Japan is believed to have contributed to a decline in gastric cancer cases (453). However, the cause of an increase in gastric cancer in young population in the United States (notably, young Hispanic men), where overall the incidence of H. pylori infection is also waning, is unexplained (458). Thus, unknown factors likely unrelated to H. pylori infection may be contributing to a rise in gastric cancer in specific populations, further complicating the perplexing etiology of the disease.
Intensive research is focusing on the elucidation of the molecular mechanisms for *H. pylori* infection and its contribution to gastric cancer to advance diagnostics, preventions, and therapies, with gastric cancer arena hallmarked by recent cutting-edge findings (459, 460). Moreover, *H. pylori* research can serve as an exemplar for studying a cancer-causing pathogen that is resident in the relatively accessibly stomach, which may provide insights to mechanisms for microbial-related carcinogenesis in less tractable organs, such as the pancreas and lung.

Herein, we discuss recent trends in the understanding of *H. pylori* and its connection to gastric cancer, namely: (i) how gastric cancer evolution dynamics can be affected by the interplay between the number of stem cells in a tissue and DNA mutational burden, and how *H. pylori* may impact this dynamic; (ii) the mechanobiology of gastric cancer, from matrix metalloproteases to the potential of 3-D organoids; (iii) the association of *H. pylori* with the microbiome and the host; and (iv) the therapeutic perspectives potentially gleaned from these current trends.

### 1.2 Gastric cancer dynamics

#### 1.2.1 Cancer dynamics and evolution in gastric tumors

Gastric neoplasms are characterized by a high degree of heterogeneity, which makes it challenging to identify which therapeutic agents should be brought to clinical trials (461). Gastric adenocarcinoma, which comprises diffuse, intestinal, and well-differentiated subtypes is responsible for approximately 90% of gastric cancer cases (461). Moreover, gastric cancer comprises four molecular subtypes, each of which involves distinct sets of molecular players and pathways. These subtypes are: (i) Epstein-Barr virus (EBV)-positive tumors that possess PIK3CA mutations and PD-L1/2 amplifications (interestingly, patients with EBV-positive gastric tumors have higher survival rates (462); (ii) genomically-stable tumors with mutations in genes encoding E-cadherin and RHO-family GTPase-activating proteins; (iii) tumors with microsatellite instability and high mutation rates that impact oncogenic proteins 3; and (iv) tumors with chromosomal instability giving rise to aneuploidies (as detected by intestinal histology), mutations in TP53, and focal amplifications of Ras proteins (461, 463). We have also previously highlighted the need to consider the BRCA1- and BRCA2-defective subgroups of gastric cancer (464).

Elucidating the molecular characteristics of a patient’s tumor in the context of cancer dynamics and evolution could provide diagnostic information to inform therapeutic actions. Molecular profiling techniques such as single-cell exome and transcriptome analysis and
spatiotemporal gene expression analysis could reveal the extent to which a gastric cancer tumor is synchronous or metachronous, which could be relevant to treatment (465). Moreover, determining whether the phenotype subtype of a gastric cancer tumor is epithelial or mesenchymal could inform the clinician how to target the most relevant molecular pathways. The epithelial subtype is associated with higher mutation and survival rates and a better response to chemotherapy, but with lower sensitivity to inhibition of IGF1/IGF1R pathway than the mesenchymal phenotype subtype (465).

Determining the critical stage at which differentiated cells become proliferative in gastric metaplasia, as assessed through human stomach biopsies, could similarly be informative for diagnostics (466). This staging could be accomplished by analyzing human stomach biopsies to (i) identify the stages at which differentiated cell structure degradation occurs as a result of decreasing mTORC1 activity and massive up-regulation of lysosomes and autophagosome; (ii) detecting metaplasia- or progenitor-associated gene induction; and (iii) identifying points of cell-cycle re-entry through reactivation of mTORC1 (a process called “paligenosis”) (466).

A recent study applied phylogenetic techniques coupled with sequencing approaches to determine factors that contribute to the localization of intraepithelial neoplasia in another glandular epithelial tissue, the pancreas (467). The authors concluded that pancreatic intraepithelial neoplasia is not necessarily spatially localized but can spread through the entire ductal system. This observation helps to explain why pancreatic cancer is such a highly recurrent disease, especially when incomplete removal of pancreatic tissue (i.e., a subtotal pancreatectomy) is applied (467). Analysis of additional patient samples is needed to determine the characteristics of precursor lesions presenting the highest risk of transformation (467), information that could potentially be useful in the diagnosis and treatment of gastric cancer as well.

1.2.2 Heterogeneity of cancer predisposition in gastric tissue

The cancer risk of a particular tissue has been suggested to depend on the number of stem cell divisions that occur in the lifetime of that tissue (468). Vogelstein and colleagues have proposed a model whereby the key factors in cancer etiology are mutations occurring during normal DNA replication, along with heredity and environmental and lifestyle factors (469). In the model, the number of cell divisions that a tissue undergoes during its lifetime determines the extent to which mutations accumulate during DNA replication. The authors suggest that this phenomenon explains the extreme variation in cancer incidence that occurs in different tissues.
It would be intriguing to explore whether this model explains the stomach’s and duodenum’s differing susceptibility to adenocarcinoma, despite both regions being exposed to \textit{H. pylori} infection. The Vogelstein model suggests that the different number of stem cell divisions that each tissue undergoes determines the mutation and therefore cancer risk, which could be assessed by quantifying cell divisions in each tissue using the Leucine-rich repeat-containing G-protein coupled receptor 5 (Lgr5) biomarker (as discussed below).

Alternatively, physiological factors that vary along the length of the stomach and pylorus such as differences in oxygen concentration, pH, mucus, and nutrient availability could play a role in determining regional cancer susceptibility. In this case, therapeutic targets could be identified by determining which molecular pathways are over- and under-activated or expressed by assaying differences in the pH between the duodenum and stomach, similar to the studies of the hypoxia-HIF1 axis for breast cancer (470). \textit{H. pylori} genes involved in pH regulation and nickel transport have been shown to be highly expressed in gastric infection (471). Moreover, the expression of the \textit{H. pylori} type IV systems secretion plasticity zone cluster has been shown to be influenced by adherence and pH and its composition impacted by differential gastric IL-8 secretion (472). These pieces of evidence align with a model proposed by Sáenz and Mills that suggests the gastric epithelium responds to injury imposed by gastric acid secretions by cellular plasticity and reprogramming: The resulting cycles of differentiation and de-differentiation then leads to an accumulation of cancer-predisposing mutations (473). On the other hand, the pyloric environment is more alkaline, which may lead to a diminished risk of cancer because of a reduced need for tissue regeneration and the associated DNA replication stress (474).

\subsection{1.2.3 Stem cells and their biomarkers in gastric cancer}

Stem cells drive the daily renewal of epithelial tissue such as the gut epithelium. As shown in the Fig. 1, gastric cancer is believed to arise from genetic mutations and chromosomal aberrations that occur in epithelial cells as a result of environmental and lifestyle factors. \textit{H. pylori} infection is one such contributing factor. In the past two decades, the role of stem cells in driving tumorigenesis has attracted increasing attention as a result of improvements in stem cell markers and Cre-mediated cell-lineage-tracing techniques, which have facilitated the identification of cells-of-origin for various types of cancers (475, 476).

The Lgr5 receptor, whose biological ligands are R-spondin proteins, has garnered particular interest because of its role as a marker of homeostatic stem cells and reserve cells in multiple tissues (477). Lgr5 may represent a promising pathological marker for gastric cancer to identify precancerous lesions or distinguish between synchronous and metachronous cancer,
a topic of recent importance (478). The stomach lining contains Lgr5+ stem cells that mark a population of reserve stem cells termed “chief cells” that reside at the base of the corpus glands of the gastric epithelium (479). Ablation of Lgr5+ cells in the mouse corpus glands has been shown to disrupt epithelial homeostasis, suggesting that chief cells play a pivotal role in maintaining the homeostatic stem cell pool in this tissue, and that they represent the cell-of-origin for gastric tumorigenesis (479). Understanding how *H. pylori* interacts with the chief cells could reveal previously unidentified mechanisms for gastric carcinogenesis.

Another study has suggested that gastric cancer stem cells originate from bone marrow-derived cells that migrate to the gastric epithelia (480). Recent advances in single-cell-sequencing technology could enable lineage-tracing studies to elucidate the extent to which gastric stem cells originate from chief cells and bone marrow-derived cells.

### 1.3 Mechanobiology

#### 1.3.1 Three-dimensional (3-D) organoid cultures for analyzing *H. pylori* interactions with gastric tissue during tumorigenesis

A major critique of *in vitro* cancer studies has been the 2-dimensional character of cell lines they have usually employed (481). Recently, developments in 3-dimensional (3-D) organoid cultures have enabled a more comprehensive mechanobiological (i.e., 3-dimensional) perspective into studies of organ development and carcinogenesis. Such studies have illuminated the role of Lgr5+ cells in 3-D organoids derived from gastric epithelia (482), gall-bladder (477), and intestine (483).

Such a system has recently been used to study the mechanisms by which *H. pylori* contributes to gastric cancer and the importance of *H. pylori* long-range interactions with gastric tissues. Sigal et al (484) used lineage tracing in a mouse model of *H. pylori* infection to show that *H. pylori* induces R-spondin in myofibroblasts underlying the antral glands, thereby promoting proliferation of an Axin2+ Lgr5+-progenitor cell residing their base. They then reconstituted these interactions in epithelial organoid cultures containing primary myofibroblasts (484). This model system and others will provide further insights to the mechanisms by which *H. pylori* activates long-range induction of gene expression in myofibroblasts, as reviewed by Waskito et al (485). More recently, researchers have developed organoids containing gastric primary cells that can be viable for more than one-year and that express Lgr5+, MUC5AC+, and combined cytokeratin- and E-cadherin-positive cells. The prolonged *H. pylori* infection of the organoids led to the creation of cellular vacuoles and cytokine overproduction (486).
Organoid cultures have been further employed to study the interaction of primary tumor epithelia with tumor-infiltrating lymphocytes (487). This system might offer mechanistic clues, at the molecular level, for the lower efficacy of immunotherapy observed so far in gastric cancer (488), with the exception of the microsatellite-unstable gastric cancer that appears to be particularly vulnerable to this therapy (463).

### 1.3.2 Matrix metalloproteases and mechanobiology

Matrix metalloproteases are molecules that function in the extracellular matrix, thus have implications for pathological processes attributed to tissue disruption such as the “leaky vessels” phenomenon (489, 490), the “leaky gut” in the context of sepsis (489, 490), increased blood-brain barrier permeability upon CNS autoimmune inflammation (491, 492), and local invasion as well as distant metastasis of tumors (493, 494). Recently, researchers have identified additional roles for matrix metalloproteases in intrinsic cell functions that regulate cell differentiation and survival. For example, the presence of MMP-9 was found to mediate HER2 oncogene transcription in gastric cancer cell lines, thereby promoting tumor malignancy [46]. The newly acknowledged nuclear localization of matrix metalloproteases implies that they play a role in activating apoptotic cell death pathways (495). On the other hand, Zhao et al (496) showed by CRISPR/Cas9 gene editing — a technique with high potential to transform medical research (reviewed, among others, in (246)) — that translocation of H. pylori’s CagA into the host is mediated by specific carcinoembryonic antigen-related cell adhesion proteins, but not integrins, the receptors that mediate cell-ECM adhesion.

As reviewed by Ladoux and Mege (497), studies of the mechanobiology of gastric tissue could provide a rationale for the observations that gastric cancer prevalence varies among different regions of the stomach, with cancers arising in the corpus potentially caused by mechanisms distinct from the other regions (498). Extracellular matrix adhesion molecules may contribute to the regional propensities for cancer within the stomach; thus, it will be worth investigating the extent to which adhesion molecules differ at the tissue and stomach sub-regional level (e.g., the antrum, cardia, or non-cardia areas).

Comparative studies of extracellular matrix adhesion molecules could provide a further mechanistic rationale for the elastic properties of the gastric epithelium. For example, certain matrix metalloproteases, notably the gelatinases (MMP-2 and MMP-9; the latter investigated as a drug target in gastric cancer) have three fibronectin repeats in their catalytic domain (499), which have long been associated with collagen specificity (500). How the action of these MMPs
differ in \textit{H. pylori}-mediated gastric neoplastic mechanisms compared to other homologues lacking the fibronectin module, such as MMP-10 (501) and MMP-3 (502), has yet to be resolved.

In the context of reproducibility in science, the subject of recent calls-to-action (503), measurements obtained from several methods assessing cell mechanics have been found to differ by up to 1,000-fold, highlighting the need for more consistently reproducible results in this field (504). However, a recent study reporting the heterogeneous elasticity of curved epithelial sheets that enclose pressured lumens may complicate such efforts (505). Such epithelial tissues, which include the stomach, were found to comprise “superelastic materials” resistant to extreme levels of strain. Although the overall tissue layer maintains a uniform tension, the superelastic tissues themselves are highly heterogeneous at the cellular level, which could confound efforts to measure the mechanical properties of such tissues.

1.4 \textit{H. pylori} and its interaction with stomach microbiome and the host tissue

1.4.1 \textit{H. pylori} in the microbiome and biofilm studies

The mechanisms by which \textit{H. pylori} colonizes humans and mediates histologic, physiological, immune, and microbiologic features of the gut has attracted considerable interest. A recent study (506) has shown that a stable colonization of \textit{H. pylori} in mice over a six-month period resulted in an increased expression of immune response genes that was conserved across the mice for a sustained duration in the stomach, as well as transiently in the lungs. The infection led to the emergence of different structures of microbial populations within the stomach and intestines. Collectively, these data indicated that \textit{H. pylori} influences host immune responses and the microbiota of not only the stomach but also of distal organs (506).

\textit{H. pylori} has been shown to influence the composition of the gastric microbiome, and it is correlated with the presence of species from the \textit{Campylobacter}, \textit{Sulfospirillum}, and \textit{Deinococcus} genus (471). Interestingly, the presence of \textit{H. pylori} and \textit{Campylobacter concisus} have been shown to be mutually exclusive, with \textit{H. pylori}-mediating protective effects in inflammatory bowel disease (457). On the other hand, among the gastric microbiome observed, \textit{H. pylori} seems to play a predominant role, reducing the microbial diversity in the region (507, 508).

The success of \textit{H. pylori}’s survival in the stomach depends on the organism’s chemotactic motility, ability to produce urease, and its resistance to a constantly changing
environment (reviewed in (509)). While the importance of biofilms in *H. pylori* colonization has yet to be determined, they may enable its survival under changing conditions. Biofilms are defined as bacterial communities that are associated with surfaces and embedded within an extracellular matrix containing extracellular polymeric substances. They contribute to an infection becoming chronic or recurrent, promote inflammation, and can make bacterial colonies resistant to antibiotics and the immune system (509).

Given that inflammation plays a pivotal role in gastric carcinogenesis and that *H. pylori* has beneficial effects to extra-gastric tissues, it may be therapeutically effective to fine-tune the presence of *H. pylori* in its active state in tissues where it is beneficial and target its inflammatory role in tissue where it is harmful. This strategy could align with studies that have shown the anti-inflammatory action of probiotics and other alimentary factors such as Mastiha gum against *H. pylori* (510). Honey and yogurt consumption have also been shown to be associated with reductions in *H. pylori* and anti-CagA IgG seroprevalence, presumably because of honey’s antimicrobial actions and yogurt’s pro- and pre-biotic activities (511). It would be intriguing to decipher if these effects can be attributed to biofilm activity, considering that quorum-sensing ability of biofilms can control microbial population density.

### 1.4.2 *H. pylori* and host tissue response

One of the crucial components of *H. pylori* is the CagA protein. CagA is part of the cag pathogenicity island, which encodes a secretion system that translocates CagA into the target cell after *H. pylori* attaches to a eukaryotic host cell. After translocation, CagA localizes to the inner surface of the cell membrane where it is phosphorylated on a tyrosine residue by Src family kinases (i.e., *Fyn* and *Lyn*). The CagA gene promotes the epithelial-mesenchymal transition, which is mediated by the effects of CagA’s EPIYA polymorphism on phosphorylation (502). Notably, Weinberg and colleagues (512, 513) have shown that the epithelial-mesenchymal transition is linked to carcinogenesis through generation of stem cell-like cells (also reviewed in (514)).

These findings have been corroborated in the clinical arena: a recent meta-analysis concluded that individuals with a CagA+ *H. pylori*+ status have an elevated risk of gastric cancer, compared to those with CagA- *H. pylori*+ status (515), suggesting that the CagA protein plays a role in gastric cancer in adults. The CagA protein also activates NF-κB by interacting with the tyrosine 3-monoxygenase / tryptophan 5-monoxygenase activation protein, epsilon (YWHAE, a member of 14-3-3 family protein) (516). Interestingly, expression of CagA in
Drosophila intestinal stem cells has been shown to promote dysregulation of the gut microbiota, resulting in cellular proliferation in the insect’s gut epithelium (517).

Another way in which H. pylori interacts with the host is by injecting CagA into eukaryotic cells using its type IV secretory apparatus (518). Type IV competence pili are a conserved mechanism for horizontal gene transfer that mediates DNA internalization by binding to DNA at their tip and then retracting (519). The type IV secretion systems of Gram-negative bacteria mediate a range of biological functions, including exchanging genetic material with other bacteria and inserting oncogenic DNA and effector proteins into eukaryotic host cells (520).

Intriguingly, H. pylori and the plant pathogen Agrobacterium tumefaciens share some similarities, such as the ability to integrate their DNA into eukaryotic host cells by the type IV system. Agrobacterium tumefaciens infection causes cancer-like malformations in plants called crown galls (521). In the case of H. pylori, this DNA integration is followed by its recognition by the Toll-like receptor 9(522). It would be interesting to determine whether in the context of H. pylori infection of human cells, the type IV secretion system mediates sharing of DNA through such a CagA protein-mediated mechanism. Moreover, it would be interesting to determine if any of the integrated genes play a functional role in gastric cancer (523).

Although the importance of horizontal gene transfer in eukaryotes remains controversial, genomic studies have begun to reveal numerous cases of bacterium-to-eukaryotic gene transfers that may have influenced adaptive evolution in eukaryotes (521). Donor bacterial species and recipient eukaryotic hosts appear to be more numerous than previously believed (521). Future interdisciplinary studies comparing H. pylori infection to that of Agrobacterium tumefaciens and other bacteria capable of integrating DNA into eukaryotic hosts may unravel intriguing mechanism shared by human cancers and other cell-proliferative phenomena such as plant crown galls.

In the evolution H. pylori’s diverse effects on human health, the CagA protein is Janus-faced: on one hand, it promotes gastric cancer and peptic ulcers, and on the other it protects against gastro-esophageal reflux and Barrett’s esophagus, a precancerous condition in the esophagus (for a review of other genes with such contradictory effects), see our review: (514). Both the harmful and beneficial effects are theorized to trace back to the Copper Age (524): indeed, the H. pylori is considered a microbial marker of human migration (525). A better understanding of the mechanisms by which CagA exerts a protective effect in one tissue and a
virulent effect in another could suggest strategies to enable the exquisite control of CagA expression. Such control could benefit patients by enabling therapeutics that balance both its beneficial and harmful activities, encapsulating the “double-edged sword” relationship of *H. pylori* with human health and disease. On the other hand, little is known about the extent to which CagA benefits the survival of the bacterium, even though it may come at the cost of the host’s survival. In this vein, Klymiuk et al (507) have suggested that CagA expression influences *H. pylori* proliferation within the stomach. Hence, in light of CagA’s importance for *H. pylori* pathogenicity, elucidating the survival advantages that CagA gene offers to *H. pylori* strains may be informative.

Recent studies have shown that in addition to CagA-mediated pathogenic mechanisms, *H. pylori* also recruits the host’s cellular machinery to circumvent the host’s defense and effectively colonize the gastric mucosa. For example, although *H. pylori* lipopolysaccharides (LPS) bear similarities to those of most bacteria, the structural and chemical properties of *H. pylori* LPS appears to be tolerogenic, as it allows its recognition by the host via the TLR2 receptor system, in contrast to the more common TLR4 system recruited by most bacteria (526). This property may reduce the immunogenicity of *H. pylori* since the TLR2 pathway mediates a relatively weak host innate immune response and acquired T-cell mediated immunity in the gastric mucosa. Moreover, under specific circumstances, *H. pylori* LPS antagonistically interferes with TLR4-mediated signaling, thus further ameliorating the host’s innate response toward the bacterium (527).

1.5 Therapeutic perspectives: from bench to bedside

Since an *H. pylori*-positive status seems protective against diseases such as asthma and gastro-esophageal reflux, clinicians may eventually use genomic or transcriptomic information from *H. pylori* clinical isolates to inform clinical actions (528). For example, therapeutic strategies could be designed to eliminate the expression of cancer-implicated *H. pylori* genes while maintaining expression of those found to be beneficial to extra-gastric tissues, an approach we propose calling “precision onco-microbiology”, in alignment with precision medicine approaches (107). Genotypic differences also need to be considered, given that *H. pylori* strains have been shown to exhibit significant genetic variation stemming from mutations and recombination events (485).

Given advances in technologies for assessing cellular DNA damage, we contend that DNA damage should be monitored in patients with *H. pylori* infection, in parallel with determining their inflammatory status (529, 530). At least theoretically, it may be possible to
simultaneously prevent gastric cancer and eradicate \textit{H. pylori} by targeting similar molecular, cellular, or genomic mechanisms. The topoisomerase enzyme is a potential example. On the one hand, bacterial topoisomerase is targeted by quinolones, which are used for \textit{H. pylori} eradication in regions with high levels of clarithromycin resistance such as Europe (531) as well as in resource-poorer settings (532). On the other hand, eukaryotic topoisomerase inhibitors are well-established anti-cancer drugs that may contribute to lower genomic instability (533).

The extended collection of gastric cancer organoids that is becoming available provides an opportunity to screen large numbers of compounds for anti-cancer effects, as well as to glean initial clues for what drugs could most effectively be repurposed for applications in gastric cancer. Promising compounds may include cyclin-dependent kinase 4 and 6 inhibitors used in breast cancer treatment, cancer cell stemness inhibitors such as Napabucasin, and ATR (ataxia telangiectasia and Rad3-related) inhibitors (534).

1.6 Conclusions

Since its first description by Warren and Marshall, \textit{H. pylori} has inspired intensive investigations to better understand its interactions with the human host at the cellular and molecular level and in the context of its pathogenicity. Its ability to promote gastric cancer is especially intriguing and warrants further investigation to understand the underlying mechanisms and why some tissues are more susceptible to \textit{H. pylori}-mediated carcinogenesis than others. Answering these questions may not only inform strategies for gastric cancer treatment and prevents, but thinking of \textit{H. pylori} as a more generalizable “role model” for microbially-induced carcinogenesis may lead to breakthroughs in the study of other forms of cancer in which microbes may play unrecognized but critical roles.
1.7 FIGURES

**Figure 1. The development of gastric cancer as induced by *H. pylori* and other factors.**

Gastric cancer stem cells arise from epithelial cells in the stomach lining, as shown in Inset A in stomach (top left) and in expanded view (top right). Within the epithelial cell lining of the stomach (Inset C, top right), multiple factors, including *H. pylori* infection, can induce the chromosomal damage and genetic mutations that give rise to cancer stem cells (bottom left and middle). The stem cells’ unique properties (bottom middle) mediate cell proliferation, immune escape, and vascularization (bottom right), all of which promote tumorigenesis.
2. *Helicobacter pylori* Infection and Metachronous Gastric Cancer: A Critical Analysis

Helicobacter pylori eradication for metachronous gastric cancer: An unsuitable methodology impeding broader clinical usage

2.1 Introduction

*Helicobacter pylori* infection is a major health concern worldwide, especially in many resource-poor countries, particularly in Africa and Latin America/Caribbean, such that more than half of the global population was infected with the pathogen *H. pylori* in 2015 (453). Gastric cancer (for which stomach adenocarcinoma accounts for around 90% of cases) is a life-threatening disease, which may be prevented by pharmacological approaches such as aspirin and non-pharmacological approaches such as gastric endoscopy (453). Basic and clinical studies have demonstrated strong associations between oncogenesis and the presence of *H. pylori* bacteria in the stomach; this includes the progression of pre-cancerous lesions (535). Remarkably, the proportion of non-cardia gastric cancer attributable to *H. pylori* increased from 74.7% to 89.0% from 2008 to 2104 (536). Furthermore, other epidemiologic factors, such as metabolic syndrome, are increasingly implicated in the etiology of gastric cancer (537). Importantly, *H. pylori* infection has also been linked to non-gastric diseases, including Parkinson’s disease(50). The eradication of *H. pylori* using antibiotic therapy may prevent gastric cancer; such treatment has been implemented with varying levels of success globally (538).

2.2 Eradication of *Helicobacter pylori* infection and risk of metachronous gastric cancer

The timing of interventions is often considered a key factor in determining whether cancer therapy is successful or not and whether *H. pylori* eradication is beneficial. A recent review of clinical studies revealed that *H. pylori* eradication is associated with a significantly lower risk of gastric cancer, particularly in patients with atrophic and non-atrophic gastritis, rather than in those with intestinal metaplasia; however, maximal benefit is obtained when eradication is performed during the early stages of infection (539). While this might be challenging because the infection is not typically targeted in childhood, a recent review of clinical studies confirmed that there is a general belief among healthcare practitioners that *H. pylori* eradication can prevent gastric cancer when it is administered in pre-cancerous or early cancerous stages (i.e., before a “point of no return”) (540).
In a remarkable and highly visible clinical study, Choi, et al. (478) reported significant reductions in the incidence of metachronous gastric cancer after *H. pylori* eradication therapy in patients with previously resected early gastric cancer, indicating that *H. pylori* infections can benefit from treatment at any stage, thus refuting the conventional concept of the “point of no return.” (Venerito, et al., 2018). This exciting finding seems important in convincing physicians, patients, and stakeholders, in favor of preventive *H. pylori* eradication, who might be otherwise skeptical of such measures; moreover, it generally aligns with the findings of similar recently published studies in the literature (541).

### 2.3 Methodological remarks in the recent clinical trial

Previous critiques of this landmark study were focused on its scientific aspects (542, 543). Unfortunately, we have identified several methodological problems in the study, which may impact the generalizability of the results and conclusion, regardless of whether the study is robust and/or can be replicated. Hence, there is a need for further evidence (or more rigorous clinical trials) regarding the promising role of *H. pylori* eradication in the prevention of metachronous gastric cancer.

More precisely, the study by Choi et al. (478) was a clinical trial in which approximately 10% of the patients developed gastric cancer, and a statistically significant difference was noted between the treated and untreated groups (P = 0.03). The authors reported a highly significant (P < 0.001) change in the atrophy grade within the corpus lesser curvature, thereby fulfilling their primary objective.

A consistent limitation of clinical studies is the inability to replicate results (frequently known as the “reproducibility crisis”); this often occurs due to low statistical power and a tendency to overinterpret statistically significant results. The researchers (478) did not report whether multiplicity corrections were used, although such statistical analyses are increasingly used in leading scientific journals (544). Combined with the reports of individual patient data, despite opposing opinions (545), we suspect that this could have helped readers to evaluate whether there is a causal association between *H. pylori* eradication and metachronous cancer reduction more accurately. A recent study demonstrated that clinical study participants are typically amenable to sharing of their individual patient data (546); the provision of such additional data would help promote detailed meta-analyses and evaluate the robustness of important results.
The corpus lesser curvature, which showed significantly less atrophy in patients who underwent *H. pylori* eradication therapy, is one of many regions where stomach adenocarcinomas exist. In the Japanese Gastric Cancer Association classification system, the corpus lesser curvature comprises 3 of the 12 possible lymph node stations; together with the corpus upper curvature, it is considered a part of the N1 region (547). Cancer reduction solely in the corpus lesser curvature will not necessarily result in fundamental changes with respect to TNM staging. According to Wu et al. (548), approximately 46% of stomach carcinomas diagnosed in the USA are located in stomach non-cardia regions; these encompass corpus lesser curvature, as well as the fundus, body, antrum, and corpus greater curvature. Non-cardia carcinoma is epidemiologically distinct from other gastric corpus cancers across different populations (498). Therefore, it is particularly notable that the authors (478) limited their analysis solely to the atrophic changes of the lesser curvature, and strong caution is advised before generalizing anti-cancer effects discerned in one form of gastric cancer from a specific population to other forms of gastric cancer and across populations. Furthermore, the authors concluded that there was a reduction in the incidence of metachronous gastric cancer and greater improvement in the grade of gastric glandular atrophy among patients who underwent *H. pylori* treatment than among patients who received placebo treatment, and this conclusion is consistent with their stated aims. However, to achieve more precise conclusions, Choi, et al. (478) should have specified anatomical limitations in their conclusion, i.e., instead of the broader term “corpus,” they could have used the term “lesser curvature”, which has a stricter definition.

A particularly surprising aspect of the study by Choi et al. (478) was that only a single study pathologist performed diagnosis and biopsy evaluations, which is in direct contrast with recent trends in the cancer arena (549). To reduce potential bias in the analysis of results, especially in cases where a global conclusion is made based on clinicopathological examinations, a robust inter-rater reliability between different independent (“blinded”) pathologists should have been reported, preferably combined with parallel reporting of their level and area of expertise; moreover, Cohen’s kappa coefficient might have been used as a statistic to measure inter-rater agreement (550). Further caution may be appropriate because this study concurrently considered the Vienna 4.2 [“non-invasive carcinoma (carcinoma *in situ*)”] and 4.3 (“suspicious for invasive carcinoma”) diagnostic categories. The latter evaluates the suspicion of cancer identification, and due to potential misclassification risks, they are not considered in this classification system typically (551). It is though important to acknowledge that the authors limited their primary outcome variable to gastric adenocarcinomas alone, thus facilitating comparisons with other studies. Diagnostic cultural differences are also well known. For example, a Japanese pathologist might classify a carcinoma based on the presence of
notable cytological alterations (carcinoma *in situ*), whereas an American pathologist might interpret this as high-grade dysplasia because invasions are absent.

Additionally, the authors of this study (478) should have provided a more comprehensive literature background for the 1-year clinical cutoff that they used to define metachronous cancer to provide a sense of comparability, regardless of whether all the results are recorded at 5 years; notable examples include studies by Nakajima, et al. (552), Park, et al. (553), Abe, et al. (554), and Boda, et al. (555). These contrast with the more commonly used cutoff of 6 months (Moertel definition) (556). More broadly, future clinical research would have been benefited if standardized criteria were used [akin to those in the medical terminology, as discussed by Mentis & Papavassiliou (171)]; this is particularly applicable for the design of large-scale clinical trials. Indeed, it is clearly not appropriate to compare studies with different criteria; this poses a problem when aggregating data from different studies, for example in meta-analyses. Therefore, it impacts the ability to translate research findings into clinical practice.

### 2.4 Additional patient and pathogen factors that should have been considered

Gastric cancer prevalence is lower in the Western hemisphere than in the Eastern hemisphere (478). An important factor associated with risk might be the *H. pylori* genotype. The CagA+, VacA s1, and VacA m1 *H. pylori* strains are associated with an increased risk of gastric cancer (515). In Asia, specific CagA polymorphisms exist; these trigger different biological mechanisms than those associated with polymorphisms found in other parts of the world (557). However, Western strains often intermix with East Asian strains; this mixing has dramatic impacts on individual disease outcome (558). Genetic screening of *H. pylori* would have been particularly useful in the study by Choi et al. (478) to help further identify individual patients who benefited from *H. pylori* eradication therapy, thus bridging precision medicine and public health (107). Interestingly, the association between CagA antibodies and gastric cancer development has been established for more than two decades (559). In parallel, any effects related to patient profiles, notably proinflammatory genetic makeup [reviewed in (560)], are largely absent in the causal analysis of the clinical trial. Collectively, these data would have supported the evaluation of the relative contributions of patient and pathogen factors to the findings reported by Choi et al. (478).

### 2.5 Conclusion

The hypothesis tested by Choi et al. (478) is critical for improved treatment; moreover, the importance of increasing knowledge regarding *H. pylori* eradication as preventive therapy for metachronous cancer can be cost-effective. Data have been generated for meta-analyses;
however, the results cannot be generalized in their current state. Well-powered trials across different populations using the latest screening and biomarker tools available to profile individual cancer cases are needed to determine the proportion of the global population for whom *H. pylori* eradication therapy may be beneficial and cost-effective. From a clinical perspective, clinicians must also consider the risk of second primary malignancies in other body parts of patients with gastric cancer (561).
3. *Helicobacter pylori* infection in patients with Parkinson’s disease: a case-control study and meta-analysis

3.1 Introduction

3.1.1 Parkinson’s disease

Parkinson’s disease (PD), first described in 1817 by James Parkinson’s as “shaking palsy”, is the second commonest neurodegenerative disease after Alzheimer disease (562, 563). PD affects 1-2 per 1000 of the population at any time, and its prevalence increases with age to affect 1% of the population above 60 years (103, 564). PD is a chronic and progressive disorder clinically characterized by motor symptoms including resting tremor, bradykinesia, rigidity, postural instability, and walking difficulties (565-567) From a neuropathological perspective, PD is hallmarked by progressive dopaminergic neuron degeneration and loss in the midbrain substantia nigra compacta, Lewy body formation, and abnormal α-synuclein accumulation (568, 569) Genetic, environmental, and neuronal aging factors appear to contribute to an increased risk of PD (563, 570, 571) but greater understanding of such risk factors would be clinically useful for treatment and prevention.(103)

3.1.2 *Helicobacter pylori*

*Helicobacter pylori* (*H. pylori*) infection is one of the commonest chronic infections worldwide. It has a prevalence approaching 50% in middle-aged adults even in developed countries, although decreasing trends in prevalence have been observed over the last century (572, 573). *H. pylori* most often causes gastritis and peptic ulcer disease but is also linked to extra-gastrointestinal disorders (572, 574, 575), namely cardio-cerebro-vascular disease; metabolic disorders (576); neurodegenerative disorders including Alzheimer disease (577), vascular dementia (578), glaucoma (described as “ocular Alzheimer disease”), (579, 580), PD (575, 581) and multiple sclerosis (582); and several other systemic disorders (575). Reduction in *H. pylori* prevalence at the population level by eradication might, therefore, inhibit the development and delay the progression of many diseases.

3.1.3 *H. pylori* and Parkinson’s disease

Even before the landmark discovery of *H. pylori* in the 1980s (583, 584), an excess of peptic ulcers was observed in patients with PD, with peptic ulcers predating the diagnosis of PD by approximately 20 years (585). Altschuler (586) first hypothesized that *H. pylori* and PD were linked, while more recently it has been proposed that neuronal damage in PD may be a response to chronic *H. pylori* infection (574). Over the last two decades, several studies have suggested
an increased prevalence of *H. pylori* infection in PD patients (587-591). Furthermore, a large population study suggested that chronic *H. pylori* infection may contribute to the pathogenesis of PD (592). Likewise, other studies have suggested that motor performance in PD patients is affected by the presence of *H. pylori* infection (588, 593).

As the duodenum is the main site for absorption of levodopa, the cornerstone of PD therapy, it is hypothesized that duodenal *H. pylori* infection might affect levodopa bioavailability by inducing inflammation, disrupting the duodenal mucosa (594), and releasing reactive oxygen species (595, 596) that reduce levodopa absorption (588, 597, 598) or inactivate the drug (590, 599). However, data are conflicting regarding the extent to, and the way in which, *H. pylori* eradication influences motor symptoms and affects levodopa absorption (588, 597, 600, 601).

Despite this potential link between *H. pylori* infection and PD, detection and eradication of *H. pylori* is not currently advocated in the management of PD. Although there appears to be some evidence of a relationship between *H. pylori* infection and PD, data are conflicting and there is significant heterogeneity in study design. A recent meta-analysis (602) examined the effects of *H. pylori* infection on PD but omitted an important study and did not examine the relationship between *H. pylori* infection and the clinical severity of PD. Hence, the aim of this case-control study and meta-analysis was to assess the prevalence of *H. pylori* infection in PD patients and to shed light on the relationship between the infection and PD clinical status.

### 3.2 Methods

#### 3.2.1 Case-control association study

*Patients and controls.* Thirty-nine PD patients were recruited from the outpatient clinics of the Department of Neurology, University General Hospital of Larissa, University of Thessaly, Central Greece. PD diagnoses were assigned using published criteria after examination by senior neurologists. Detailed demographic data are shown in Table 1. Serum samples were also obtained from 68 healthy controls (HCs). HCs were selected from a number of women and their partners attending the Gynecological and Obstetrics Department outpatients’ clinic of the same hospital. Eligibility criteria for HCs were no family history of PD and no neurological or other autoimmune disorders. HCs were also free of other major comorbidities such as cancer, chronic cardiovascular disease, diabetes, depression, or
hypertension. Both patients and HCs were white Caucasians. All patients had previously provided informed written consent.

**Laboratory methods.** Blood samples were obtained by standard venipuncture. Serum was separated from blood samples and kept in aliquots at −80°C until used. Samples were tested for IgG anti-HP antibodies using enzyme-linked immunosorbent assay (ELISA) (Euroimmun AG, Lübeck, Germany), according to the manufacturer’s instructions. Values of >20 relative units per ml (RU/ml) were considered positive.

**Statistical analysis.** Differences between groups were estimated with chi-square tests for categorical variables. Statistical analysis was carried out in SPSS version 17.0 for Windows (SPSS Inc., Chicago, Illinois, USA). Statistical significance was set at p≤0.05.

### 3.2.2 Meta-analysis

Search strategy. Eligible case-control studies and randomized clinical trials (RCTs) were selected by searching the PubMed database using combinations of the following terms: “Helicobacter pylori” and “Parkinson’s disease”. The complete search algorithm is available in [Supplementary file 1](#). The final literature search was performed on June 16th, 2017. The references in all the retrieved publications were also reviewed to identify any studies missed in our initial search.

### 3.2.2.1 Data extraction.

Eligible articles were first sorted by screening titles and abstracts and assessing the full text of eligible studies. Data extracted directly from the articles were: study design, *H. pylori* infection detection method, clinical assessment scale, number of participants, number of infected patients and controls, sex ratio, and mean age. The values of the aforementioned data were presented as in the original article. If those data were not provided, they were calculated using related study data. Moreover, the number of participants and the means (and standard deviations [SD]) of unified PD rating scale (UPDRS) scores were extracted from the RCTs. Two reviewers (E.D. and Z.T.) independently conducted data extraction and evaluated the methodological quality.

### 3.2.2.2 Inclusion criteria

*H. pylori infection in PD patients vs. healthy controls.* This was the first part of the meta-analysis. The studies selected for meta-analysis met the following criteria: a) case-control studies including PD patients (i.e., cases) and healthy controls (i.e., controls); b) both PD patients and healthy controls tested for *H. pylori* infection; c) studies reporting either cardinal
number or percentage of subjects regarding *H. pylori* infection for both groups; d) English language; e) studies conducted in humans. Studies with incomplete or not mentioned original data were excluded from the meta-analysis.

**Mean UPDRS score between *H. pylori* infected and non-infected PD patients.** Studies that were eligible for the second part of the meta-analysis fulfilled the following criteria: a) case-control studies including *H. pylori* infected PD patients (i.e., cases) and *H. pylori* non-infected PD patients (i.e., controls); b) all PD patients tested for *H. pylori* infection; c) mean UPDRS score (and SD) provided separately for both groups of patients; d) English language; e) studies conducted in humans. Studies with incomplete or not reported data were excluded from the meta-analysis.

### 3.2.3 Statistical analysis

**H. pylori infection in PD patients and healthy controls.** The association between *H. pylori* infection and PD was determined by calculating the 95% confidence interval (95% CI) and the pooled odds ratio (OR). The Z test was used to determine the statistical significance of the OR, and significance was set at p<0.05.

**Mean UPDRS scores between *H. pylori* infected and non-infected PD patients.** To examine the association between *H. pylori* infection and mean UPDRS scores in PD patients, a second meta-analysis was performed. The association was determined by calculating the standardized mean difference (SMD), and the 95% CI was estimated by using the Z test. Statistical significance was set at p<0.05.

**Tests of heterogeneity.** Cochran’s Q and I² indices were calculated to estimate the statistical heterogeneity of the studies. In cases of substantial heterogeneity (P<0.10 or I²>75%), a random effects model was applied. Otherwise, the fixed effects model was used. Funnel plots were used to assess for possible publication bias. All statistical analyses were performed in Review Manager (RevMan) Version 5.3 software (The Nordic Cochrane Centre, The Cochrane Collaboration, Copenhagen, Denmark [http://tech.cochrane.org/revman]). PRISMA guidelines for reporting reviews and meta-analyses were applied.(605).

### 3.2.4 Results

#### 3.2.4.1 Case-control association study

Thirty-nine PD patients (51.3% females) with a mean age of 68.7±9.4 years and 68 HCs (58.8% females) with a mean age of 47.4±16.9 years were included. Fourteen (35.9%) PD patients were
anti-HP seropositive, while anti-HP antibodies were present in 33 (48.5%) of HCs. However, there was no statistically significant difference in the frequency of *H. pylori* infection between PD patients and HCs (p=0.21).

### 3.2.4.2 Meta-analysis

**Selection and characteristics of the included studies.** The initial PubMed database search yielded 64 studies published between November 1996 and June 2017. After reviewing the titles and abstracts, reviews, letters, case reports, and articles not eligible for the meta-analysis were excluded, resulting in 24 potentially eligible studies. Three were studies with no case-control analysis (588, 597, 606) and were therefore excluded from the meta-analysis. Furthermore: a) a case-control study written in Polish was excluded (607); b) two studies without estimates of the mean UPDRS score were excluded (566, 606); and c) three studies were excluded, as they did not provide *H. pylori* seropositivity results and mean UPDRS scores for *H. pylori* positive and negative PD patients (608-610). In one protocol, the mean (and SD) UPDRS scores of *H. pylori* positive and negative groups were calculated from the pooled mean and SD of the two positive and two negative groups, respectively (Supplementary file 2). Finally, fifteen studies were included in the quantitative meta-analysis (579, 581, 587, 590-593, 600, 601, 611-616) involving in total 5043 PD patients and 23449 healthy controls in the first part, and 170 *H. pylori*-infected and 283 non-infected PD patients in the second part. The selection algorithm of eligible studies is shown in Figure 1.

*H. pylori infection in PD patients and healthy controls.* The studies used in the first part of meta-analysis were published between 1999 and 2017 (587, 591, 611, 613, 614), and their main characteristics are presented in Table 2. All ten were case-control studies. With respect to *H. pylori* detection, serology by enzyme-linked immunosorbent assay (ELISA) was used in our present study and in four other studies (587, 611, 614, 615), while in the other two studies 13C-urea breath testing (UBT) was performed (591, 613). From the remaining protocols, one was a population-based study and *H. pylori* infection status was determined by medical prescription records (592), one used gastroscopy and histological examination (579), and one determined positivity by detecting *H. pylori*-specific DNA by RT-PCR (616). The cases in all ten studies were diagnosed with PD, and all the subjects used as controls were healthy. The case and control groups were age- and sex-matched in five studies (587, 591, 592, 613, 614) and only age-matched in one study (611).

*Mean UPDRS score between *H. pylori* infected and non-infected PD patients.* The studies used in the second part of the meta-analysis were published between 2008 and 2017 (581, 590, 591, 593, 600, 601, 612, 613). Their main characteristics are shown in Table 3. In
all eight studies, case-control analysis was conducted between *H. pylori* infected and non-infected PD patients. Regarding *H. pylori* detection, UBT was used in six studies (581, 590, 591, 593, 612, 613), while the antigen stool test was used in one study (600) and ELISA in another (601). Several parts of UPDRS were used for clinical assessment of the patients. Specifically, in six studies (581, 590, 591, 600, 601, 613), only section III of the UPDRS was evaluated, and in the other two studies (593, 612) the total UPDRS score was used to assess clinical status. All the studies provided mean UPDRS scores with SD. Furthermore, seven of the eight studies presented mean age and sex ratios for both *H. pylori* infected and non-infected patients (581, 590, 600, 601, 612, 613) while one did not (591).

### 3.2.4.2.1 Subgroup analysis

*H. pylori infection in PD patients and healthy controls.* The fixed-effects model was applied to the first group of studies due to low heterogeneity ($I^2=31\%$, $P=0.16$). *H. pylori* infection was significantly more prevalent among PD patients than among controls [OR (95% CI): 1.47 (1.27, 1.70); $P<0.00001$]. The results of the meta-analysis for the prevalence of *H. pylori* infection in PD patients are presented in Figure 2. The funnel plot presented in Figure 3 did not reveal any significant asymmetry.

*Mean UPDRS scores between *H. pylori* infected and non-infected PD patients.* The random-effects model was applied to the second group of studies due to high heterogeneity ($I^2=66\%$, $P=0.004$). There was no significant association between *H. pylori* infection and higher mean UPDRS scores [standardized mean difference, SMD (95% CI): 0.15 (-0.20, 0.50), $P=0.39$]. The results of the meta-analysis for the mean UPDRS score between *H. pylori* infected and non-infected PD patients are presented in Figure 4. The funnel plot presented in Figure 5 revealed significant asymmetry.

### 3.2.5 Discussion

Here we present the results of a new case-control study and a meta-analysis examining the association between *H. pylori* infection status and PD. Our case-control study revealed no significant differences in prevalence of anti-*H. pylori* seropositivity between cases and controls, consistent with the variability seen in previous studies and probably due to sample size and other variabilities such as sex, race, or age of patients and controls. However, and highlighting the value of meta-analysis, we revealed a higher prevalence of *H. pylori* infection in PD patients than in healthy controls. *H. pylori* infection was determined by serology or urea breath testing. The second part of this meta-analysis revealed no significant association between *H. pylori* infection and mean UPDRS score. *H. pylori* infection was determined by urea breath testing or by stool antigen testing.
It is important to note that the H. pylori detection method used represents a significant potential confounder and a key element of heterogeneity between studies. Both Nafisah et al. (591) and (613) evaluated H. pylori infection using UBT, which has higher sensitivity and specificity than ELISA. Consequently, both studies suggested that the use of serology may overestimate the prevalence of infection. Also, ELISA could remain positive months after potential H. pylori spontaneous elimination (573). In addition, even though all studies indicated higher H. pylori infection prevalence in PD patients, only two were statistically significant (587, 614).

Charlett et al. (587) reported that H. pylori seropositivity was two-fold higher in patients with idiopathic parkinsonism than disease-free controls (587). It was reported as early as 1965 that peptic ulcers were more common in PD patients (585), and this relationship was also hypothesized by Altschuler in 1996 (586). Furthermore, several studies have revealed a high prevalence of H. pylori infection in PD patients. Dobbs et al. (606) found that the UBT was positive in 48% of patients, and Lee et al.,(590) Dobbs et al.,46 and Tan et al. (581) obtained positive H. pylori results in 53%, 70%, and 32% of PD patients, respectively. Specifically, Dobbs et al. 51 reported that the prevalence of H. pylori seropositivity was higher in parkinsonism before the age of 72.5 years. Host susceptibility, alone or combined with specific H. pylori strain(s) were suggested to have been responsible for the higher seropositivity in PD patients before the eight decade of life. PD patients were also three times more likely to be H. pylori-seropositive than controls, and the patients’ siblings were also more likely to be both seropositive and show signs of parkinsonism (587), potentially due to familial transmission (573, 617).

In a population-based case-control study, Nielsen et al. (592) showed that prescription of H. pylori eradication therapy and proton pump inhibitors five or more years prior to the diagnosis of PD were associated with a 45% and 23% increase in PD risk, respectively. They concluded that chronic H. pylori infection and/or gastritis contribute to PD pathogenesis or that these are PD-related pathologies that precede motor symptoms (592).

Even though most studies suggest that the prevalence of H. pylori infection is higher in PD than in disease-free status, considerable controversy remains on whether the clinical status of H. pylori-infected and non-infected PD patients differs. In this meta-analysis, the UPDRS score, which reflects disease motor severity (618, 619), tended towards being higher in infected PD patients, although this did not reach statistical significance. Five relevant studies (581, 590,
suggested that *H. pylori*-infected patients tended to have increased motor fluctuations and worse mean UPDRS scores. Although Hashim et al. (593) reported a significantly better motor performance and mean total UPDRS score in *H. pylori* non-infected than infected patients, no other study was significant. On the contrary, Nafisah et al. (591) reported no association between *H. pylori* infection and mean UPDRS score, and Rahne et al. (612) reported that infected patients had less motor fluctuation and better mean UPDRS IV scores (p=0.005). In addition, Fasano et al. (613) suggested synergy between *H. pylori* infection and small intestinal bacterial overgrowth (SIBO) that led to increased motor fluctuations in positive PD patients, underscoring hence the gut microbiome’s role in neurology (620).

Nevertheless, *H. pylori* eradication has been shown to ameliorate symptoms of PD (609). *H. pylori* eradication has been suggested in PD patients treated with levodopa, as it may improve drug bioavailability and decrease motor fluctuations (590, 593, 612, 613). Interestingly, there is greater mortality from PD in livestock farmers infected with *H. suis*, the most common zoonotic *Helicobacter* in man (616). This observation is reinforced by the finding that a significantly higher frequency of *H. suis* has been observed in patients with idiopathic parkinsonism than in controls (616). Remarkably, the higher frequency of *H. suis* in idiopathic parkinsonism appears to be exaggerated following *H. pylori* eradication (616).

Regarding *H. pylori*’s potential involvement in PD pathophysiology, systemic inflammation might be important. There is evidence that interleukin-1β (IL-1β), IL-6, tumor necrosis factor-α (TNF-α), other proinflammatory cytokines, and C-reactive protein play important roles in neurodegenerative disorders including PD (621, 622). Consistent with this association, Bu et al. (614) recently showed that *H. pylori*, cytomegalovirus, herpes simplex virus type 1, *Chlamydia pneumoniae*, and *Borrelia burgdorferi* seropositivities (together defined as “infectious burden”) were associated with PD. Moreover, a higher infectious burden was associated with more severe motor symptoms and higher serum levels of proinflammatory cytokines and α-synuclein (614).

*H. pylori* induces the release of large amounts of IL-1β, -6, -8, -10, -13 and TNF-α, eicosanoids, and acute phase proteins and activates monocytes, which may result in disruption of the blood brain barrier (BBB), microglial activation, and deleterious effects on the nigrostriatal dopaminergic system (518, 623, 624). In this respect, neuroinflammation seems to contribute to neurodegeneration with activation of microglia (625). Being the brain resident immune cells, microglia may respond to a proinflammatory stimulus and release neurotoxic substances, thereby contributing to the neurodegeneration observed in PD patients (624, 626). Apoptosis, rather than necrotic microglia-associated nerve cell death, appears to underlie a
number of common neurological conditions including PD associated with *H. pylori* infection. Indeed, besides the Fas-FasL apoptotic pathway, *H. pylori* can induce apoptosis through the mitochondrial apoptotic pathway or through inducible nitric oxide, leading to neurodegeneration (627).

Peripheral inflammation due to *H. pylori* infection can activate microglia via the humoral pathway when circulating proinflammatory cytokines or leukocytes cross the disrupted BBB (624, 628, 629). Alternatively, noxious chemicals produced by *H. pylori* can be transmitted through the vagal axonal afferent pathway and affect brainstem neurons, as indicated by animal and human studies (624, 630, 631). Conversely, α-synuclein has been recently reported to reach the gastric wall via preganglionic vagal route, suggesting brain-to-stomach interactions with implications in PD (632). Moreover, *H. pylori* might access the brain via the oral-nasal-olfactory pathway or via circulating monocytes (infected with *H. pylori* due to defective autophagy) through the disrupted BBB, leading to neurodegeneration (633, 634).

Furthermore, auto-antibodies against dopaminergic neurons found in the cerebrospinal fluid and/or the blood of PD patients (635) might represent molecular mimicry from *H. pylori* infection (636). Molecular mimicry between *H. pylori* and host antigens might trigger auto-antibody production, consequentially leading to a chronic inflammatory state, which contributes to the pathogenic process in both gastric and systemic diseases including PD (566). In this respect, persistence of serum *H. pylori* cytotoxin CagA antibodies appears to be predictive of PD and is associated with a poor prognosis (610); however, it is difficult to attribute this association exclusively to a specific *H. pylori*-related virulence factor (e.g., CagA) given their large number (637).

Therefore, our outcomes from the first part of the meta-analysis are consistent with many studies indicating a higher prevalence of *H. pylori* infection in PD patients than in healthy controls. In contrast with the majority of included studies (591, 611, 613), our results were highly significant (p<0.00001). Regarding the second part of the meta-analysis, our results were consistent with most of the included studies (581, 590, 600, 613), showing a trend (albeit insignificant) towards an association between *H. pylori* infection and the clinical status (i.e., mean UPDRS score) of PD patients. Therefore, further large-scale studies are warranted to further establish the latter relationship.

This meta-analysis has certain limitations. First, we included studies regardless of their *H. pylori* detection method. While both ELISA and UBT reflect the infection status of patients,
there are several differences between these methods (638), especially that serology tends to overestimate the prevalence of infection. Second, we observed significant heterogeneity between the scales used to assess the clinical status of PD patients. Although UPDRS was used in most of the studies, different sections of the scale were applied in different disease states. Homogeneous data could contribute to more accurate results. Also, it is important to note that, due to the low availability of data, we could not conduct a meta-analysis to evaluate the clinical status of PD patients before and after *H. pylori* eradication. Third, *H. pylori* status could rather be a marker of hygiene status, making the associations more correlational than causal (639).

### 3.3 Conclusions

Given the higher prevalence in PD patients, *H. pylori* infection may contribute to the pathogenesis and prognosis of PD. As a global, efficient, and inexpensive regimen, *H. pylori* eradication therapy may positively influence the clinical outcome of PD. However, further large-scale studies with unified clinical data are required to accurately estimate the impact of the infection and elucidate the detailed mechanisms by which *H. pylori* is involved in PD pathogenesis.
### 3.4 TABLES

**Table 1. Demographic data of PD patients and healthy controls included in case-control study**

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>PD (n=39)</th>
<th>HC (n=68)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex (F/M)</td>
<td>20 (51.3%)/19 (48.7%)</td>
<td>40 (58.8%)/28 (41.2%)</td>
<td>-</td>
</tr>
<tr>
<td>Age (mean±SD)</td>
<td>68.7 ± 9.4</td>
<td>47.4 ± 16.9</td>
<td>-</td>
</tr>
<tr>
<td>Anti-HP positive</td>
<td>14 (35.9%)</td>
<td>33 (48.5%)</td>
<td>0.21</td>
</tr>
</tbody>
</table>
Table 2. Characteristics of the studies included in the meta-analysis of *H. pylori* infection in PD patients and healthy controls.

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Study design</th>
<th>H. pylori detection</th>
<th>PD patients</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>H. pylori (+/-)</td>
<td>Mean age (F/M)</td>
<td>H. pylori (+/-)</td>
</tr>
<tr>
<td>Charlett et al.</td>
<td>1999</td>
<td>Case-Control</td>
<td>ELISA</td>
<td>23/10</td>
<td>75</td>
</tr>
<tr>
<td>Dobbs et al.</td>
<td>2000</td>
<td>Case-Control</td>
<td>ELISA</td>
<td>50/55</td>
<td>74</td>
</tr>
<tr>
<td>Charlett et al.</td>
<td>2009</td>
<td>Cross-sectional</td>
<td>ELISA</td>
<td>57/120</td>
<td>NM</td>
</tr>
<tr>
<td>Nafisah et al.</td>
<td>2013</td>
<td>Cross-sectional</td>
<td>Urea Breath Test</td>
<td>14/15</td>
<td>64.1</td>
</tr>
<tr>
<td>Fasano et al.</td>
<td>2013</td>
<td>Case-Control</td>
<td>Urea Breath Test</td>
<td>11/22</td>
<td>67.8</td>
</tr>
<tr>
<td>Blaecher et al.</td>
<td>2013</td>
<td>Case-Control</td>
<td>RT-PCR</td>
<td>17/60</td>
<td>62</td>
</tr>
<tr>
<td>Bu et al.</td>
<td>2015</td>
<td>Case-Control</td>
<td>ELISA</td>
<td>60/71</td>
<td>67</td>
</tr>
<tr>
<td>Tsolaki et al.</td>
<td>2015</td>
<td>Case-Control</td>
<td>Gastroscopy and histological examination</td>
<td>6/9</td>
<td>NM</td>
</tr>
<tr>
<td>Present study</td>
<td>2017</td>
<td>Case-Control</td>
<td>ELISA</td>
<td>14/39</td>
<td>68.7</td>
</tr>
</tbody>
</table>

ELISA, enzyme-linked immunosorbent assay; PD, Parkinson’s disease; NM, Not measured
3.5 FIGURES

Figure 1. The selection process of eligible studies.
Figure 2. Forest plot: meta-analysis of the prevalence of *H. pylori* infection in the PD patient group compared with the healthy control group.

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>PD Patients</th>
<th>Control</th>
<th>Odds Ratio M-H, Fixed, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Charlotte 1999</td>
<td>23 33</td>
<td>19 78</td>
<td>3.49 [1.46, 8.32]</td>
</tr>
<tr>
<td>Dodd 2000</td>
<td>50 105</td>
<td>83 210</td>
<td>1.39 [0.87, 2.23]</td>
</tr>
<tr>
<td>Charlotte 2009</td>
<td>57 120</td>
<td>77 196</td>
<td>1.40 [0.88, 2.21]</td>
</tr>
<tr>
<td>Nielsen 2012</td>
<td>128 4434</td>
<td>505 22416</td>
<td>1.32 [1.14, 1.67]</td>
</tr>
<tr>
<td>Nafissah 2013</td>
<td>14 29</td>
<td>5 23</td>
<td>3.36 [1.30, 9.14]</td>
</tr>
<tr>
<td>Fasano 2013</td>
<td>11 33</td>
<td>8 39</td>
<td>1.98 [0.68, 4.07]</td>
</tr>
<tr>
<td>Blaecher 2013</td>
<td>17 60</td>
<td>42 256</td>
<td>2.01 [1.03, 3.87]</td>
</tr>
<tr>
<td>Bu 2015</td>
<td>60 131</td>
<td>44 141</td>
<td>1.80 [1.12, 3.05]</td>
</tr>
<tr>
<td>Tsolaki 2015</td>
<td>6 9</td>
<td>14 31</td>
<td>2.43 [0.51, 11.51]</td>
</tr>
<tr>
<td>Present study</td>
<td>14 39</td>
<td>33 68</td>
<td>0.59 [0.26, 1.33]</td>
</tr>
</tbody>
</table>

**Total (95% CI)**: 5043 (23449) 100.0% 1.47 [1.27, 1.70]

Heterogeneity: Chi² = 13.08, df = 9 (p = 0.16); I² = 32%
Test for overall effect: Z = 5.28 (p < 0.00001)
Figure 3. Funnel plot to detect publication bias in the studies reporting *H. pylori* infection in PD patients and healthy controls.
Figure 4. Forest plot: standardized mean difference of mean UPDRS scores between *H. pylori* infected (Hp+) and non-infected (Hp-) PD patients.

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Hp+ Mean</th>
<th>SD</th>
<th>Total Mean</th>
<th>SD</th>
<th>Total</th>
<th>Weight</th>
<th>Std. Mean Difference IV, Random, 95% Cl</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lee 2008</td>
<td>22.9</td>
<td>7.4</td>
<td>35</td>
<td>23.1</td>
<td>7</td>
<td>30</td>
<td>14.0% -0.03 [-0.52, 0.46]</td>
</tr>
<tr>
<td>Nafisah 2013</td>
<td>19.43</td>
<td>8.65</td>
<td>14</td>
<td>19.2</td>
<td>11.66</td>
<td>15</td>
<td>10.5% 0.02 [-0.71, 0.75]</td>
</tr>
<tr>
<td>Rahne 2013</td>
<td>4.8</td>
<td>3.0</td>
<td>20</td>
<td>7.3</td>
<td>3.5</td>
<td>55</td>
<td>13.4% -0.73 [-1.26, -0.21]</td>
</tr>
<tr>
<td>Fasano 2013</td>
<td>40.61</td>
<td>7.98</td>
<td>11</td>
<td>13.53</td>
<td>12.8</td>
<td>19</td>
<td>10.5% 0.30 [-0.42, 1.02]</td>
</tr>
<tr>
<td>Naranzanska 2014</td>
<td>38.8</td>
<td>19.7</td>
<td>18</td>
<td>37.6</td>
<td>12.8</td>
<td>19</td>
<td>11.6% 0.07 [-0.57, 0.72]</td>
</tr>
<tr>
<td>Hashim 2014</td>
<td>85.05</td>
<td>25.48</td>
<td>21</td>
<td>63.98</td>
<td>29.17</td>
<td>55</td>
<td>13.5% 0.74 [0.22, 1.26]</td>
</tr>
<tr>
<td>Tan 2015</td>
<td>33.97</td>
<td>13.04</td>
<td>33</td>
<td>27.32</td>
<td>10.05</td>
<td>69</td>
<td>15.0% 0.59 [0.17, 1.02]</td>
</tr>
<tr>
<td>Mridula 2017</td>
<td>54.1</td>
<td>5.2</td>
<td>18</td>
<td>52.3</td>
<td>11</td>
<td>18</td>
<td>11.5% 0.20 [-0.45, 0.86]</td>
</tr>
</tbody>
</table>
| **Total (95% CI)** | **170** | **283**| **100.0%** | **0.15**| **[-0.20, 0.50]** | **-1** | **-0.5** | **0** | **0.5** | **Favours [Hp+]** | **Favours [Hp-]**

Heterogeneity: Tau² = 0.16; Chi² = 29.86, df = 7 (P = 0.004); I² = 66%

Test for overall effect: Z = 0.86 (P = 0.39)
Figure 5. Funnel plot to detect publication bias in the studies reporting mean UPDRS scores in *H. pylori* infected and non-infected PD patients.
Chapter 4 — The Contribution of Microbiome to Neurological Diseases: The case of CSF Microbiome and Virome in Multiple Sclerosis

1. Viruses and Multiple Sclerosis: From Mechanisms and Pathways to Translational Research Opportunities

1.1 Multiple Sclerosis: the clinical and historical context

Multiple sclerosis (MS) is a chronic disease of the central nervous system (CNS) and the commonest cause of disability in the young adult population (640). MS is an on-going and worsening clinical problem: the prevalence of MS has increased 10-fold over the last fifty years in some regions (641), and MS reduces life expectancy by up to seven years in high prevalence regions (642-644). Nearly 2.5 million individuals are affected by MS globally (645, 646), and the disease incurs an annual cost in Europe – where the majority of MS patients reside – exceeding 15 billion US dollars (640). There is still no cure for MS, with treatment relying on managing symptoms, accelerating recovery from relapse, and reducing the number and severity of relapses (647). There is, therefore, a need to develop strategies to identify at-risk individuals with the view to prevention and to develop therapies that permanently prevent relapse.

Patients with MS have variable clinical courses that follow one of three main patterns: relapsing remitting (RR; 85% of cases), primary progressive (PP) or secondary progressive (SP) (647), with sensory and motor disturbances, vision abnormalities, and cognitive impairment common clinical features (647). While central nervous system (CNS) oligodendrocytes (the myelin-forming cells in the brain) are thought to be the main target in MS by autoimmune CD4+ T helper (Th1) cells, it is increasingly apparent that CD8+ T cells are important effectors since they are commonly present in the periventricular demyelinating plaques in the brain, brain stem, optic nerve, and spinal cord (648, 649).

The earliest description of MS dates to the 14th century, but it was not until 1868 that Jean-Martin Charcot documented the correlation between the clinical and pathological changes of myelin loss in MS, which was from then onwards perceived as a distinct nosological entity (650). In spite of intense research on MS over the last 150 years, a principal etiology has yet to be determined. MS is multifactorial and thought to arise from a complex interaction between genetic (notably immunogenetic), autoimmune, and environmental factors (17, 651-653). Genetic associations include many immune-related loci including human leukocyte antigen (HLA) and non-HLA (e.g., interleukin (IL)-2 and IL-7) loci (654) and, together with the fact that the observed discordance in MS incidence in monozygotic twins cannot be
attributed to genomic, transcriptomic, or epigenomic factors alone (655), the environment is recognized as important in MS pathophysiology (656). About 90% of MS patients have detectable oligoclonal IgG bands in their cerebrospinal fluid (CSF) (657). Given that these antibodies are directed towards antigenic targets in the CSF and are usually only present in infectious CNS disorders, one or more infectious agents are implicated in MS (657). Although myelin is the main biochemical target in MS, the distribution of lesions in the brain is patchy, and many myelinated areas remain unaffected suggesting an etiology with specific tropism.

MS does not harbor pathognomonic histological features for a specific virus in the brain parenchyma in contrast to, for example, the “owl’s eye” appearance of inclusion bodies of cytomegalovirus (CMV) (658). Other infectious agents implicated in MS include various bacteria and protozoa (659). Nevertheless, there is now accumulated evidence from modern epidemiological, animal, and human studies and immunohistochemistry and molecular biology-based analyses supporting the role of viruses as a trigger for MS (652). Several studies now show a strong association between several viruses (especially double-stranded DNA viruses (according to the Baltimore classification) (660)) and MS, notably EBV, human herpesvirus-6 (HHV-6) (661), CMV, human endogenous retrovirus (HERV), the measles virus, influenza virus, and varicella zoster virus (VZV) (651, 652, 662). These associations can only be described as casual rather than causal in humans, and exactly which viruses trigger or cause MS and how this happens is still unresolved.

Here, we discuss various theories on the mechanisms and pathways underpinning the viral etiology of MS. The aim is to provide a thorough overview of relevant virus-related MS pathophysiology to highlight possible therapeutic and diagnostic opportunities and the work that still needs to be done to expedite clinical translation.

1.2 Virus-induced MS: from evolution to mechanistic insights from animal models

Two main hypotheses on the viral etiology of MS bear consideration: first, that there is a single causative agent, analogous to Treponema pallidum causing neurosyphilis, with MS representing either a rare manifestation of a common virus or a common manifestation of a rare virus in regions with similar prevalence. The competing hypothesis is that multiple agents disrupt the multiple pathophysiological pathways implicated in MS. A hallmark of MS is the destruction of myelin and axonal damage, so a general framework for understanding how viruses cause demyelination includes: (i) direct lysis of virus-infected oligodendrocytes by virus; (ii) direct lysis of virus-infected oligodendrocytes by the host immune response to the
virus; (iii) lysis of uninfected oligodendrocytes by a virus-triggered autoimmune response; and (iv) lysis of oligodendrocytes by a nonspecific immune response triggered by a virus(659). Details of the immunological mechanisms underlying these processes are detailed below.

From the experimental modeling perspective, it is helpful to note that myelin’s components are evolutionarily closer to those of jawed vertebrates (“gnathostomata”) than other phyla (663). Adaptive immunity appeared 450 million years ago in parallel with the separation of jawed and jawless (“agnatha”) vertebrates (664). Thus, models of MS across the entire gnathostomata family (i.e., primates, rodents, and even zebrafish ((665)) are likely to share clinically relevant features. Relevant to this discussion, some of these animal models rely on infection with viruses and include, in rodents, Theiler’s murine encephalomyelitis virus (TMEV)-induced demyelinating murine models, mouse Semliki Forest virus (SFV) infection, and infection with mouse hepatitis virus. These virus-induced models and other important animal models used in MS research are summarized in Table 1, and interested readers are referred to(666-669) for excellent and comprehensive reviews on the use of animal models in MS.

The TMEV model probably represents the most useful and translatable virus-induced model since (i) the inflammatory demyelination seen in susceptible strains of mice is secondary to axonal degeneration (668, 670); (ii) the pattern of disease is biphasic under the right conditions, similar to human disease; and (iii) the mechanism appears to be primarily autoimmune. TMEV infection triggers T cell and macrophage infiltration into the CNS and results in a delayed-type (IV) hypersensitivity reaction (668). The T cells infiltrating the brains of TMEV-infected mice produce a humoral response against the virus in the early, acute phase of infection and also cause late chronic demyelinating disease, at least with low virulence strains (670). In the context of the non-spontaneous appearance of MS-like disease in murine models, CD4+ Th1 cells, which predominate in demyelinating lesions, only manifest on prior exposure of the murine brain to TMEV (668). The model does not adequately model the role of CD8+ T cells since TMEV-infected mice with deficient CD8+ cells are still highly susceptible to the virus (Benner et al., 2016). On the other hand, a transgenic mouse study found that non-structural Theiler’s virus genes induce innate immunity, suppressing initial viral replication and demyelination. If also true for viruses with tropism for human MS, these findings suggest that targeted therapies against non-structural viral proteins may be viable treatment options. Furthermore, since certain low virulence TMEV strains (DA or BeAn) but not others (GDVII) and mouse strains with certain genetic background produce distinct phenotypes including biphasic disease that mirrors multifactorial human disease (671), this model might provide an
opportunity to apply modern virology and “systems genetics” approaches to help to decipher the contribution of genetic and viral factors. Such approaches are described in further detail below.

Another virus model for MS might be Japanese macaque rhadinovirus (JMRV) infection in Japanese macaques (JM), the first described spontaneous demyelination animal model. This gamma-2 herpesvirus was isolated from JMs and was responsible for causing an inflammatory, demyelinating encephalomyelitis. The latter closely resembles MS in terms of both the clinical and pathological features including two or more oligoclonal bands in the CSF and lesions positive for interleukin-17 and CD4+ and CD8+ T cells (672). This finding corroborates the role of a viral agent in demyelinating diseases in a natural vertebrate host and is even more intriguing considering the close evolutionary relationship of this primate to humans (673).

1.3 Humans as models for multiple sclerosis

Human demyelinating diseases of known viral etiology can also serve as models for MS. Progressive multifocal leukoencephalopathy (PML), a disease with high mortality, is the only known human demyelinating disease with an established viral etiology(657): after long periods of latency, the John Cunningham virus (JC virus) becomes active during periods of relative immunosuppression and induces astrocytic infection and oligodendrocyte lysis(674). This neurovirulence could be due to interviral recombination with EBV(675). PML supports the hypothesis of an interaction between virus and host, and it suggests a probable viral trigger for autoimmune reactions. In this regard, the TMEV model does not follow the same pathogenesis but instead causes (sub-)acute demyelination (674) that is more reminiscent of MS, which manifests with slower clinical deterioration than PML.

Other human diseases also serve as models for certain aspects of MS. For example, subacute sclerosing panencephalitis (SSPE) is a rare brain infection caused by a mutated measles virus(676). Moreover, progression of acute disseminated encephalomyelitis (ADEM) – which is better modeled by experimental autoimmune encephalomyelitis (EAE) (677) – to MS might be triggered by viral infection, although these two diseases have different response to therapy and prognoses (678). In parallel, the cytokine-induced destruction of the CNS by human T-lymphotropic virus (HTLV)-1 resembles the demyelination seen in MS and also suggests that a virus triggers immune-mediated host response (657). The application of modern high-throughput technologies and analyses (such as next-generation sequencing) to these distinct clinical cohorts might help to decipher the complex host-virus factors that predispose individuals to not only these specific neurodegenerative diseases but also MS.
1.4 Stem cells-based modeling of multiple sclerosis: the role of viruses

More recently, induced pluripotent stem cells (iPSCs) have provided a useful in vitro model of aspects of MS pathophysiology that complement animal or human models. Although a detailed description is out-with the scope of the current review, iPSCs have been useful for studying active disease pathways prior to or at the early stage of the disease, the effects of specific genetic variants, or those of environmental factors (including viruses) on differentiation into neurons, astrocytes, and oligodendrocytes, which are generally inaccessible in humans (679). iPSCs for MS research were first derived in 2012 from the skin fibroblasts of a patient with RRMS (680), with current developments including the so-called heterotypic model the “reproducible iPSC-derived human 3D brain microphysiological system” (for further details, see (681)). From the virus perspective, it is of note that persistent infection with certain viral variants (e.g., the JHM neurotropic strain of mouse hepatitis virus) seem to constitute a model of both demyelination and remyelination capacity of neural precursor cells (reviewed in (682)). Furthermore, stem cell models have allowed investigation into whether neural (683) or mesenchymal stem cells (684) are neuroregenerative or at least neuroprotective in MS.

1.5 The viral pathophysiology of Multiple Sclerosis: a cornucopia of theories

No single theory or hypothesis captures the totality of MS pathogenesis, but a number of virus-related mechanisms are thought to be involved in orchestrating the immunology of MS. The main theories are molecular mimicry, bystander activation, epitope spreading, fertile field, and viral “déjà vu”. These are illustrated in Figure 1.

1.5.1 Molecular mimicry

Viruses can trigger an immune response by molecular mimicry, which is the development of an immune reaction against structurally similar self and foreign peptides(661). In molecular mimicry, professional antigen presenting cells (APCs) process virus antigens similar to self-antigens (e.g., myelin basic protein) via MHC class II molecules to activate an inflammatory response against CNS myelin. Indeed, a recent rare example of a patient sensitized with their own brain cells resulted in an MS-like disease that fulfilled all the pathological diagnostic criteria, direct evidence that self-peptides can induce MS-like disease (685). The efficacy of this mechanism depends on the APCs and the peptide obtained from a pathogen that resembles the self-myelin peptide (674). The first proof-of-concept molecular mimicry study was in a murine model of autoimmune herpes stromal keratitis, in which UL-6, a herpes simplex virus (HSV)-1 protein, was recognized by autoreactive T cells, while mutant viral strains not recognized by these T cells failed to induce the disease (686).
Studies supporting the role of molecular mimicry in MS suggest that some CD8+ T cells may be cross-reactive for viruses and self-antigens, consistent with the CD8+ infiltrates seen in demyelinating plaques (687). Human heterogeneous nuclear ribonucleoprotein L (HNRNPL) was identified as an MS autoantigen that cross-reacts with EBV EBNA-1 (688). Furthermore, antibodies against EBNA-1 antigen and Mycobacterium avium subsp. paratuberculosis peptides cross-react with myelin basic protein (MBP), which could possibly lead to autoimmunity (689). Polyreactive antibodies have also been identified that can bind to viral antigens as well as to the autoantigens (690, 691). T-cell epitope redundancy can facilitate molecular mimicry and can be responsible for autoimmune reactions following infections (Moise et al., 2016). This cross-reactivity of MBP, in addition to components of the EBV virus, has also been demonstrated with peptides isolated from CMV (692). The Synechococcus phage might also lead to molecular mimicry and contribute to MS pathogenesis (693).

Viruses generally adopt a variety of strategies to escape intracellular sensors of exogenous nucleic acids: modifying their DNA or RNA, sequestration, or post-translational modification of adaptor proteins, among others (694). For example, HHV-6 viruses synthesize proteins similar to endogenous cellular proteins as a kind of mimicry during infection. This phenomenon may facilitate the immune escape of the viruses but at a cost of “redirecting” the immune response to self-proteins so might contribute to molecular mimicry in susceptible individuals in MS (695).

Interestingly, molecular mimicry may be “delayed”, consistent with cases of MS developing many years after the initial viral infection. A repertoire of autoreactive memory T-cells can develop in response to viral infection with certain herpesviruses. These cells, potentially persisting for years, may be reactivated in response to CNS injury following infection with an unknown pathogen and induce autoimmune neurological disease. The risk of developing MS depends on the intensity of the immune response against the “injury” rather than the pathogenic event; this model has been called the “response-to-damage” model (696).

1.5.2 Bystander activation

Although remaining a possible contender, “bystander activation” is no longer thought to be a predominant cause of virus-induced MS. In bystander activation, viral products and their inflammatory mediators activate immature tolerogenic dendritic cells presenting self-peptides to produce a clone of (unrelated) autoreactive T cells. Thus, during an infection, autoreactive T cells also induce direct inflammation and cell death and cause demyelinating disease by targeting other antigens in the absence of molecular mimicry. Autoreactive lymphocytes that
infiltrate the CNS become important in disease progression and might become activated in the future to cause relapse (661, 674, 697). Viral proteins with sequence or structural similarity to CNS self-proteins can prime genetically susceptible hosts. Any immunological challenge in the form of an infection can then activate cytokines and cause disease by bystander activation (698).

Nevertheless, a recent study on myelin oligodendrocyte protein (whose auto-antibodies are implicated in ADEM) revived this theory, at least for enveloped viruses. More specifically, the B cell receptor can “co-capture” a self-antigen (membrane protein) and viral antigen from a virus-infected cell (in this case, influenza); afterwards, it can “fraudulently” gain T-cell help from a viral antigen-specific T cell, leading to the secretion of autoantibodies (699).

### 1.5.3 Epitope spreading

Both inter- and intra-molecular myelin peptide antigens can trigger T and B cell-mediated autoimmune reactions by epitope spreading with or without bystander activation and molecular mimicry. In epitope spreading, viral peptides first activate macrophages which destroy neuronal myelin resulting in the further release of self-tissue peptides, which are themselves then processed and presented on APCs, resulting in autoimmunity (700). A non-pathogenic TMEV peptide demonstrating molecular mimicry with MBP, when inoculated in SJL mice, has been shown to activate autoreactive T cells. As a result, the myelin-specific T cells attacking the virus may themselves initiate demyelination through epitope spreading. Epitope spreading is thought to be initiated by naive T cells that enter the inflamed CNS and are subsequently activated by APCs (661, 674, 701).

### 1.5.4 Fertile field theory

Another intriguing hypothesis is the “fertile field” theory, according to which any viral infection in the body produces a transient, short-lived immunological state called the “fertile field”. The fertile field theory states that an initial inflammatory reaction caused by viral infection creates a clone of autoreactive T cells via any of the other mechanisms described (e.g., molecular mimicry), which are then subsequently activated by another antigenic stimulus to cause neuronal inflammation. This field can vary depending upon the type of virus, anatomical location, and the virus-induced inflammatory response. Any antigen exposure during this period, either viral or self, can modulate the immune system through bystander activation or molecular mimicry to produce autoreactive T cells, resulting in clinically observable disease (702).

### 1.5.5 Viral déjà vu

131
Finally, although only described in animal models, “viral déjà vu” states that an initial viral infection, although initially cleared, can produce T cell clones that are subsequently reactivated by a second “precipitating” virus and neuronal inflammation. In animal models, attenuated LCMV was eliminated from most tissues except for the CNS by the innate immune system of neonatal mice. However, cytotoxic T cells specific for the virus also persisted that were efficiently triggered by a second viral infection in adulthood to cause disease. Conceptually, the cytotoxic cells targeted the epitope shared by the predisposing virus and the precipitating virus. Interestingly, molecular mimicry and bystander activation are not implicated in this conceptual model (703).

1.6 Peripheral vs. central progression in multiple sclerosis

It is unknown whether a putative infection starts in the periphery and then moves to the CNS or the infecting agent directly attacks the CNS. This is important since peripheral-central progression would implicate a different set of mechanisms and viruses than a solely central mechanism. Based on emerging evidence that the CNS is subject to immune surveillance, viral infections may be able to trigger a local inflammatory response, leading to demyelination and progression of disease as a consequence of autoreactive lymphocytes. T and B cell interactions can produce an immune response that results in degeneration of neurons (304) (653).

On the other hand, inflammation might start in the periphery. More precisely, it was previously thought that the blood-brain barrier protects the CNS and there are no central immunocompetent immune cells. However, recent findings indicate that the CNS is less immune privileged that previously thought and is subject to immune surveillance. What remains obscure, however, is why the brain is a target for inflammation. Macrophages or the microglial cells in the CNS are capable of mounting immune reactions, including those that result in neurodegeneration, in response to immunological challenges. In addition, the observed trafficking of activated immune cells from the periphery to the CNS may also challenge the notion that the CNS is “immuno-privileged” (322). A recent attractive and simple finding is that dura mater CNS lymphatic vessels drain interstitial fluid, macromolecules, and cells from the brain to the deep cervical nodes, suggesting that the CNS is susceptible to immune reactions via local resident immune cells as well as via peripheral immune-activated cells that travel to the CNS (306). A recent study reported a strong role for CD4+ cells in cell trafficking; these cells transported antibodies to peripheral sites of action to limit the spread of infection and the INF-γ secreted by CD4+ cells facilitated the transport of antibodies by increasing the vascular permeability (704).
Thus, as recently proposed, it seems that there is evidence for bidirectional traffic: immune cells from the CNS may travel to secondary lymphoid tissues via the lymphatic drainage, where affinity maturation may take place; or, B cells found in the CNS might actually mature in the draining cervical nodes. These studies suggest that immune cells may move freely “in and out” of the CNS across tissue barriers (Louveau et al., 2015, Stern et al., 2014) and that immune reactions occurring in other parts of the body may be able to cause MS, especially the spleen and lung as key immunomodulatory organs (Engelhardt et al., 2016).

1.7 New theories on viral etiology: the enteric virome and multiple sclerosis

In terms of which extra-CNS sites might trigger MS, there is a strong association between the gut microbiome and immunological functions. Immune responses are modulated by both host and microbe interactions and so-called “transkingdom interactions” between the gut bacteria, fungi, parasites, and viruses, with the latter both affecting and being affected by such interactions (705). In this context, an intriguing hypothesis supported by empirical evidence states that gut commensals, if dysregulated, may increase susceptibility to MS (706, 707), especially if – as noted above – the CNS lacks immune privilege and peripheral immune reactions can drive CNS pathology.

In support of this hypothesis, the gut microbiota has been shown to enhance EAE in mice compared to those maintained in germ-free conditions (708) (707, 709). Furthermore, patients with MS have a distinct gut microbiota (710, 711). The composition of the gut microbiome is dysregulated by risk factors repeatedly associated with MS such as smoking, alcohol, diet, and vitamin D deficiency (712). It has also been observed, consistent with the molecular mimicry theory, that an MS immunogen resembles non-pathogenic gut bacteria such as Bacteroides, Lactobacilli, Bifidobacterium, and Clostridium (713). However, a non-specific polyclonal bystander activation of the innate immune response (either locally or systemically) might also play an important role (709). Recently, a reduction in species of concrete Clostridia and Bacteroidetes clusters was implicated in MS, albeit not assessed with regard to clinical course. Although dysregulation of the “enteric virome” is implicated as a cause of certain autoimmune disorders such as Crohn’s disease, its relevance to a disease such as MS without an established antigenic etiology requires further study (714). Furthermore, this needs doing in the context of the other microorganisms in the gut since there appears to be significant crosstalk between the two. For instance, the intestinal virome is largely composed of bacteriophages (715), and the “phageome”, i.e., the sum of phages in gut microbes, might also be implicated in MS, with phages attracting the interest of metagenomicists over the last few years (716). Endogenous retrovirus levels appear to be dependent on the microbiota (717), which is
particularly important since endogenous retroviruses are themselves implicated in MS pathogenesis (718). Delineating how these extra-CNS factors contribute to MS will be complicated given the co-existence of microorganisms, the immune system, the unspecified role of the enteric nervous system, and complex metabolism of microbial substances.

1.8 “-Omites” and “systems medicine”: future perspectives in the MS-virus arena

Deciphering the epidemiological contribution of viruses to MS, how the virus and host interact, and their pathogenic mechanisms may help to develop diagnostic tests to determine individual susceptibility; diagnostic tests to monitor progression, treatment effect, and relapse; and targeted therapies to treat the disease, prevent relapses, and maintain remission. The preceding discussion highlights that viruses are likely to be important in MS but also identifies several knowledge gaps. Specifically: (i) there are currently no biomarkers that identify at-risk individuals who might be suitable candidates for preventative treatment; (ii) direct evidence of specific viruses in MS is lacking; (iii) the contribution and interaction between host genetic factors and viruses in MS have yet to be fully elucidated; (iv) the contribution of immune regulation at other, extra-CNS sites (such as the gut) to MS pathogenesis deserves further consideration. These are considered in turn below and in Box 1.

1.9 Susceptibility biomarkers and next-generation sequencing

As noted above, there is a rich history of genetics research into the etiology of MS and many HLA- and non-HLA loci have been discovered as risk factors (654). Due to the limitations of the technologies or size of the cohorts analyzed, these mainly represent common variants with minor allele frequencies of ≥5%. SNP analyses in genome-wide association studies, together with hypothesis-driven approaches, have identified over 100 risk loci (654), which together have been useful for understanding pathogenesis but together only account for about 27% of total heritability (Lill, 2014). This “missing heritability” (654) is at least in part due to the heterogeneity of patient cohorts analyzed, insufficiently powered studies, high false-positive detection rates, and the focus on the weak effects of common alleles rather than the stronger effects of rare alleles.

Although newer technologies such as NGS do not fully overcome these limitations, they do provide the opportunity to identify functional rare variants in causative genes (719) and, since high-throughput sequencing is not limited solely to DNA, transcriptomic and epigenetic pathways disrupted before, during, and after disease onset (Sanders et al., 2016). These studies that capture genome-wide information from carefully selected clinical cohorts should shed
further light on which immune risk factors and pathways are virus-related. However, perhaps the real strength of NGS in viral etiology research lies in the area of coupling it with metagenomics – identifying genetic material from known and novel pathogens in environmental samples, in this case, clinical samples collected from humans. A recent metagenomics study found VZV sequence in one out of twelve patients with idiopathic inflammatory demyelinating disease; its analysis, however, was heavily restricted due to high contamination levels (720). A later study, using similar techniques noted high levels of environmental contamination as above, but EBV, CMV, and parvovirus – albeit in low number of reads and in few MS patients – were still identified in silico (721). NGS is now approaching clinical “prime-time”, with one recent study employing NGS to detect replacement mutations in a VH4 gene to identify individuals who have or are likely to convert to RRMS (722).

However, the “big data” generated by integrating exome and whole genome data from large numbers of patients is a bioinformatic challenge and likely to produce a large number of false-positive results; advances in “statistical genomics” and their associated computational tools are required in tandem with experimental advances. One way to circumvent the bioinformatics challenges is to combine modern technologies with hypothesis-driven studies on rationally selected patient groups – such as affected and unaffected twins – to deconvolute risk factors in tightly defined patient populations.

1.10 The search for direct evidence of viruses in MS

Direct evidence of viruses in the brains of MS patients is lacking, and no specific microbe has yet been isolated from the brains of MS patients and confirmed according to modified Koch’s postulates (657); however, recent studies have noted the predominance of certain bacterial phyla in progressive MS (723). This is perhaps unsurprising since several of the mechanistic hypotheses detailed above rely on secondary activation of immune pathways; the virus has, in effect, “hit and run”. Nevertheless, new technologies offer the opportunity to find pathogens that may hitherto have been missed. NGS studies of MS post-mortem biopsies have only detected a single case of viral infection, namely, GBV-C (hepatitis G) (724) and, as mentioned above, VZV sequence was found in one out of twelve patients with idiopathic inflammatory demyelinating disease (720). The causative agent might manifest as a low-level infection, fall under the detection threshold of existing techniques, or the active lesions that harbor virus might be difficult to isolate, so technologies such as laser capture microdissection might be useful to enrich for lesional cells when performing experimental pathology studies (661, 725), (662). As noted, searching for viral sequences in complex clinical samples via the metagenomic approach
provides a further opportunity to search for causative pathogens. We contend that these types of studies require a coordinated, multicenter effort to generate a tissue biorepository to underpin translational tissue-based genomics studies.

1.11 Deciphering host-virus interactions

The observations that certain mouse and virus strains (e.g., TMEV DA or BeAn) but not others (GDVII) produce distinct phenotypes (671) and that certain viruses cause neuroinflammation in some humans (e.g., PML, SSPE, ADEM) provide direct models of virus-host interactions that are directly applicable to MS. These experimental and human models provide an opportunity to apply modern virology and “systems genetics” approaches to help to decipher the contribution of genetic and viral factors, i.e., analysis at all levels of biological complexity in both host (genome, transcriptome, epigenome, proteome) and virus (virome). Such a multi-level approach would leverage all available comparative information from very well-defined experimental groups (in the case of TMEV) or rare patient groups (in the case of PML) to discover relevant causative pathways. Furthermore, such a “systems-based” approach could be applied to the analysis of B and T cell receptor repertoires (“immunosequencing”) in MS samples in relation to viral detection and burden following enrichment of immuno-sequencing databases with extensive experimental data on the repertoires induced by a range of human viruses.

1.12 Beyond the brain: the contribution of extra-CNS immunity to MS

The recent intriguing data the gut microbiota enhances EAE in mice compared to those maintained in germ-free conditions (707-709) and that patients with MS have a distinct gut microbiota (710, 711) raises the prospect that the gut might be an important disease-initiating or disease modifying site. These initial, proof-of-principle “applied” metagenomic studies now need to focus on the virome to determine what, if any, role the enteric virome plays in disease pathogenesis and whether the enteric virome affects MS or vice versa. Similar to the experiments performed with bacteria, virus-inoculated and pathogen-free MS mice should be compared; second, any differences in enteric viromes should be examined in mouse models of MS vs. normal inbred strains. Other possibilities include investigating other components of the brain-gut axis, for instance, the vagus nerve, which has recently been shown to inhibit cytokine production on stimulation to modulate disease severity in rheumatoid arthritis (726) and might indirectly affect the microbiome via its effect on gut motility (727) or through other ways (e.g., brain-to-gut transmission of proteins, like in α-synucleinopathies (632)). Given the relative ease
with which the gut microbiota can be therapeutically modulated either directly (e.g., with enemas) or systematically (with anti-virals) (at least compared with therapeutic targets in the CNS), this may prove to be an attractive means by which MS can be treated.

1.13 Viruses in the treatment of multiple sclerosis

Whilst MS is not usually an acute life-threatening disease, MS reduces life expectancy by up to seven years in certain regions. It is, however, a highly debilitating disease from both the neurological and neuro-psychiatric perspectives and the social and psychosocial perspectives.

The therapeutic armamentarium has been enriched by the “disease modifying therapies” (DMTs) such as certain monoclonal antibodies (e.g., alemtuzumab, anti-CD52; daclizumab, anti-CD25; natalizumab, anti-α4-integrin), cytokine therapies (e.g., the interferons Anovex, Betaferon), and small molecules (terifluonimide, dimethyl fumarate, fingolimod) which are immunomodulatory in RR-MS rather than etiologically driven (as reviewed elsewhere, e.g., in (728) and (96)). Thus, they tend not to be effective in progressive MS, which principally shows features of neurodegeneration rather than inflammation, posing therapeutic challenges (729).

Further elucidating the viral etiology of MS (besides their apparent role in disease relapse) will therefore potentially help to develop (i) prognostic biomarkers (i.e., for relapse in RR and/or deterioration in the PP subtype) and (ii) more etiologically-driven therapeutic approaches (i.e., targeted antiviral therapies, therapeutic vaccines). In this regard, there have been promising phase II trials of a monoclonal antibody targeting multiple sclerosis-associated retrovirus (MSRV) envelope protein (MSRV-Env) (730), which belongs to a human endogenous retrovirus implicated in MS and originally characterized in 1997 (731).

1.14 Concluding remarks

In summary, the exact role of viruses in the pathophysiology of MS is still far from fully understood. We have reviewed the current literature on the role of viruses in MS pathophysiology and identify four key areas for focused research: (i) biomarkers to identify at-risk individuals; (ii) the search for direct evidence of specific viruses in MS; (iii) establishing the contribution and interaction between host genetic factors and viruses in MS; (iv) the contribution of immune regulation at other, extra-CNS sites (such as the gut) to MS pathogenesis. Research in these areas will be facilitated by the application of high-throughput technologies, the development of systems-based bioinformatic approaches, careful selection of
experimental and human models, and the acquisition of high-quality clinical material for tissue-based translational research.
FIGURES

Figure 1. The viral pathophysiology of Multiple Sclerosis. A number of theories have been proposed to explain a viral etiology for multiple sclerosis. In molecular mimicry, virus antigens similar to self-antigens (in this case, myelin basic protein) are processed by professional antigen presenting cells via MHC class II molecules to activate an inflammatory response against CNS myelin. In bystander activation, immature tolerogenic dendritic cells presenting self-peptides are activated by viral products and their inflammatory mediators to produce a clone of autoreactive T cells. In epitope spreading, viral peptides first activate macrophages (1) which destroy neuronal myelin (2) resulting in the further release of self-tissue peptides which are themselves then processed and presented on APCs, resulting in autoimmunity. The fertile field theory states that an initial inflammatory reaction caused by viral infection creates a clone of autoreactive T cells (1) via any of the other mechanisms described, which are then subsequently activated by another antigenic stimulus (2) to cause neuronal damage (3). Viral infection causes a transient and local increase in susceptibility to antigenic exposure via this mechanism, or a “fertile field”. Finally, although only described in animal models, viral déjà vu states that an initial viral infection, although initially cleared, can produce T cell clones that are subsequently reactivated by a second “precipitating” virus and neuronal inflammation.
Box 1

- How do we harness the power of next-generation sequencing to decipher the viral etiology of MS, and do we have the bioinformatics tools to analyze “big data”?
- Given a lack of tissue or samples for translational research, is there scope to coordinate interested stakeholders to facilitate this research, perhaps by establishing carefully curated biorepositories?
- How can we integrate *big data* from the host and viruses to better understand how they interact?
- Is the gut microbiome the “missing link” in the pathoetiology of MS?
### Table 1. Important animal models used in multiple sclerosis research

<table>
<thead>
<tr>
<th>Model</th>
<th>Animal(s)</th>
<th>Features</th>
<th>Pathology</th>
<th>Strengths</th>
<th>Limitations</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mouse SFV infection (mutant M9 and avirulent A7 strains)</td>
<td>Rat, mouse</td>
<td>Subclinical disease</td>
<td>CD8+ T cell-mediated demyelination</td>
<td>Might represent molecular mimicry mechanism</td>
<td>Might not be indicative of autoimmunity</td>
<td>(732-734)</td>
</tr>
<tr>
<td>TMEV</td>
<td>Mouse</td>
<td>Flaccid hind limb paralysis</td>
<td></td>
<td>Good model for studying immunosuppressive therapies</td>
<td>Requires direct injection of virus to CNS</td>
<td>(735)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Biphasic infection with low virulence strains</td>
<td>CD4+ T cell-mediated autoimmunity to myelin via molecular mimicry or epitope spreading</td>
<td>Good model of virus-induced myelin and axonal damage</td>
<td>Limited to the mouse</td>
<td></td>
</tr>
<tr>
<td>Mouse hepatitis virus (neurovirulent strain MHV-A59)</td>
<td>Rat, mouse</td>
<td>Acute encephalitis</td>
<td>T cell and macrophage-mediated demyelination and axonal damage from viral infection and immune-mediated attack</td>
<td>Axonal damage and remyelination</td>
<td>Requires direct injection of virus to CNS</td>
<td>(736, 737)</td>
</tr>
<tr>
<td>Japanese macaque rhadinovirus</td>
<td>Macaque</td>
<td>Inflammatory demyelinating disease</td>
<td>Active lesions containing CD4+ T cells (Th1, Th17), CD8+ T cells, and positive CSF</td>
<td>Clinical and biological features that closely resemble human MS</td>
<td>Primate model, limited availability</td>
<td>(672)</td>
</tr>
<tr>
<td>Transgenic zebrafish</td>
<td>Zebrafish</td>
<td>Demyelination/remyelination</td>
<td>Oligodendrocyte ablation</td>
<td>Good for functional genomics studies</td>
<td>Not closely related to humans</td>
<td></td>
</tr>
<tr>
<td>---------------------</td>
<td>----------</td>
<td>-----------------------------</td>
<td>--------------------------</td>
<td>------------------------------------</td>
<td>-----------------------------</td>
<td></td>
</tr>
<tr>
<td>EAE</td>
<td>Rat, mouse, rhesus monkey, marmoset, Guinea pig.</td>
<td>Spectrum of disease: hyperacute, acute, chronic, chronic relapsing, secondary progressive, spontaneous</td>
<td>Spectrum of changes from minimal demyelination to marked gliosis, demyelination, remyelination, and axonal and neuronal loss</td>
<td>Spectrum of disease represented depending on animal and autoantigen</td>
<td>Mainly affects spinal cord white matter rather than cerebral and cerebellar cortex</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CD4+ and CD8+ T cell-mediated disease via both Th1 and Th2 pathways</td>
<td></td>
<td>Some models produce chronic relapsing disease and secondary progressive disease (ABH mouse)</td>
<td>CD4 T cell focus via stimulation with self-peptides rather than the dominant CD8+ T cells seen in human MS</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mechanism of relapse uncertain</td>
<td></td>
<td>Specific mechanisms can be probed with genetically engineered animals</td>
<td>Do not address the role of B cells</td>
<td></td>
</tr>
<tr>
<td>Transgenic mice, e.g., myelin targets (MOG, TCR); chemokines/cytokines (e.g., TNFα, IL12)</td>
<td>Mouse</td>
<td>Dependent on model, e.g., MOG knockout prevents chronic relapsing disease</td>
<td>Dependent on model</td>
<td>Allows contribution of specific target to pathogenesis to be determined</td>
<td>Redundancy in biological systems produces unexpected results</td>
<td></td>
</tr>
<tr>
<td>Toxin-induced demyelination, e.g., LPC, EB, bacterial endotoxin, cuprizone</td>
<td>Rat, mouse</td>
<td>Dependent on model but include direct membrane damage (LPC, LPS), secondary inflammation, mitochondrial dysfunction (cuprizone)</td>
<td>Dependent on model but focal demyelination (e.g., with LPC, EB) or stereotypical demyelination of sensitive brain regions with cuprizone</td>
<td>Some models easy to administer (e.g., cuprizone)</td>
<td>Some models require specialist equipment (e.g., stereotactic equipment)</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: SFV, Semliki Forest virus. TMEV, Theiler’s murine encephalomyelitis virus. EAE, experimental autoimmune encephalitis. MOG, myelin oligodendrocyte glycoprotein. TCR, T cell receptor. TNF, tumor necrosis factor. IL, interleukin. LPC, lysophosphatidylcholine. EB, ethidium bromide.
2. Viruses and Endogenous Retroviruses in Multiple Sclerosis: From Correlation to Causation

2.1 Introduction

The prevalence of multiple sclerosis (MS) has steadily increased over the last five decades. This high socio-economic burden, together with its challenging management especially when chronic and progressive, underscores the need for further research to determine its exact aetiology. MS is a multifactorial disease that arises from a complex interaction between genetic (notably immunogenetic), autoimmune, and environmental factors (17, 651-653, 674). Environmental-immune system interactions are increasingly recognised as important in MS pathophysiology (99, 656) and are likely to explain the discordant MS incidence in monozygotic twins that cannot be attributed to genomic, transcriptomic, or epigenomic factors alone (655). Furthermore, the environment represents a modifiable factor in contrast to the genomic landscape, so it is of particular interest from the perspective of prevention.

MS was initially proposed to be of infectious origin at the end of the 19th century, but the development of the experimental allergic encephalitis (EAE) model in 1934 shifted attention away from microorganisms and toward an allergy-related and then autoimmune basis for the disease. However, a myelin-targeting autoimmune model does not fully explain the segmental distribution of lesions since myelin is ubiquitous in the central nervous system (CNS) (745), and auto-antigens are neither pathognomonic nor universal in MS (746). In addition, some authors have suggested that the EAE model might more closely represent immunologically-induced encephalomyelitis rather than demyelination (677, 747).

However, the microbial aetiological theory – in which viruses take center stage – has not been abandoned but has flourished in light of mainly indirect discoveries of different viruses in MS (Venkatesan and Johnson, 2014). Although direct evidence for causative viruses in MS has generally been lacking, accumulated evidence from human and animal studies support a role for viruses as at least a trigger for MS (652). Epidemiological evidence in support of this theory include observations of MS epidemics in the Faroe Islands in the 80s, and more recently, MS clusters in Ottawa, Canada (748).

Although the evidence for a causative viral aetiology for MS in humans remains inconclusive, viruses appear to play a role in modulating the neuro-immunological system of genetically susceptible individuals to cause MS. For instance, IgG antibodies against several viruses including varicella-zoster virus (VZV), cytomegalovirus (CMV), measles, rubella,
mumps, and herpes simplex virus (HSV-1) have been identified in the cerebrospinal fluid (CSF) of MS patients (674, 749). More recently, other viruses have attracted attention including Saffold virus (a novel human cardiovirus) (750, 751). With this in mind, this review provides an in-depth discussion of the viruses implicated in MS pathogenesis. We first consider viruses with the greatest evidence-base, namely Epstein-Barr virus (EBV), human herpesvirus (HHV-6), VZV, human endogenous retroviruses (HERV), and then go on to describe the potential roles for “minor” viruses in MS. We focus on the connection between viruses and MS pathophysiology rather than its clinical progression and we highlight the limitations of existing studies and possible future research directions.

2.2 Search methods and selection criteria

The PubMed and Google Scholar databases were searched for articles published (or appeared “Epub ahead of print”) between January 1, 2006 to December 31, 2016 and the bibliographies examined. For initial screening, “multiple sclerosis”, “infectious cause”, “virus”, or “viral model” were applied through the Boolean operators “AND” and “OR”. More specific terms were applied for different sections of the review based on their relevance. If and when an infectious agent had more than one name, all relevant search terms were applied. Priority was given to original research articles and systematic reviews/meta-analyses over case reports or hypothesis/viewpoint articles and the most recent papers as applicable. Some references prior to the above time period were included given their historical importance. Studies referring to pediatric MS, infectious agents other than viruses, and those not published in English were excluded.

2.3 Epstein-Barr virus (EBV)

There is a lively on-going debate on the role of EBV, the prevailing MS infectious risk factor, and MS pathogenesis (696, 752, 753). One hypothesis suggests that MS is caused by a genetically predisposed deficiency in eliminating previous EBV infection; EBV then persistently accumulates or even establishes itself in the brains of such patients (99, 754). Consistent with this theory, EBV might exercise a strong influence on the number of naïve and/or memory B cells and their differentiation status (755). A competing hypothesis is that abnormal responses to EBV infection are secondary to and not a cause of MS (756).

At the epidemiological level, several systematic reviews clearly support an association between MS and EBV seropositivity (757-759). Practically all MS patients are EBV
seropositive, raising the question of whether EBV-seronegative MS patients even exist (760). EBV seropositivity confers double the risk of MS than infectious mononucleosis (IM) (OR = 4.56 vs. OR = 2.17, respectively) (17), and IM appears to have a stronger genetic component than EBV infection (761). However, the reasons for this difference in risk between EBV seropositivity and IM might be due to: (i) reporting bias for IM; (ii) the molecular stochasticity of EBV-induced downstream events; (iii) the role of EBV latency; or, importantly, (iv) subclinical infection. High Epstein-Barr virus nuclear antigen (EBNA) IgG titers are associated with other MS risk factors such as non-HLA gene loci and the HLADRB1*15 allele (the most important genetic factor in MS) (99, 762). T cells restricted to the HLADRB1*15 allele and linked to MS-related antigens seem to cross-react with the immunological response induced by the EBNA-1 sequence (763). However, the latest meta-analysis revealed an additive but not synergistic effect between the two risk factors, corroborating that HLADRB1*15 carriage is not a confounding factor for EBV and MS (764, 765).

A highly synergistic increase (14-fold) in MS risk was reported for EBV detection or IM combined with obesity, notably during adolescence (99). However, there are conflicting results on the interaction between EBV and other well-established MS risk factors (reviewed also in (766)). For instance, a prospective study found a positive association between smoking and MS development only in older patients and a negative one in patients less than around 30 years old (767), whereas a later case-control study reported a negative, multiplicative interaction between IM history and a prior history of smoking on MS risk (768). With regard to vitamin D, some studies have failed to detect a statistically significant interaction (765, 769) while others have reported an interaction with either EBV antibodies or DNA load (770). Mechanistically, observations that there is overlap between EBNA-2a and vitamin D receptor (VDR) binding sites within MS-associated genomic regions and that EBNA-3 binds to the VDR may provide further insights (756, 771).

In neuroimaging studies, MRI (magnetic resonance imaging) markers of MS activity and grey matter atrophy were found to be associated with anti-EBV antibody levels (772, 773). At the cellular level, CD8+ T cells specific for EBV lytic and latent antigens were more frequent in patients with active and inactive MS, respectively (774). Deep sequencing of T cell receptor-β genes (“immunosequencing”) showed intrathecally-enriched EBV-reactive CD8+ T cells that were specific to MS patients (775). Furthermore, in animal models using lymphocryptovirus (LCV), which is a close relative of EBV, LCV-infected B-cells lost their ability to process the extrinsic pathogenic CD8+ T cell epitope in myelin oligodendrocyte glycoprotein (MOG). In
doing so, they cross-presented this epitope to auto-aggressive cytotoxic T lymphocytes, a reaction that can initiate an autoimmune reaction and demyelination (776).

With regard to humoral immunity, the high antibody titers against EBNA proteins in MS patients might be due to intrathecal synthesis but it has yet to be clarified whether they result from high frequency latent EBV-infected cells or, alternatively, have a concrete pathogenic role (777). Conversely, patients with IM showed activation of MOG (Myelin oligodendrocyte glycoprotein)-specific memory B cells (778).

Furthermore, EBV genetic material has been identified in the CSF and perivascular infiltrates of brain and spinal cord white matter, and, more recently, in the cortical grey matter and cervical lymph nodes of patients with MS (777, 779). EBV brain infection is likely to be limited to only a small number of B cells (approximately 5 to 3000 per 10⁷ memory B cells) (780, 781). This could explain why histological studies for their detection are difficult, and it underscores the need for technologies such as massively-parallel single cell sequencing to detect these rare events in the future (782). Dual infection with EBV types 1 and 2 is more common in MS patients compared to single infection (783).

Mechanistically, EBV might act as both an environmental trigger or by attacking the CNS (784). With respect to the former, an EAE model with the murine EBV homologue gamma-herpes virus 68 showed more pronounced MS-like clinicopathological features that were dependent on the latent life cycle of the virus (785). There are a number of theories with regards to the latter mechanism of direct CNS destruction by the virus: (i) cross-reactivity of EBV-infected T-cells with self-antigens (“molecular mimicry”) causes destruction of CNS tissue but does not explain the presence of EBV-infected B cells in the brain; (ii) the bystander damage hypothesis proposes that immune responses in the CNS are directed towards EBV antigens but does not explain the autoimmune component of the disease and the failure to eliminate these cells; (iii) MS results from EBV infection of autoreactive B-cells, which in turn produces pathogenic autoantibodies (777); (iv) the “mistaken self” hypothesis based on proteomic analyses shows a higher frequency of a peptide corresponding to an EBNA-1 region sharing homology with the N-terminus of αB-crystallin in MS patients (786). Overall, understanding these mechanisms paves the way for novel anti-MS strategies, notably EBV-specific adoptive immunotherapy (754).

It is also mechanistically intriguing how EBV plays a role in both cancer – a disorder of cellular proliferation – and MS – a disease characterized by neuronal cell death; however, recent
reports of a genetic overlap between the EBV-related Hodgkin lymphoma and MS could shed some light on this (787, 788). In parallel, dogma that EBV cannot possibly be found in glial cells or neurons, the host immune response must remain the focus of studies (787), or that EBV latency status underpins virus-mediated pathogenesis (756) should be re-examined in light of recent observations that EBV can cause lytic infection in human primary neurons (789).

To summarize, in the context of discordance between the high rates of EBV infection vs. low rates of MS worldwide, EBV is likely to be necessary but not sufficient to cause MS (752). Future studies on shared polygenic risk from genome-wide association studies (GWAS) on MS cases with those with markers of increased EBV levels (e.g., EBNA-1 (790)) are likely to shed further light on such host-pathogen interactions.

2.4 Human herpesvirus 6 (HHV–6)

A recent, inconclusive, non-systematic summary of evidence on the role of HHV-6 in MS (791) highlighted the need for a formal meta-analysis on this topic. Furthermore, although HHV-6 has been detected mostly in acute demyelinating brain lesions in MS, detection rates are highly variable (HHV-6 DNA in the CSF ranging from 3 to 46% of patients) (792). Additionally, other markers such as B- or T-cell reactivity, higher antibody responses, or higher viral loads have not been consistently observed in MS patients’ serum in different ethnic groups or prospective studies (791, 793).

Specific single nucleotide polymorphisms (SNPs; e.g., in CD46 and MHC2TA) are strongly associated with active replication of HHV-6 and, together, with worse clinical prognosis in MS (794). At the edge of such gene-environment interactions lie the human endogenous retroviruses (HERVs) (see below). One of their subtypes (HERV-K18) was shown to be activated by HHV-6A, mainly in cell lines productively infected with the virus and followed by those with latently infected virus. These observations reinforce the notion that there is a common HERV-mediated pathway downstream of viral infection in MS (795), which might be therapeutically exploitable (730).

The marmoset (Callithrix jacchus) HHV-6 model has been used to study viral neurotropism (796). Interestingly, in contrast to how the virus seems to gain entry to the human CNS via olfactory pathways (797), findings in marmosets revealed that only those with intravenous (and not intranasal) inoculation of HHV-6A (and not HHV-6B) developed neurological disease (796). Furthermore, in contrast to the global seroprevalence of >95% for
HHV-6B, HHV-6A is more frequent in MS patients than HHV-6B, which is certainly worthy of further investigation (791). HHV-6A infection leads to apoptosis in the brain, induces autoimmunity in several ways (695), and activates antiviral genes in human astrocytes including some genes upregulated in MS (798).

2.5 Varicella zoster virus (VZV)

Varicella zoster virus (VZV)-induced encephalomyelitis is characterized by demyelination similar to that seen in MS, so VZV is suggested as an MS-triggering factor (799). However, while some epidemiological studies reported no association between a history of varicella infection in childhood and MS risk (800), others have observed an association, most notably for relapse-remitting (RR) and secondary progressive types (801). A four-fold increase in MS risk in the year following herpes zoster infection has been observed in a region with a low MS prevalence (802).

Regrettably, serological and molecular studies have not helped much in this area. VZV seropositivity was not significantly higher in MS patients vs. controls in two studies (803, 804). Moreover, while VZV DNA was identified in the CSF of MS patients (particularly of RR type) in some studies (805, 806), others failed to confirm these findings in the CSF, blood, or in acute MS lesions (792, 807-809). More consistently, however, are the observations that the high levels of VZV DNA in CSF and PBMCs during relapse ultimately disappear during clinical remission (810, 811). Interestingly, the progressive MS type has been associated with VZV DNA at levels between those found during the relapse and remission periods of the RR-MS type (811, 812).

The median fraction of intrathecal VZV-specific IgG of total IgG can differentiate MS patients from those with VZV reactivation (35-fold higher in the latter) (813). This observation implies that low-level infection is present in at least some MS cases. It also helps address whether or not VZV detection in MS is due to reactivated, previously latent VZV infection; that is, “centripetal infection” from the neural ganglia towards the CNS (814). Another theory suggests that VZV in MS is purely epiphenomenal due to leakage from destroyed sensory neurons; however, experimental evidence is lacking (815). Also, VZV has not been identified in “traditional” autoimmune diseases, implying a more specific connection with MS (814). Finally, VZV antigens induced and maintained activity of HERVs in peripheral lymphocytes from MS patients compared to controls; retroviruses, as explained below, are implicated as causal in MS (816).
2.6 Human endogenous retroviruses (HERVs)

Although initially both were implicated in MS pathology since the 80s, subsequent studies have continued to support a role for endogenous rather than exogenous retroviruses in MS (817). HERVs were integrated into the human genome relatively recently in evolutionary terms, i.e., some 30-40 million years ago, as a result of ancestral retroviral infections. In humans, they form up to 8% of the genome and constitute a notable category of long terminal repeat (LTR) retrotransposons. These transposable elements, also known as “jumping genes”, change position within a genome and have repetitive sequences, explaining why it is more difficult to investigate their inheritance with classical genetics approaches (818). Despite these difficulties, the estimated 320,000 transcription factor binding sites (TFBSs) regulated by HERVs underscore their genome-wide role. Deciphering their pathophysiological roles will offer further insights into the molecular basis of disease beyond that offered by focusing exclusively on the exome (819, 820).

Putative mechanisms of HERV-related pathophysiology in MS are illustrated in Figure 1. For example, several SNPs are associated with MS corresponding to genes implicated in immunological or vitamin D regulation. These SNPs occur more often in the vicinity of HERV-related open reading frames (ORFs) than non-MS related SNPs (821). Conversely, SNPs in regions around HERVs (such as the X-linked HERV-Fc1) are associated with MS, primarily the RR and secondary progressive types (822, 823). Furthermore, HERVs might “bridge” the environmental-genetic interaction in MS given that any trigger (including viruses) may reactivate HERVs and enhance their expression (824). Of recent note – given that interferon signalling is implicated in MS – is the observation that HERVs contain binding sites for interferon-γ-induced transcription factors and therefore affect the expression of other genes, notably ones with immune function (825).

There is also an established body of evidence that the envelope protein of the “MS-associated retrovirus” (MSRV-Env) in the HERV-W family is causal for MS (826). Initially observed in leptomeningeal cells, the MRSV-Env protein has been detected in MS plaques containing macrophages, microglia, and perivascular cells in actively demyelinating lesions and in the astrocytes of inactive areas but not in control brains (827). High MRSV-Env DNA copy number, transcript, and antigen levels have recently been detected in the blood of over 70% of MS patients (828); the increased DNA copy number is indicative of HERV-related reverse transcriptase activity. Earlier studies suggested potential MRSV-Env selectivity for the MS brain after observations of viral genetic material present at higher levels in the brain than in the blood of the same patients (829).
HERV-W Env expression is also increased on the surface of B cells and monocytes during the active phase of MS and parallels MS exacerbations (830). This protein, a toll-like receptor 4 (TLR4) agonist, stimulates immune cells and enhances expression of markers of leukocyte adhesion to endothelial cells. The above raises interesting questions about the effect of MRSV-Env on blood-brain barrier integrity (831). In parallel, HERV-W Env impairs remyelination by inhibiting the differentiation of oligodendrocyte precursors to myelin-producing oligodendrocytes, potentially due to nitrosative stress (832). The HERV-W glycoprotein syncytin-1 also seems to be implicated in MS via a similar mechanism; it causes an endoplasmic reticulum stress sensor to induce inducible nitric oxygen synthase (iNOS) and, concomitantly, the release of oligodendrocyte cytotoxins by astrocytes (for further details, see (833)). Also, HERVs can induce EAE in mice, implying a role upstream of other mediators (834). Therefore, HERVs seem to be implicated in both the neuroinflammatory and neurodegenerative component of the disease, rendering them promising therapeutic targets.

The autoimmune mechanism may also lie in the fact that common viruses (including but not limited to HSV-1, HHV-6, EBV, or influenza) can activate HERV proteins (816, 835) (Table 1). As a “dual infection”, EBV may be an exogenous and delayed cause for MS, with HERV-W acting as a precipitant (836). However, the mechanisms of transcriptional activation of HERVs are generally obscure, as are the downstream events in human cells. A general framework might be that HERVs and, more broadly, endogenous transposons act as a genomic defence response to external stimuli (837). Only a few studies have failed to find differences in the presence of HERV nucleic acids or antibodies between MS cases and controls (838, 839). In (840), no difference was detected in HERV-K113 levels between MS patients and healthy controls, but this study did not investigate the retroviral families most related to MS. However, other studies favour a relationship between HERVs and MS pathobiology. For instance, MOG shares similarities with five regions in the envelope protein (ERVWE2), with one region consisting of B and T cell epitopes capable of mediating antibody production and T cell function in vivo, respectively (841).

2.7 Human immunodeficiency virus (HIV)

Human immunodeficiency virus (HIV) is an exogenous retrovirus and HIV infection contributes to HERV activation, possibly via TLR-4 stimulation (842). This association is exemplified by post-mortem studies of brains from HIV patients and their epitope cross-reactivity in T cell responses to HIV (843). There have been, to our knowledge, less than twenty HIV cases reported that describe demyelinating CNS diseases including MS, with a disturbance in the CD8+ cytotoxic T cell and CD4+ T regulatory cell ratio implicated as causal (844, 845).
This rarity of documented cases of HIV and MS is consistent with the largest relevant record linkage study, in which HIV patients – all presumed to have undergone highly-active antiretroviral therapy (HAART) therapy – were at a statistically significant reduced risk (relative risk = 0.38)) for developing MS, with this relative risk including all recorded time intervals from first HIV record to the first MS record (846). One explanation for this finding could be that HIV-induced immunodeficiency is protective against MS. Alternatively, HAART usually employs competitive or non-competitive reverse transcriptase (RT) inhibitors and, due to suspected similarity between the HIV RTs and those of other viruses like HERVs, these inhibitors might suppress expression of the latter (847).

2.8 Cytomegalovirus (CMV)

The majority of epidemiological studies on CMV in MS are underpowered and inconclusive (99, 848). Two synchronous but different meta-analyses suggested a protective role for CMV seropositivity in MS (849, 850).

At the molecular level, it seems that CMV is present in the CNS including in some MS cases, but both exacerbating and protective roles are proposed (848). For example, CMV- and brain-specific B cells are correlated in MS patients (851), while concurrently CMV infection might indirectly exacerbate MS by inducing specific T cells with proinflammatory properties (848). Conversely, some studies have shown that higher anti-CMV antibody titres in MS patients are positively associated with improved MS-related neuroimaging and disability status markers (852). In addition, human CMV-induced natural killer cell expansion reduces the risk of disability progression in MS patients (853).

In animal models, cross-reactivity between human CMV peptide and MOG has been detected, while secondary CMV infection following vaccinia virus infection can worsen T cell autoreactivity and white matter lesions. In contrast, murine CMV infection prior to Theiler’s murine encephalomyelitis virus (TMEV) infection in the TMEV murine model of MS appears to improve symptoms both clinically (i.e., motor performance) and histologically (i.e., the severity of the inflammatory cell infiltrate) (854).

Finally, CMV (betaherpesvirinae subfamily) and EBV (gammaherpesvirinae subfamily) might oppose each other with regard to the downstream immune cascade (the so-called “immune response competition”), which might explain their inverse epidemiological patterns.
in MS (850). It has been also suggested that these herpesviridae viruses could both be required to elicit a “primate-specific autoimmune pathway” (848).

2.9 Measles and other morbilliviruses

The association between the measles virus and MS has been investigated for over fifty years, with MS postulated to be a host response to later measles infection. However, measles vaccination is not associated with MS, indicating that the measles virus is probably not connected with MS and supporting the evidence that measles vaccines are safe despite unjustified and well-publicised claims to the contrary (855). However, it is worth mentioning that two CNS complications of measles virus infection manifest with features of demyelination: acute disseminated encephalomyelitis (ADEM), a differential diagnosis of paediatric MS, and the very rare subacute sclerosing panencephalitis (SSPE) (856).

To our knowledge, recent research in this area has focused on the association between virus-specific CSF-to-serum antibody indices (AIs) and MS, not on virus detection using molecular techniques. The AIs for measles, rubella, and VZV, which form the “MRZ reaction” – high-specificity markers for “ruling-in” MS (reviewed in (857)) – are twofold higher than that for EBV (858). In particular, the measles AI is higher in patients with ≥6 lesions on MRI than those with fewer lesions in early MS (859). Another study showed that anti-measles virus antibody titers in the serum and CSF of MS patients increase according to the age and duration of the disease (860).

The phylogenetically-close rinderpest virus has not been shown to be demyelinating or even neurotropic in its ruminant hosts (856). In contrast, infection with the more distant canine distemper virus (CDV) causes CDV demyelinating leukoencephalitis and serves as an established animal model of MS. In that model, demyelinating lesions and initial and later phases are characterised by direct infection of astrocytes and excess inflammation with myelin loss, respectively (856, 861). Interestingly, axonal damage precedes demyelination, prompting questions on the role of inflammation and astrocytes as intermediate players (862).

2.10 Lymphocytic choriomeningitis virus (LCMV)

Lymphocytic choriomeningitis virus (LCMV) can affect the human CNS to cause paralysis and reduced consciousness. However, investigating its role in MS is more difficult due to low titers and short presence of LCMV in the CSF (863). In our opinion, this might indicate a “hit-and-
run” mechanism. On the other hand, recent in silico predictions show high sequence and structural similarity between LCMV’s nucleoprotein and specific myelin basic protein (MBP) residues (864).

Murine models of chronic LCMV infection have given rise to two Nobel Prizes (863). The virus is thought to activate microglia and astrocytes in the CNS via a TLR2-mediated cascade (865). Moreover, LCMV blocks induction of type-1 interferon and consequential upregulation of HLA class II. This observation supports a potential virus-induced disturbance in the interferon-tumour necrosis factor balance, which is already known to trigger autoimmunity (864).

Interestingly, LCMV infection limited to the periphery with concurrent CNS measles virus infection can induce CNS pathology via LCMV-specific CD8+ T cell recruitment to the brain without the need for LCMV replication. The underlying reason why the brain, broadly considered “immuno-privileged”, attracts these mis-recruited cells needs further exploration (866).

2.11 Coronavirus

In rodents, certain coronavirus-family mouse hepatitis virus strains are neurotropic, disrupt the blood-brain barrier, and cause immune-mediated demyelinating-like lesions (867). Human coronaviruses (HCoV) predominantly cause upper respiratory tract infections and are also neurotropic. Recent epidemiological studies are lacking, while molecular analyses have shown the HCoV-specific surface glycoprotein acts as a trigger for programmed cell death in a murine model of neurodegeneration. In addition, HCoV-229E/MBP cross-reactive T cells have been isolated from MS patients in single cell analyses, implying a molecular mimicry mechanism (for a review, see (868)). In a mouse model of encephalomyelitis/demyelination induced by glia-tropic murine coronavirus, the initial activation and accumulation of self-reactive CD4+ T cells was followed by a mechanism of host-mediated suppression that consequentially led to their decline, thus diminishing autoimmune phenotypes (869).

2.12 Saffold virus

Saffold virus (SAFV), a picornaviridae family member identified in 2007, was the first human virus in the Cardiovirus genus to be described (870). SAFV has a seroprevalence of over 90% in the adult population and is known to cause infection early in life (871). SAFV is associated
with both enteric and extra-intestinal diseases and, due to homology with TMEV, is implicated in MS (870).

However, its ubiquity has created difficulties in deciphering any association between SAFV and MS (750). SAFV was not detected in CSF samples from MS patients (870). One hypothesis is that SAFV might cause low-grade persistent infection followed by inflammation rather than act as a “hit-and-run” trigger for autoimmunity. However, a recent study failed to find any SAFV in MS brains and only rare SAFV-specific oligoclonal bands in MS patients and not different from controls (750). Limitations of existing data.

Several methodological issues could explain the described inconsistencies between studies in the MS-virus arena: (1) not choosing appropriate healthy matched controls following a specific study design but instead samples simply available at the time of study (i.e., an “opportunistic” approach); (2) even though quantification of viral load by real time-PCR is helpful, there seems to be a failure to use positive PCR or serology to distinguish active from latent infection (i.e., earlier infection during childhood in the case of serology). The enigmatic nature of MS poses challenges in the interpretation of the results since, according to some authors, detecting some antibodies under certain circumstances, i.e., in worsening MS, could be due to a hyperactivated immune system and not real infection (872); (3) conversely, interpreting absence of evidence of virus infection as evidence of absence, especially in the genomic era, may be a mistake (873).

2.13 Conclusions and suggestions for future research

There is, therefore, accumulated evidence that viruses may trigger or cause MS, with these organisms and the immune system interacting in several, potentially overlapping, ways. Deciphering the epidemiological contribution of viruses to MS along with their pathogenic mechanisms may help in the development of effective targeted therapies to develop vaccines, treat the disease, prevent relapses, and maintain remission.

Possible future research avenues include prospectively studying and monitoring carefully defined groups of patients, such as comparing patients with clinically isolated syndrome (CIS) who went on to convert to MS with those that did not. Although EBV has been studied in such cases, a broader causative role for viruses would be strengthened if any marker of viral presence (i.e., increased viral load and/or higher antiviral response) was observed in the first category. Furthermore, the B and T cell receptor repertoires in MS samples need to be fully characterised, preferably in relation to viral detection and burden and perhaps using newer high-
throughput technologies such as deep sequencing. This would be facilitated by the enrichment
of immunosequencing databases with extensive experimental data on the repertoires induced
by different human viruses. It would also be sensible to examine latent-to-lytic switching of
potentially existing viruses in MS biopsies. To complement previous efforts focusing on EBV-
specific markers (780), it would be interesting to analyse more recently proposed markers of
cellular antiviral response with respect to the above switch (874). Finally, given that many
viruses, not least EBV, express several proteins during different viral life cycle stages, the full
spectrum of antibody responses to viruses over their infective course needs further exploration,
perhaps using protein arrays methods for novel antigen discovery to overcome the limitations
of current techniques (875).
### Table 1. The association between environmental viruses and HERV elements and the downstream effects

<table>
<thead>
<tr>
<th>Virus</th>
<th>HERV element</th>
<th>Downstream effect</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSV-1</td>
<td>Matrix protein Gag protein</td>
<td>Oligodendrotoxic and immunopathogenic</td>
<td>Ruprecht et al., 2006 (876)</td>
</tr>
<tr>
<td>HHV-6</td>
<td>HERV-W Env and pol proteins</td>
<td>Synergy; interaction with HHV-6 U94/rep and DNA-pol</td>
<td>Nexo et al., 2016 (877); Perron et al., 2010 (878)</td>
</tr>
<tr>
<td>EBV</td>
<td>HERV-W genes</td>
<td>Increased HERV-W Env transcripts in PBMCs of IM patients; correlation of EBNA IgG levels with HERV gene expression levels in healthy, latently-infected individuals (i.e., with anti-EBNA-1 titers &gt;600)</td>
<td>Perron et al., 2010 (878)</td>
</tr>
<tr>
<td>EBV</td>
<td>HERV-K18 env protein</td>
<td>Endothelial permeability; Proinflammatory reactions</td>
<td>Tai et al., 2008 (879)</td>
</tr>
</tbody>
</table>
2.15 FIGURES

Figure 1. Putative mechanisms of human endogenous retrovirus (HERV)-related autoimmunity in Multiple Sclerosis.

(A) HERV-encoded RNAs with intact open reading frames (ORFs) can be translated into proteins. Some of these (e.g., HERV-K, HCML-ARV) are in close proximity to SNPs shown to be associated with MS in genome-wide association studies and representing genes involved in immune responses and vitamin D metabolism. (B) Some HERV proteins, notably MASP-3, HERV-H, and HERV-W, are expressed on the surface of normal cells including B cells. This serological response may be associated with autoimmunity, although causality has yet to be established. (C) HERVs are integral to the human genome but are epigenetically inactivated under normal conditions. HERV expression may be induced by environmental triggers including HSV-1, HHV-6, VZV, and EBV viruses to stimulate an immune response and autoimmunity. (D) The MRSV-Env protein has been identified in MS plaques and is brain-selective and immunopathogenic so may directly stimulate an autoimmune response (1). Furthermore, this protein inhibits differentiation of oligodendrocyte precursors so may have a negative feedback effect in the brains of MS patients (2).
3. Experimental analysis of CSF Microbiome and Virome and associated transcriptomics profiles

Multiple sclerosis (MS), a chronic inflammatory disorder of the brain and spinal cord, is the commonest cause of neurological disability in young adults; however, its aetiology remains enigmatic. An appealing theory has been that MS has an infectious aetiology, or that viruses and microbes may be environmental triggering factors for MS.

In this section, the primary goal was to analyze the microbiome and virome (i.e., the sum of viral genomic information) in the cerebro-spinal fluid (CSF) and gut of MS patients and controls using both microbial genomic/computational; I also investigated, at the gene expression level, the association between the hosts (patients with MS) and their CSF microbiome.

Interestingly, CSF, a protein-rich fluid which is anatomically closely to the brain interstitial fluid, is considered a proxy for brain tissue when tissue biopsies are unavailable (880). CSF is believed to provide comprehensive molecular information about processes in the brain (881); however, it requires an invasive lumbar puncture for sampling, and it is prone to contamination analysis. Because of this, it is essential to use several technical controls to avoid misinterpretations of findings (720). Case individuals can be further subdivided into three disease sub-categories: Sub-category 1: Primary progressive MS; Sub-category 2: Relapse-remitting MS; and, Sub-category 3: Other demyelination status (CIS, transverse myelitis). These sub-categories contain seven, five, and three samples, respectively, i.e., equal, or higher than the minimum accepted number of biological replicates (three) allowing meaningful comparisons between groups. Additionally, four negative technical control samples (with no RNA input) were processed using the same protocol.

A strength of this approach is its use of CSF samples from patients and controls, in whom lumbar puncture is performed as part of their diagnostic work-up, as well as, the use of two types of technical controls. NGS-based analyses have been successfully applied to CSF samples in several published studies (882) (883) (884); therefore, a body of literature supporting these investigations is already present.

3.1 Methods applied and tools used for the Next-generation sequencing experiments
All wet-lab experiments have been already performed in collaboration with the Johns Hopkins University School of Medicine Microarray and Deep Sequencing Facility on a Next-Seq instrument (Illumina) operating in High Output Mode to generate 2x150 bp paired-end reads, and 50 million reads for each sample, on two different time points (two samples and two technical controls as initial experimentation, followed by 19 samples and two technical controls). All chemistry and protocols were applied in a standard manner. Briefly, total RNA extraction was performed using appropriate kits (Qiagen, Hilden, Germany) following the manufacturer’s protocol for bodily fluid and for low-input clinical samples. The RNA extraction was followed by preparing a library using the Ovation Single Cell RNA-Seq System (NuGEN, San Carlos, CA) designed for low nucleic acid-content samples. In particular, next-generation RNA-Seq to determine the transcriptomes of bacteria, viruses, and other microbes has entailed a workflow, including among others: extraction of total RNA, reverse transcription to double stranded DNA, adapter ligation, clonal amplification, and high throughput sequencing. In parallel, standard operating procedures to avoid contamination have been applied throughout the entire workflow, and laboratory workers were blinded to the identity of control and experimental samples. In addition, samples were processed in random order, and clinical samples ran separately and not multi-plexed. Two types of no-template negative controls were processed identically and in parallel: one used to control for contaminants potentially introduced throughout the entire process, and the other for contaminants introduced only during library preparation steps.

3.2 Methods applied and tools used for the analysis of CSF virome: Description of commands used and modules

3.2.1 The pipeline

a. Downloading of 24 samples Quality control check, filtering (screening for adapters and phi x, rna spike-ins, quality trimming, and computationally removing human, cat, dog, mouse, common microbial lab contaminants, and ribosomal sequences)

b. Assembly using Megahit (v1.1.1)

c. Creation of Statistics

d. Clustering (de-replication)

e. Clustering with Isolate viruses and coverage (38,701 uniq contigs)

f. Blast against IMG/VR database (DNA and RNA)
g. Read Mapping against uniq contigs

3.2.2 Computer Code for Bioinformatics Analysis

Quality Control of reads and Metatranscriptomics assembly:

a. Interleave “.fastaq.gz” files

*Pre-requisites:
module load bbtools

*Script: reformat.sh

reformat.sh in=read1.fastq.gz in2=read2.fastq.gz out=reads.fastq.gz

b. Filter the interleave file to remove contaminants (screening for adapters and phi x, rna spike-ins, quality trimming, and removing human, cat, dog, mouse, common microbial lab contaminants, and ribosomal sequences)

*Pre-requisites:
set -e
set -o pipefail
module load blast+
module load bbtools

*Script: rqcfilter.sh

rqcfilter.sh in=<reads>.fastq.gz path=<script_location> rna=t trimfragadapter=t qtrim=t trimq=0 maxns=1 maq=10 minlen=51 mlf=0.33 phix=t filterk=25 removeribo=t removehuman=t removedog=t removecat=t removemouse=t khist=t removemicrobes=t mtst=t sketch outribo==<alexios_samples>_cont.fastq.gz clumpify=t barcodefilter=t > filter.log 2>&1

c. Running Megahit on the filtered outputs (i.e. “.anqrpht.fastq.gz” files)

*Pre-requisites:
module load megahit/1.1.1

i. *Script: megahit

megahit --k-list 23,43,63,83,103,123 --12 INPUT_interleaved_input.anqrpht.fastq.gz

TOTAL Scaffolds assembled: 46,873 (breakdown per sample and statistics in Table “Echantillon List”)

d. Run stats on megahit outputs (i.e. “final.contigs.fa” files)

*Pre-requisites:
Module load bbtools

*Script: stats.sh

stats.sh assembled_sample_final.contigs.fa

3.2.2.1 Clustering

I blasted all-vs-all assembled scaffolds (46,873), removed self-hits, and clustered the sequences into groups based on a 90% identity over 75% length. This has helped to explore the same sequences (members of the same cluster) across multiple samples. This way one can explore sequences only found in controls vs only found in diseased samples. I clustered 9,231 scaffolds into 487 clusters, and 37,642 scaffolds remained singletons (clustering table). Total diversity of scaffolds is: (487+37,642) =38,129. The longest representative of each cluster was selected as seed per group for further analyses.

3.2.2.2 Blast against databases

blastn -query <sample>_final.contigs.fa -db <DNA or RNA db> -outfmt '6 std qlen slen' -out <output>_final.contigs.blout -evalue 1.0e-08 -perc_identity 85

3.2.2.3 Read mapping

All read mapping works followed the same steps.

I used bowtie2 to create and index our reference dataset:

bowtie2-build <references_dereplicated> reference_dereplicated_index

3.2.2.4 Aligning reads to the indexed database (creating a SAM file):

bowtie2 -x reference_dereplicated_index -1 <reads_1.fq> -2 <reads_2.fq> -S eg2.sam

Converting the SAM file into a coverage table using pileup.sh script from the BBMap Suite tools:

module load bbtools

pileup.sh in=<SAM> out=<output>
3.2.3 Metagenomic Results of the pipeline

Blast against DNA virus dataset

43 hits to a viral contig (human buccal mucosa). From most samples and NO controls.

Streptococcus mitis

Blast output
Read mapping

Top 10k contigs with highest coverage overall
Read mapping only diseased samples

Contigs present across ALL diseased samples AND no present across controls
Global read mapping

- Index IMG/VR database
- Read mapping of all samples vs. all IMG/VR
  - Suspicious viral contig covered ~40-60% with average fold (~20-1,300 X) on all samples checked.

Human bile duct microbial communities from gallstone patients in Hangzhou, China – B4 bile
3.2.4 Host RNA-Seq results

3.2.4.1 Up-regulated KEGG pathway in Multiple Sclerosis vs. Healthy Controls

<table>
<thead>
<tr>
<th>Term</th>
<th>Count</th>
<th>%</th>
<th>PValue</th>
</tr>
</thead>
<tbody>
<tr>
<td>hsa00190: Oxidative phosphorylation</td>
<td>19</td>
<td>2.60989011</td>
<td>1.01E-05</td>
</tr>
<tr>
<td>hsa05012: Parkinson's disease</td>
<td>17</td>
<td>2.335164835</td>
<td>2.89E-04</td>
</tr>
<tr>
<td>hsa03018: RNA degradation</td>
<td>12</td>
<td>1.648351648</td>
<td>3.41E-04</td>
</tr>
<tr>
<td>hsa04380: Osteoclast differentiation</td>
<td>15</td>
<td>2.06043956</td>
<td>0.001163013</td>
</tr>
<tr>
<td>hsa04932: Non-alcoholic fatty liver disease (NAFLD)</td>
<td>16</td>
<td>2.197802198</td>
<td>0.001655485</td>
</tr>
<tr>
<td>hsa05010: Alzheimer's disease</td>
<td>17</td>
<td>2.335164835</td>
<td>0.001823574</td>
</tr>
<tr>
<td>hsa03040: Spliceosome</td>
<td>14</td>
<td>1.923076923</td>
<td>0.003823765</td>
</tr>
<tr>
<td>hsa04660: T cell receptor signaling pathway</td>
<td>12</td>
<td>1.648351648</td>
<td>0.00386647</td>
</tr>
<tr>
<td>hsa04145: Phagosome</td>
<td>15</td>
<td>2.06043956</td>
<td>0.004964701</td>
</tr>
<tr>
<td>hsa05016: Huntington's disease</td>
<td>17</td>
<td>2.335164835</td>
<td>0.006847995</td>
</tr>
<tr>
<td>hsa05140: Leishmaniasis</td>
<td>9</td>
<td>1.236263736</td>
<td>0.009534325</td>
</tr>
<tr>
<td>hsa05166: HTLV-I infection</td>
<td>20</td>
<td>2.747252747</td>
<td>0.011231663</td>
</tr>
<tr>
<td>hsa04141: Protein processing in endoplasmic reticulum</td>
<td>15</td>
<td>2.06043956</td>
<td>0.01163839</td>
</tr>
<tr>
<td>hsa04120: Ubiquitin mediated proteolysis</td>
<td>13</td>
<td>1.785714286</td>
<td>0.012515233</td>
</tr>
<tr>
<td>hsa04130: SNARE interactions in vesicular transport</td>
<td>6</td>
<td>0.824175824</td>
<td>0.013058032</td>
</tr>
<tr>
<td>hsa05211: Renal cell carcinoma</td>
<td>8</td>
<td>1.098901099</td>
<td>0.018820105</td>
</tr>
<tr>
<td>hsa04664: Fc epsilon RI signaling pathway</td>
<td>8</td>
<td>1.098901099</td>
<td>0.023571023</td>
</tr>
</tbody>
</table>

Institutional Repository - Library & Information Centre - University of Thessaly
01/11/2023 00:13:51 EET - 35.160.27.221
<table>
<thead>
<tr>
<th>Human Pathway ID</th>
<th>Pathway Description</th>
<th>Disease Score</th>
<th>p-Value</th>
<th>q-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>hsa05150</td>
<td>Staphylococcus aureus infection</td>
<td>7</td>
<td>0.961538462</td>
<td>0.025070024</td>
</tr>
<tr>
<td>hsa00620</td>
<td>Pyruvate metabolism</td>
<td>6</td>
<td>0.824175824</td>
<td>0.025185125</td>
</tr>
<tr>
<td>hsa04662</td>
<td>B cell receptor signaling pathway</td>
<td>8</td>
<td>1.098901099</td>
<td>0.025322969</td>
</tr>
<tr>
<td>hsa05323</td>
<td>Rheumatoid arthritis</td>
<td>9</td>
<td>1.236263736</td>
<td>0.031312175</td>
</tr>
<tr>
<td>hsa01100</td>
<td>Metabolic pathways</td>
<td>65</td>
<td>8.928571429</td>
<td>0.034873133</td>
</tr>
<tr>
<td>hsa05152</td>
<td>Tuberculosis</td>
<td>14</td>
<td>1.923076923</td>
<td>0.035591939</td>
</tr>
<tr>
<td>hsa05310</td>
<td>Asthma</td>
<td>5</td>
<td>0.686813187</td>
<td>0.035637559</td>
</tr>
<tr>
<td>hsa05168</td>
<td>Herpes simplex infection</td>
<td>14</td>
<td>1.923076923</td>
<td>0.044657696</td>
</tr>
<tr>
<td>hsa01130</td>
<td>Biosynthesis of antibiotics</td>
<td>15</td>
<td>2.06043956</td>
<td>0.062258461</td>
</tr>
<tr>
<td>hsa04666</td>
<td>Fc gamma R-mediated phagocytosis</td>
<td>8</td>
<td>1.098901099</td>
<td>0.062852239</td>
</tr>
<tr>
<td>hsa05164</td>
<td>Influenza A</td>
<td>13</td>
<td>1.785714286</td>
<td>0.062861056</td>
</tr>
<tr>
<td>hsa04722</td>
<td>Neurotrophin signaling pathway</td>
<td>10</td>
<td>1.373626374</td>
<td>0.065673811</td>
</tr>
<tr>
<td>hsa00310</td>
<td>Lysine degradation</td>
<td>6</td>
<td>0.824175824</td>
<td>0.06648168</td>
</tr>
<tr>
<td>hsa04010</td>
<td>MAPK signaling pathway</td>
<td>17</td>
<td>2.335164835</td>
<td>0.070359357</td>
</tr>
<tr>
<td>hsa04650</td>
<td>Natural killer cell mediated cytotoxicity</td>
<td>10</td>
<td>1.373626374</td>
<td>0.071331112</td>
</tr>
<tr>
<td>hsa04668</td>
<td>TNF signaling pathway</td>
<td>9</td>
<td>1.236263736</td>
<td>0.077667119</td>
</tr>
<tr>
<td>hsa05133</td>
<td>Pertussis</td>
<td>7</td>
<td>0.961538462</td>
<td>0.095160276</td>
</tr>
<tr>
<td>hsa04612</td>
<td>Antigen processing and presentation</td>
<td>7</td>
<td>0.961538462</td>
<td>0.099878244</td>
</tr>
</tbody>
</table>
### 3.2.4.2 Down-regulated KEGG pathway in Multiple Sclerosis vs. Healthy Controls

<table>
<thead>
<tr>
<th>Term</th>
<th>Count</th>
<th>%</th>
<th>PValue</th>
</tr>
</thead>
<tbody>
<tr>
<td>hsa00190: Oxidative phosphorylation</td>
<td>10</td>
<td>2,457002457</td>
<td>4.76E-04</td>
</tr>
<tr>
<td>hsa05012: Parkinson's disease</td>
<td>10</td>
<td>2,457002457</td>
<td>7.66E-04</td>
</tr>
<tr>
<td>hsa04062: Chemokine signaling pathway</td>
<td>9</td>
<td>2,211302211</td>
<td>0.015340838</td>
</tr>
<tr>
<td>hsa05016: Huntington's disease</td>
<td>9</td>
<td>2,211302211</td>
<td>0.018245345</td>
</tr>
<tr>
<td>hsa03040: Spliceosome</td>
<td>7</td>
<td>1,71990172</td>
<td>0.027845189</td>
</tr>
<tr>
<td>hsa04932: Non-alcoholic fatty liver disease (NAFLD)</td>
<td>7</td>
<td>1,71990172</td>
<td>0.047187443</td>
</tr>
<tr>
<td>hsa05010: Alzheimer's disease</td>
<td>7</td>
<td>1,71990172</td>
<td>0.071653922</td>
</tr>
<tr>
<td>hsa05202: Transcriptional misregulation in cancer</td>
<td>7</td>
<td>1,71990172</td>
<td>0.071653922</td>
</tr>
</tbody>
</table>
### 3.2.4.3 Gene Ontology – Up-regulated genes

<table>
<thead>
<tr>
<th>Term</th>
<th>Count</th>
<th>%</th>
<th>PValue</th>
</tr>
</thead>
<tbody>
<tr>
<td>GO:0006397–mRNA processing</td>
<td>18</td>
<td>2.47</td>
<td>3.76E-04</td>
</tr>
<tr>
<td>GO:0006120–mitochondrial electron transport, NADH to ubiquinone</td>
<td>9</td>
<td>1.23</td>
<td>4.08E-04</td>
</tr>
<tr>
<td>GO:0032981–mitochondrial respiratory chain complex I assembly</td>
<td>10</td>
<td>1.37</td>
<td>5.02E-04</td>
</tr>
<tr>
<td>GO:0000398–mRNA splicing, via spliceosome</td>
<td>20</td>
<td>2.74</td>
<td>6.44E-04</td>
</tr>
<tr>
<td>GO:0042776–mitochondrial ATP synthesis coupled proton transport</td>
<td>6</td>
<td>0.82</td>
<td>8.65E-04</td>
</tr>
<tr>
<td>GO:0038095–Fc-epsilon receptor signaling pathway</td>
<td>17</td>
<td>2.33</td>
<td>1.00E-04</td>
</tr>
<tr>
<td>GO:0006479–protein methylation</td>
<td>6</td>
<td>0.82</td>
<td>0.00134481</td>
</tr>
<tr>
<td>GO:0050852–T cell receptor signaling pathway</td>
<td>14</td>
<td>1.92</td>
<td>0.003530977</td>
</tr>
<tr>
<td>GO:0008380–RNA splicing</td>
<td>15</td>
<td>2.06</td>
<td>0.003645057</td>
</tr>
<tr>
<td>GO:0006754–ATP biosynthetic process</td>
<td>6</td>
<td>0.82</td>
<td>0.003948732</td>
</tr>
<tr>
<td>GO:0043928–exonucleolytic nuclear-transcribed mRNA catabolic process involved in deadenylation-dependent decay</td>
<td>6</td>
<td>0.82</td>
<td>0.003948732</td>
</tr>
<tr>
<td>GO:0006338–chromatin remodeling</td>
<td>10</td>
<td>1.37</td>
<td>0.004596958</td>
</tr>
<tr>
<td>GO:006974–cellular response to DNA damage stimulus</td>
<td>17</td>
<td>2.33</td>
<td>0.004862396</td>
</tr>
<tr>
<td>GO:0007005–mitochondrion organization</td>
<td>9</td>
<td>1.23</td>
<td>0.007685108</td>
</tr>
<tr>
<td>GO:0048679–regulation of axon regeneration</td>
<td>3</td>
<td>0.41</td>
<td>0.007869699</td>
</tr>
<tr>
<td>GO:0016192–vesicle-mediated transport</td>
<td>13</td>
<td>1.78</td>
<td>0.011290196</td>
</tr>
<tr>
<td>GO:0070534–protein K63-linked ubiquitination</td>
<td>6</td>
<td>0.82</td>
<td>0.011365643</td>
</tr>
<tr>
<td>GO:0071569–protein ufmylation</td>
<td>3</td>
<td>0.41</td>
<td>0.01279426</td>
</tr>
<tr>
<td>GO:0043488–regulation of mRNA stability</td>
<td>10</td>
<td>1.37</td>
<td>0.014456084</td>
</tr>
<tr>
<td>GO:0002223</td>
<td>stimulatory C-type lectin receptor signaling pathway</td>
<td>10</td>
<td>3.73626</td>
</tr>
<tr>
<td>GO:0043968</td>
<td>histone H2A acetylation</td>
<td>4</td>
<td>0.549451</td>
</tr>
<tr>
<td>GO:1904874</td>
<td>positive regulation of telomerase RNA localization to Cajal body</td>
<td>4</td>
<td>0.549451</td>
</tr>
<tr>
<td>GO:0006301</td>
<td>postreplication repair</td>
<td>4</td>
<td>0.549451</td>
</tr>
<tr>
<td>GO:0043001</td>
<td>Golgi to plasma membrane protein transport</td>
<td>5</td>
<td>0.686813</td>
</tr>
<tr>
<td>GO:0070979</td>
<td>protein K11-linked ubiquitination</td>
<td>5</td>
<td>0.686813</td>
</tr>
<tr>
<td>GO:0035493</td>
<td>SNARE complex assembly</td>
<td>3</td>
<td>0.412088</td>
</tr>
<tr>
<td>GO:0046939</td>
<td>nucleotide phosphorylation</td>
<td>3</td>
<td>0.412088</td>
</tr>
<tr>
<td>GO:0061418</td>
<td>regulation of transcription from RNA polymerase II promoter in response to hypoxia</td>
<td>5</td>
<td>0.686813</td>
</tr>
<tr>
<td>GO:0043066</td>
<td>negative regulation of apoptotic process</td>
<td>27</td>
<td>3.708791</td>
</tr>
<tr>
<td>GO:0006915</td>
<td>apoptotic process</td>
<td>32</td>
<td>4.395604</td>
</tr>
<tr>
<td>GO:0045429</td>
<td>positive regulation of nitric oxide biosynthetic process</td>
<td>6</td>
<td>0.824176</td>
</tr>
<tr>
<td>GO:0016236</td>
<td>macroautophagy</td>
<td>8</td>
<td>1.098901</td>
</tr>
<tr>
<td>GO:0032508</td>
<td>DNA duplex unwinding</td>
<td>6</td>
<td>0.824176</td>
</tr>
<tr>
<td>GO:0061025</td>
<td>membrane fusion</td>
<td>6</td>
<td>0.824176</td>
</tr>
<tr>
<td>GO:006376</td>
<td>mRNA splice site selection</td>
<td>4</td>
<td>0.549451</td>
</tr>
<tr>
<td>GO:0050860</td>
<td>negative regulation of T cell receptor signaling pathway</td>
<td>4</td>
<td>0.549451</td>
</tr>
<tr>
<td>GO:0044344</td>
<td>cellular response to fibroblast growth factor stimulus</td>
<td>5</td>
<td>0.686813</td>
</tr>
<tr>
<td>GO:0045454</td>
<td>cell redox homeostasis</td>
<td>8</td>
<td>1.098901</td>
</tr>
<tr>
<td>GO:0031295</td>
<td>T cell costimulation</td>
<td>8</td>
<td>1.098901</td>
</tr>
<tr>
<td>GO:0006412</td>
<td>translation</td>
<td>17</td>
<td>2.335165</td>
</tr>
<tr>
<td>GO ID</td>
<td>Description</td>
<td>Count</td>
<td>Adjusted p-value</td>
</tr>
<tr>
<td>--------------------</td>
<td>-------------------------------------------------------------------------------</td>
<td>-------</td>
<td>------------------</td>
</tr>
<tr>
<td>GO:0031668</td>
<td>cellular response to extracellular stimulus</td>
<td>4</td>
<td>0.549451</td>
</tr>
<tr>
<td>GO:0006413</td>
<td>translational initiation</td>
<td>11</td>
<td>1.510989</td>
</tr>
<tr>
<td>GO:0006909</td>
<td>phagocytosis</td>
<td>6</td>
<td>0.824176</td>
</tr>
<tr>
<td>GO:2001244</td>
<td>positive regulation of intrinsic apoptotic signaling pathway</td>
<td>5</td>
<td>0.686813</td>
</tr>
<tr>
<td>GO:1903071</td>
<td>positive regulation of ER-associated ubiquitin-dependent protein catabolic process</td>
<td>3</td>
<td>0.412088</td>
</tr>
<tr>
<td>GO:0071499</td>
<td>cellular response to laminar fluid shear stress</td>
<td>3</td>
<td>0.412088</td>
</tr>
<tr>
<td>GO:0015992</td>
<td>proton transport</td>
<td>6</td>
<td>0.824176</td>
</tr>
<tr>
<td>GO:0006457</td>
<td>protein folding</td>
<td>13</td>
<td>1.785714</td>
</tr>
<tr>
<td>GO:007266</td>
<td>Rho protein signal transduction</td>
<td>6</td>
<td>0.824176</td>
</tr>
<tr>
<td>GO:0050900</td>
<td>leukocyte migration</td>
<td>10</td>
<td>1.373626</td>
</tr>
<tr>
<td>GO:0071404</td>
<td>cellular response to low-density lipoprotein particle stimulus</td>
<td>3</td>
<td>0.412088</td>
</tr>
<tr>
<td>GO:0048096</td>
<td>chromatin-mediated maintenance of transcription</td>
<td>3</td>
<td>0.412088</td>
</tr>
<tr>
<td>GO:0000122</td>
<td>negative regulation of transcription from RNA polymerase II promoter</td>
<td>37</td>
<td>5.082418</td>
</tr>
<tr>
<td>GO:0006366</td>
<td>transcription from RNA polymerase II promoter</td>
<td>28</td>
<td>3.846154</td>
</tr>
<tr>
<td>GO:1900087</td>
<td>positive regulation of G1/S transition of mitotic cell cycle</td>
<td>4</td>
<td>0.549451</td>
</tr>
<tr>
<td>GO:0015986</td>
<td>ATP synthesis coupled proton transport</td>
<td>4</td>
<td>0.549451</td>
</tr>
<tr>
<td>GO:0007162</td>
<td>negative regulation of cell adhesion</td>
<td>5</td>
<td>0.686813</td>
</tr>
<tr>
<td>GO:0061077</td>
<td>chaperone-mediated protein folding</td>
<td>5</td>
<td>0.686813</td>
</tr>
<tr>
<td>GO:0060333</td>
<td>interferon-gamma-mediated signaling pathway</td>
<td>7</td>
<td>0.961538</td>
</tr>
<tr>
<td>GO:0051436</td>
<td>negative regulation of ubiquitin-protein ligase activity involved in mitotic cell cycle</td>
<td>7</td>
<td>0.961538</td>
</tr>
<tr>
<td>GO:2000406</td>
<td>positive regulation of T cell migration</td>
<td>3</td>
<td>0.412088</td>
</tr>
<tr>
<td>GO:0001731</td>
<td>formation of translation preinitiation complex</td>
<td>4</td>
<td>0.549451</td>
</tr>
<tr>
<td>GO:0006281</td>
<td>DNA repair</td>
<td>15</td>
<td>2.06044</td>
</tr>
<tr>
<td>GO:0051602</td>
<td>response to electrical stimulus</td>
<td>4</td>
<td>0.549451</td>
</tr>
<tr>
<td>GO:0034976</td>
<td>response to endoplasmic reticulum stress</td>
<td>7</td>
<td>0.961538</td>
</tr>
<tr>
<td>GO:0010941</td>
<td>regulation of cell death</td>
<td>3</td>
<td>0.412088</td>
</tr>
<tr>
<td>GO:0001886</td>
<td>endothelial cell morphogenesis</td>
<td>3</td>
<td>0.412088</td>
</tr>
<tr>
<td>GO:0032786</td>
<td>positive regulation of DNA-templated transcription, elongation</td>
<td>3</td>
<td>0.412088</td>
</tr>
<tr>
<td>GO:0051437</td>
<td>positive regulation of ubiquitin-protein ligase activity involved in regulation of mitotic cell cycle transition</td>
<td>7</td>
<td>0.961538</td>
</tr>
<tr>
<td>GO:0015949</td>
<td>nucleobase-containing small molecule interconversion</td>
<td>4</td>
<td>0.549451</td>
</tr>
<tr>
<td>GO:0030168</td>
<td>platelet activation</td>
<td>9</td>
<td>1.236264</td>
</tr>
<tr>
<td>GO:0050821</td>
<td>protein stabilization</td>
<td>10</td>
<td>1.373626</td>
</tr>
<tr>
<td>GO:0010629</td>
<td>negative regulation of gene expression</td>
<td>10</td>
<td>1.373626</td>
</tr>
<tr>
<td>GO:0008286</td>
<td>insulin receptor signaling pathway</td>
<td>7</td>
<td>0.961538</td>
</tr>
<tr>
<td>GO:000245</td>
<td>spliceosomal complex assembly</td>
<td>4</td>
<td>0.549451</td>
</tr>
<tr>
<td>GO:0031629</td>
<td>synaptic vesicle fusion to presynaptic active zone membrane</td>
<td>3</td>
<td>0.412088</td>
</tr>
<tr>
<td>GO:0006164</td>
<td>purine nucleotide biosynthetic process</td>
<td>3</td>
<td>0.412088</td>
</tr>
<tr>
<td>GO:0042102</td>
<td>positive regulation of T cell proliferation</td>
<td>6</td>
<td>0.824176</td>
</tr>
<tr>
<td>GO:1990418</td>
<td>response to insulin-like growth factor stimulus</td>
<td>2</td>
<td>0.274725</td>
</tr>
<tr>
<td>GO:0060502</td>
<td>epithelial cell proliferation involved in lung morphogenesis</td>
<td>2</td>
<td>0.274725</td>
</tr>
<tr>
<td>GO:0042590</td>
<td>antigen processing and presentation of exogenous peptide antigen via MHC class I</td>
<td>2</td>
<td>0.274725</td>
</tr>
<tr>
<td>GO:0031145</td>
<td>anaphase-promoting complex-dependent catabolic process</td>
<td>7</td>
<td>0.961538</td>
</tr>
</tbody>
</table>
GO:1904837~beta-catenin-TCF complex assembly 5 0.686813 0.074591852
GO:0006888~ER to Golgi vesicle-mediated transport 11 1.510989 0.075397169
GO:0043161~proteasome-mediated ubiquitin-dependent protein catabolic process 13 1.785714 0.076675497
GO:0035264~multicellular organism growth 7 0.961538 0.076914282
GO:0032008~positive regulation of TOR signaling 4 0.549451 0.07727605
GO:0006635~fatty acid beta-oxidation 5 0.686813 0.079758027
GO:0045892~negative regulation of transcription, DNA-templated 26 3.571429 0.079868324
GO:0006977~DNA damage response, signal transduction by p53 class mediator resulting in cell cycle arrest 6 0.824176 0.080056505
GO:0000956~nuclear-transcribed mRNA catabolic process 3 0.412088 0.082042408
GO:0043518~negative regulation of DNA damage response, signal transduction by p53 class mediator 3 0.412088 0.082042408
GO:0002479~antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-dependent 6 0.824176 0.084443746
GO:0032436~positive regulation of proteasomal ubiquitin-dependent protein catabolic process 6 0.824176 0.088955086
GO:0006369~termination of RNA polymerase II transcription 6 0.824176 0.088955086
GO:0001701~in utero embryonic development 12 1.648352 0.089745312
GO:0010942~positive regulation of cell death 4 0.549451 0.091451753
GO:0030307~positive regulation of cell growth 7 0.961538 0.092246972
GO:0000380~alternative mRNA splicing, via spliceosome 3 0.412088 0.093435419
GO:0030490~maturation of SSU-rRNA 3 0.412088 0.093435419
GO:0042110~T cell activation 5 0.686813 0.096280159
GO:0051017~actin filament bundle assembly 4 0.549451 0.098941377
### 3.2.4.4 Gene Ontology – Down-regulated genes

<table>
<thead>
<tr>
<th>Term</th>
<th>Count</th>
<th>%</th>
<th>PValue</th>
</tr>
</thead>
<tbody>
<tr>
<td>GO:0006351–transcription, DNA-templated</td>
<td>57</td>
<td>14.00491</td>
<td>5.78E-04</td>
</tr>
<tr>
<td>GO:0030154–cell differentiation</td>
<td>20</td>
<td>4.914005</td>
<td>0.001045099</td>
</tr>
<tr>
<td>GO:0006283–transcription-coupled nucleotide-excision repair</td>
<td>7</td>
<td>1.719902</td>
<td>0.002478516</td>
</tr>
<tr>
<td>GO:0036498–IRE1-mediated unfolded protein response</td>
<td>6</td>
<td>1.474201</td>
<td>0.004667171</td>
</tr>
<tr>
<td>GO:0006457–protein folding</td>
<td>9</td>
<td>2.211302</td>
<td>0.01915878</td>
</tr>
<tr>
<td>GO:0048147–negative regulation of fibroblast proliferation</td>
<td>4</td>
<td>0.982801</td>
<td>0.019283613</td>
</tr>
<tr>
<td>GO:0070125–mitochondrial translational elongation</td>
<td>6</td>
<td>1.474201</td>
<td>0.020782375</td>
</tr>
<tr>
<td>GO:0030335–positive regulation of cell migration</td>
<td>9</td>
<td>2.211302</td>
<td>0.021585844</td>
</tr>
<tr>
<td>GO:0070126–mitochondrial translational termination</td>
<td>6</td>
<td>1.474201</td>
<td>0.021742178</td>
</tr>
<tr>
<td>GO:1902600–hydrogen ion transmembrane transport</td>
<td>5</td>
<td>1.228501</td>
<td>0.026385575</td>
</tr>
<tr>
<td>GO:0030490–maturation of SSU-rRNA</td>
<td>3</td>
<td>0.737101</td>
<td>0.026865161</td>
</tr>
<tr>
<td>GO:0007275–multicellular organism development</td>
<td>17</td>
<td>4.176904</td>
<td>0.034121263</td>
</tr>
<tr>
<td>GO:0007568–aging</td>
<td>8</td>
<td>1.965602</td>
<td>0.034192575</td>
</tr>
<tr>
<td>GO:0006940–regulation of smooth muscle contraction</td>
<td>3</td>
<td>0.737101</td>
<td>0.034579048</td>
</tr>
<tr>
<td>GO:0009890–negative regulation of biosynthetic process</td>
<td>2</td>
<td>0.4914</td>
<td>0.036699514</td>
</tr>
<tr>
<td>GO:1903894–regulation of IRE1-mediated unfolded protein response</td>
<td>2</td>
<td>0.4914</td>
<td>0.036699514</td>
</tr>
<tr>
<td>GO:0006355–regulation of transcription, DNA-templated</td>
<td>38</td>
<td>9.336609</td>
<td>0.045715234</td>
</tr>
<tr>
<td>GO:0071880–adenylate cyclase-activating adrenergic receptor signaling pathway</td>
<td>3</td>
<td>0.737101</td>
<td>0.047524403</td>
</tr>
<tr>
<td>GO:0060828–regulation of canonical Wnt signaling pathway</td>
<td>3</td>
<td>0.737101</td>
<td>0.052173986</td>
</tr>
</tbody>
</table>
GO:0003151~outflow tract morphogenesis
GO:0000398~mRNA splicing, via spliceosome
GO:0045600~positive regulation of fat cell differentiation
GO:0035914~skeletal muscle cell differentiation
GO:0007420~brain development
GO:0090398~cellular senescence
GO:0000715~nucleotide-excision repair, DNA damage recognition
GO:0071883~activation of MAPK activity by adrenergic receptor signaling pathway
GO:0032811~negative regulation of epinephrine secretion
GO:0035625~epidermal growth factor-activated receptor transactivation by G-protein coupled receptor signaling pathway
GO:0000122~negative regulation of transcription from RNA polymerase II promoter
GO:1901796~regulation of signal transduction by p53 class mediator
GO:0046487~glyoxylate metabolic process
GO:0032148~activation of protein kinase B activity
GO:0032355~response to estradiol
GO:0032091~negative regulation of protein binding
GO:0018027~peptidyl-lysine dimethylation
GO:0009653~anatomical structure morphogenesis
GO:0001569~patterning of blood vessels
GO:0018146~keratan sulfate biosynthetic process
3.2.4.5 Volcano Plot
3.2.4.6 Differential Expression Gene Distribution

DEG distribution across chromosomes

DEG distribution across biotypes
3.2.4.7 Boxplot before Normalization
3.2.4.8 Correlation of expression between different samples (Correlogram)
3.2.4.9 Filtered genes per chromosome and biotype
3.2.4.10 MDS Plot

![MDS Plot Image]

Institutional Repository - Library & Information Centre - University of Thessaly
01/11/2023 00:13:51 EET - 35.160.27.221
3.2.4.11 RNA-Seq Reads Noise
3.2.5 Interpretation of CSF Virome and RNA-Expression results

My initial bioinformatics analysis of 21 samples from MS vs. healthy controls support the feasibility and power of the bioinformatics approach, and it suggests (based on *de novo* assembly of viruses through MEGAHIIT, followed by hierarchical clustering) that a *distinct viral pattern* between MS cases and healthy controls exists but remains to be fully deciphered, alongside with concrete gene expression patterns between MS cases and healthy controls. Future studies should aim to decipher more thoroughly the role of specific viruses in these samples, as well as, the major biological pathways implicated in MS sub-categories under investigation.
Chapter 5 — The ELSI (Ethical, Legal, and Social Implications) of Neurological and Biomedical Research: A critical viewpoint

1. Neurological and Biomedical Research in times of social and financial austerity

1.1 Biomedical Research: Lessons from the Last Decade’s Crisis and Austerity-stricken Small Countries to the Current COVID-19-related Crisis

Many countries were afflicted by the most recent, decade-long, financial crisis and its accompanying austerity measures. In Greece, Spain, Portugal, and others, funding scarcity has greatly impeded the performance of especially expensive biomedical research. The latter field was particularly hit, because it took place while there was, at the same period, an explosion of costly, resource-expensive studies on biological pathways, precision medicine, big-data science, super-resolution imaging, robotics, and high-throughput experimental technologies.

There are several long-standing programs that support research in low- and middle-income countries. For instance, such countries could benefit from the Research4Life programs AGORA, HINARI, OARE, ARDI and GOALI, or could be entitled to request waivers of full tuition fees for their graduate students in leading foreign academic institutions. These countries face fundamental difficulties of their own, and such programs are sorely needed. In contrast, when small high-income developed countries are stricken by decreases of their gross domestic product (GDP), they cannot benefit from the developmental policies and remedial programs available to developing countries. Therefore, they could be fairly described as research resource-poor countries.

Should developed but crisis-stricken countries receive an extramural research boost as an act of solidarity and science diplomacy, or, instead, consider themselves a lost case? Using our country, Greece, as an example, we argue for a third alternative, one that many other countries around the world would find applicable now and in the future: to look locally for attractive niches and hidden pearls of added value to global biomedical research (Tables 1 and 2). This notion becomes more important during the post-austerity years, considering Greece’s situation as part of a broader policy gap, in the sense that while a remedial mode for middle- and low-income countries exists, there is not one suitable for small high-income, research resource-poor countries. Such support is especially important for Greece and many other equally unprepared or not well-equipped countries, not only as a consequence of the recent austerity period, but also because another economic breakdown will likely follow the current
COVID-19 pandemic, likely imposing fiscal budget shifts from the research (oft-considered luxurious in epidemics, unless epidemic-related) to the health sector.

Greece, –like many other countries, – has become a place of remarkable contradiction. It has a high number of well-trained scientists, but the vast majority of them are unemployed, under-employed or seeking work abroad. Policy makers have promised to address several long-standing structural problems in research funding,(885) yet, the total spending on R&D remains quite limited(886). In particular, research funding in Greece has been consistently limited, especially for basic research, and of note, unequally distributed, because of: a) barely existent structural changes oriented towards innovation; b) lack of matching between the university curricula and the skills required by the generally meager industry; c) indecisive policies on research priorities (leading, in turn, to disintegration of strong research groups); d) not considering cost-effectiveness during policy formulation; e) a vast, albeit improving bureaucracy; f) high corruption indices (inversely linked to innovation performance); g) lagging behind in alignment with the Sustainable Development Agenda and World Health Organization Goals (e.g., supporting research on traditional, environmentally burdensome energy forms instead of long-term planning for the upcoming climate crisis); and, h) frequent political party-guided unmeritocratic science policies. These conditions intensified during the austerity years, leading research performance to even lower capacities (661, 887-893).

Despite the above conditions, Greece has managed to survive scientifically and score high in terms of academic productivity of its scientists, –such as publication metrics and citations– in part because of the strong mindset of its human capital, namely its stamina for research despite adversities and complying with the concept of the Greek philotimon, describing a set of several human virtues. This is in contrast to a relative paucity in the filing for new patents, based on the locally prevailing idea of considering science purely as a search for the truth that is incompatible with generation of personal profit.

In light of such traits, we should recognize that research resource-poor countries, by their own nature, cannot contribute equally to highly complicated, costly projects requiring advanced expensive experimentation; thus, mutual respect and appropriately balanced acknowledgement of intellectual properties and co-authorships in inter-country collaborations are crucial for success. Treating these countries through an equity rather than an equality perspective is warranted, let alone during epidemics. Furthermore, diversity issues embracing inclusion of resource-poor countries in the eligibility criteria for major international grant proposals should be advocated. Bold ideas could potentially emerge in every research corner
of the world; these, coupled with the possibility of addressing complex experiments to be conducted on a collaborative or outsourcing basis, could potentially yield truly impactful results (894). Building up scientific collaborations, in which research resource-poor countries could offer a competitive advantage (e.g., based on their scientists and research niches) can lead potentially to scientific breakthroughs. This notion could strengthen science equity in the global research agenda. Especially in light of recent actions in other diversity-inclusive policies (e.g., gender equity (895)), publishers of high-caliber journals, including multidisciplinary ones, should be encouraged to develop country diversity indices, as the importance of publication metrics cannot be overstated.

In addition, young scientists could be involved in exchange programs between research resource-rich and -poor countries; notably, after completion of their studies in the former countries, as part of brain regain efforts (such as the so-called 2017 Hellenic Pasteur Institute Declaration in Greece, a document signed by notable scientists of the country and its diaspora to advocate for political interventions on brain regain), they could be asked to return to secure jobs in their home countries, allowing them to transfuse the knowledge obtained abroad into new regional research hubs. Based on appropriate strategies supporting young scientists should include protected and sustained, albeit competitive funding from grant applications and establishing a solid formal network of collaborations with resource-rich countries (896). Doing so can be the best antidote to brain drain and brain deficit (for these terms, see (897)). We call this approach twinning of laboratories, akin to twinning of towns, which—far from scientific colonialism—could be mutually beneficial in various scientific fields (e.g., research on endemic infections and local rare diseases, cultural anthropology, and so on).

The above examples, stemming from a range of research fields, hopefully illustrate how a small, resource-poor country’s well-educated research force and rich natural diversity represent a model to deal with the global research agenda during crises and austerity times. Providing resources to local and foreign collaborating scientists should be sustained, and quality checkpoints assessing alignment to international scientific standards (including bioethical, biosecurity, biosafety, and biobanking principles and guidelines, as well as, commonly accepted standards of reproducibility (such as integrity, validation methodologies, and data openness (352)), should continue to be met. Ultimately, these examples will pave the way for other small countries to look into their own human and natural resources to develop their competitive advantages in the emerging economically harsh global era.
2. The ethical aspects of neurological research

2.1 The case of Epigenomic Engineering for Down Syndrome

2.1.1 Introduction

Down syndrome (DS) (or trisomy 21) affects approximately 250,000 individuals in the United States (US) alone and represents the commonest genetic cause of mental disability (898). Despite longstanding interest in developing gene therapies for disorders such as cystic fibrosis or muscular dystrophy, other non-Mendelian genetic disorders such as DS have not received the same attention, partly due to their biological and phenotypic complexity. There have, however, been advances in the treatment of DS-related morbidity - for example, cardiac surgery for congenital heart disease - resulting in impressive increases in life expectancy from 25 years in 1983 to 49 years in 1997 in the US, with overall life expectancy in the developed world now averaging over 55 years (899, 900). Nevertheless, there remains room for improvement in both quality of life and mortality outcomes in these individuals.

Current DS treatment is directed toward mitigating neonatal complications (congenital cardiac and gastrointestinal problems and respiratory infections) and preventative health measures (899). Mental retardation and dementia, the commonest neurological complications of DS, remain a therapeutic challenge. Children with DS have impaired language skills, learning difficulties, and both short- and long-term memory deficits (901), although phenotypes are highly variable. In later life, many patients suffer from early-onset dementia from Alzheimer’s disease (902). Although this latter phenotype is thought to be caused by a gene dosage effect caused by the extra copy of the amyloid precursor protein (APP) gene on the third chromosome 21 to produce the characteristic amyloid plaques seen in Alzheimer’s disease, only 50-70% of individuals with DS develop the disease (902). The heterogeneous phenotypes seen in DS suggest that, in spite of triplication of chromosome 21 (HSA21) in all patients, gene action in DS may not simply be a product of gene dosage independent of the environment, other genes, or transcriptional regulation, but instead result from epigenomic phenomena, i.e., falls under the control of the molecular regulators of gene function such as DNA and histone proteins. In support of this, only 22% of genes expressed in DS lymphoblastoid cells show 1.5-times increase in expression expected by trisomy alone, with the majority (56%) showing repression (903), and microarray-based gene expression profiling of adult DS brains implicates chromatin remodeling genes in DS (904). Further elegant studies in monozygotic twins discordant for DS have shown genome-wide alterations in gene expression extending beyond chromosome 21 and suggesting chromatin-related epigenetic modifications (905). It, therefore,
follows that the epigenetic modifications that affect the global transcriptome in DS may be suitable targets for therapy.

Here we provide a short overview of the role of epigenetics in the pathophysiology of DS and how it might be exploited for therapeutic benefit. In particular, we focus on how new gene editing systems can be used therapeutically against epigenetic targets to overcome the neurological deficits seen in individuals with DS. We also outline some of the anticipated challenges in using this technology in the future for clinical purposes.

2.1.2 Epigenetics and the neurobiology of Down syndrome

The exponential growth of epigenome research and data from high-throughput approaches has opened up novel therapeutic avenues for several diseases (906), including neurological and genetic disorders. It is evident that learning and memory, which are both impaired in many individuals with DS, can be modulated by epigenetic mechanisms (for an excellent review on the topic, see Dekker, De Deyn (907)).

Epigenetics is defined as a heritable phenotype resulting from chromosomal changes without alterations in the DNA sequence (908). Chromatin is formed from DNA, histones, and non-histone proteins, with each nucleosomal unit formed from 146 DNA base pairs wrapped around a histone octamer (two of H3, H4, H2A, and H2B) (909). Since the local structure of chromatin determines accessibility to the gene expression control machinery (which is generally accessible in open euchromatin and inaccessible in compact heterochromatin), mechanisms that remodel the chromatin orchestrate transcription. There are five main remodeling processes: DNA methylation, post-translational histone modifications, nucleosomal positioning, histone variant incorporation, and the action of small noncoding RNAs (microRNAs (miRNAs) and long noncoding RNAs (lncRNAs)) (910). Of these, recent “cognitive epigenetic” research (911) has shown that DNA methylation, post-translational histone modifications, and small noncoding RNAs, in particular, participate in neurodevelopment, synaptic plasticity, and learning and memory (907).

All of these mechanisms have been shown to play a role in DS and might, therefore, provide suitable epigenetic targets. Genome-wide DNA hypermethylation is implicated in DS (912, 913) and might represent a biomarker for DS-related neurodegeneration (914); DNA methylation signatures are associated with premature aging and negative neurodevelopmental effects in DS (915-917). Further, 5-methyl-cytosine and 5-hydroxy-methyl-cytosine are implicated in CpG island methylation in DS (918). In terms of specific targets, DNMT3L is
found on HSA21, and its downstream effectors DNMT3a and 3b are DNA methylators that might account for the hypermethylated phenotype and cognitive deficits. Although the causal mechanism for this is uncertain, DNA methylation levels are positively correlated with cognitive function in DS, as measured by the Dalton Brief Praxis test (919). Other DNA methylation targets include the methyl donor as S-adenosyl methionine (SAM), which is present at reduced levels in DS due to overexpression of cystathionine beta-synthase (CBS) (920). Instead of resulting in hypermethylation, CBS overexpression, and consequent SAM deficiency are thought to lead to mitochondrial dysfunction via hypomethylation of mitochondrial DNA (907, 920).

Histone modifications represent a major type of epigenetic alteration that alters gene expression by modifying chromatin structure. Histone methylation is usually associated with transcriptional repression (some notable exceptions being methylation of lysine 4 on H3 and arginine residues on H3 and H4, which result in transcriptional activation), so histone demethylase inhibitors have been used both experimentally and clinically to reverse histone demethylation for transcriptional repression (921). Although the post-translational histone landscape has yet to be systematically characterized in DS, the DSCR (Down syndrome critical region; a genomic region shared by individuals with a given phenotype) contains the gene DYRK1A, whose downstream targets include direct phosphorylation of the SIRT1 histone deacetylase and the cyclic AMP response element-binding protein (CREB) and indirect regulation of gene expression via neuron-restrictive silencer factor (NRSF) (907). The former may deteriorate cognitive function by promoting deacetylation of histone tails or recruiting the CBP/P300 histone acetyltransferase to promote CREB-related gene expression (922). NSRF expression is decreased in DS and has been shown to be a “master regulator” of neurogenesis (923).

Although less is known about the function and role of histone core variants and chromatin proteins in neurobiology and DS, in particular, histone core variant pseudogenes are expressed on chromosome 21 (H2AFZP and H2BFS) and might, therefore, be worthy of further investigation (914). Further, CHAF1B and HMGNJ, which encode two constitutive histone proteins, are found in the DSCR and recruit H3 and H4 (924) and regulate the expression of MeCP2 (925, 926), a learning disability-related protein, respectively.

Finally, various small noncoding RNAs are found on HSA21 and might therefore contribute to DS pathogenesis (miRNA-99a, miRNA-125b-2, miRNA-155, miRNA-802, and let-7c) (914, 927), with some of these implicated in neurodevelopmental diseases such as
miRNA-802 in Rett syndrome, which again targets MeCP2 (914, 928). Even less is known about the role of IncRNAs in DS, but several are known to reside on HSA21 (929).

2.1.3 Epigenetic targets in Down syndrome: indicators from pre-clinical studies

The preceding discussion, therefore, identifies a number of epigenetic molecules and pathways that might be therapeutically targeted to treat the cognitive defects seen in DS. Given that epigenetic drugs - “epidrugs” - are already in clinical use for psychiatric (e.g., valproic acid) or neoplastic diseases (e.g., decitabine, vorinostat), drug repositioning may be viable for individuals with DS (907). Epigenetic targets are particularly attractive since – in contrast to removing or silencing an entire aneuploid chromosome – epigenetic marks are reversible.

Although trial results have been variable, there is clinical precedent for using epidrugs in humans, particularly DNA methyltransferase inhibitors and histone deacetylase (HDAC) inhibitors to treat cancer (930). Clues from experimental models suggest that pharmacological interventions can improve cognitive function, albeit by targeting critical neuroinhibitory pathways (such as gamma-aminobutyric acid (GABA) blockade) rather than specific epigenetic pathways. The two best-known mouse models of DS are the Ts65Dn mouse, which is trisomic for orthologs of about half of HSA21 genes, and the Ts1Cje mouse, which is trisomic for about a third of the Ts65Dn mouse trisomies (901). A “transchromosomic” mouse has also been developed that carries a nearly intact copy of human HSA21, called the Tc1 mouse, which shows many of the developmental defects seen in DS (931). All three models, along with complete HSA21 sequencing, have been invaluable for studying genotype-phenotype correlations and the effects of therapies.

For instance, chronic pentylenetetrazole administration, which reduces GABA-mediated inhibition in the hippocampus, restores spatial cognition in the Ts65Dn mouse model of DS (932). Perinatal choline supplementation during pregnancy and lactation significantly improves attention in trisomic adult mouse offspring in the same model (933), while learning deficits are reversed with a norepinephrine prodrug, L-threo-3,4-dihydroxyphenylserine (xamoterol), a partial beta 1-adrenergic receptor agonist (934). Vitamin E supplementation and memantine, a non-competitive N-methyl-D-aspartic acid receptor antagonist, improve learning in mice (900, 935). However, epigenetic pathways in DS mouse models have yet to be specifically modulated, although experimental studies in other neurological diseases suggest that the epidrug approach may be successful. For instance, histone acetylation was modulated in a mouse model of the haploinsufficient form of Rubinstein-Taybi syndrome, which is characterized by moderate to severe intellectual disability: knockout of the CBP
acetyltransferase and the consequent impaired histone acetylation reduced memory formation (936), while increasing histone acetylation with HDAC inhibitors improved memory formation (937). With respect to histone methylation, mice lacking Mll, an H3K4-specific methyltransferase, have impaired contextual fear conditioning (938), and conditional neuronal deletion of the HMT G9a/GLP results in mice with motor and severe learning and memory defects (939).

There is, however, currently no treatment that prevents the neurocognitive defects of DS, and the majority of progress for children and adults with DS has been social (940). Donepezil, although FDA approved for symptomatic treatment of Alzheimer’s disease, did not affect the cognition of children aged between 10 and 17 with DS (941). In humans with DS, vitamin supplementation has been proposed as a way to improve thyroid and immune function and cognitive abilities (942).

Therefore, the epidrug approach is showing some promise, at least in the experimental setting, and there appear to be relevant epigenetic targets in DS. Clinical trials of epigenetic drugs in DS are certainly warranted. However, as with all systemically administered drugs with broad-spectrum effects on cellular processes, side-effects are common and to be expected. For instance, in cancer trials, HDAC inhibitors have been associated with severe cardiotoxicity (943), and given the high frequency of cardiac abnormalities in DS, their use is likely to be contraindicated in this group of patients. An ideal therapeutic strategy in DS would be to target the primary cause of DS-mediated effects, namely the dysregulated “dosage-sensitive” genes affected by the aneuploidy, accepting the limitation noted above that not all HSA21 genes are 1.5-fold upregulated in DS (944) and that there has yet to be systematic network analysis of the key “nodes” regulating epigenomic function in DS. Pharmacological intervention with highly specific small molecules or small interfering RNA (siRNA)-mediated inhibition of dosage-sensitive genes might represent a useful means to reverse DS phenotypes at the organism-wide level. Recent advances in gene editing may provide a promising and specific new solution to this problem that avoids some of the limitations of other gene knockdown strategies (such as siRNAs) or pharmacological approaches.

2.1.4 CRISPR-Cas9 for epigenomic editing in Down syndrome: opportunities and limitations

Targeted epigenetic editing to silence (combinations) of individual genes on HSA21 (such as DYRK1A, DNMT3L, and CBP described above) would represent a highly efficacious and specific strategy to target the epigenome. Such a strategy would be expected to overcome off-
target effects, as at least partially exemplified by the successful use of targeted therapies in the oncology arena (e.g., trastuzumab to target HER2 in breast cancer (945)).

The epigenetic editing approach is not without precedent. Jiang, Jing (946) used zinc finger nucleases – an earlier genomic targeting method that specifically recognizes 18-nucleotide regions – to introduce an inducible XIST transgene into the DYRK1A locus in pluripotent stem cells in vitro. Since XIST endogenously silences the second X chromosome in females, the approach successfully proved concept that the entire chromosome could be epigenetically inactivated: stable heterochromatin modifications were induced, there were chromosome transcriptional silencing and DNA methylation, and Barr body formation. Further, HSA21 target gene transcription was suppressed, e.g., DYRK1A and APP, and neurogenesis was promoted (946).

However, zinc finger assemblies are difficult to construct, and 70% fail to bind to their target sequence, so editing efficacy is poor (947). Further, the whole chromosome approach is not selective for dose-sensitive genes and may, therefore, have unintended or unexpected consequences. As a result, this approach has been supplanted by clustered regularly interspaced short palindromic repeats (CRISPR) technology, which is now coming of age for gene deletion or selective modification of DNA methylation (948, 949) either of individual or multiple targets in several diseases (950). CRISPR uses a short and specific guide RNA to target the bacterially derived Cas9 DNA endonuclease to a genomic sequence of interest. By altering the 5’ end of the guide RNA, almost any genomic locus can be targeted, including non-coding loci such as enhancers and IncRNAs. Furthermore, for site-specific epigenetic editing, a catalytically inactive modified Cas9 (dCas9) can be used to introduce an epigenetic-modifying domain (e.g., a DNA methyltransferase or TET) without cutting the DNA (i.e., an epigenetic “toggle” switch) (951). This latter approach sees target specificity retained using guide RNAs and the targeting domain of Cas9 but the endonuclease function of Cas9 substituted for a function specific to correct the defect of interest, such as a methyltransferase (949). For a detailed description of the CRISPR-Cas9 system, see (952). CRISPR-Cas9 seems to be a highly and uniquely specific tool to alter the expression of epigenetic targets in DS either via the silencing of the entire chromosome (perhaps via XIST as in (946)) or deletion of overexpressed HSA21 genes via Cas9 endonuclease cutting or the toggle switch action of DS-specific enhancers in a cell type-specific way (953).

Although promising, the field is still in its relative infancy, and several technical limitations with respect to efficacy, specificity, and applicability need to be overcome, first in
the experimental setting and later for clinical application. First, many studies have reported only modest effects on gene expression, with the interaction between the guide RNA and Cas9 the main determinant of CRISPR effectiveness (954); optimal guide RNA design is essential. However, CRISPR studies measuring phenotype are reassuring, in that modest transcriptional effects can still translate into significant phenotypic changes (955), so complete gene silencing may not always be necessary. Second, CRISPR suffers from off-target effects due to mismatches between guide RNAs and genomic sequences (956); again, optimal selection of unique target sequences and optimizing the guide RNA and Cas9 is imperative. Third, the CRISPR/Cas9 system is large (around 4.2 kb), making delivery in vivo challenging. Adeno-associated virus vectors (AAVs) have been used to deliver the CRISPR/Cas system into adult mouse brain by stereotactic injection to edit single (MeCP2) and multiple (Dnmt1, 3a, and 3b) epigenetic targets, which is promising with respect to delivery to the site of interest (i.e., the brain) in DS (957). Other options for delivery might include larger viruses such as adenoviral and lentiviral vectors, or non-viral vectors such as cationic lipid- or polymer-based vectors (958). Other sophisticated delivery methods are also being developed, such as using the chemical ligand anisamide to coat nanoparticles containing modified RNAs to target ovarian cancers (959), and similar tissue-homing strategies could be adapted to carry CRISPR-Cas9 to sites of interest. Fourth, epigenetic suppression of epigenetic targets may itself create unwanted phenotypes. For example, DSCR1 is itself a tumor suppressor, so deletion might increase the risk of solid tumors in DS patients (960).

Finally, and despite the range of possible targets identified above, the main window of opportunity for preventing cognitive impairment in DS is likely to occur well before birth (given that the brain is already significantly altered by the beginning of the second trimester) (961). Delivering prenatal therapies via the mother is challenging even in the case of small molecules, let alone complex gene editing systems, with toxicity, teratogenicity, and crossing the placental and blood-brain barriers all potential difficulties – both practical and ethical. However, there may be opportunities to combine ex vivo approaches with CRISPR-Cas9, for instance by extracting neural stem cells from mice, subjecting them to epigenetic editing, and reinjecting them back into sites of interest such as the hippocampus to reverse DS-related Alzheimer’s disease changes. Indeed, there seems to progress in this regard in the cancer arena in which the first clinical trial for ex vivo CRISPR editing of T cells in patients with cancer has just been approved (962).

2.1.5 Conclusions
Down syndrome patients harbor a range of epigenetic alterations that drive the neurocognitive phenotype and its variability. Since epigenetic changes are reversible, they are exciting targets for therapy; however, systemic drugs are usually non-specific with off-target effects. Epigenomic engineering is a promising technique that might overcome some of the off-target problems and provide novel therapeutic avenues for DS, with CRISPR-Cas9 and its variants the current leading platforms. However, the application of CRISPR-Cas9 is likely to be confined to in vitro animal models for the foreseeable future, which is likely to reap many rewards in terms of understanding the pathogenesis of DS and its epigenetic basis, especially given DS’s status as the exemplar for aneuploidies. As well as the technical challenges of applying epigenomic engineering to DS, there are also likely to be ethical hurdles, especially if treatments are to be administered in utero.
3. Neurological and Biomedical Research and its psychological implications: The Example of Imposter syndrome as threat to diversity

Oh Tim, I’ve just had a most ghastly weekend because I felt so unworthy (963). Surprising as it may seem, this quote was directed by Paul Nurse to Tim Hunt shortly after they both learned that they received the Nobel Prize, clearly illustrating the so-called Impostor Phenomenon (or Impostor Syndrome). Originally described in 1978, this phenomenon refers to feelings of intellectual phoniness or self-doubt about one’s accomplishments and skills, despite factual evidence indicating otherwise. Syndromal impostors often believe that they have fooled their peers into overrating their abilities and professional competence, or they simply attribute their success to luck; hence, they have an innate fear of being discovered as a fraud or non-deserving professional (964).

Affecting up to two out of three people in certain professional settings (965), the prevalence of the impostor syndrome in academia might be grossly underestimated in the predominant culture of silence in higher education (966). Highly demanding families and professional environments, psychological traits, such as perfectionism, and social inequality, are all putative contributors to the impostor syndrome (967-969), while potential neurobiological underpinnings cannot be excluded. Although the impostor syndrome in Academia affects people at all stages of their careers in the highly competitive academic environments (970), it is seemingly more prevalent in high-achievers (5, 9), women (especially those regarding their mentor as super-hero or being discourage by female mentors (301)), and underrepresented racial/ethnic/religious and other minorities (968, 970-972); in many cases, this phenomenon manifests as early as high school and college (973). Strikingly, mental health problems in racial minority college students were better predicted by impostor feelings than by the stress associated with their minority status (974).

At an individual level, impostorism can lead to psychological distress, emotional suffering, and serious mental health disorders, including chronic dysphoric stress, anxiety, depression, drug abuse, and so on (975). Moreover, by constantly downplaying their own accomplishments, syndromal impostors may be self-sabotaging their own career and their dreams of succeeding in academia and beyond (968). At the societal level, the impostor syndrome may also explain the higher drop-out rates of women and minorities from the STEM pipeline (301, 887). Consequently, syndromal impostors may have a deep impact on their own academic success, affecting negatively both their overall productivity and quality of research, as well their quality of life (976).
The individual and social consequences of the impostor syndrome call for both mental health coaching and supportive organizational policies. The solution may be in increasing visibility of the problem, providing professional and peer coaching, and promoting the role of mentors as counselors. We should encourage our students and peers to focus on factual evidence regarding their academic performance, set realistic expectations, and embrace themselves unconditionally. We also need to openly acknowledge that toning down the expectations concerning own goals in order to justify potential failures can create vicious cycles of self-discomfort (977).

Granted that so many suffer from syndromal impostorism at some point, we should promote open discussions about the impostor syndrome at the institutional level to finally break the current culture of silence, nonacceptance, shame, and guilt regarding mental health in academia (966). As experts suggest, these discussions should put a name on these feelings, and normalize them as common experiences rather than pathologizing them (301). Hopefully, an inclusive environment could be formed, within which, all members of the scientific community, regardless of their background or seniority, share their experiences and obtain peer support. Indeed, group peer mentoring, beyond one-to-one mentoring, may be a good strategy, as mentees gradually change roles to become independent scientists and mentors (978). Finally, because some mentors may lack the first aid skills to support mentees with mental health problems, we must consider educating mentors to become well-informed counselors, who will recognize the negative consequences of the impostor syndrome and assist affected individuals by showing compassion and by providing proper guidance (in parallel to conducting research on how to prepare best mentors (979)).

As higher education institutions, worldwide, are promoting the diversification of both graduate school admissions and hiring of academics, adopting equity and diversity-encompassing policies, building a system supportive of women and underrepresented minorities, and providing inspiration by key mentors that belong to the above groups, will be major steps forward to mitigate individual burdensome feelings and to build a healthy scientific community for all (968, 979).
Chapter 6 — Conclusions

In the *meta-umbrella* study, it was reported that: 

a) Mediterranean diet was associated with lower risk of dementia, Alzheimer disease (AD), cognitive impairment, and stroke;  
b) In Parkinson’s disease (PD) and AD/dementia, coffee consumption and physical activity were protective factors;  
c) Low serum uric acid levels were associated with increased risk of PD;  
d) Smoking was associated with elevated risk of multiple sclerosis and dementia but with a lower risk of PD, and hypertension with lower risk of PD but higher risk of dementia;  
e) Chronic occupational exposure to lead was associated with higher risk of Amyotrophic lateral sclerosis (ALS); and,  

A) Late-life depression was associated with higher risk of AD and any form of dementia. This study highlights different non-genetic factors spanning several neurological diseases with relevance to clinical practice, health promotion, and health policy. Importantly, these results offer new perspectives and novel research questions in secondary research fields.

In the conceptual framework for AD, I elaborated on the existence and roles of the meningeal lymphatic system in normal and pathological brain function which have been a long-standing enigma, and a reviewed the evidence linking the meningeal lymphatic system with human AD. Novel findings suggest that the recently described meningeal lymphatic vessels could be linked to, and possibly drain, the efferent paravascular glial lymphatic (glymphatic) system carrying cerebrospinal fluid, after solute and immune cell exchange with brain interstitial fluid. In so doing, the lymphatic system could contribute to the export of toxic solutes and immune cells from the brain (an exported fluid we wish to describe as glymph, similarly to lymph) to the meningeal lymphatic system; the latter, by being connected with downstream anatomic regions, carries the glymph to the conventional cervical lymphatic vessels and nodes. Thus, abnormal function in the meningeal lymphatic system could potentially lead to the accumulation of amyloid-beta, cellular debris, inflammatory mediators, and immune cells in the brain, resulting in damage of the brain parenchyma and leading to cognitive and other neurologic dysfunctions. In addition, I provided novel insights into APOE4—the leading genetic risk factor for AD—and its relation to the meningeal lymphatic system. In this regard, I have reanalyzed previously published RNA-Seq data to show that induced pluripotent stem cells carrying the APOE4 allele (either as APOE4 knock-in or stemming from APOE4 patients) express lower levels of a) genes associated with lymphatic markers, and b) genes for which well-characterized missense mutations have been linked to peripheral lymphedema. Taking into account this evidence, we propose a new conceptual framework in which APOE4 could play a novel role in the premature shrinkage of the meningeal lymphatic vessels (meningeal lymphosclerosis), leading to abnormal meningeal lymphatic functions (meningeal lymphedema), and reduction in the clearance of amyloid-beta and other...
macromolecules, inflammatory mediators, and immune cells from the brain. Altogether, this physiological cascade may lead to the exacerbation of AD manifestations and progression of the disease. These findings and their interpretation hold great potential for novel diagnostic tools and therapeutic approaches in patients with AD.

In the combined case-control study and meta-analysis for *H. pylori* and Parkinson’s disease, I elaborated on initial evidence supporting *H. pylori* infection as a trigger or driving event, although detection and eradication of *H. pylori* are not part of PD management. By conducting a case-control study and meta-analysis I aimed to determine: (a) the prevalence of *H. pylori* infection in PD patients; (b) the associations between *H. pylori* infection and clinical status; and (c) possible differences in motor status in PD patients before and after *H. pylori* eradication. In the ten studies included, *H. pylori* infection prevalence was higher in PD patients than in HCs while in seven studies reporting UPDRS scores, there was a significant association between *H. pylori* infection and mean UPDRS scores. Regarding *H. pylori* eradication, five studies reported a significant reduction in UPDRS-III scores after treatment. In conclusion, the present meta-analysis revealed a higher prevalence of *H. pylori* infection in PD patients, suggesting that *H. pylori* may contribute to PD pathophysiology. In addition, the significantly lower UPDRS scores in non-infected PD patients and in patients after *H. pylori* eradication therapy demonstrate that the infection may worsen the clinical severity of the disease.

With regards to the role of CSF virome-microbiome in MS, I initially reviewed the main viruses implicated in MS pathogenesis and provided conceptual input on why the absence of evidence should not be translated into evidence of absence, especially given the myriad of viruses yet to be discovered (which have been aptly described as *viral dark matter*). I also posited that, to our knowledge, except for a case of a spontaneous inflammatory demyelination in a Japanese macaque infected with Japanese macaque rhadinovirus (a novel herpesvirus), only humans present with MS, indicating a human-specific causal factor. I also attempted to address critical questions relating to MS pathogenesis, that is, do viruses and/or microbes and phages (i.e., viruses infecting bacteria) contribute to the aetiology of MS? Are they at least environmental risk factors and, in being so, potential biomarkers for MS? To this end, I analyzed primary data for the microbiome/virome (i.e., the sum of viral genomic information) in the CSF of patients with MS vs. healthy controls. The initial bioinformatics analysis of 21 samples from MS vs. healthy controls support the feasibility and power of the bioinformatics approach and suggests (based on *de novo* assembly of viruses through MEGAHIT, followed by hierarchical clustering) that a distinct viral pattern between MS cases and healthy controls exists but remains to be fully deciphered, alongside with concrete gene expression patterns.
between MS cases and healthy controls. Future studies should aim to thoroughly elucidate the role of specific viruses in these samples as well as the major biological pathways implicated in MS subcategories under investigation.

Last, but not least, to justify the historic(al) roots of a Ph.D. thesis as Doctor Philosophicus, I presented a critical analysis of the so-called ELSI (Ethical, Legal, and Societal Implications) of Neurological Research. Specifically, I argued that: a) the research niches of countries with limited resources for research (yet often classified as wealthy) can be regarded as hidden pearls that can boost local research pipelines in times of austerity, and to this end, I provide specific examples from patient population and biotechnological resources in Greece; b) the post-CRISPR/Cas9 genome editing era represents a tremendous opportunity as potential therapeutic armamentarium even for complex disorders with neurological manifestations, such as Down’s syndrome (Trisomy 21); however, rigorous bioethics regulations should be set prior to clinical trials, especially if the latter are applied during the embryological stage of patients with Down’s syndrome; and c) stronger public advocacy is needed for syndromes with neuropsychiatric manifestations, such as the Imposter syndrome, whose high prevalence and psychological burden has had a huge impact in societally underrepresented groups.

Collectively, my Ph.D. thesis aims to shed some light into the causality of certain neurological disorders. I anticipate that the results and conclusions presented here will open an avenue to future research questions; prior to these future steps, however, addressing the ELSI is pivotal to ensure bioethics compliance and societal acceptance of neurological research.
1. Chapter 1

1.1 Supplementary File – Update on Search Results

We performed an updated, supplementary search in PubMed on umbrella reviews, using “umbrella review” [ti] as search string, and publication dates from September 21st, 2018, and until December 31st, 2019. This search returned a good number of results. After independent analysis by four reviewers, twenty-three umbrella reviews were considered of potential relevance to our research questions.


## 1.2 Appendix 1

### Appendix 1. Search strategy

<table>
<thead>
<tr>
<th>Database</th>
<th>Hits</th>
<th>Relevant</th>
</tr>
</thead>
<tbody>
<tr>
<td>PUBMED/MEDLINE</td>
<td>1178</td>
<td>21</td>
</tr>
<tr>
<td>SCOPUS</td>
<td>1023</td>
<td>6</td>
</tr>
<tr>
<td>Web of Science</td>
<td>165</td>
<td>17</td>
</tr>
<tr>
<td>CINAHL</td>
<td>106</td>
<td></td>
</tr>
<tr>
<td>Cochrane Database of Systematic Reviews</td>
<td>17</td>
<td>0</td>
</tr>
<tr>
<td>JBI Database of Systematic Reviews and Implementation Reports</td>
<td>84</td>
<td>1</td>
</tr>
<tr>
<td>DARE</td>
<td>31</td>
<td>0</td>
</tr>
<tr>
<td>PROSPERO register</td>
<td>193</td>
<td>5</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>2797</strong></td>
<td><strong>50</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Search terms used alone or in combination</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
</tbody>
</table>
Supplementary Table S1. Characteristics and quantitative synthesis of the 76 eligible meta-analyses of non-genetic (environmental) risk and protective factors for non-communicable diseases.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Neurological disorder</th>
<th>Level of comparison</th>
<th>Total number of participants</th>
<th>Number of primary studies</th>
<th>Effect size metric</th>
<th>Random-effects summary effect size (95% CI)</th>
<th>P random</th>
<th>95% PI</th>
<th>I²</th>
<th>Small-study Effects / Excess statistical significance</th>
<th>Level of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grosso, 2017</td>
<td>Parkinson’s disease</td>
<td>High vs. Low</td>
<td>N/A</td>
<td>9</td>
<td>RR</td>
<td>0.67 (0.57–0.80)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verones, 2018</td>
<td>Stroke</td>
<td>High vs. low</td>
<td>322,732</td>
<td>5</td>
<td>RR</td>
<td>0.858 (0.786–0.937)</td>
<td>&lt;0.001</td>
<td>0.74–0.99</td>
<td>0</td>
<td>N/A</td>
<td>IV</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grosso, 2017</td>
<td>Alzheimer’s disease</td>
<td>High vs. Low</td>
<td>15,761</td>
<td>4</td>
<td>RR</td>
<td>0.73 (0.55–0.97)</td>
<td></td>
<td></td>
<td></td>
<td>No/No</td>
<td>III</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Parkinson’s disease</td>
<td>High vs. low</td>
<td>187,740</td>
<td>4</td>
<td>RR</td>
<td>0.70 (0.56–0.88)</td>
<td></td>
<td></td>
<td></td>
<td>No/No</td>
<td>III</td>
</tr>
</tbody>
</table>

A. RISK AND PROTECTIVE FACTORS

CAFFEINE

CHOCOLATE

COFFEE CONSUMPTION
<table>
<thead>
<tr>
<th>Reference</th>
<th>Neurological disorder</th>
<th>Level of comparison</th>
<th>Total number of participants</th>
<th>Number of primary studies</th>
<th>Effect size metric</th>
<th>Random-effects summary effect size (95% CI)</th>
<th>P random</th>
<th>95% PI</th>
<th>I²</th>
<th>Small-study Effects / Excess statistical significance</th>
<th>Level of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poole 2017</td>
<td>Parkinson’s disease</td>
<td>High vs. low</td>
<td>894,568</td>
<td>7</td>
<td>RR</td>
<td>0.64 (0.53–0.76)</td>
<td>16%</td>
<td>N/A</td>
<td></td>
<td>Suggestive</td>
<td>II</td>
</tr>
<tr>
<td></td>
<td>Alzheimer’s disease</td>
<td>Any vs. No</td>
<td>NP</td>
<td>3</td>
<td>RR</td>
<td>1.02 (0.99–1.05)</td>
<td>16%</td>
<td>Yes/Yes</td>
<td></td>
<td></td>
<td>IV</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>DIETARY FIBER</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Veronese, 2018</td>
<td>Stroke</td>
<td>High vs Low</td>
<td>325,167</td>
<td>14</td>
<td>RR</td>
<td>0.81 (0.73–0.90)</td>
<td>9 × 10⁻⁵</td>
<td>0.61–1.12</td>
<td>38.7</td>
<td>III</td>
<td></td>
</tr>
<tr>
<td>McRae, 2017</td>
<td>Stroke</td>
<td>High vs Low</td>
<td>498,487</td>
<td>6</td>
<td>RR</td>
<td>0.86 (0.45–0.92)</td>
<td>&lt;0.005</td>
<td>NR</td>
<td>0%</td>
<td>III</td>
<td></td>
</tr>
<tr>
<td><strong>MEDITERRANEAN DIET</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Galbete, 2018</td>
<td>Ischemic Stroke</td>
<td>High vs. low</td>
<td>181,295</td>
<td>5</td>
<td>RR</td>
<td>0.82 (0.73–0.92)</td>
<td>N/A</td>
<td>0.67–1.00</td>
<td>0</td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mild cognitive impairment incidence</td>
<td>High vs. low</td>
<td>27,667</td>
<td>5</td>
<td>RR</td>
<td>0.83 (0.75–0.93)</td>
<td>N/A</td>
<td>0.7–1.0</td>
<td>0</td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td>Reference</td>
<td>Neurological disorder</td>
<td>Level of comparison</td>
<td>Total number of participants</td>
<td>Number of primary studies</td>
<td>Effect size metric</td>
<td>Random-effects summary effect size (95% CI)</td>
<td>( P ) random</td>
<td>95% PI</td>
<td>( I^2 )</td>
<td>Small-study Effects / Excess statistical significance</td>
<td>Level of evidence</td>
</tr>
<tr>
<td>--------------------</td>
<td>-----------------------</td>
<td>---------------------</td>
<td>-----------------------------</td>
<td>---------------------------</td>
<td>-------------------</td>
<td>---------------------------------------------</td>
<td>--------------</td>
<td>-------</td>
<td>--------</td>
<td>------------------------------------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Alzheimer’s disease</td>
<td>High vs. low</td>
<td>7609</td>
<td>5</td>
<td>RR</td>
<td>0.60</td>
<td>(0.48–0.77)</td>
<td>N/A</td>
<td>0.41–0.89</td>
<td>0</td>
<td></td>
<td>IV</td>
</tr>
<tr>
<td>MEDITERRANEAN DIET</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dinu, 2017</td>
<td>Stroke</td>
<td>High vs. low</td>
<td>171,138</td>
<td>9</td>
<td>RR</td>
<td>0.58 (0.353–0.953)</td>
<td>0.032</td>
<td>0.36–1.59</td>
<td>52%</td>
<td>N/A</td>
<td>Weak III</td>
</tr>
<tr>
<td></td>
<td>Cognitive impairment</td>
<td>High vs low</td>
<td>11,877</td>
<td>9</td>
<td>RR</td>
<td>0.51 (0.309–0.843)</td>
<td>0.009</td>
<td>0.49–0.91</td>
<td>60%</td>
<td>N/A</td>
<td>Suggestive II</td>
</tr>
<tr>
<td></td>
<td>Alzheimer’s disease</td>
<td>High vs. low</td>
<td>6111</td>
<td>5</td>
<td>RR</td>
<td>0.60 (0.48–0.77)</td>
<td>&lt;0.0001</td>
<td>0.44–0.82</td>
<td>0%</td>
<td>N/A</td>
<td>Convincing I</td>
</tr>
<tr>
<td></td>
<td>Dementia</td>
<td>High vs. low</td>
<td>8174</td>
<td>5</td>
<td>RR</td>
<td>0.69 (0.57–0.84)</td>
<td>0.00002</td>
<td>0.51–0.95</td>
<td>0%</td>
<td>N/A</td>
<td>Convincing I</td>
</tr>
<tr>
<td>MELATONIN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posadzki, 2018</td>
<td>Cognitive impairment</td>
<td>High vs low</td>
<td>121</td>
<td>2</td>
<td>MD</td>
<td>–2.64 (–5.99 to 0.71)</td>
<td>0.123</td>
<td>N/A</td>
<td>68.6</td>
<td></td>
<td>IV</td>
</tr>
<tr>
<td>SERUM URIC ACID</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference</td>
<td>Neurological disorder</td>
<td>Level of comparison</td>
<td>Total number of participants</td>
<td>Number of primary studies</td>
<td>Effect size metric</td>
<td>Random-effects summary effect size (95% CI)</td>
<td>$P$ random</td>
<td>95% PI</td>
<td>$I^2$</td>
<td>Small-study Effects / Excess statistical significance</td>
<td>Level of evidence</td>
</tr>
<tr>
<td>-----------</td>
<td>----------------------</td>
<td>---------------------</td>
<td>-----------------------------</td>
<td>--------------------------</td>
<td>-------------------</td>
<td>---------------------------------------------</td>
<td>------------</td>
<td>-------</td>
<td>------</td>
<td>-------------------------------------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Li, 2017</td>
<td>Alzheimer’s disease</td>
<td>General SUA (mg/dl)</td>
<td>3617</td>
<td>21</td>
<td>RR</td>
<td>0.29 (0.11–0.76)</td>
<td>0.012</td>
<td>0.01–8.97</td>
<td>97%</td>
<td></td>
<td>IV</td>
</tr>
<tr>
<td></td>
<td>Dementia or cognitive impairment</td>
<td>General SUA (mg/dl)</td>
<td>7021</td>
<td>31</td>
<td>RR</td>
<td>0.58 (0.41–0.83)</td>
<td>0.003</td>
<td>0.08–4.48</td>
<td>89%</td>
<td></td>
<td>IV</td>
</tr>
<tr>
<td></td>
<td>Parkinson’s disease</td>
<td>Hyper vs. normal</td>
<td>33,185</td>
<td>6</td>
<td>RR</td>
<td>0.65 (0.43–0.97)</td>
<td>0.04</td>
<td>0.24–1.77</td>
<td>42</td>
<td></td>
<td>IV</td>
</tr>
<tr>
<td></td>
<td>Multiple sclerosis</td>
<td>General SUA (mg/dl)</td>
<td>2216</td>
<td>19</td>
<td>RR</td>
<td>0.49 (0.27–0.87)</td>
<td>0.02</td>
<td>0.05–4.96</td>
<td>92</td>
<td></td>
<td>IV</td>
</tr>
<tr>
<td></td>
<td>NMO</td>
<td>General SUA (mg/dl)</td>
<td>1137</td>
<td>3</td>
<td>RR</td>
<td>0.22 (0.10–0.45)</td>
<td>0.00009</td>
<td>0.02–3.14</td>
<td>82</td>
<td></td>
<td>IV</td>
</tr>
<tr>
<td>Reference</td>
<td>Neurological disorder</td>
<td>Level of comparison</td>
<td>Total number of participants</td>
<td>Number of primary studies</td>
<td>Effect size metric</td>
<td>Random-effects summary effect size (95% CI)</td>
<td>( P ) random</td>
<td>95% PI</td>
<td>( I^2 )</td>
<td>Small-study Effects / Excess statistical significance</td>
<td>Level of evidence</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------------</td>
<td>---------------------</td>
<td>-----------------------------</td>
<td>--------------------------</td>
<td>-------------------</td>
<td>---------------------------------------------</td>
<td>----------------</td>
<td>--------</td>
<td>------</td>
<td>--------------------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Amyotrophic lateral sclerosis (ALS)</td>
<td>Patients with ALS vs. Controls SUA (mg/dl)</td>
<td>826</td>
<td>3</td>
<td>RR</td>
<td>0.21 (0.14–0.32)</td>
<td>6.3 ( \times ) 10(^{-13} )</td>
<td>0.04–1.05</td>
<td>51</td>
<td></td>
<td></td>
<td>IV</td>
</tr>
<tr>
<td>Theodoratou, 2014</td>
<td>Alzheimer’s disease</td>
<td>High vs low</td>
<td>1005</td>
<td>7</td>
<td>RR</td>
<td>0.08 (0.01–0.63)</td>
<td>0.02</td>
<td>N/A</td>
<td>98</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cognition</td>
<td>High vs low</td>
<td>9004</td>
<td>7</td>
<td>RR</td>
<td>0.42 (0.34–0.53)</td>
<td>2.2 ( \times ) 10(^{-13} )</td>
<td>N/A</td>
<td>56</td>
<td></td>
<td>III/IV</td>
</tr>
<tr>
<td>B. NEUROLOGICAL DISEASES</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Belbasis, 2015</td>
<td>Serum uric acid</td>
<td>High vs. Low</td>
<td>515</td>
<td>3</td>
<td>OR</td>
<td>0.31 (0.18–0.52)</td>
<td>1.27 ( \times ) 10(^{-5} )</td>
<td>0.001–1.37</td>
<td>71.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \beta )-Carotene intake</td>
<td>2500 ( \mu )g/day increase</td>
<td>1,052,482</td>
<td>9</td>
<td>RR</td>
<td>0.92 (0.87–0.97)</td>
<td>2.24 ( \times ) 10(^{-3} )</td>
<td>0.86–0.98</td>
<td>0</td>
<td></td>
<td>II</td>
</tr>
<tr>
<td>Reference</td>
<td>Neurological disorder</td>
<td>Level of comparison</td>
<td>Total number of participants</td>
<td>Number of primary studies</td>
<td>Effect size metric</td>
<td>Random-effects summary effect size (95% CI)</td>
<td>( P ) random</td>
<td>95% PI</td>
<td>( I^2 )</td>
<td>Small-study Effects / Excess statistical significance</td>
<td>Level of evidence</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------------</td>
<td>---------------------</td>
<td>------------------------------</td>
<td>--------------------------</td>
<td>-------------------</td>
<td>-------------------------------------------</td>
<td>----------------</td>
<td>--------</td>
<td>--------</td>
<td>----------------------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>n-3 fatty acids intake</td>
<td>1 g/day increase</td>
<td>1,055,843</td>
<td>8</td>
<td>RR</td>
<td>0.71 (0.59–0.85)</td>
<td>1.95 \times 10^{-4}</td>
<td>0.49–1.01</td>
<td>0</td>
<td></td>
<td></td>
<td>III</td>
</tr>
<tr>
<td>Farming</td>
<td>Exposed vs. not exposed</td>
<td>2,605,321</td>
<td>10</td>
<td>OR</td>
<td>1.42 (1.17–1.73)</td>
<td>3.75 \times 10^{-4}</td>
<td>0.90–2.26</td>
<td>42.1</td>
<td></td>
<td></td>
<td>III</td>
</tr>
<tr>
<td>Pesticides</td>
<td>Exposed vs. not exposed</td>
<td>3,716,144</td>
<td>15</td>
<td>OR</td>
<td>1.44 (1.22–1.70)</td>
<td>2.04 \times 10^{-5}</td>
<td>0.94–2.20</td>
<td>41.2</td>
<td></td>
<td></td>
<td>III</td>
</tr>
<tr>
<td>Other heavy metals</td>
<td>Exposed vs. not exposed</td>
<td>558</td>
<td>4</td>
<td>OR</td>
<td>2.13 (1.33–3.41)</td>
<td>1.60 \times 10^{-3}</td>
<td>0.48–9.52</td>
<td>26</td>
<td></td>
<td></td>
<td>II</td>
</tr>
<tr>
<td>Lead</td>
<td>Exposed vs. not exposed</td>
<td>1544</td>
<td>9</td>
<td>OR</td>
<td>1.81 (1.39–2.35)</td>
<td>8.72 \times 10^{-6}</td>
<td>1.14–2.88</td>
<td>12.7</td>
<td></td>
<td></td>
<td>I</td>
</tr>
<tr>
<td>Extremely low-frequency electromagnetic fields</td>
<td>Exposed vs. not exposed</td>
<td>9,902,859</td>
<td>17</td>
<td>OR</td>
<td>1.29 (1.03–1.62)</td>
<td>0.03</td>
<td>0.64–2.59</td>
<td>58.9</td>
<td></td>
<td></td>
<td>III</td>
</tr>
</tbody>
</table>

DEMENTIA
<table>
<thead>
<tr>
<th>Reference</th>
<th>Neurological disorder</th>
<th>Level of comparison</th>
<th>Total number of participants</th>
<th>Number of primary studies</th>
<th>Effect size metric</th>
<th>Random-effects summary effect size (95% CI)</th>
<th>( P ) random</th>
<th>95% PI</th>
<th>( I^2 )</th>
<th>Level of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bellou, 2017</td>
<td>Alcohol intake</td>
<td>Light or moderate vs never</td>
<td>11,784</td>
<td>6</td>
<td>RR</td>
<td>0.72 (0.61–0.86)</td>
<td>2.0 ( \times ) 10^{-4}</td>
<td>0.44–1.18</td>
<td>56.4</td>
<td>III</td>
</tr>
<tr>
<td></td>
<td>Physical activity</td>
<td>High vs. Low</td>
<td>39,639</td>
<td>21</td>
<td>RR</td>
<td>0.76 (0.66–0.86)</td>
<td>1.9 ( \times ) 10^{-5}</td>
<td>0.49–1.15</td>
<td>68.9</td>
<td>III</td>
</tr>
<tr>
<td></td>
<td>Depression at any age</td>
<td>Diseased vs. not diseased</td>
<td>441,491</td>
<td>33</td>
<td>RR</td>
<td>1.99 (1.84–2.16)</td>
<td>8.0 ( \times ) 10^{-62}</td>
<td>1.65–2.40</td>
<td>27.8</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td>Early-life depression</td>
<td>Diseased vs. not diseased</td>
<td>28,383</td>
<td>9</td>
<td>RR</td>
<td>1.63 (1.27–2.11)</td>
<td>1.5 ( \times ) 10^{-4}</td>
<td>1.01–2.64</td>
<td>16.2</td>
<td>III</td>
</tr>
<tr>
<td></td>
<td>Late-life depression</td>
<td>Diseased vs. not diseased</td>
<td>51,353</td>
<td>25</td>
<td>RR</td>
<td>1.85 (1.67–2.05)</td>
<td>3.1 ( \times ) 10^{-32}</td>
<td>1.66–2.06</td>
<td>0</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td>Low-frequency electromagnetic fields</td>
<td>Exposed vs. not exposed</td>
<td>10,446,519</td>
<td>25</td>
<td>RR</td>
<td>1.74 (1.37–2.21)</td>
<td>5.9 ( \times ) 10^{-6}</td>
<td>0.77–3.91</td>
<td>55.2</td>
<td>III</td>
</tr>
<tr>
<td></td>
<td>Type 2 diabetes mellitus</td>
<td>Diseased vs. not diseased</td>
<td>5,291,160</td>
<td>21</td>
<td>RR</td>
<td>1.54 (1.39–1.72)</td>
<td>3.1 ( \times ) 10^{-15}</td>
<td>1.37–1.73</td>
<td>0</td>
<td>I</td>
</tr>
<tr>
<td>Reference</td>
<td>Neurological disorder</td>
<td>Level of comparison</td>
<td>Total number of participants</td>
<td>Number of primary studies</td>
<td>Effect size metric</td>
<td>Random-effects summary effect size (95% CI)</td>
<td>P random</td>
<td>95% PI</td>
<td>I²</td>
<td>Small-study Effects / Excess statistical significance</td>
</tr>
<tr>
<td>--------------------</td>
<td>----------------------------------------</td>
<td>---------------------</td>
<td>------------------------------</td>
<td>----------------------------</td>
<td>-------------------</td>
<td>--------------------------------------------</td>
<td>----------</td>
<td>--------</td>
<td>----</td>
<td>------------------------------------------------------</td>
</tr>
<tr>
<td>Vitamin C intake</td>
<td>High vs. Low</td>
<td>13,468</td>
<td>6</td>
<td>RR</td>
<td>0.85 (0.74–0.96)</td>
<td>0.011</td>
<td>0.71–1.01</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vitamin E intake</td>
<td>High vs. low</td>
<td>14,509</td>
<td>7</td>
<td>RR</td>
<td>0.80 (0.67–0.95)</td>
<td>0.011</td>
<td>0.52–1.24</td>
<td>46.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chlamydia pneumonia infection</td>
<td>Diseased vs. not diseased</td>
<td>510</td>
<td>11</td>
<td>OR</td>
<td>6.0 (1.93–18.66)</td>
<td>2.0 × 10⁻³</td>
<td>0.19–193.8</td>
<td>73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spirochetal infection</td>
<td>Diseased vs. not diseased</td>
<td>555</td>
<td>13</td>
<td>OR</td>
<td>10.65 (3.4–33.4)</td>
<td>5.0 × 10⁻⁵</td>
<td>0.41–279.5</td>
<td>51.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>Low vs High</td>
<td>54,301</td>
<td>16</td>
<td>RR</td>
<td>1.82 (1.36–2.43)</td>
<td>5.5 × 10⁻⁵</td>
<td>0.55–6.05</td>
<td>90.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Midlife BMI</td>
<td>Obese vs. normal weight</td>
<td>17,812</td>
<td>5</td>
<td>RR</td>
<td>1.81 (1.22–2.69)</td>
<td>3.0 × 10⁻³</td>
<td>0.52–6.29</td>
<td>63.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mild traumatic brain injury</td>
<td>Exposed vs. not exposed</td>
<td>28,761</td>
<td>19</td>
<td>OR</td>
<td>1.40 (1.03–1.90)</td>
<td>0.034</td>
<td>0.39–4.98</td>
<td>85.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference</td>
<td>Neurological disorder</td>
<td>Level of comparison</td>
<td>Total number of participants</td>
<td>Number of primary studies</td>
<td>Effect size metric</td>
<td>Random-effects summary effect size (95% CI)</td>
<td>( P ) random</td>
<td>95% PI</td>
<td>( I^2 )</td>
<td>Small-study Effects / Excess statistical significance</td>
</tr>
<tr>
<td>-----------</td>
<td>----------------------</td>
<td>---------------------</td>
<td>-----------------------------</td>
<td>--------------------------</td>
<td>-------------------</td>
<td>---------------------------------------------</td>
<td>----------------</td>
<td>-------</td>
<td>--------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>Statins</td>
<td>Exposed vs. not exposed</td>
<td>766,626</td>
<td>13</td>
<td>RR</td>
<td>0.72 (0.59–0.89)</td>
<td>1.9 × 10^{-3}</td>
<td>0.39–1.35</td>
<td>54.7</td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td>Herpesviridae infection</td>
<td>Diseased vs. not diseased</td>
<td>2895</td>
<td>33</td>
<td>OR</td>
<td>1.38 (1.14–1.65)</td>
<td>7.3 × 10^{-4}</td>
<td>0.86–2.21</td>
<td>20.3</td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td>Aspirin</td>
<td>Ever vs. Never</td>
<td>18656</td>
<td>11</td>
<td>RR</td>
<td>0.77 (0.63–0.95)</td>
<td>0.014</td>
<td>0.42–1.42</td>
<td>55.5</td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td>Non-Aspirin NSAIDS</td>
<td>Ever vs. Never</td>
<td>18,048</td>
<td>9</td>
<td>RR</td>
<td>0.65 (0.49–0.86)</td>
<td>2.3 × 10^{-3}</td>
<td>0.29–1.45</td>
<td>59.1</td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td>NSAIDS</td>
<td>Ever vs. Never</td>
<td>281,491</td>
<td>16</td>
<td>RR</td>
<td>0.74 (0.64–0.86)</td>
<td>6.9 × 10^{-5}</td>
<td>0.45–1.22</td>
<td>70</td>
<td>III</td>
<td></td>
</tr>
<tr>
<td>Aluminum</td>
<td>Exposed vs. not exposed</td>
<td>10,567</td>
<td>8</td>
<td>OR</td>
<td>1.72 (1.33–2.21)</td>
<td>3.1 × 10^{-5}</td>
<td>1.16–2.54</td>
<td>6.2</td>
<td>III</td>
<td></td>
</tr>
<tr>
<td>Fish intake</td>
<td>High vs. Low</td>
<td>21,941</td>
<td>5</td>
<td>RR</td>
<td>0.88 (0.79–0.98)</td>
<td>0.022</td>
<td>0.63–1.22</td>
<td>63.4</td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td>Stroke</td>
<td>Diseased vs. not diseased</td>
<td>14,730</td>
<td>6</td>
<td>HR</td>
<td>1.59 (1.25–2.02)</td>
<td>1.7 × 10^{-4}</td>
<td>1.13–2.23</td>
<td>0</td>
<td>II</td>
<td></td>
</tr>
<tr>
<td>Reference</td>
<td>Neurological disorder</td>
<td>Level of comparison</td>
<td>Total number of participants</td>
<td>Number of primary studies</td>
<td>Effect size metric</td>
<td>Random-effects summary effect size (95% CI)</td>
<td>( P ) random</td>
<td>( 95% ) PI</td>
<td>( I^2 )</td>
<td>Level of evidence</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------------</td>
<td>---------------------</td>
<td>-----------------------------</td>
<td>--------------------------</td>
<td>-------------------</td>
<td>-------------------------------------------</td>
<td>---------------</td>
<td>--------------</td>
<td>--------</td>
<td>----------------</td>
</tr>
<tr>
<td>Frequency of social contacts</td>
<td>Low vs. high level</td>
<td>15,762</td>
<td>8</td>
<td>RR</td>
<td>1.57 (1.32–1.85)</td>
<td>1.9 × 10^-7</td>
<td>1.27–1.93</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Loneliness</td>
<td>High vs. Low</td>
<td>3,252</td>
<td>3</td>
<td>RR</td>
<td>1.58 (1.19–2.09)</td>
<td>1.5 × 10^-3</td>
<td>0.25–9.78</td>
<td>0</td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td>Social participation</td>
<td>Low vs high</td>
<td>7,714</td>
<td>6</td>
<td>RR</td>
<td>1.41 (1.13–1.75)</td>
<td>2.0 × 10^-3</td>
<td>0.85–2.34</td>
<td>31.2</td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td>Antihypertensive drugs</td>
<td>Ever vs. never</td>
<td>1,523,417</td>
<td>11</td>
<td>HR</td>
<td>0.84 (0.75–0.94)</td>
<td>1.7 × 10^-3</td>
<td>0.60–1.16</td>
<td>73.4</td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td>Hypertension</td>
<td>Diseased vs. not diseased.</td>
<td>8,123</td>
<td>6</td>
<td>HR</td>
<td>1.59 (1.20–2.11)</td>
<td>1.4 × 10^-3</td>
<td>0.78–3.21</td>
<td>36.3</td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td>Smoking</td>
<td>Ever vs never smokers</td>
<td>886,794</td>
<td>8</td>
<td>RR</td>
<td>1.26 (1.05–1.50)</td>
<td>0.013</td>
<td>0.79–2.00</td>
<td>43.9</td>
<td>IV</td>
<td></td>
</tr>
</tbody>
</table>

**MULTIPLE SCLEROSIS**
<table>
<thead>
<tr>
<th>Reference</th>
<th>Neurological disorder</th>
<th>Level of comparison</th>
<th>Total number of participants</th>
<th>Number of primary studies</th>
<th>Effect size metric</th>
<th>Random-effects summary effect size (95% CI)</th>
<th>( P ) random</th>
<th>95% PI</th>
<th>( I^2 )</th>
<th>Small-study Effects / Excess statistical significance</th>
<th>Level of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belbasis, 2015</td>
<td>Anti-EBNA IgG seropositivity</td>
<td>High vs Low</td>
<td>7,308</td>
<td>30</td>
<td>OR</td>
<td>4.46 (3.26–6.09)</td>
<td>1.5 ( \times ) 10(^{-19} )</td>
<td>1.46–13.62</td>
<td>43%</td>
<td>0</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td>Infectious mononucleosis</td>
<td>Diseased vs, not diseased</td>
<td>35,655</td>
<td>18</td>
<td>OR</td>
<td>2.17 (1.97–2.39)</td>
<td>3.1 ( \times ) 10(^{-50} )</td>
<td>2.06</td>
<td>0</td>
<td>0</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td>Smoking</td>
<td>Ever vs never smoker</td>
<td>460,571</td>
<td>14</td>
<td>OR</td>
<td>1.52 (1.39–1.66)</td>
<td>1.7 ( \times ) 10(^{-18} )</td>
<td>1.37–1.68</td>
<td>0</td>
<td>0</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td>Appendectomy at age ( \leq ) 20 years</td>
<td>Exposed vs not exposed</td>
<td>417,242</td>
<td>7</td>
<td>OR</td>
<td>1.17 (1.02–1.34)</td>
<td>0.02</td>
<td>0.99–1.38</td>
<td>0</td>
<td>0</td>
<td>II</td>
</tr>
<tr>
<td></td>
<td>Diphtheria vaccination</td>
<td>Exposed vs. not exposed</td>
<td>524</td>
<td>3</td>
<td>OR</td>
<td>0.60 (0.40–0.91)</td>
<td>0.02</td>
<td>0.08–8.57</td>
<td>0</td>
<td>0</td>
<td>II</td>
</tr>
<tr>
<td></td>
<td>EBV DNA in mononuclear cells and serum</td>
<td>Diseased vs not diseased</td>
<td>729</td>
<td>6</td>
<td>OR</td>
<td>1.84 (1.02–3.20)</td>
<td>0.04</td>
<td>0.39–8.63</td>
<td>0</td>
<td>0</td>
<td>II</td>
</tr>
<tr>
<td>Reference</td>
<td>Neurological disorder</td>
<td>Level of comparison</td>
<td>Total number of participants</td>
<td>Number of primary studies</td>
<td>Effect size metric</td>
<td>Random-effects summary effect size (95% CI)</td>
<td>( P ) random</td>
<td>95% PI</td>
<td>( I^2 )</td>
<td>Level of evidence</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------------</td>
<td>---------------------</td>
<td>----------------------------</td>
<td>--------------------------</td>
<td>------------------</td>
<td>------------------------------------------</td>
<td>----------------</td>
<td>--------</td>
<td>-------</td>
<td>-----------------</td>
<td></td>
</tr>
<tr>
<td>Serum vitamin B(_{12})</td>
<td>High vs low</td>
<td>730</td>
<td>8</td>
<td>OR</td>
<td>0.64 (0.44–0.93)</td>
<td>0.02</td>
<td>0.26–1.60</td>
<td>38</td>
<td>II</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tetanus vaccination</td>
<td>Exposed vs. not exposed</td>
<td>4132</td>
<td>8</td>
<td>OR</td>
<td>0.71 (0.57–0.88)</td>
<td>0.002</td>
<td>0.47–1.07</td>
<td>17</td>
<td>II</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tonsillectomy at age (\leq) 20 years</td>
<td>Exposed vs. not exposed</td>
<td>8836</td>
<td>12</td>
<td>OR</td>
<td>1.32 (1.09–1.61)</td>
<td>0.005</td>
<td>0.80–2.18</td>
<td>44%</td>
<td>II</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Traumatic injury</td>
<td>Exposed vs. not exposed</td>
<td>2965</td>
<td>12</td>
<td>OR</td>
<td>1.41 (1.03–1.92)</td>
<td>0.03</td>
<td>0.60–3.29</td>
<td>42%</td>
<td>II</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anti-EBV IgG seronegativity</td>
<td>Negative vs. positive</td>
<td>2360</td>
<td>7</td>
<td>OR</td>
<td>0.13 (0.05–0.32)</td>
<td>2.0 \times 10^{-5}</td>
<td>0.01–1.39</td>
<td>52%</td>
<td>III/IV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anti-VCA IgG seropositivity</td>
<td>Positive vs negative</td>
<td>6,325</td>
<td>24</td>
<td>OR</td>
<td>4.52 (2.85–7.15)</td>
<td>3.4 \times 10^{-10}</td>
<td>0.87–23.42</td>
<td>58</td>
<td>III/IV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMD in femoral neck</td>
<td>High vs low</td>
<td>1420</td>
<td>10</td>
<td>OR</td>
<td>0.36 (0.21–0.61)</td>
<td>1.3 \times 10^{-4}</td>
<td>0.06–2.26</td>
<td>81</td>
<td>III/IV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMD in lumbar spine</td>
<td>High vs low</td>
<td>1490</td>
<td>11</td>
<td>OR</td>
<td>0.34 (0.24–0.50)</td>
<td>1.07 \times 10^{-8}</td>
<td>0.11–1.12</td>
<td>67%</td>
<td>III/IV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference</td>
<td>Neurological disorder</td>
<td>Level of comparison</td>
<td>Total number of participants</td>
<td>Number of primary studies</td>
<td>Effect size metric</td>
<td>Random-effects summary effect size (95% CI)</td>
<td>( P ) random</td>
<td>95% PI</td>
<td>( I^2 )</td>
<td>Small-study Effects / Excess statistical significance</td>
<td>Level of evidence</td>
</tr>
<tr>
<td>------------</td>
<td>-----------------------</td>
<td>---------------------</td>
<td>-----------------------------</td>
<td>--------------------------</td>
<td>-------------------</td>
<td>---------------------------------------------</td>
<td>----------------</td>
<td>--------</td>
<td>-------</td>
<td>---------------------------------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>BMD in hip</td>
<td>High vs low</td>
<td>1299</td>
<td>9</td>
<td>OR</td>
<td>0.33 (0.18–0.60)</td>
<td>2.95 ( \times ) ( 10^{-4} )</td>
<td>0.04–2.65</td>
<td>86%</td>
<td></td>
<td></td>
<td>III/IV</td>
</tr>
<tr>
<td><em>Chlamydia pneumoniae</em> DNA in CSF</td>
<td>Diseased vs. not diseased</td>
<td>1510</td>
<td>19</td>
<td>OR</td>
<td>3.22 (1.20–8.59)</td>
<td>0.02</td>
<td>N/A</td>
<td>88%</td>
<td></td>
<td></td>
<td>III/IV</td>
</tr>
<tr>
<td>Chronic cerebrospinal venous insufficiency</td>
<td>Diseased vs. not diseased</td>
<td>2149</td>
<td>19</td>
<td>OR</td>
<td>8.45 (3.47–20.56)</td>
<td>3.5 ( \times ) ( 10^{-4} )</td>
<td>0.33–2.17</td>
<td>80%</td>
<td></td>
<td></td>
<td>III/IV</td>
</tr>
<tr>
<td>Intrathecal production of Ig for <em>Chlamydia pneumoniae</em></td>
<td>Diseased vs. not diseased</td>
<td>628</td>
<td>6</td>
<td>OR</td>
<td>3.84 (1.32–11.21)</td>
<td>0.01</td>
<td>NA</td>
<td>55%</td>
<td></td>
<td></td>
<td>III/IV</td>
</tr>
<tr>
<td>Organic solvents</td>
<td>Exposed vs. not exposed</td>
<td>1,142,801</td>
<td>15</td>
<td>OR</td>
<td>1.54 (1.03–2.29)</td>
<td>0.03</td>
<td>0.37–6.39</td>
<td>77%</td>
<td></td>
<td></td>
<td>III/IV</td>
</tr>
<tr>
<td>Serum vitamin D</td>
<td>High vs. low</td>
<td>1836</td>
<td>11</td>
<td>OR</td>
<td>0.44 (0.24–0.78)</td>
<td>0.005</td>
<td>0.05–3.70</td>
<td>89%</td>
<td></td>
<td></td>
<td>III/IV</td>
</tr>
<tr>
<td>Reference</td>
<td>Neurological disorder</td>
<td>Level of comparison</td>
<td>Total number of participants</td>
<td>Number of primary studies</td>
<td>Effect size metric</td>
<td>Random-effects summary effect size (95% CI)</td>
<td>( P ) random</td>
<td>95% PI</td>
<td>( I^2 )</td>
<td>Small-study Effects / Excess statistical significance</td>
<td>Level of evidence</td>
</tr>
<tr>
<td>-----------</td>
<td>----------------------</td>
<td>---------------------</td>
<td>-----------------------------</td>
<td>--------------------------</td>
<td>------------------</td>
<td>--------------------------------------------</td>
<td>----------------</td>
<td>------</td>
<td>-------</td>
<td>---------------------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Serum homocysteine</td>
<td>High vs. low</td>
<td>1069</td>
<td>8</td>
<td>OR</td>
<td>4.57 (1.40–14.89)</td>
<td>0.01</td>
<td>0.06–338</td>
<td>96%</td>
<td>III/IV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serum uric acid</td>
<td>High vs. low</td>
<td>N/A</td>
<td>8</td>
<td>OR</td>
<td>0.28 (0.14–0.57)</td>
<td>4.1 \times 10^{-4}</td>
<td>0.02–3.29</td>
<td>87%</td>
<td>III/IV</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**PARKINSON’S DISEASE**

<table>
<thead>
<tr>
<th>Reference</th>
<th>Neurological disorder</th>
<th>Level of comparison</th>
<th>Total number of participants</th>
<th>Number of primary studies</th>
<th>Effect size metric</th>
<th>Random-effects summary effect size (95% CI)</th>
<th>( P ) random</th>
<th>95% PI</th>
<th>( I^2 )</th>
<th>Small-study Effects / Excess statistical significance</th>
<th>Level of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bellou 2016</td>
<td>Alcohol intake</td>
<td>High vs. Low</td>
<td>9994</td>
<td>33</td>
<td>RR</td>
<td>0.75 (0.66–0.85)</td>
<td>5.0 \times 10^{-6}</td>
<td>0.44–1.25</td>
<td>52.3</td>
<td>III</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Coffee intake</td>
<td>High vs. Low</td>
<td>5801</td>
<td>19</td>
<td>RR</td>
<td>0.67 (0.58–0.76)</td>
<td>3.4 \times 10^{-9}</td>
<td>0.45–1.00</td>
<td>42.9</td>
<td>III</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Smoking</td>
<td>Ever vs. never smokers</td>
<td>19,537</td>
<td>67</td>
<td>RR</td>
<td>0.64 (0.60–0.69)</td>
<td>1.3 \times 10^{-4}</td>
<td>0.45–0.92</td>
<td>49.6</td>
<td>II</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Physical activity</td>
<td>High vs. Low level</td>
<td>1348</td>
<td>5</td>
<td>HR</td>
<td>0.66 (0.57–0.78)</td>
<td>3.0 \times 10^{-7}</td>
<td>0.55–0.80</td>
<td>0</td>
<td>I</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Welding</td>
<td>Exposed vs not exposed</td>
<td>8198</td>
<td>9</td>
<td>RR</td>
<td>0.86 (0.80–0.92)</td>
<td>3.0 \times 10^{-5}</td>
<td>0.79–0.94</td>
<td>0</td>
<td>III</td>
<td></td>
</tr>
</tbody>
</table>

218
<table>
<thead>
<tr>
<th>Reference</th>
<th>Neurological disorder</th>
<th>Level of comparison</th>
<th>Total number of participants</th>
<th>Number of primary studies</th>
<th>Effect size metric</th>
<th>Random-effects summary effect size (95% CI)</th>
<th>( P ) random</th>
<th>95% PI</th>
<th>( I^2 )</th>
<th>Small-study Effects / Excess statistical significance</th>
<th>Level of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydrocarbon</td>
<td>Exposed vs not exposed</td>
<td>4483</td>
<td>14</td>
<td>OR</td>
<td>1.36</td>
<td>0.001</td>
<td>0.88–2.08</td>
<td>28.1</td>
<td></td>
<td></td>
<td>IV</td>
</tr>
<tr>
<td>Farming</td>
<td>Exposed vs not exposed</td>
<td>9533</td>
<td>38</td>
<td>OR</td>
<td>1.30</td>
<td>5.7 ( \times ) 10^{-6}</td>
<td>0.86–1.98</td>
<td>37.3</td>
<td></td>
<td></td>
<td>III</td>
</tr>
<tr>
<td>Organic solvents</td>
<td>Exposed vs not exposed</td>
<td>3811</td>
<td>18</td>
<td>OR</td>
<td>1.22</td>
<td>0.036</td>
<td>0.72–2.08</td>
<td>43.6</td>
<td></td>
<td></td>
<td>IV</td>
</tr>
<tr>
<td>Pesticides</td>
<td>Exposed vs not exposed</td>
<td>7151</td>
<td>39</td>
<td>OR</td>
<td>1.62</td>
<td>1.1 ( \times ) 10^{-10}</td>
<td>0.81–3.23</td>
<td>63.7</td>
<td></td>
<td></td>
<td>III</td>
</tr>
<tr>
<td>Rural living</td>
<td>Exposed vs not exposed</td>
<td>4306</td>
<td>31</td>
<td>OR</td>
<td>1.32</td>
<td>1.7 ( \times ) 10^{-6}</td>
<td>0.84–2.10</td>
<td>78.6</td>
<td></td>
<td></td>
<td>III</td>
</tr>
<tr>
<td>Well water drinking</td>
<td>Exposed vs not exposed</td>
<td>5037</td>
<td>28</td>
<td>RR</td>
<td>1.21</td>
<td>0.011</td>
<td>0.66–2.21</td>
<td>70.6</td>
<td></td>
<td></td>
<td>IV</td>
</tr>
<tr>
<td>Vitamin C intake</td>
<td>High vs. Low</td>
<td>936</td>
<td>7</td>
<td>OR</td>
<td>0.81</td>
<td>0.028</td>
<td>0.63–1.04</td>
<td>0</td>
<td></td>
<td></td>
<td>IV</td>
</tr>
<tr>
<td>Dairy products</td>
<td>High vs. Low</td>
<td>1083</td>
<td>7</td>
<td>RR</td>
<td>1.40</td>
<td>2.4 ( \times ) 10^{-5}</td>
<td>1.08–1.81</td>
<td>8.2</td>
<td></td>
<td></td>
<td>III</td>
</tr>
<tr>
<td>Reference</td>
<td>Neurological disorder</td>
<td>Level of comparison</td>
<td>Total number of participants</td>
<td>Number of primary studies</td>
<td>Effect size metric</td>
<td>Random-effects summary effect size (95% CI)</td>
<td>P random</td>
<td>95% PI</td>
<td>I²</td>
<td>Small-study Effects / Excess statistical significance</td>
<td>Level of evidence</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------------</td>
<td>---------------------</td>
<td>-----------------------------</td>
<td>--------------------------</td>
<td>------------------</td>
<td>---------------------------------------------</td>
<td>----------</td>
<td>-------</td>
<td>----</td>
<td>---------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Carbohydrate intake</td>
<td>High vs. Low</td>
<td>1482</td>
<td>8</td>
<td>RR</td>
<td>1.24 (1.05–1.48)</td>
<td>0.014</td>
<td>1.00–1.54</td>
<td>0</td>
<td></td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td>Energy intake</td>
<td>High vs. Low</td>
<td>1415</td>
<td>8</td>
<td>RR</td>
<td>1.39 (1.01–1.92)</td>
<td>0.042</td>
<td>0.50–3.90</td>
<td>83.8</td>
<td></td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td>Constipation</td>
<td>Exposed vs. unexposed</td>
<td>11,242</td>
<td>9</td>
<td>RR</td>
<td>2.30 (2.02–2.63)</td>
<td>3.5 × 10⁻³⁵</td>
<td>1.76–2.96</td>
<td>18.2</td>
<td></td>
<td>I</td>
<td></td>
</tr>
<tr>
<td>Head injury</td>
<td>Exposed vs. unexposed</td>
<td>35,799</td>
<td>22</td>
<td>OR</td>
<td>1.55 (1.33–1.81)</td>
<td>2.2 × 10⁻⁸</td>
<td>0.93–2.58</td>
<td>61</td>
<td></td>
<td>II</td>
<td></td>
</tr>
<tr>
<td>Anxiety or Depression</td>
<td>Exposed vs. unexposed</td>
<td>16,211</td>
<td>13</td>
<td>RR</td>
<td>1.86 (1.64–2.10)</td>
<td>2.6 × 10⁻²²</td>
<td>1.30–2.66</td>
<td>7.7</td>
<td></td>
<td>II</td>
<td></td>
</tr>
<tr>
<td>Hypertension</td>
<td>Exposed vs. unexposed</td>
<td>5933</td>
<td>12</td>
<td>RR</td>
<td>0.75 (0.61–0.90)</td>
<td>0.003</td>
<td>0.40–1.40</td>
<td>76.4</td>
<td></td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td>Serum Vitamin D</td>
<td>Exposed vs. unexposed</td>
<td>1008</td>
<td>7</td>
<td>OR</td>
<td>0.16 (0.05–0.50)</td>
<td>0.002</td>
<td>0.003–10.09</td>
<td>97.7</td>
<td></td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td>Nigral volume</td>
<td>High vs low</td>
<td>193</td>
<td>8</td>
<td>OR</td>
<td>0.31 (0.17–0.55)</td>
<td>8.3 × 10⁻⁵</td>
<td>0.06–1.46</td>
<td>47.4</td>
<td></td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td>Reference</td>
<td>Neurological disorder</td>
<td>Level of comparison</td>
<td>Total number of participants</td>
<td>Number of primary studies</td>
<td>Effect size metric</td>
<td>Random-effects summary effect size (95% CI)</td>
<td>P random</td>
<td>95% PI</td>
<td>I²</td>
<td>Small-study Effects / Excess statistical significance</td>
<td>Level of evidence</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------------</td>
<td>---------------------</td>
<td>-----------------------------</td>
<td>--------------------------</td>
<td>------------------</td>
<td>------------------------------------------</td>
<td>--------</td>
<td>-------</td>
<td>----</td>
<td>--------------------------------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Serum urate</td>
<td>High vs low</td>
<td>594</td>
<td>6</td>
<td>RR</td>
<td>0.65 (0.43–0.97)</td>
<td>0.034</td>
<td>0.23–1.82</td>
<td>42.1</td>
<td>IV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serum uric acid</td>
<td>High vs Low</td>
<td>1217</td>
<td>6</td>
<td>OR</td>
<td>0.39 (0.27–0.57)</td>
<td>6.8 × 10⁻²</td>
<td>0.13–1.22</td>
<td>75.9</td>
<td>II</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serum vitamin B₁₂</td>
<td>High vs Low</td>
<td>735</td>
<td>10</td>
<td>OR</td>
<td>0.50 (0.40–0.63)</td>
<td>4.7 × 10⁻⁹</td>
<td>0.31–0.82</td>
<td>23.8</td>
<td>IV</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

OR, odds ratio; RR, relative risk; HR, hazard ratio at follow-up or mean difference (MD); 95% CI, 95% confidence interval; PI, prediction interval.
Supplementary Table S2. Methodological quality of included umbrella reviews based on the AMSTAR criteria and score.

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Q1</th>
<th>Q2</th>
<th>Q3</th>
<th>Q4</th>
<th>Q5</th>
<th>Q6</th>
<th>Q7</th>
<th>Q8</th>
<th>Q9</th>
<th>Q10</th>
<th>Q11</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dinu, 2018</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>8</td>
</tr>
<tr>
<td>Poole, 2017</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>7</td>
</tr>
<tr>
<td>Grosso, 2017</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>8</td>
</tr>
<tr>
<td>Li, 2017</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>8</td>
</tr>
<tr>
<td>Belbasis, 2016</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>8</td>
</tr>
<tr>
<td>Bellou, 2016</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>8</td>
</tr>
<tr>
<td>Veronese, 2018 (Am J Clin Nutr)</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>8</td>
</tr>
<tr>
<td>Posadzki, 2018</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>9</td>
</tr>
<tr>
<td>Bellou, 2017</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>8</td>
</tr>
<tr>
<td>McRae, 2017</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>8</td>
</tr>
<tr>
<td>Belbasis, 2015</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>8</td>
</tr>
<tr>
<td>Theodoratou, 2014</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>8</td>
</tr>
<tr>
<td>Galbete, 2018</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>8</td>
</tr>
<tr>
<td>Veronese, 2018 (Clin Nutr)</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>8</td>
</tr>
</tbody>
</table>

AMSTAR: A Measurement tool to Assess Systematic Reviews [please see Ref.:18]; # Q1: A-priori design; Q2: Duplicate study selection and data extraction; Q3: Search comprehensiveness; Q4: Inclusion of gray literature; Q5: Included and excluded studies provided; Q6: Characteristics of the included studies provided; Q7: Scientific quality of the primary studies assessed and documented; Q8: Scientific quality of included studies used appropriately in formulating conclusions; Q9: Appropriateness of methods used to combine studies’ findings; Q10: Likelihood of publication bias was assessed; Q11: Conflict of interest—potential sources of support were clearly acknowledged in both the systematic review and the included studies.
2. Chapter 2

2.1 Supplementary File 1. Excel table with the studies identified by systematic query, their GEO accession number, the decision on their inclusion or exclusion from being reanalyzed (and the accompanying rationale), and the number of samples per each study.

<table>
<thead>
<tr>
<th>GEO accession</th>
<th>Inclusion or Exclusion (if exclusion, reason for exclusion) or Overlapping</th>
<th>Number of samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSE117589</td>
<td>Overlapping (Super Series containing series GSE117588, which are Included in the analysis, see below)</td>
<td>See below</td>
</tr>
<tr>
<td>GSE6677</td>
<td>Included (Data used for visualization only. No replication and no statistical analysis are performed)</td>
<td>6</td>
</tr>
<tr>
<td>GSE48350</td>
<td>Included</td>
<td>253</td>
</tr>
<tr>
<td>GSE125050</td>
<td>Included</td>
<td>113</td>
</tr>
<tr>
<td>GSE106241</td>
<td>Included</td>
<td>60</td>
</tr>
<tr>
<td>GSE102956</td>
<td>Included</td>
<td>26</td>
</tr>
<tr>
<td>GSE29652</td>
<td>Included</td>
<td>18</td>
</tr>
<tr>
<td>GSE117588</td>
<td>Included</td>
<td>6</td>
</tr>
<tr>
<td>GSE24290</td>
<td>Excluded (Sural nerve of diabetic neuropathy patients, no APOE status, APOE mentioned as affected gene)</td>
<td>NA</td>
</tr>
<tr>
<td>GSE39420</td>
<td>Excluded (Only 3 samples with APOE4 allele)</td>
<td>21 (in total)</td>
</tr>
<tr>
<td>GSE67333</td>
<td>Excluded (No samples contain APOE4 allele)</td>
<td>NA</td>
</tr>
<tr>
<td>GSE5281</td>
<td>Excluded (No information on APOE genotype status is provided with the dataset)</td>
<td>161</td>
</tr>
<tr>
<td>GSE31163</td>
<td>Excluded (No APOE status, APOE mentioned as affected gene)</td>
<td>NA</td>
</tr>
<tr>
<td>GSE11061</td>
<td>Excluded (No APOE status, APOE mentioned as affected gene)</td>
<td>NA</td>
</tr>
<tr>
<td>GSE60824</td>
<td>Excluded (No APOE status, APOE is mentioned in reference to ApoE deficient mice used in IL-17A inhibition experiment)</td>
<td>NA</td>
</tr>
<tr>
<td>GSE12696</td>
<td>Excluded (No APOE status, APOE is mentioned in reference to ApoE deficient mice used in atherosclerosis model)</td>
<td>NA</td>
</tr>
<tr>
<td>GSE44461</td>
<td>Excluded (No APOE status, APOE is mentioned in reference to ApoE deficient mice used in a migration experiment)</td>
<td>NA</td>
</tr>
<tr>
<td>GSE32136</td>
<td>Excluded (Lymphoblastoid cell lines)</td>
<td>NA</td>
</tr>
<tr>
<td>GSE26326</td>
<td>Excluded (Expression data derived from iliac artery of cynomolgus monkeys. No APOE status, APOE mentioned as affected gene)</td>
<td>NA</td>
</tr>
<tr>
<td>GSE110226</td>
<td>Excluded (E4 containing samples come from a mix of disease states (AD, fronto-temporal dementia, Huntington's) and severity)</td>
<td>20 (in total)</td>
</tr>
<tr>
<td>GSE48782</td>
<td>Excluded (Dataset profiles melanoma cell line treatment. No APOE status, APOE mentioned as affected gene)</td>
<td>NA</td>
</tr>
<tr>
<td>GSE79795</td>
<td>Excluded (Dataset profiles head and neck squamous cell carcinoma samples)</td>
<td>NA</td>
</tr>
<tr>
<td>GSE25425</td>
<td>Excluded (APOE status refers to mouse samples only)</td>
<td>NA</td>
</tr>
</tbody>
</table>
2.2 Supplementary File 2. Table with characteristics of the studies submitted to re-analysis

### 2.2.1 Features of Experimental design, GEO Platform Reference, Platform’s type, and Platform Name from the studies submitted to re-analysis.

<table>
<thead>
<tr>
<th>GEO accession</th>
<th>Experimental design</th>
<th>GEO Platform Reference</th>
<th>Platform’s type</th>
<th>Platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSE106241</td>
<td>71 human temporal cortical tissue samples were divided into 7 groups based on Braak staging (degree of AD-related neurofibrillary pathology) indicating the degree of disease severity. From 60 of these samples, transcripts were identified by microarray analysis.</td>
<td>GPL24170</td>
<td>Microarray</td>
<td>Agilent-044312 Human 8x60K Custom Exon array</td>
</tr>
<tr>
<td>GSE48350</td>
<td>Microarray data from normal controls (aged 20-99 years) and Alzheimer's disease cases, from 4 brain regions: hippocampus, entorhinal cortex, superior frontal cortex, post-central gyrus.</td>
<td>GPL570</td>
<td>Microarray</td>
<td>Affymetrix Human Genome [HG-U133_Plus_2]</td>
</tr>
<tr>
<td>GSE6677</td>
<td>The hippocampus was dissected at the time of autopsy and matching 100-200 mg portions of CA 1-4 was removed and used for RNA isolation and expression studies. From AD and normal individuals.</td>
<td>GPL4</td>
<td>SAGE (Serial Analysis of Gene Expression)</td>
<td>SAGE:10:NlaIII:Homo sapiens [HG-U133_Plus_2]</td>
</tr>
<tr>
<td>GSE29652</td>
<td>Astrocytes were isolated by laser micro-dissection from age, sex, and brain pH-matched cases of AD at different Braak stages.</td>
<td>GPL570</td>
<td>Microarray</td>
<td>Affymetrix Human Genome U133 Plus 2.0 Array</td>
</tr>
<tr>
<td>GSE117588</td>
<td>iPSCs were derived from individuals with sporadic AD APOE e4/e4 genotype. Parental iPSCs lines were edited to e3/e3 genotype and cerebral organoids were generated from each of the lines.</td>
<td>GPL16791</td>
<td>Next-generation sequencing</td>
<td>Illumina HiSeq 2500 (Homo sapiens)</td>
</tr>
</tbody>
</table>
iPSCs were derived from unaffected subject carrying APOE3 alleles. Cells were gene edited using CRISPR/Cas9 to generate APOE4 iPSCs from parental APOE3 cells. Isogenic iPSCs were then differentiated into neurons, astrocytes, and microglia-like cells.

RNA from purified cell types from AD and control post-mortem frozen superior frontal gyrus of AD and control patients. AD patients all had Braak stages V or VI and were also pathologically confirmed to have amyloid plaque.

<table>
<thead>
<tr>
<th>Experiment ID</th>
<th>Platform</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSE102956</td>
<td>GPL11154</td>
<td>Next-generation sequencing Illumina HiSeq 2000 (Homo sapiens), Illumina NextSeq 500 (Homo sapiens)</td>
</tr>
<tr>
<td>GSE125050</td>
<td>GPL18573</td>
<td>Next-generation sequencing Illumina HiSeq 2500 (Homo sapiens)</td>
</tr>
<tr>
<td></td>
<td>GPL16791</td>
<td>Next-generation sequencing Illumina HiSeq 2500 (Homo sapiens)</td>
</tr>
</tbody>
</table>
### 2.2.2 Features of Tissue, Sample size, Additional Variables, and Comments from the studies submitted to re-analysis

<table>
<thead>
<tr>
<th>GEO accession</th>
<th>Tissue</th>
<th>Sample size</th>
<th>Additional variables</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSE106241</td>
<td>Inferior temporal cortex</td>
<td>No ApoE4 copies n=30, one ApoE4 copy n= 30, two ApoE4 copies n=6</td>
<td>Braak stage, sex, age, rna quality,</td>
<td></td>
</tr>
</tbody>
</table>
Hippocampus, entorhinal cortex, superior frontal cortex, post-central gyrus.

Out of 253 samples in the dataset, 80 are affected by AD. The 80 AD samples are distributed into the following number of APOE genotype states: E2/E4 n=3, E3/E3 n=27, E3/E4 n=32, E4/E4 n=18. These are further non-uniformly distributed between Braak stages and brain regions.

Gender, age, brain region, Braak stage.

Removed all Normal samples and samples without Braak stage annotation. Model fitted for APOE status, gender, age, Braak stage and brain region (77 samples in total, of them 54 samples with at least 2 E4 allele).
| GSE6677 | Hippocampus, entorhinal cortex, superior frontal cortex, post-central gyrus. | Total n=4; One of each: Control ApoE3/E3, AD ApoE3/E3, AD ApoE3/E4 and AD ApoE4.E4. | None | No replication, no statistical analysis is performed. Data used for visualization only. |
| GSE29652 | Temporal cortex isolated astrocytes | Total n=18; n=3 APOE e4+ and 3 e4- for each of Braak groups: I-II, III-IV, and V-VI | Braak stage | Does not distinguish between e3/e4 and e4/e4, both considered e4+ |
| GSE117588 | iPSCs-derived cerebral organoids | Total n=6; ApoE3/E3 n=3; ApoE4/E4 n=3; | None | |

<table>
<thead>
<tr>
<th>GSE102956</th>
<th>iPSCs-derived neurons, astrocytes, and microglia-like cells</th>
<th>For neurons and astrocytes n=3, for microglia-like cells n=4 for each APOE3 and APOE4 cells.</th>
<th>None</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSE125050</td>
<td>Neuron, astrocyte, myeloid and endothelial cells from frontal gyrus.</td>
<td>Only AD patients have E4 allele. 2/3 n=2, 3/3 n=24; 3/4 n=20, 4/4 n=2</td>
<td>Disease status, sex, Braak stage (only V and VI), cell type, age</td>
</tr>
</tbody>
</table>

iPSCs were derived from a single individual and therefore replicates cannot be considered truly biological. It is not clear from the methods description at what point the replication was carried out, i.e. where replicates were generated from independently differentiated cells (separate plates), from the same cell pellet, same RNA sample or same library was sequenced multiple times. Replication can, therefore, be considered somewhere between biological and technical.

Only 13 samples with library size > 10M, of them only one is E4. Filtering by library diversity instead - require the library to have at least 10000 genes with at least 5 counts. This yields 19 E4 and 81 noE4.
<table>
<thead>
<tr>
<th>Note: In case of multiple papers being cited, Author name, Year and Journal are given for the oldest paper, the rest are cited in PMID field.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Splitting by cell type, each contains between 3 and 8 E4 samples and between 16 and 30 noE4 samples. Removing &quot;Control&quot; samples also leaves only Braak stages V and VI - no fitting on Braak stage is performed.</td>
</tr>
</tbody>
</table>
### 2.2.3 Features on First author, Year, Journal, PMID(s), Data of Publication, Last Update Date and Species from the studies submitted to re-analysis

<table>
<thead>
<tr>
<th>GEO accession</th>
<th>1st Author</th>
<th>Year</th>
<th>Journal</th>
<th>PMID(s)</th>
<th>Public on</th>
<th>Last update date</th>
<th>Species</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSE29652</td>
<td>Simpson JE</td>
<td>2011</td>
<td>Aging</td>
<td>21705112</td>
<td>Jun 01, 2011</td>
<td>Mar 25, 2019</td>
<td>Homo sapiens</td>
</tr>
<tr>
<td>GSE102956</td>
<td>Lin YT</td>
<td>2018</td>
<td>Neuron</td>
<td>29861287</td>
<td>May 31, 2018</td>
<td>May 15, 2019</td>
<td>Homo sapiens</td>
</tr>
</tbody>
</table>
2.3 **Supplementary File 3.** Excel table with cell counts’ numbers of the single-cell next-generation sequencing study by Mathys *et al.* (2019) reanalyzed.

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Astrocyte</th>
<th>Endothelial</th>
<th>Excitatory Neuron</th>
<th>Inhibitory Neuron</th>
<th>Microglia</th>
<th>Oligodendrocyte</th>
<th>Oligodendrocyte precursor</th>
<th>Pericyte</th>
</tr>
</thead>
<tbody>
<tr>
<td>23</td>
<td>706</td>
<td>29</td>
<td>8982</td>
<td>2349</td>
<td>449</td>
<td>4321</td>
<td>652</td>
<td>38</td>
</tr>
<tr>
<td>33</td>
<td>1760</td>
<td>59</td>
<td>18076</td>
<td>4546</td>
<td>812</td>
<td>9565</td>
<td>1322</td>
<td>63</td>
</tr>
<tr>
<td>34</td>
<td>499</td>
<td>9</td>
<td>4514</td>
<td>1430</td>
<td>390</td>
<td>1911</td>
<td>369</td>
<td>44</td>
</tr>
<tr>
<td>44</td>
<td>339</td>
<td>16</td>
<td>2602</td>
<td>661</td>
<td>235</td>
<td>2090</td>
<td>221</td>
<td>17</td>
</tr>
</tbody>
</table>
2.4 Supplementary Figure 1. Box and Whisker plot showing gene expression levels in the study with GEO accession number “GSE48350”
2.5 Supplementary Figure 2. Box and Whisker plot showing gene expression levels in the study with GEO accession number “GSE106241”
2.6 Supplementary Figure 3. Box and Whisker plot showing gene expression levels in the study with GEO accession number “GSE125050”.

![Box and Whisker plot showing gene expression levels](image)
2.7 Supplementary Figure 4. Box and Whisker plot showing gene expression levels in the study with GEO accession number “GSE29652”
2.8 Supplementary Figure 5. Scatter plot showing gene expression levels in the study with GEO accession number “GSE6677”
3. **Chapter 3**

3.1 **Supplementary file 1**

*Complete Search Algorithm*

("helicobacter pylori"[MeSH Terms] OR ("helicobacter"[All Fields] AND "pylori"[All Fields]) OR "helicobacter pylori"[All Fields]) AND ("parkinson's disease"[MeSH Terms] OR ("parkinson's"[All Fields] AND "disease"[All Fields]) OR "parkinson's disease"[All Fields])

3.2 **Supplementary File 2**

According to Cochrane Handbook for Systematic Reviews of Interventions (Version 5.1.0 [updated March 2011]) * supposing that given group means are $M_1$ and $M_2$, given SDs are $SD_1$ and $SD_2$, and the number of observations in each group is $N_1$ and $N_2$, respectively, then the pooled mean $(M_{1+2})$ is calculated by:

$$M_{1+2} = \frac{N_1 M_1 + N_2 M_2}{N_1 + N_2}$$

Given there is no overlap between the groups, the pooled SD $(SD_{1+2})$ of the new group is estimated by:

$$SD_{1+2} = \sqrt\frac{(N_1 - 1)SD_1^2 + (N_2 - 1)SD_2^2 + \frac{N_1 N_2}{N_1 + N_2} (M_1^2 + M_2^2 - 2M_1 M_2)}{N_1 + N_2 - 1}$$

REFERENCES


61. Solmi M, Correll CU, Carvalho AF, Ioannidis JPA. The role of meta-analyses and umbrella reviews in assessing the harms of psychotropic medications: beyond qualitative synthesis. Epidemiology and psychiatric sciences. 2018;27(6):537-42.


197. Fischer R, Maier O. Interrelation of oxidative stress and inflammation in neurodegenerative disease: role of TNF. Oxid Med Cell Longev. 2015;2015:610813-.


254


256


Tomasetti C, Vogelstein B. Cancer etiology. Variation in cancer risk among tissues can be explained by the number of stem cell divisions. Science. 2015;347(6217):78-81.


265


Cristea IA, Ioannidis JPA. P values in display items are ubiquitous and almost invariably significant: A survey of top science journals. PloS one. 2018;13(5):e0197440.


780. Lassmann H, Niedobitek G, Middeldorp JM, NeuroproMiSe EBVWG. Epstein-Barr virus in the multiple sclerosis brain: a controversial issue--report on a focused workshop held in the Centre for Brain Research of the Medical University of Vienna, Austria. Brain. 2011;134(Pt 9):2772-86.


Kop MLvd. Does antiretroviral therapy for HIV reduce the risk of developing multiple sclerosis? Neurol Neurosurg Psychiatry. 2015;86(3).


Matullo CM, O'Regan KJ, Curtis M, Rall GF. CNS recruitment of CD8+ T lymphocytes specific for a peripheral virus infection triggers neuropathogenesis during polymicrobial challenge. PLoS Pathog. 2011;7(12):e1002462.


284


Mentis AA, Chrousos GP. Springboard to Science from the "Worn Out Niche" of a Research Resources-limited Environment: Greece Must Get Serious about Research and Development. European journal of clinical investigation. 2019;e13167.

Press WH. Presidential address. What's so special about science (and how much should we spend on it?). Science. 2013;342(6160):817-22.

Syriza may have lost the election, but Greece's research reforms deserve to stay. Nature. 2019;572(7768):153.


Smith DR. The outsourcing and commercialization of science: Is the lab CEO the future of academic research? EMBO reports. 2015;16(1):14-6.


