



**Université  
de Lille**

EL Missaoui Imane  
Second year masters in Health Data science

## Masters Thesis

---

### **Investigating Maternal Inflammation During Pregnancy and its Association with Autism Spectrum Disorder in Offspring**

---

June 25, 2024

Jury members

PR BENJAMIN GINOUEHYA	Supervisor
DR DJAMEL ZITOUNI	President
DR DHIFLI WAJDI	Assessor

Promotion Thomas BAYES (1702-1761)

Faculty of Health Engineering and Management - ILIS 42 rue Ambroise Paré 59120 LOOS

# ACKNOWLEDGEMENTS

I would like to take this opportunity to express my gratitude to my supervisors Mr Guinhouya Benjamin and Mr Zitouni for their patience and support. I would also like to acknowledge the time and effort that went into the completion of this work. I extend my thanks to my friend Maky, my family and my cat for their encouragement and support throughout this project. In completing this project, I have had the opportunity to engage with a variety of resources and perspectives. I am grateful for the chance to learn and contribute to the ongoing research in this field. Thank you for the opportunity to share it.

Sincerely,  
Imane

# SUMMARY

<b>1</b>	<b>Introduction</b>	<b>7</b>
<b>2</b>	<b>Materials and Methods</b>	<b>9</b>
2.1	First pipeline . . . . .	9
2.1.1	Data Retrieval . . . . .	9
2.1.2	Natural language processing . . . . .	10
2.1.3	Semantic Enrichment . . . . .	11
2.1.4	Integration of External Knowledge . . . . .	12
2.1.5	Biomedical Knowledge Graph Deployment using Neo4j . . . . .	12
2.2	Second Pipeline . . . . .	12
2.2.1	Data Retrieval . . . . .	13
2.2.2	Splitting the Dataset into chunks . . . . .	13
2.2.3	Prompt engineering . . . . .	13
2.2.4	Natural Language Processing . . . . .	13
2.2.5	Entity1-Relation-Entity2 table . . . . .	14
2.2.6	Biomedical Knowledge graph deployment . . . . .	14
<b>3</b>	<b>Results</b>	<b>15</b>
3.1	DO Knowledge graph . . . . .	15
3.1.1	DO Knowledge graph analysis . . . . .	17
3.1.1.1	Key Nodes and Connections . . . . .	19
3.1.1.2	Immune Dysregulation and ASD . . . . .	19
3.1.1.3	Detailed Pathways to Autism . . . . .	19
3.2	LLama3 Knowledge graph . . . . .	20
3.2.1	LLama3 Knowledge graph analysis . . . . .	21
3.2.1.1	Connectivity and Topology . . . . .	21
3.2.1.2	Node Structure and Distribution . . . . .	21
3.2.1.3	Color-Coding of Nodes . . . . .	22
3.2.1.4	Detailed Analysis of the ASD Node . . . . .	22
<b>4</b>	<b>Discussion</b>	<b>26</b>
<b>5</b>	<b>Conclusion</b>	<b>29</b>
<b>Bibliography</b>		<b>34</b>
<b>A</b>	<b>Appendix</b>	<b>35</b>
I	LLM Prompts . . . . .	35
I.1	System Prompt . . . . .	35
I.2	User Prompt . . . . .	35
II	Neo4j Cipher Queries for DO in Order . . . . .	36
III	Neo4j Cipher Queries for LLM in Order . . . . .	37
IV	Dataset.txt . . . . .	38
V	LLM Output.csv . . . . .	45

# LIST OF FIGURES

2.1	First Pipeline using Spacy for NLP task . . . . .	9
2.2	Steps of literature search and test for eligibility of screened papers [1, 2, 3, 4, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20] . . . . .	10
2.3	Second pipeline using the LLama3 LLM . . . . .	13
2.4	Source relation target example dataframe . . . . .	14
3.1	DO labeled knowledge graph . . . . .	15
3.2	(a) DO Knowledge graph, (b) Color code of the DO graph . . . . .	16
3.3	Neo4j UI: Excerpt of all the relationships. . . . .	16
3.4	Disease ontology autism node view . . . . .	17
3.5	LLama3:8b-instruct-q8_0 labeled Knowledge graph . . . . .	20
3.6	LLama3:8b-instruct-q8_0 (a) Labeled and Cleaned. (b) Color code. . . . .	22
3.7	LLama3:8b-instruct-q8_0, Main ASD node and its connections(Central Node view) .	23

# LIST OF TABLES

2.1	Extracted Entity Types . . . . .	11
2.2	Wikidata property IDs and their corresponding relations. . . . .	11
3.1	All outgoing relationships of the autism node. . . . .	18

# GLOSSARY

**API** Application Programming Interface.

**ASD** Autism Spectrum Disorder.

**BBB** Blood-Brain Barrier.

**CNS** Central Nervous System.

**DO** Disease Ontology.

**GI** Gastro Intestinal.

**ID** Intellectual disability.

**IFN- $\gamma$**  Interferon gamma.

**IG** immunoglobulin.

**IL-1 $\alpha$**  Interleukin 1 alpha.

**IL-6** Interleukin 6.

**LLM** Large Language Model.

**MIA** Maternal Immune Activation.

**NER** Named-entity Recognition.

**NLP** Natural Language Processing.

**RDF** Resource Description Framework.

**RE** Relationship Extraction.

**SPARQL** SPARQL Protocol and RDF Query Language.

# INTRODUCTION

Autism Spectrum Disorder (ASD) is a complex neurodevelopmental disorder characterized by impairments in social interaction, communication, and repetitive behaviors. The exact etiology of ASD is not well understood, but it is thought to be a combination of genetic and environmental factors. One mechanism through which environmental factors may trigger the risk of ASD in children is maternal inflammation during pregnancy.

Several studies have shown that maternal infection during pregnancy is associated with an increased risk of ASD in offspring. For example, a systematic review and meta-analysis by Jiang et al. [1] found that maternal infections, particularly viral infections, during pregnancy are associated with an increased risk of ASD in children. Similarly, a study by Lee et al. [2] found that maternal hospitalization due to infection during pregnancy is linked to a 30% increase in the risk of autism spectrum disorders in the offspring.

Maternal inflammation during pregnancy has also been linked to alterations in fetal neurodevelopment. Kwon et al [3] reviewed the literature on maternal inflammation and its impact on fetal neurodevelopment, and found that maternal inflammation can lead to changes in fetal brain structure and function, including altered white matter development and reduced connectivity between brain regions. These changes may contribute to the development of ASD and other neurodevelopmental disorders.

In addition to maternal infection, other factors such as maternal obesity, stress, and autoimmune disorders have also been linked to an increased risk of ASD in offspring. A systematic review by Han et al. [4] found that maternal acute and chronic inflammation in pregnancy is associated with common neurodevelopmental disorders, including ASD.

Given the potential link between maternal inflammation and ASD, it is important to understand the key elements involved. For this thesis project, we aimed to explore the hypothesis “Maternal inflammatory status during pregnancy is a risk factor of ASD in offspring”. The primary objective is to conduct an in-depth analysis of existing literature on investigating this potential association between maternal inflammation and the incidence of ASD in offspring and identify causal pathways and key elements from various fields, including epidemiology, immunology, and neurobiology, that are involved in this association.

This thesis employs a knowledge graph approach to investigate the complex relationship between maternal inflammation during pregnancy and the risk of autism spectrum disorder (ASD). The need for a comprehensive, systemic, and holistic understanding of the processes involved necessitates a tool that can integrate and analyze data from diverse sources. A knowledge graph fulfills this requirement by consolidating data from scientific literature, databases, and clinical records, thereby providing a more encompassing view of the topic. The visual representation of a knowledge graph simplifies the analysis of inflammation, a multifaceted process, in the context of ASD and will allow us to gain insights into the inflammatory status of pregnant women and its implications on fetal development. Furthermore, the ability to identify patterns and generate new hypotheses is particularly valuable in a field where the exact mechanisms are still not fully understood. According to a study by Hänsel et al. [5], knowledge graphs have been successfully used to integrate and analyze biomedical data, leading to new insights and discoveries. Lastly, a knowledge graph serves as a platform for collaboration and knowledge sharing among researchers, potentially accelerating the pace of research and discovery.

In building knowledge graphs, numerous studies have employed a variety of libraries and explored various methodologies. For instance, a study by Milosevic et al. [6] compared several rule-based

and machine learning-based methods, including Naive Bayes, Random Forests, and transformer-based models like DistilBERT and PubMedBERT, for scalable relationship extraction from biomedical literature. This research highlighted the effectiveness of transformer-based models in handling both small and unbalanced datasets, with PubMedBERT-based models showing the best performance in terms of F1-score.

Additionally, another study expanded a database-derived biomedical knowledge graph by employing multi-relation extraction techniques from biomedical abstracts. This study utilized data programming to quickly annotate large datasets using multiple noisy signals, enhancing the process of knowledge extraction and graph creation in the biomedical domain. [7].

A considerable amount of research is currently focused on how to construct knowledge graphs using large language models. For example, Arsenyan et al. [8] investigated the potential of LLMs in automating the extraction of entities and relationships from unstructured text to facilitate biomedical knowledge graph construction, and demonstrated the effectiveness of these models in identifying interactions among medical entities.

Our objective is to evaluate the use of large language models (LLMs) for natural language processing (NLP) tasks in creating a biomedical knowledge graph. By comparing traditional NLP methods with advanced LLMs, we aim to identify potential improvements in efficiency and overall effectiveness. This approach will help us determine if LLMs can make the task of extracting knowledge from biomedical text more efficient, thus enhancing our investigation of the association between maternal inflammation during pregnancy and the risk of autism spectrum disorder.

# MATERIALS AND METHODS

We implemented two distinct pipelines to build a knowledge graph. The primary distinction between these pipelines lies in the approach to the Natural Language Processing (NLP) tasks. In the first pipeline, we used the SciSpaCy library to handle text tokenization, named entity recognition (NER), and other related tasks. In contrast, the second pipeline leveraged the capabilities of Large language models model for the NLP tasks.

## 2.1 First pipeline

In this first pipeline, we retrieved relevant biomedical literature from [Pubmed](#), a comprehensive and widely-used database of life sciences and biomedical research, then we parsed medical articles abstracts from PubMed using [BeautifulSoup](#) and we preprocessd the text using natural language processing (NLP) techniques, such as tokenization, and named entity recognition (NER), using the [Scispacy](#) library, a popular and efficient NLP tool. And to further enrich the semantics of the extracted entities, we used the [Disease Ontology](#), which provides standardized and comprehensive vocabularies for diseases and related concepts. We also integrated external knowledge from Wikidata, a large and freely-available knowledge graph, using the [Pywikibot](#) library, a Python interface for Wikimedia APIs. Finally, we deployed the resulting biomedical knowledge graph using [Neo4J](#), a powerful and scalable graph database, to gain insights into the potential association between maternal inflammatory status and the incidence of ASD in offspring.

The figure below provides an overview of the different steps in the pipeline and the tools and techniques that were used.

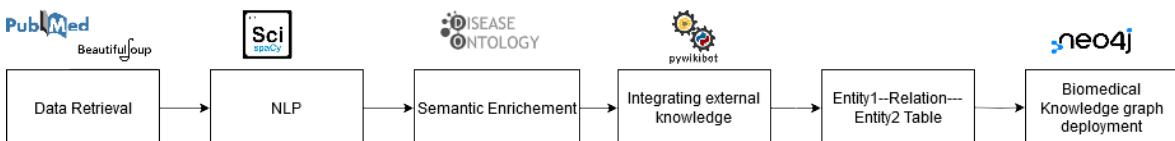


Figure 2.1: First Pipeline using Spacy for NLP task

### 2.1.1 Data Retrieval

The data collection process did not involve a pre-existing dataset; instead, information was gathered through a broad literature search in PubMed to identify relevant publications on the topic of association between maternal inflammation and the risk of ASD in offspring. The electronic search was restricted to the period 2013 to present, with only articles written in english reviewed. For the search we used keywords such as "Maternal Immune activation", "Pregnancy," "Inflammation," "gestational inflammation" and "Autism Spectrum Disorder". The search results were then screened using the [Rayyan](#) software tool to identify relevant studies. The selected studies were then used as collected data and for building our bibliography in [Zotero](#). Afterwards the abstracts of these studies were parsed using BeautifulSoup and stored in a single text file. Refer to Dataset.txt (Appendix IV).

The screening process resulted in the selection of 16 articles for their relevance to our research question. The screening process is illustrated bellow :

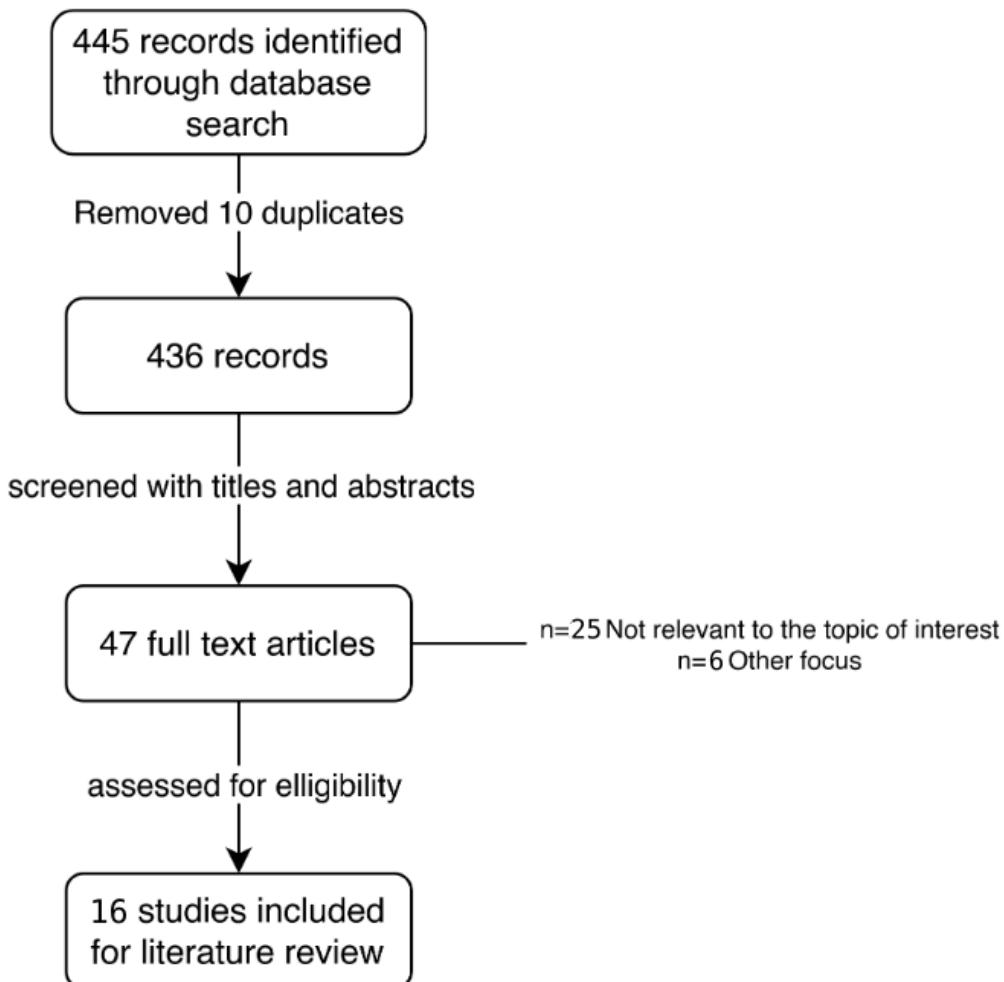


Figure 2.2: Steps of literature search and test for eligibility of screened papers [1, 2, 3, 4, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20]

### 2.1.2 Natural language processing

Following parsing of abstracts, our pipeline proceeded with the identification of entities and their semantic relationships. Utilizing SciSpaCy for Named Entity Recognition (NER), we extracted entities including proteins, diseases, chemicals, and biological processes. To facilitate the analysis of textual data, we leveraged the SciSpaCy library to extract relevant information from the abstracts. Specifically, we utilized the `en_core_sci_md` model provided by spaCy, which is well-suited for processing biomedical and scientific text data. Additionally, we implemented supplementary preprocessing and postprocessing steps to enhance the accuracy of the NER results. Our analysis yielded a total of 16 distinct entity types, encompassing:

Table 2.1: Extracted Entity Types

DISEASE	GENE	RNA	DNA
DRUG	CHEMICAL	CELL_LINE	CELL_TYPE
PROTEIN	PATHWAY MUTATION	ORGAN TISSUE	SYMPTOM
SPECIES	BIOLOGICAL_PROCESS	ENVIRONMENTAL_FACTOR	ANATOMY

### 2.1.3 Semantic Enrichment

To semantically enrich the extracted entities, we used the [DO ontology\(Ontology IRI\)](#). The Disease Ontology(DO) is a comprehensive and standardized sources of biomedical knowledge that has been developed as a standardized ontology for human disease with the purpose of providing the biomedical community with consistent, reusable and sustainable descriptions of human disease terms, phenotype characteristics and related medical vocabulary disease concepts.

We used the [rdflib](#) library and SPARQL queries to extract semantic relationships between entities and mapped them to their corresponding P-values using the DO ontology. These P-values are called property identifiers or properties for short in the format of "P[num]", and they are used in the Wikidata knowledge base to describe the different types of relationships between entities (items) in the database. This allowed us to infer additional information such as the relationships between the entities and their functional annotations. Our extraction yielded a total of 31 distinct relation types, including:

Table 2.2: Wikidata property IDs and their corresponding relations.

P31	(instance of)	P1554	(Orphanet ID)
P279	(subclass of)	P486	(MeSH ID)
P2293	(is a medical condition)	P699	(Disease Ontology ID)
P1050	(medical treatment)	P2176	(drug used for treatment)
P2176	(drug used for treatment)	P2888	(drug interaction)
P828	(has gene location)	P668	(gene)
P1199	(drug target)	P2670	(has parts of the class)
P780	(symptoms)	P2175	(clinical features)
P1120	(affected by)	P491	(ICD-10)
P493	(ICD-9)	P527	(has part)
P131	(located in)	P150	(pathway)
P2006	(treats)	P2116	(mode of action)
P805	(has function)	P144	(derived from)
P694	(regulates)	P224	(develops from)
P1552	(has quality)	P2289	(manifestation of)
P301	(method)	P518	(applies to)
P921	(main subject)		

### 2.1.4 Integration of External Knowledge

To further enrich the entities, we integrated external knowledge from Wikidata, using the Pywikibot library, which is a Python interface for the Wikimedia APIs. We linked the extracted entities to the corresponding Wikidata items and mapped them to their corresponding item IDs in the format of "Q[num]", which allowed us to retrieve additional information such as the aliases and descriptions of the entities. We also used the Wikidata SPARQL query service and the Wikidata RDF dumps to extract additional knowledge from Wikidata. A rule-based approach is employed to identify these relations, relying on the DO Ontology as a knowledge base.

### 2.1.5 Biomedical Knowledge Graph Deployment using Neo4j

To deploy the resulting biomedical knowledge graph, we used the Neo4j graph database [Neo4J website](#), which is a powerful and scalable tool for storing and querying graph data. We created nodes and edges in the graph to represent the entities and their relationships, and we stored additional information such as the ontology annotations and the Wikidata links in the node and edge properties. We also used the Neo4j Cypher query language and the Neo4j Browser to query and explore the graph.

## 2.2 Second Pipeline

In the second pipeline of our research, we aimed to extract entities and relations from our dataset (Dataset.txt) using a Large Language model and build a knowledge graph afterwards. To achieve this, we evaluated various open-source models based on their performance on biomedical texts. From the Models we chose the most recent and the ones that seemed to deliver the best performance for our use case of Named-entity Recognition (NER) and Relationship Extraction (RE). We considered primarily BioLlama, BioMistral, and Llama 3.

Among these, Meta Llama 3 stood out as an open-source large language model available in 8B and 70B parameter sizes. (pre-trained or instruction-tuned), delivering superior performance. Llama 3 instruction-tuned models are fine-tuned and optimized for dialogue/chat use cases and outperform many of the available open-source chat models on common benchmarks. [5] [21] [22]

We specifically chose to use the LLama3:8b-instruct-q8\_0 model, with 8B parameter sizes, instruction tuned and quantized on 8 bits. This choice was mainly due to our limited access to hardware as we used a RTX-3070ti graphics card with 8GB of VRAM.

The following figure illustrates the steps involved in the second pipeline:

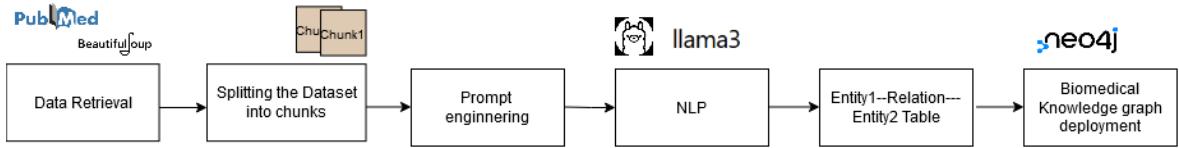


Figure 2.3: Second pipeline using the LLama3 LLM

### 2.2.1 Data Retrieval

The same data that was retrieved in subsection 2.1.1 (Dataset.txt) was used for the second pipeline in order to facilitate a fair basis for comparison.

### 2.2.2 Splitting the Dataset into chunks

To ensure that we could efficiently process the large amount of data we had collected, we decided to split the dataset into smaller chunks. This allowed us to run the LLM on each chunk separately, which not only reduced the overall processing time but also helped us to manage the memory requirements of the model. By carefully optimizing the chunk size, we were able to strike a balance between processing speed and memory usage, ultimately enabling us to extract high-quality information from the data.

### 2.2.3 Prompt engineering

In this step of the second pipeline, we focused on creating effective prompts for the LLM model. Prompts were the input text that we provided to the model, and the quality of the prompts can significantly impact the model's performance. We experimented with different prompt formats and lengths to determine the most effective way to elicit the desired information from the model. The optimal prompts can be viewed in the appendix. I.1 for the system prompt and I.2 for the associated user prompt.

### 2.2.4 Natural Language Processing

In this step, we utilized the LLM model to perform various NLP tasks, such as named entity recognition and relation extraction. To extract entities and relations from the text, we first specified the properties and entity types of interest to the model. These included various biomedical concepts such as diseases, genes, drugs, and biological processes, among others. We then provided the model with a prompt, which contained the text data we wanted to extract entities and relations from. We used Ollama <https://ollama.com/> to run the model that can be accessed through the following link [llama3:8b-instruct-q8\\_0](#) on the prompt and extract the relevant entities and relations. The Ollama tool is a command-line interface that allows users to easily apply pre-trained language models to their own text data. It is highly customizable and allows for the extraction of a wide range of entities and relations.

### 2.2.5 Entity1-Relation-Entity2 table

The extracted entities and relations were subsequently integrated into a structured data framework, specifically a pandas DataFrame, serving as a repository for the compiled information. The resulting DataFrame features distinct columns for storing entities, relationships, and their corresponding property values. An example of the source target relation table is presented below (Figure 2.4).

This intermediate structure is subsequently leveraged to construct a knowledge graph utilizing Neo4j, a graph database management system capable of efficiently storing and querying complex networks of interconnected data.

	source_name	source_q	rel_p	rel_name	target_name	target_q
0	autism spectrum disorder	Q1436063	P31	instance_of	neurodevelopmental disorder	Q3450985
1	autism spectrum disorder	Q1436063	P31	instance_of	disability	Q12131
2	autism spectrum disorder	Q1436063	P31	instance_of	class of disease	Q112193867
3	autism spectrum disorder	Q1436063	P279	subclass_of	pervasive developmental disorder	Q6691991
4	autism spectrum disorder	Q1436063	P31	instance_of	neurodevelopmental disorder	Q3450985

Figure 2.4: Source relation target example dataframe

### 2.2.6 Biomedical Knowledge graph deployment

The previous table was used to deploy the knowledge graph on a [Neo4J Sandbox](#) free instance for further exploration and visualization. In our project, we utilized the Cypher query language to interact with and retrieve data from our Neo4j graph database. That enabled us to visualize the graph and interact with it through performing a wide range of operations, including creating and deleting nodes and edges, querying for specific patterns and relationships, and aggregating and analyzing data.

We started off by visualizing the Knowledge graph in the Neo4J interface, then we proceeded to delete all duplicate relationships and to label them by category based on their type key. We chose not to label the nodes that were categorized as "node" in our dataset. These nodes have different key types and a small count per type, which made it less important to distinguish them visually. The Cypher Queries are noted and explained for each following step in section II of the Appendix.

# RESULTS

## 3.1 DO Knowledge graph

The Knowledge graph obtained by this pipeline is illustrated in Fig 3.1 or Fig3.2 bellow :

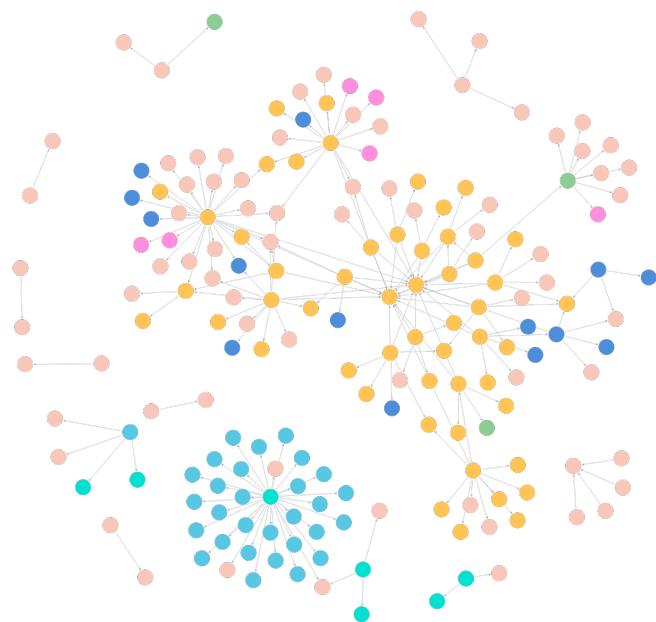


Figure 3.1: DO labeled knowledge graph

## CHAPTER 3

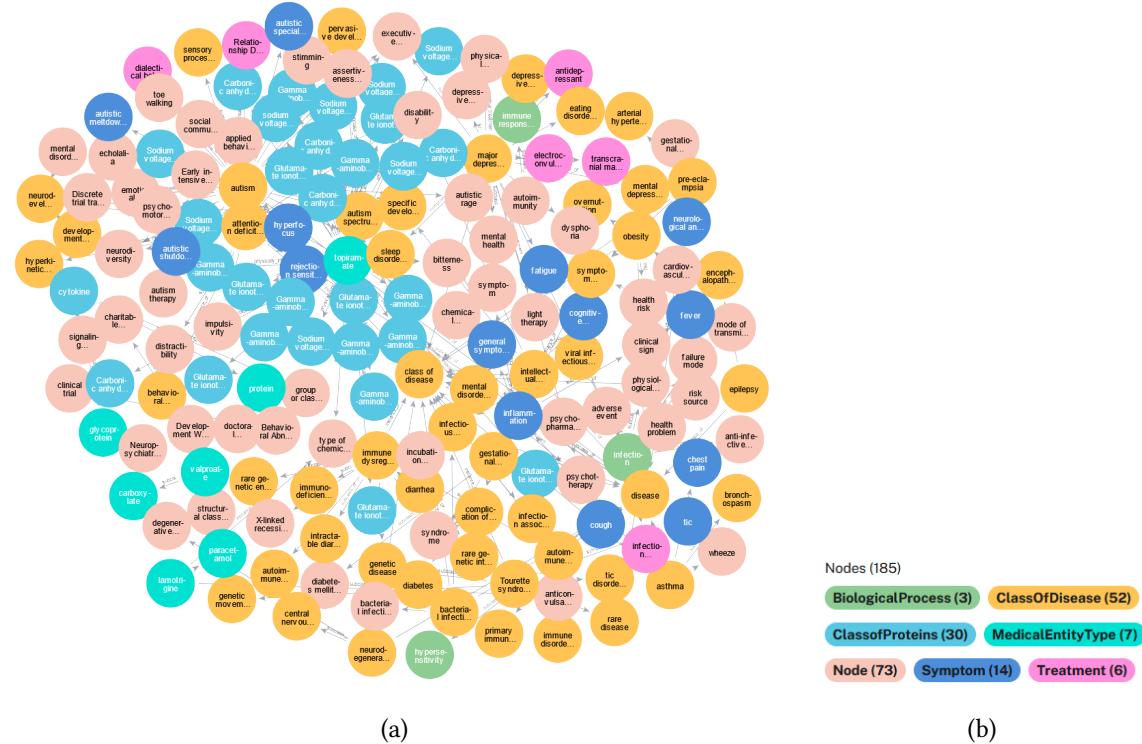


Figure 3.2: (a) DO Knowledge graph, (b) Color code of the DO graph

The graph can be visualized in the Neo4j web interface as a Table like this :



Figure 3.3: Neo4j UI: Excerpt of all the relationships.

Here for example we visualize all the nodes that connect to the main autism node for a clearer visual:



Figure 3.4: Disease ontology autism node view

Subsequently we showed a small excerpt of the relationships found in this graph in the following Source-Relation-Target Table 3.1.

### 3.1.1 DO Knowledge graph analysis

The knowledge graph obtained with the DO ontology conducted in this study revealed a complex network of 185 nodes and 211 relationships, providing a comprehensive view of the various biological processes, diseases, proteins, medical entities, symptoms, and treatments involved. The nodes in the graph were labeled with six different types, allowing for a more nuanced understanding of the relationships between them. For instance, the "contributing\_factor\_of" relationship was used to indicate that a particular biological process or medical entity contributes to the development of a disease, while the "medical\_treatment" relationship was used to indicate that a particular drug or therapy is used to treat a disease.

Table 3.1: All outgoing relationships of the autism node.

<b>Source</b>	<b>Relation</b>	<b>Target</b>
autism	instance_of	class of disease
autism	instance_of	pervasive developmental disorder
autism	instance_of	disability
autism	subclass_of	neurodevelopmental disorder
autism	subclass_of	disease
autism	subclass_of	autism spectrum disorder
autism	subclass_of	neurodiversity
autism	medical_treatment	autism therapy
autism	medical_treatment	Relationship Development Intervention
autism	medical_treatment	psychomotor education
autism	medical_treatment	Discrete trial training
autism	medical_treatment	assertiveness training
autism	medical_treatment	Early intensive behavioral intervention (EIBI) for young children with autism spectrum disorders (ASD)
autism	medical_treatment	applied behavior analysis
autism	medical_treatment	dialectical behavior therapy
autism	symptoms	stimming
autism	symptoms	sleep disorder
autism	symptoms	autistic shutdown
autism	symptoms	autistic rage
autism	symptoms	social communication disorder
autism	symptoms	toe walking
autism	symptoms	autistic meltdown
autism	symptoms	hyperfocus
autism	symptoms	autistic special interest
autism	symptoms	echolalia
autism	symptoms	executive dysfunction
autism	symptoms	sensory processing differences

### 3.1.1.1 Key Nodes and Connections

#### Central Nodes:

- Autism: This node connects to various developmental, behavioral, and psychological conditions.
- Inflammation: Linked to various health risks, diseases, and symptoms.

#### Relevant Connections to Autism:

- Behavioral and Developmental Issues: Autism is connected to nodes like "cognitive," "emotional," and "behavioral," which are common issues associated with ASD. Specific Developmental Disorders: Autism is directly linked to "specific developmental disorders," indicating its broad impact on various developmental aspects.
- Immune System: While the direct link to "immune dysregulation" isn't explicitly visible, nodes like "autoimmune disorders" and "primary immunodeficiency" suggest immune-related influences.

#### Connections to Inflammation:

- Obesity, Asthma, Cardiovascular Issues: These health risks are linked to inflammation and can affect maternal health during pregnancy.
- Pregnancy-Related Complications: Conditions like gestational diabetes and pre-eclampsia are connected to inflammation and impact pregnancy outcomes.

### 3.1.1.2 Immune Dysregulation and ASD

**Immune Dysregulation Node:** This node links to:

- Autoimmune Disorders: Autoimmune conditions often involve chronic inflammation and immune system dysfunction.
- Primary Immunodeficiency: These conditions indicate a compromised immune system, which can lead to increased inflammation.

### 3.1.1.3 Detailed Pathways to Autism

- **Maternal Inflammation:** Maternal inflammation can result from autoimmune disorders or infections during pregnancy or conditions such as gestational diabetes and pre-eclampsia are inflammatory conditions that can impact fetal development.

- **Impact on Fetal Development:**

Neurodevelopmental Disorders: Chronic inflammation during pregnancy can affect fetal brain development, potentially leading to neurodevelopmental disorders like ASD.

Behavioral Issues: Immune dysregulation and inflammation can lead to developmental delays and behavioral problems associated with ASD.

- **Direct Evidence in Graph:**

Inflammation to ASD Pathway: Inflammation is linked to various health risks and conditions during pregnancy.

Immune dysregulation, indicated by nodes like "autoimmune disorders" and "primary immunodeficiency," is associated with inflammation.

Autism is connected to various developmental and cognitive conditions, suggesting a pathway where maternal inflammation impacts fetal neurodevelopment, leading to ASD.

The autism cluster was found to be closely connected to the inflammation cluster, suggesting that inflammation may play a role in the development of autism.

## 3.2 LLama3 Knowledge graph

The use of the LLama3:8b-instruct-q8\_0 model and the Ollama tool in our second pipeline allowed us to efficiently extract entities and relations from the parsed abstracts. The resulting knowledge graph provided us with valuable insights into the relationships between various biomedical concepts involved in Maternal inflammation's association with Autism Spectrum Disorder. Appendix V

The presented graph is a complex network visualization displaying nodes and their interconnections.

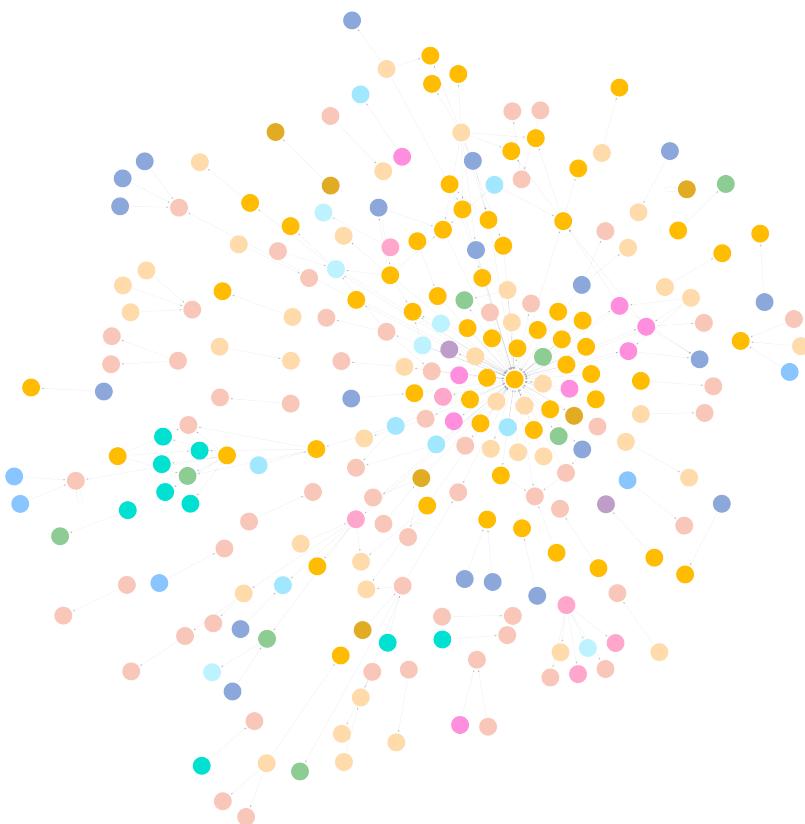


Figure 3.5: LLama3:8b-instruct-q8\_0 labeled Knowledge graph

### 3.2.1 LLama3 Knowledge graph analysis

#### 3.2.1.1 Connectivity and Topology

The graph exhibits a hierarchical structure with a clear distinction between the highly connected central cluster and the more loosely connected peripheral clusters and isolated nodes. The central cluster's high degree of connectivity suggests it is a critical component of the network, potentially responsible for the main flow of information or interactions. Peripheral clusters and isolated nodes indicate the presence of sub-communities or less central entities within the network, with limited direct interaction with the central cluster.

#### 3.2.1.2 Node Structure and Distribution

**Central Cluster** At the core of the graph, there is a densely connected central cluster. This cluster contains a high concentration of nodes interconnected with multiple edges, indicating a significant level of connectivity among these nodes. Within this central cluster, there are several nodes acting as hubs, each having numerous connections to other nodes. These hubs are likely central figures or key components in the network, playing a critical role in maintaining the network's connectivity.

**Peripheral Cluster** Surrounding the central cluster, there are multiple smaller clusters of nodes, each typically comprising a few nodes with limited interconnections. These peripheral clusters are sparsely connected to the central cluster or to each other, suggesting a more decentralized and less integrated structure in these areas of the network. Each peripheral cluster often has a star-like configuration, with one central node connected to a few surrounding nodes.

**Isolated Nodes and Small Groups** On the outermost edges of the graph, there are several isolated nodes and small groups of nodes, each group containing between two to five nodes. These nodes are either completely isolated or connected in a linear or star-like manner.

### 3.2.1.3 Color-Coding of Nodes

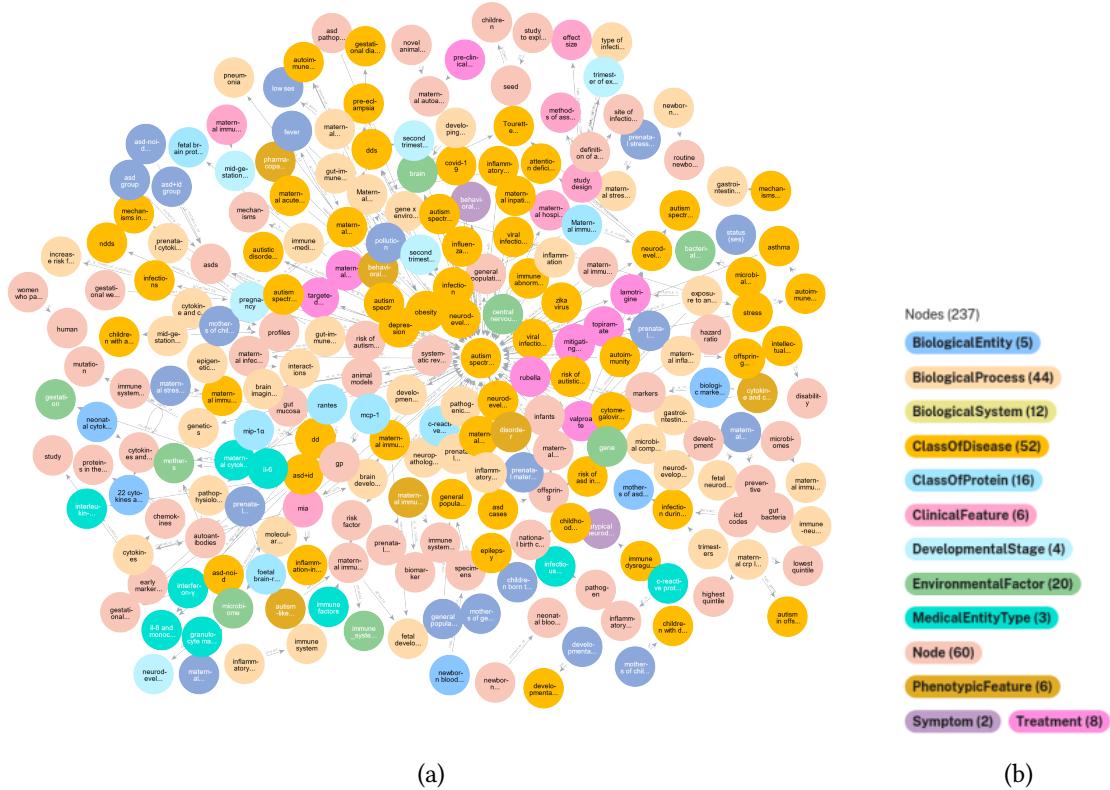


Figure 3.6: LLama3:8b-instruct-q8\_0 (a) Labeled and Cleaned. (b) Color code.

Nodes are color-coded based on their type, which helps in identifying and differentiating the various entities or concepts represented in the network and illustrates their interconnections through directed edges with descriptive labels. The central cluster contains a mix of different colored nodes, indicating a diverse interaction among various types of entities. Peripheral clusters often have nodes of similar colors, suggesting specialized sub-communities or related concepts.

### 3.2.1.4 Detailed Analysis of the ASD Node

The graph excerpt focuses on a central hub of the "autism spectrum disorder" (ASD) node and its surrounding nodes. The graph uses color coding to distinguish between different types of entities and illustrates their interconnections through directed edges with descriptive labels.



Figure 3.7: LLama3:8b-instruct-q8\_0, Main ASD node and its connections(Central Node view)

**Central Node: Autism Spectrum Disorder (ASD)** The central node in this excerpt is "autism spectrum disorder," represented in yellow. This node is highly connected, serving as the focal point for numerous related nodes and concepts.

**Key Surrounding Nodes and Their Connections** The "Inflammation" node is connected to the node "autism spectrum disorder" which indicates a recognized link between inflammation and ASD.

The "Maternal Immune Activation" (MIA) node is linked to ASD, highlighting the impact of maternal immune responses during pregnancy on the development of ASD. It is also connected to various other nodes like "maternal immune activation" and "cytokines," emphasizing its role in prenatal influences.

Nodes like "maternal infection," "maternal antibodies," and "maternal inflammation" are connected

to ASD, indicating the significant role of maternal health and prenatal conditions on ASD outcomes.

A node labeled "genetics" is connected to ASD, which underscores the genetic component in the etiology of ASD.

"Pollution" and "obesity" are linked to ASD, suggesting environmental and lifestyle factors contributing to ASD risk.

Nodes such as "stress," "viral infections," and "autoimmunity" are connected to ASD, indicating comorbid conditions or factors exacerbating the risk of ASD.

Nodes like "valproate," "lamotrigine," and "topiramate" are connected to ASD, highlighting the impact of specific medications taken during pregnancy.

Nodes such as "viral infections" and specific viruses like "rubella" are connected to ASD, illustrating the role of infections in the development of ASD.

**3.2.1.4.1 Directionality and Relationships** The directed edges show the relationship type, with descriptions like "is connected with," "is a medical condition connected with," or "may interact with." For example, "stress" and "asthma" are connected to "Autism Spectrum Disorder" via the edge "is associated with" indicating their role as contributing factors. Whereas some bidirectional relationships for example "Autism Spectrum Disorder (ASD)" and "pregnancy" might imply a more complex interaction of risks associated with ASD.

**3.2.1.4.2 Main insights** This knowledge graph allowed visualizing the relationships between various factors and the risk of autism spectrum disorder (ASD). Here are some key insights from the graph:

Several clusters of factors are associated with the risk of ASD. These include:

- **Maternal Health and Immune Response**

Maternal infection, Maternal acute and chronic inflammation during pregnancy, Maternal immune activation (MIA), Maternal hospitalization with infection during pregnancy, Maternal inflammation, Maternal immune dysfunction, Second Trimester infection, Second Trimester infection accompanied by fever, Pregnancy, C-reactive protein, mip-1 $\alpha$ , rantes, Neonatal cytokines and chemokines (IL-6, Il-1 $\alpha$ , interferon- $\gamma$  [markers of inflammation]).

- **Infections and Immune System**

Viral infection, immune abnormalities, Influenza, covid-19, Rubella, Zika virus, cytomegalovirus, Infection during pregnancy

- **Environmental and External Influences**

Pollution, Stress, Obesity

- **Developmental Factors**

Prenatal exposure, Brain imaging, Neurodevelopment, Autism-specific, Cognitive Foetal brain-reactive antibodies

- **Medications and Chemicals factors**

Valproate, Prenatal Topiramate, Lamotrigine

- **Additional Factors**

gut-immune-brain paradigm, Asthma, Genetics, Microbiome alteration.

- **Treatment**

Targeted therapeutics

## DISCUSSION

The graphs generated by the two pipelines exhibit notable differences in structure and content, reflecting the distinct methodologies employed in their creation. The first pipeline, which utilized the SciSpaCy library along with Disease Ontology, produced a densely interconnected graph with a high level of detail in specific biomedical entities and their relationships. The use of Disease Ontology provided a standardized vocabulary, ensuring consistency and comprehensiveness in the representation of diseases and related concepts. The entities in this graph are well-defined, particularly in the context of diseases and medical conditions, demonstrating the efficacy of the Disease Ontology in enriching the semantic content. Named entity recognition (NER) is precise, capturing relevant biomedical terms accurately. Overall, this graph offers a robust and detailed map of biomedical knowledge, specifically tailored to the domain of diseases, which can be highly beneficial for exploring the association between maternal inflammation and ASD.

In contrast, the second pipeline, which leveraged the capabilities of Llama 3 for NLP tasks, resulted in a graph that is more expansive and appears to capture a broader range of entities and relationships. The structure is slightly more spread out, indicating a potentially more comprehensive extraction of knowledge. Leveraging the advanced NLP capabilities of Llama 3, the graph incorporates a wider variety of terms and connections, reflecting the model's ability to understand context and infer relationships. This model's ability to understand context and infer relationships might result in richer and more diverse connections. The resulting graph suggests a broader and potentially more nuanced understanding of the biomedical text, providing a wider array of entities and connections that can be explored for insights into maternal inflammation and ASD.

For future research, it is crucial to complement these qualitative assessments with quantitative measures to evaluate the effectiveness of the pipelines. One approach is to use precision, recall, and F1-score to evaluate the accuracy of entity recognition and relationship extraction by comparing the results against a manually curated gold standard dataset. Precision measures the proportion of correctly identified items, recall measures the proportion of true items that were correctly identified, and the F1-score provides a balance between the two. This method is well-documented in the field of information retrieval [23].

Another approach is to use network analysis metrics to evaluate the structural properties of the knowledge graphs. Metrics such as node degree, clustering coefficient, and average path length can provide insights into the centrality, cohesiveness, and efficiency of information spread within the graph. These metrics are standard in network analysis and can help quantify the structural characteristics of the generated graphs [24].

Assessing ontology coverage and consistency is also important. This involves evaluating the extent to which the generated graphs cover the relevant biomedical ontologies and maintain consistency with established biomedical terminologies. The OBO Foundry provides a coordinated approach to ontology development that can be used as a reference for this assessment [25].

The findings of this thesis highlight the significant role of maternal inflammation during pregnancy in the development of Autism Spectrum Disorder (ASD) in offspring. A significant strength of this thesis is the construction of a knowledge graph to illustrate the complex relationships between maternal inflammation and ASD. This innovative approach allows for a more integrative and comprehensible visualization of the data, which enhances our understanding of how various factors interact to influence ASD risk. Compared to existing literature [1, 2, 17] which has also explored the role of maternal infections and inflammation in ASD, this study provides a unique methodological perspective that synthesizes and visualizes these relationships more effectively and provides a

holistic view of the association between maternal inflammation and ASD.

The knowledge graphs generated in this project have provided valuable insights into the mechanisms underlying the association between maternal inflammation and Autism Spectrum Disorder (ASD) and the key elements involved in this association. However, there are several potential improvements that could be made to further enhance the quality and accuracy of the results and gain a more comprehensive understanding of this association.

Firstly, it may be more informative to perform NLP on the full-text of articles, rather than just the abstracts. This could provide a more comprehensive representation of the results and conclusions discussed in the articles. Studies have shown that full-text mining can capture additional knowledge and context that abstracts alone may miss, thus improving the accuracy and depth of biomedical information extraction. [26]

Medical text can be complex and difficult to analyze using NLP techniques. Therefore, it may be necessary to preprocess the text before analysis. Medical text is usually preprocessed in three steps, medical text processing, feature extraction and finally ensemble learning model for text classification.[27]. Regarding singular texts, this could include removing stop words, stemming or lemmatizing words, and correcting spelling and grammar errors[28]. Cleaning the text by reducing noise rectifying structural anomalies and standardising text formats.[29]

Feature engineering is an important step in the NLP process, as it can significantly impact the accuracy and precision of entity and relation extraction. Therefore, it may be beneficial to explore different feature engineering techniques and preselect the target entity types and relations to be extracted. Chen, et. al. [30] demonstrate the effectiveness of a combined approach on various corpora. This could help to reduce noise and improve the overall quality of the extracted knowledge graph. Deep learning models allow for automatic feature extraction [29]. Which can be incorporated as lexical features and classifiers to enhance the performance of entity extraction and relationship classification[31]. Yang, et. al [32] have also shown that Deep Learning based approaches can outperform traditional approaches by model depth and accuracy.

One of the main strengths of the first pipeline we used in this thesis is its ability to leverage the rich and comprehensive knowledge of the DO ontology to identify and extract relevant biomedical concepts and relationships from unstructured text. While the DO ontology is a valuable resource for biomedical research, it may not cover all relevant concepts and relationships related to the research question. Therefore, it may be beneficial to explore other ontologies, such as the Gene Ontology (GO) [33] or the Neuro Behavior Ontology (NBO) [34], and integrate them into the knowledge graph.

The extracted knowledge graph may contain a wealth of information about the relationships between different biomedical concepts, including potential causal relationships [35]. Therefore, it may be beneficial to conduct a causation study within the knowledge graph to identify and validate potential causal mechanisms underlying the research question. This study could be conducted by using tools like CasualKG [36] for enhancing explanability in causal relationships within our knowledge graphs. Additionally large scale knowledge bases like CauseNet [37] may be used for reference of properly labelled causal relationships. Blomqvist et. al. [38] present a possible approach for explicit causal knowledge extraction. This approach should help distinguish correlation from causation and provide a transparent framework for research.

Finally, in the second pipeline we used the LLama3 model to perform NLP tasks. The main strength of this pipeline is that it leverages the latest advances in language modeling and transfer learning to provide more accurate and context-aware entity and relation extraction. However, the second pipeline has some potential weaknesses. For example, the LLama3 model is relatively large and

resource-intensive, which may make it more difficult to deploy and scale in practice. Additionally the quality of the output isn't consistent over the entire dataset and hallucinations still occur if the sys-prompt and user-prompt aren't tweaked for each model properly. Furthermore, while the model is highly accurate, it may not always be transparent or interpretable. Fine-tuning the Large language model on the specific dataset used for the research question may improve the accuracy and precision of entity and relation extraction. Therefore, it may be beneficial to build a dataset that is specifically designed to the research question and use it to fine-tune the NLP model. This could help to reduce noise and improve the overall quality of the extracted knowledge graph. [39, 40, 41, 42].

## CONCLUSION

In this thesis, we explored the complex relationship between maternal inflammation during pregnancy and the risk of autism spectrum disorder (ASD) in offspring through the construction of a biomedical knowledge graph. By employing two distinct NLP pipelines, one using SciSpaCy library with Disease Ontology and the other utilizing the Llama 3 model, we were able to compare different methodologies for their effectiveness in extracting and representing biomedical knowledge.

Overall, this thesis provides new and substantial insights into the role of maternal inflammation in the development of ASD, using a knowledge graph model to demonstrate the critical connections. By integrating data from multiple studies, the knowledge graph offered a comprehensive view of how maternal inflammation, immune response proteins, and genetic factors influence ASD risk. The knowledge graph analysis helped identify potential biological pathways and interactions that might be overlooked in traditional analyses. Furthermore, it allowed for the visualization of complex relationships and dependencies between various risk factors, facilitating a deeper understanding of ASD etiology.

This opens new avenues for future research to explore the complex interplay of various factors influencing ASD risk and highlights the importance of maternal health management during pregnancy in mitigating the risk of ASD. The use of a knowledge graph represents a significant methodological advancement, providing a more comprehensive and integrative approach to understanding the multifactorial nature of ASD. However, there are several potential improvements that could be made to further enhance the quality and accuracy of the results. By implementing these improvements, we can gain more accurate, comprehensive, and actionable insights into the underlying mechanisms of Autism spectrum Disorder.

Ultimately, this research underscores the potential of advanced NLP models and knowledge graph methodologies in uncovering complex biological relationships, thereby contributing to a more informative and more effective biomedical research. As we move forward, it is crucial to address the challenges identified and leverage the opportunities for improvement to push the boundaries of what medical text analysis can achieve.

# BIBLIOGRAPHY

1. Jiang Hy, Xu Ll, Shao L, Xia Rm, Yu Zh, Ling Zx, Yang F, Deng M, and Ruan B. Maternal infection during pregnancy and risk of autism spectrum disorders: A systematic review and meta-analysis. en. *Brain, Behavior, and Immunity* 2016 Nov; 58. systematic review and meta analysis to keep:165–72. doi: [10.1016/j.bbi.2016.06.005](https://doi.org/10.1016/j.bbi.2016.06.005) PMID: 27287966. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0889159116301544> [Accessed on: 2023 Dec 27]
2. Lee BK, Magnusson C, Gardner RM, Blomström Å, Newschaffer CJ, Burstyn I, Karlsson H, and Dalman C. Maternal hospitalization with infection during pregnancy and risk of autism spectrum disorders. *Brain, behavior, and immunity* 2015 Feb; 44. to keep:100–5. doi: [10.1016/j.bbi.2014.09.001](https://doi.org/10.1016/j.bbi.2014.09.001) PMID: 25218900. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4418173/> [Accessed on: 2023 Dec 27]
3. Kwon HK, Choi GB, and Huh JR. Maternal inflammation and its ramifications on fetal neurodevelopment. English. *Trends in Immunology* 2022 Mar; 43. Review, to keep:230–44. doi: [10.1016/j.it.2022.01.007](https://doi.org/10.1016/j.it.2022.01.007) PMID: 35131181. Available from: [https://www.cell.com/trends/immunology/abstract/S1471-4906\(22\)00007-2](https://www.cell.com/trends/immunology/abstract/S1471-4906(22)00007-2) [Accessed on: 2023 Nov 17]
4. Han VX, Patel S, Jones HF, Nielsen TC, Mohammad SS, Hofer MJ, Gold W, Brilot F, Lain SJ, Nassar N, and Dale RC. Maternal acute and chronic inflammation in pregnancy is associated with common neurodevelopmental disorders: a systematic review. *Translational Psychiatry* 2021 Jan; 11. to keep systematic review:71. doi: [10.1038/s41398-021-01198-w](https://doi.org/10.1038/s41398-021-01198-w) PMID: 33479207. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7820474/> [Accessed on: 2023 Dec 27]
5. Hänsel K, Dudgeon SN, Cheung KH, Durant TJS, and Schulz WL. From Data to Wisdom: Biomedical Knowledge Graphs for Real-World Data Insights. en. *Journal of Medical Systems* 2023 May; 47:65. doi: [10.1007/s10916-023-01951-2](https://doi.org/10.1007/s10916-023-01951-2). Available from: <https://link.springer.com/10.1007/s10916-023-01951-2> [Accessed on: 2024 Jun 10]
6. Milosevic N and Thielemann W. Comparison of biomedical relationship extraction methods and models for knowledge graph creation. *Journal of Web Semantics* 2023 Jan; 75. arXiv:2201.01647 [cs]:100756. doi: [10.1016/j.websem.2022.100756](https://doi.org/10.1016/j.websem.2022.100756). Available from: [http://arxiv.org/abs/2201.01647](https://arxiv.org/abs/2201.01647) [Accessed on: 2024 Jun 14]
7. Nicholson DN, Himmelstein DS, and Greene CS. Expanding a database-derived biomedical knowledge graph via multi-relation extraction from biomedical abstracts. en. *BioData Mining* 2022 Oct; 15:26. doi: [10.1186/s13040-022-00311-z](https://doi.org/10.1186/s13040-022-00311-z). Available from: <https://biodatamining.biomedcentral.com/articles/10.1186/s13040-022-00311-z> [Accessed on: 2024 Jun 14]
8. Arsenyan V, Bughdaryan S, Shaya F, Small K, and Shahnazaryan D. Large Language Models for Biomedical Knowledge Graph Construction: Information extraction from EMR notes. en. arXiv:2301.12473 [cs]. 2023 Dec. Available from: [http://arxiv.org/abs/2301.12473](https://arxiv.org/abs/2301.12473) [Accessed on: 2024 Jun 12]
9. McLellan J, Kim DHJ, Bruce M, Ramirez-Celis A, and Water J Van de. Maternal Immune Dysregulation and Autism—Understanding the Role of Cytokines, Chemokines and Autoantibodies. *Frontiers in Psychiatry* 2022; 13. Minireview, to keep. doi: [PMID35722542](https://doi.org/10.3389/fpsyg.2022.935722). Available from:

- <https://www.frontiersin.org/articles/10.3389/fpsyg.2022.834910> [Accessed on: 2023 Dec 27]
10. Tioleco N, Silberman AE, Stratigos K, Banerjee-Basu S, Spann MN, Whitaker AH, and Turner JB. Prenatal maternal infection and risk for autism in offspring: A meta-analysis. en. *Autism Research* 2021 Jun; 14. meta analysis,to keep:1296–316. doi: [10.1002/aur.2499](https://doi.org/10.1002/aur.2499) PMID: [33720503](https://pubmed.ncbi.nlm.nih.gov/33720503/). Available from: <https://onlinelibrary.wiley.com/doi/10.1002/aur.2499> [Accessed on: 2023 Dec 27]
  11. Shuid AN, Jayusman PA, Shuid N, Ismail J, Kamal Nor N, and Mohamed IN. Association between Viral Infections and Risk of Autistic Disorder: An Overview. *International Journal of Environmental Research and Public Health* 2021 Mar; 18. review,infectiongraph:2817. doi: [10.3390/ijerph18062817](https://doi.org/10.3390/ijerph18062817) PMID: [33802042](https://pubmed.ncbi.nlm.nih.gov/33802042/). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7999368/> [Accessed on: 2023 Dec 27]
  12. Ravaccia D and Ghafourian T. Critical Role of the Maternal Immune System in the Pathogenesis of Autism Spectrum Disorder. en. *Biomedicines* 2020 Dec; 8. litterature review, to keep:557. doi: [10.3390/biomedicines8120557](https://doi.org/10.3390/biomedicines8120557) PMID: [33271759](https://pubmed.ncbi.nlm.nih.gov/33271759/). Available from: <https://www.mdpi.com/2227-9059/8/12/557> [Accessed on: 2023 Dec 27]
  13. Croen LA, Qian Y, Ashwood P, Zerbo O, Schendel D, Pinto-Martin J, Fallin MD, Levy S, Schieve LA, Yeargin-Allsopp M, Sabourin KR, and Ames JL. Infection and Fever in Pregnancy and Autism Spectrum Disorders: Findings from the Study to Explore Early Development. *Autism research : official journal of the International Society for Autism Research* 2019 Oct; 12. to keep:1551–61. doi: [10.1002/aur.2175](https://doi.org/10.1002/aur.2175) PMID: [31317667](https://pubmed.ncbi.nlm.nih.gov/31317667/). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7784630/> [Accessed on: 2023 Dec 27]
  14. Beversdorf DQ, Stevens HE, Margolis KG, and Water J Van de. Prenatal stress and maternal immune dysregulation in autism spectrum disorders- potential points for intervention. *Current pharmaceutical design* 2019; 25. to keep:4331–43. doi: [10.2174/1381612825666191119093335](https://doi.org/10.2174/1381612825666191119093335) PMID : [31742491](https://pubmed.ncbi.nlm.nih.gov/31742491/). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7100710/> [Accessed on: 2023 Dec 27]
  15. Azhari A, Azizan F, and Esposito G. A systematic review of gut-immune-brain mechanisms in Autism Spectrum Disorder. en. *Developmental Psychobiology* 2019 Jul; 61. To keep:752–71. doi: [10.1002/dev.21803](https://doi.org/10.1002/dev.21803) PMID: [30523646](https://pubmed.ncbi.nlm.nih.gov/30523646/). Available from: <https://onlinelibrary.wiley.com/doi/10.1002/dev.21803> [Accessed on: 2023 Dec 27]
  16. Jones KL, Croen LA, Yoshida CK, Heuer L, Hansen R, Zerbo O, DeLorenze GN, Kharrazi M, Yolken R, Ashwood P, and Water J Van de. Autism with Intellectual Disability is Associated with Increased Levels of Maternal Cytokines and Chemokines During Gestation. *Molecular psychiatry* 2017 Feb; 22. to keep:273–9. doi: [10.1038/mp.2016.77](https://doi.org/10.1038/mp.2016.77) PMID: [27217154](https://pubmed.ncbi.nlm.nih.gov/27217154/). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5122473/> [Accessed on: 2023 Dec 27]

17. Zerbo O, Qian Y, Yoshida C, Grether JK, Van De Water J, and Croen LA. Maternal Infection During Pregnancy and Autism Spectrum Disorders. en. *Journal of Autism and Developmental Disorders* 2015 Dec; 45(12):4015–25. doi: [10.1007/s10803-013-2016-3](https://doi.org/10.1007/s10803-013-2016-3) PMID: 24366406. Available from: <http://link.springer.com/10.1007/s10803-013-2016-3> [Accessed on: 2023 Dec 27]
18. Zerbo O, Yoshida C, Grether JK, Water J Van de, Ashwood P, Delorenze GN, Hansen RL, Kharrazi M, and Croen LA. Neonatal cytokines and chemokines and risk of Autism Spectrum Disorder: the Early Markers for Autism (EMA) study: a case-control study. *Journal of Neuroinflammation* 2014 Jun; 11(1):113. doi: [10.1186/1742-2094-11-113](https://doi.org/10.1186/1742-2094-11-113) PMID: 24951035. Available from: <https://doi.org/10.1186/1742-2094-11-113> [Accessed on: 2023 Dec 27]
19. Brown AS, Sourander A, Hinkka-Yli-Salomäki S, McKeague IW, Sundvall J, and Surcel HM. Elevated Maternal C-Reactive Protein and Autism in a National Birth Cohort. *Molecular psychiatry* 2014 Feb; 19(2):259–64. doi: [10.1038/mp.2012.197](https://doi.org/10.1038/mp.2012.197) PMID: 23337946. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3633612/> [Accessed on: 2023 Dec 27]
20. Hernández-Díaz S, Straub L, Bateman BT, Zhu Y, Mogun H, Wisner KL, Gray KJ, Lester B, McDougle CJ, DiCesare E, Pennell PB, and Huybrechts KF. Risk of Autism after Prenatal Topiramate, Valproate, or Lamotrigine Exposure. eng. *The New England Journal of Medicine* 2024 Mar; 390(12):1069–79. doi: [10.1056/NEJMoa2309359](https://doi.org/10.1056/NEJMoa2309359)
21. Soman K, Rose PW, Morris JH, Akbas RE, Smith B, Villouta-Reyes C, Cerono G, Shi Y, Rizk-Jackson A, Israni S, Nelson CA, Huang S, and Baranzini SE. Biomedical knowledge graph-optimized prompt generation for large language models. en
22. Huang W, Ma X, Qin H, Zheng X, Lv C, Chen H, Luo J, Qi X, Liu X, and Magno M. How Good Are Low-bit Quantized LLaMA3 Models? An Empirical Study. en. arXiv:2404.14047 [cs]. 2024 Apr. Available from: [http://arxiv.org/abs/2404.14047](https://arxiv.org/abs/2404.14047) [Accessed on: 2024 Jun 12]
23. Manning C, Raghavan P, and Schuetze H. *Introduction to Information Retrieval*. en. 2009
24. Newman MEJ. *Networks: an introduction*. en. OCLC: ocn456837194. Oxford ; New York: Oxford University Press, 2010
25. The OBI Consortium, Smith B, Ashburner M, Rosse C, Bard J, Bug W, Ceusters W, Goldberg LJ, Eilbeck K, Ireland A, Mungall CJ, Leontis N, Rocca-Serra P, Ruttenberg A, Sansone SA, Scheuermann RH, Shah N, Whetzel PL, and Lewis S. The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. en. *Nature Biotechnology* 2007 Nov; 25(11):1251–5. doi: [10.1038/nbt1346](https://doi.org/10.1038/nbt1346). Available from: <https://www.nature.com/articles/nbt1346> [Accessed on: 2024 Jun 14]
26. Cohen KB, Johnson HL, Verspoor K, Roeder C, and Hunter LE. The structural and content aspects of abstracts versus bodies of full text journal articles are different. en. *BMC Bioinformatics* 2010 Dec; 11(492). doi: [10.1186/1471-2105-11-492](https://doi.org/10.1186/1471-2105-11-492). Available from: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-11-492> [Accessed on: 2024 Jun 14]

27. Zhang F, Wang H, Barhamgi M, Chen L, and Zhou R, eds. Web Information Systems Engineering – WISE 2023: 24th International Conference, Melbourne, VIC, Australia, October 25–27, 2023, Proceedings. en. Vol. 14306. Lecture Notes in Computer Science. Singapore: Springer Nature Singapore, 2023. doi: [10.1007/978-981-99-7254-8](https://doi.org/10.1007/978-981-99-7254-8). Available from: <https://link.springer.com/10.1007/978-981-99-7254-8> [Accessed on: 2024 Jun 10]
28. Anandarajan M, Hill C, and Nolan T. Practical Text Analytics: Maximizing the Value of Text Data. en. Vol. 2. Advances in Analytics and Data Science. Cham: Springer International Publishing, 2019. doi: [10.1007/978-3-319-95663-3](https://doi.org/10.1007/978-3-319-95663-3). Available from: <https://link.springer.com/10.1007/978-3-319-95663-3> [Accessed on: 2024 Jun 10]
29. Luo B, Cheng L, Wu ZG, Li H, and Li C, eds. Neural Information Processing: 30th International Conference, ICONIP 2023, Changsha, China, November 20–23, 2023, Proceedings, Part XII. en. Vol. 1966. Communications in Computer and Information Science. Singapore: Springer Nature Singapore, 2024. doi: [10.1007/978-981-99-8148-9](https://doi.org/10.1007/978-981-99-8148-9). Available from: <https://link.springer.com/10.1007/978-981-99-8148-9> [Accessed on: 2024 Jun 10]
30. Chen Y, Yang W, Wang K, Qin Y, Huang R, and Zheng Q. A neuralized feature engineering method for entity relation extraction. en. *Neural Networks* 2021 Sep; 141:249–60. doi: [10.1016/j.neunet.2021.04.010](https://doi.org/10.1016/j.neunet.2021.04.010). Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0893608021001453> [Accessed on: 2024 Jun 10]
31. Zhu Y, Li X, Wang Z, Li J, Yan C, and Zhang Y. ER-LAC: Span-Based Joint Entity and Relation Extraction Model with Multi-Level Lexical and Attention on Context Features. en. *Applied Sciences* 2023 Sep; 13:10538. doi: [10.3390/app131810538](https://doi.org/10.3390/app131810538). Available from: <https://www.mdpi.com/2076-3417/13/18/10538> [Accessed on: 2024 Jun 10]
32. Yang Y, Wu Z, Yang Y, Lian S, Guo F, and Wang Z. A Survey of Information Extraction Based on Deep Learning. en. *Applied Sciences* 2022 Sep; 12:9691. doi: [10.3390/app12199691](https://doi.org/10.3390/app12199691). Available from: <https://www.mdpi.com/2076-3417/12/19/9691> [Accessed on: 2024 Jun 10]
33. Gene Ontology Resource. Available from: <http://geneontology.org/> [Accessed on: 2024 Jun 10]
34. Neuro Behavior Ontology - Summary | NCBO BioPortal. Available from: <https://bioportal.bioontology.org/ontologies/NBO> [Accessed on: 2024 Jun 10]
35. Jaimini U, Henson C, and Sheth AP. CausalDisco: Causal discovery using knowledge graph link prediction. en. arXiv:2405.02327 [cs]. 2024 Apr. Available from: [http://arxiv.org/abs/2405.02327](https://arxiv.org/abs/2405.02327) [Accessed on: 2024 Jun 10]
36. Jaimini U and Sheth A. CausalKG: Causal Knowledge Graph Explainability Using Interventional and Counterfactual Reasoning. en. *IEEE Internet Computing* 2022 Jan; 26:43–50. doi: [10.1109/MIC.2021.3133551](https://doi.org/10.1109/MIC.2021.3133551). Available from: <https://ieeexplore.ieee.org/document/9706608/> [Accessed on: 2024 Jun 10]

37. Heindorf S, Scholten Y, Wachsmuth H, Ngonga Ngomo AC, and Potthast M. CauseNet: Towards a Causality Graph Extracted from the Web. en. *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*. Virtual Event Ireland: ACM, 2020 Oct :3023–30. doi: [10.1145/3340531.3412763](https://doi.org/10.1145/3340531.3412763). Available from: <https://doi.org/10.1145/3340531.3412763> [Accessed on: 2024 Jun 10]
38. Blomqvist E, Alirezaie M, and Santini M. Towards Causal Knowledge Graphs - Position Paper. en
39. Wei C, Chen Y, Wang K, Qin Y, Huang R, and Zheng Q. APRE: Annotation-Aware Prompt-Tuning for Relation Extraction. en. *Neural Processing Letters* 2024 Feb; 56:62. doi: [10.1007/s11063-024-11437-y](https://doi.org/10.1007/s11063-024-11437-y). Available from: <https://doi.org/10.1007/s11063-024-11437-y> [Accessed on: 2024 Jun 10]
40. Wadhwa S, Amir S, and Wallace BC. Revisiting Relation Extraction in the era of Large Language Models. en. arXiv:2305.05003 [cs]. 2023 May. Available from: [http://arxiv.org/abs/2305.05003](https://arxiv.org/abs/2305.05003) [Accessed on: 2024 Jun 10]
41. Anutariya C and Bonsangue MM, eds. Data Science and Artificial Intelligence: First International Conference, DSAI 2023, Bangkok, Thailand, November 27–29, 2023, Proceedings. en. Vol. 1942. Communications in Computer and Information Science. Singapore: Springer Nature Singapore, 2023. doi: [10.1007/978-981-99-7969-1](https://doi.org/10.1007/978-981-99-7969-1). Available from: <https://doi.org/10.1007/978-981-99-7969-1> [Accessed on: 2024 Jun 10]
42. Zeng J, Qin P, Jing W, Song X, and Lu Z, eds. Data Science: 7th International Conference of Pioneering Computer Scientists, Engineers and Educators, ICPCSEE 2021, Taiyuan, China, September 17–20, 2021, Proceedings, Part II. en. Vol. 1452. Communications in Computer and Information Science. Singapore: Springer Singapore, 2021. doi: [10.1007/978-981-16-5943-0](https://doi.org/10.1007/978-981-16-5943-0). Available from: <https://doi.org/10.1007/978-981-16-5943-0> [Accessed on: 2024 Jun 10]

# APPENDIX

## I LLM Prompts

### I.1 System Prompt

```
'You are a knowledge graph expert who extracts terms and their relations from a given context.'
'You are provided with a context chunk (delimited by ``) Your task is to extract the medical and clinical entities of DISEASE, GENE, RNA, DNA, DRUG, CHEMICAL, CELL_LINE, CELL_TYPE, PROTEIN, PATHWAY, MUTATION,ORGAN, TISSUE , SYMPTOM, SPECIES, BIOLOGICAL_PROCESS,ENVIRONMENTAL_FACTOR,PHENOTYPIC_FEATURE,ANATOMY,HORMONE, CYTOKINE, INFLAMMATORY_RESPONSE, PREGNANCY, FETUS, GESTATION, MATERNAL_HEALTH, IMMUNE_SYSTEM, TREATMENT, THERAPY.'
'MEDICAL_PROCEDURE, DEVELOPMENTAL_STAGE, NUTRITION, TOXIN, RISK_FACTOR these should be writting into entity_type_1/entity_type_2'
'mentioned in the given context. These terms should represent the key concepts as per the context.'
'Choose the edge from this list: ' + p_dc + '.''
Thought 1: While traversing through each sentence, Think about the key terms mentioned in it.\n"
"\tTerms may include protein,medcalentity,disease,symptoms,treatment or one of the entity types mentionned before\n"
"\tcondition, concept, object, entity, cause, effect etc.\n"
"\tTerms should be as atomistic as possible\n\n"
Thought 2: Think about how these terms can have one on one relation with other terms.\n"
"\tTerms that are mentioned in the same sentence or the same paragraph are typically related to each other.\n"
"\tTerms can be related to many other terms\n\n"
Thought 3: Find out the relation between each such related pair of terms. \n\n"
"Format your output as a list of json. Each element of the list contains a pair of terms"
'NEVER change the value of the chunk_ID as defined in this prompt'
'Give Result ONLY in this JSON format: '
>[
{
  "chunk_id": "CHUNK_ID_GOES_HERE",
  "node_1": "A concept from extracted ontology",
  "entity_type_1": "The Entity Type for node_1",
  "node_2": "A related concept from extracted ontology",
  "entity_type_2": "The Entity Type for node_2",
  "edge": "relationship between the two concepts, they should be a maximum of 3 words"
}, {
  ...As many relationships as you can find...
]
'
```

Where p\_dc is expressed as:

```
instance_of, subclass_of, is_a_medical_condition, Disease_Ontology_ID, medical_treatment,
drug_used_for_treatment, drug_interaction, has_gene_location, gene, drug_target,
has_parts_of_the_class, symptoms, has_gene_location, clinical_features, affected_by,
diseases, ICD-10, ICD-9, located_in, pathway, derived_from, regulates, develops_from,
has_characteristic, is_method, applies_to, physically_interacts_with,
contributing_factor_of, expressed_in
```

### I.2 User Prompt

```
Hello please do not comment on your output just give me the JSON file . Here is my data:
``"+ input +```\n output:
```

Where input is expressed as the current chunk of data to be analyzed.

## II Neo4j Cipher Queries for DO in Order

Visualize all Relationships.

```
MATCH (a)-[b]-(c)
RETURN a,b,c;
```

After which all duplicates were removed.

```
MATCH (s)-[r]->(t)
WITH s, t, COLLECT(r) AS rs
FOREACH (x IN rs [1..] | DELETE x);
```

To be able to visualize the nodes with their corresponding categories, we labeled them by their key type using this cypher query :

```
MATCH (n:Node)
FOREACH (ignoreMe IN CASE WHEN n.type IN ['disease', 'class of disease', 'disability', 'rare disease', 'major depressive disorder', 'health problem', 'neurodevelopmental disorder', 'behavioral disorder', 'genetic disease', 'health risk', 'disease attributes', 'second-order class', 'designated intractable/rare disease', 'infectious disease', 'developmental disorder'] THEN [1] ELSE [] END |
SET n:ClassOfDisease REMOVE n:Node)
FOREACH (ignoreMe IN CASE WHEN n.type IN ["structural class of chemical entities", "type of chemical entity"] THEN [1] ELSE [] END |
SET n:MedicalEntityType REMOVE n:Node)
FOREACH (ignoreMe IN CASE WHEN n.type IN ['protein', 'group or class of proteins'] THEN [1] ELSE [] END |
SET n:ClassofProteins REMOVE n:Node)
FOREACH (ignoreMe IN CASE WHEN n.type IN ["structural class of chemical entities", "type of chemical entity"] THEN [1] ELSE [] END |
SET n:MedicalEntityType REMOVE n:Node)
FOREACH (ignoreMe IN CASE WHEN n.type IN ["symptom", "symptom or sign", "clinical sign"] THEN [1] ELSE [] END |
SET n:Symptom REMOVE n:Node)
FOREACH (ignoreMe IN CASE WHEN n.type IN ["autism therapy", "preventive medicine", "treatment of mental disorders", "drug class", "procedure"] THEN [1] ELSE [] END |
SET n:Treatment REMOVE n:Node)
FOREACH (ignoreMe IN CASE WHEN n.type IN ["failure mode", "biological process"] THEN [1] ELSE [] END |
SET n:BiologicalProcess REMOVE n:Node)
```

To visualize the node with the name "autism" and all the nodes and relationships attached to it in the Neo4j database, we used the following Cypher query:

```
MATCH (a {name: "autism"})-[r*]->(b)
RETURN a, b, r, labels(a) as a_labels, labels(b) as b_labels;
```

The corresponding source-relation-target dataframe is obtained using this query :

```
MATCH (a)-[r]->(b)
WHERE a.name = ""autism"
RETURN a.name AS source, type(r) AS relation, b.name AS target;
```

### III Neo4j Cipher Queries for LLM in Order

Load generated .csv file from public hosted Server with their respective extracted attributes.

```
LOAD CSV WITH HEADERS FROM 'http://publictemporaryserver:port/filepath.csv' AS row
CALL apoc.merge.node(["Node"], {name: coalesce(row.node_1, "Unknown")},{type: coalesce(
    toLower(row.entity_type_1),"Unkown")}) YIELD node AS source
CALL apoc.merge.node(["Node"], {name: coalesce(row.node_2, "Unknown")},{type: coalesce(
    toLower(row.entity_type_2),"Unkown")}) YIELD node AS target
WITH source,target,row
CALL apoc.create.relationship(source, row.edge, {}, target) YIELD rel
RETURN rel
```

Set Node Labels from type.

```
MATCH (n:Node)
FOREACH (ignoreMe IN CASE WHEN n.type IN ["condition","condition or state",
    "medical_condition","disease or syndrome","disorder","neurodevelopmental disorders",
    "medical condition","conditions","disease","state_of_being"] THEN [1] ELSE [] END |
    SET n:ClassOfDisease REMOVE n:Node
)
FOREACH (ignoreMe IN CASE WHEN n.type IN ["chemical_substance","chemicals","chemokines",
    "biologic markers","substances","specific cytokines and chemokines","chemical"] THEN [1]
    ELSE [] END |
    SET n:MedicalEntityType REMOVE n:Node
)
FOREACH (ignoreMe IN CASE WHEN n.type IN ["biological entity","biological system",
    "components of biological systems","immune system in pregnancy","gene",
    "microbial_community","organ","organism","immune_system"] THEN [1] ELSE [] END |
    SET n:BiologicalSystem REMOVE n:Node
)
FOREACH (ignoreMe IN CASE WHEN n.type IN ["medication","drug","treatment"] THEN [1] ELSE []
    END |
    SET n:Treatment REMOVE n:Node
)
FOREACH (ignoreMe IN CASE WHEN n.type IN ["behavior or habits","behavior or habit",
    "environmental factor","sociodemographic characteristic","conditions",
    "mechanism","difference in immune response","biological_process","pathway"] THEN [1] ELSE [] END |
    SET n:BiologicalProcess REMOVE n:Node
)
FOREACH (ignoreMe IN CASE WHEN n.type IN ["environmental_factor","population"] THEN [1]
    ELSE [] END |
    SET n:EnvironmentalFactor REMOVE n:Node
)
FOREACH (ignoreMe IN CASE WHEN n.type IN ["phenotypic_feature"] THEN [1] ELSE [] END |
    SET n:PhenotypicFeature REMOVE n:Node
)
FOREACH (ignoreMe IN CASE WHEN n.type IN ["biological_entity"] THEN [1] ELSE [] END |
    SET n:BiologicalEntity REMOVE n:Node
)
FOREACH (ignoreMe IN CASE WHEN n.type IN ["clinical_feature","clinical_features"] THEN [1]
    ELSE [] END |
    SET n:ClinicalFeature REMOVE n:Node
)
FOREACH (ignoreMe IN CASE WHEN n.type IN ["protein","chemokine"] THEN [1] ELSE [] END |
    SET n:ClassOfProtein REMOVE n:Node
)
FOREACH (ignoreMe IN CASE WHEN n.type IN ["symptom"] THEN [1] ELSE [] END |
    SET n:Symptom REMOVE n:Node
)
FOREACH (ignoreMe IN CASE WHEN n.type IN ["developmental_stage","disease_ontology_id"] THEN
    [1] ELSE [] END |
    SET n:DevelopmentalStage REMOVE n:Node
)
```

Selective labeling/re-labeling by node names. In this case we removed the nodes from the label "MedicalEntityType" and added them to the label "ClassOfProtein" based on their names.

```
MATCH (n:MedicalEntityType)
FOREACH (ignoreMe IN CASE WHEN n.name IN ["chemokines","cytokines","autoantibodies","proteins in the developing brain","il-6","interleukin α-1 (il-α-1)","interferon γ-","granulocyte macrophage colony-stimulating factor","il-8 and monocyte chemotactic protein -1"] THEN [1] ELSE [] END |
SET n:ClassOfProtein REMOVE n:MedicalEntityType
)
```

Merge selected duplicate nodes into one. In this example the node "autism spectrum disorder" is merged into the node "asd".

```
MATCH (n WHERE n.name="autism spectrum disorder") , (o WHERE o.name="asd")
CALL apoc.refactor.mergeNodes([n,o]) YIELD node
RETURN node
```

Delete duplicate relationships

```
MATCH (s)-[r]->(t)
WITH s, t, COLLECT(r) AS rs
FOREACH (x IN rs[1..] | DELETE x);
```

Visualize the Graph.

```
MATCH (a)-[b]->(c) RETURN a,b,c;
```

## IV Dataset.txt

PubMed ID: 35722542

Title: Maternal Immune Dysregulation and Autism—Understanding the Role of Cytokines, Chemokines and Autoantibodies

Abstract: Autism spectrum disorder (ASD) is acknowledged as a highly heterogeneous, behaviorally defined neurodevelopmental disorder with multiple etiologies. In addition to its high heritability, we have come to recognize a role for maternal immune system dysregulation as a prominent risk factor for the development of ASD in the child. Examples of these risk factors include altered cytokine/chemokine activity and the presence of autoantibodies in mothers that are reactive to proteins in the developing brain. In addition to large clinical studies, the development of pre-clinical models enables the ability to evaluate the cellular and molecular underpinnings of immune-related pathology. For example, the novel animal models of maternal autoantibody-related (MAR) ASD described herein will serve as a preclinical platform for the future testing of targeted therapeutics for one 'type' of ASD. Identification of the cellular targets will advance precision medicine efforts toward tailored therapeutics and prevention. This minireview highlights emerging evidence for the role of maternal immune dysregulation as a potential biomarker, as well as a pathologically relevant mechanism for the development of ASD in offspring. Further, we will discuss the current limitations of these models as well as potential avenues for future research.

PubMed ID: 35131181

Title: Maternal inflammation and its ramifications on fetal neurodevelopment

Abstract: Exposure to heightened inflammation in pregnancy caused by infections or other inflammatory insults has been associated with the onset of neurodevelopmental and psychiatric disorders in children. Rodent models have provided unique insights into how this maternal immune activation (MIA) disrupts brain development. Here, we discuss the key immune factors involved, highlight recent advances in determining the molecular and cellular pathways of MIA, and review how the maternal immune system affects fetal development. We also examine the roles of microbiomes in shaping maternal immune function and the development of autism-like phenotypes. A comprehensive understanding of the gut bacteria-immune-neuro interaction in MIA is essential for developing

diagnostic and therapeutic measures for high-risk pregnant women and identifying targets for treating inflammation-induced neurodevelopmental disorders.

PubMed ID: 33720503

Title: Prenatal maternal infection and risk for autism in offspring: A meta-analysis

Abstract: While prenatal maternal infection has received attention as a preventable and treatable risk factor for autism, findings have been inconsistent. This paper presents the results of a meta-analysis to determine whether the weight of the evidence supports such an association. Studies with a categorical diagnosis of autism as the outcome and an assessment of its association with prenatal maternal infection or fever (or the data necessary to compute this association) were included. A total of 36 studies met these criteria. Two independent reviewers extracted data on study design, methods of assessment, type of infectious agent, site of infection, trimester of exposure, definition of autism, and effect size. Analyses demonstrated a statistically significant association of maternal infection/fever with autism in offspring (OR = 1.32; 95% CI = 1.20–1.46). Adjustment for evident publication bias slightly weakened this association. There was little variation in effect sizes across agent or site of infection. Small differences across trimester of exposure were not statistically significant. There was some evidence that recall bias associated with status on the outcome variable leads to differential misclassification of exposure status. Nonetheless, the overall association is only modestly reduced when studies potentially contaminated by such bias are removed. Although causality has not been firmly established, these findings suggest maternal infection during pregnancy confers an increase in risk for autism in offspring. Given the prevalence of this risk factor, it is possible that the incidence of autism would be reduced by 12%–17% if maternal infections could be prevented or safely treated in a timely manner. LAY SUMMARY: This study is a meta-analysis of the association of maternal infection during pregnancy and subsequent autism in offspring. In combining the results from 36 studies of this association we find that a significant relationship is present. The association does not vary much across the types of infections or when they occur during pregnancy. We conclude that the incidence of autism could be substantially reduced if maternal infections could be prevented or safely treated in a timely manner.

PubMed ID: 33802042

Title: Association between Viral Infections and Risk of Autistic Disorder: An Overview

Abstract: Autism spectrum disorder (ASD) is a neurodevelopmental condition of the central nervous system (CNS) that presents with severe communication problems, impairment of social interactions, and stereotypic behaviours. Emerging studies indicate possible associations between viral infections and neurodegenerative and neurobehavioural conditions including autism. Viral infection during critical periods of early in utero neurodevelopment may lead to increased risk of autism in the offspring. This review is aimed at highlighting the association between viral infections, including viruses similar to COVID-19, and the aetiology of autism. A literature search was conducted using Pubmed, Ovid/Medline, and Google Scholar database. Relevant search terms included "rubella and autism", "cytomegalovirus and autism", "influenza virus and autism", "Zika virus and autism", "COVID-19 and autism". Based on the search terms, a total of 141 articles were obtained and studies on infants or children with congenital or perinatal viral infection and autistic behaviour were evaluated. The possible mechanisms by which viral infections could lead to autism include direct teratogenic effects and indirect effects of inflammation or maternal immune activation on the developing brain. Brain imaging studies have shown that the ensuing immune response from these viral infections could lead to disruption of the development of brain regions and structures. Hence, long-term follow up is necessary for infants whose mothers report an inflammatory event due to viral infection at any time during pregnancy to monitor for signs of autism. Research into the role of viral infection in the development of ASD may be one avenue of improving ASD outcomes in the future. Early screening and diagnosis to detect, and maybe even prevent ASD are essential to reduce the burden of this condition.

PubMed ID: 33479207

Title: Maternal acute and chronic inflammation in pregnancy is associated with common neurodevelopmental disorders: a systematic review

Abstract: Inflammation is increasingly recognized as a cause or consequence of common problems of humanity including obesity, stress, depression, pollution and disease

states such as autoimmunity, asthma, and infection. Maternal immune activation (MIA), triggered by both acute and systemic chronic inflammation, is hypothesized to be one of the mechanisms implicated in the pathogenesis of neurodevelopmental disorders (NDD). Although there is substantial preclinical evidence to support the MIA hypothesis, the human evidence is disparate. We performed a systematic review on human studies examining associations between maternal inflammatory states and offspring NDDs (autism spectrum disorder - ASD, attention deficit hyperactivity disorder - ADHD, Tourette syndrome - TS). 32 meta-analyses and 26 additional individual studies were identified. Maternal states associated with ASD include obesity, gestational diabetes mellitus, pre-eclampsia, pollution, stress, depression, autoimmune diseases, and infection. Maternal states associated with ADHD include obesity, pre-eclampsia, smoking, low socioeconomic status (SES), stress, autoimmune disease, and asthma. Maternal states associated with TS include low SES, depression, and autoimmune diseases. Diverse maternal inflammatory states in pregnancy are associated with common offspring NDDs. Given the increased prevalence of NDDs, there is urgent need to explore relative and cumulative maternal risk factors and disease mechanisms. Defining preventable risk factors in high-risk pregnancies could mitigate the expression and severity of NDDs.

PubMed ID: 33271759

Title: Critical Role of the Maternal Immune System in the Pathogenesis of Autism Spectrum Disorder

Abstract: Autism spectrum disorders (ASD) are a group of neurodevelopmental disorders characterised by impairments in communication, social interaction, and the presence of restrictive and repetitive behaviours. Over the past decade, most of the research in ASD has focused on the contribution of genetics, with the identification of a variety of different genes and mutations. However, the vast heterogeneity in clinical presentations associated with this disorder suggests that environmental factors may be involved, acting as a "second hit" in already genetically susceptible individuals. To this regard, emerging evidence points towards a role for maternal immune system dysfunctions. This literature review considered evidence from epidemiological studies and aimed to discuss the pathological relevance of the maternal immune system in ASD by looking at the proposed mechanisms by which it alters the prenatal environment. In particular, this review focuses on the effects of maternal immune activation (MIA) by looking at foetal brain-reactive antibodies, cytokines and the microbiome. Despite the arguments presented here that strongly implicate MIA in the pathophysiology of ASD, further research is needed to fully understand the precise mechanisms by which they alter brain structure and behaviour. Overall, this review has not only shown the importance of the maternal immune system as a risk factor for ASD, but more importantly, has highlighted new promising pathways to target for the discovery of novel therapeutic interventions for the treatment of such a life-changing disorder.

PubMed ID: 31317667

Title: Infection and Fever in Pregnancy and Autism Spectrum Disorders: Findings from the Study to Explore Early Development

Abstract: Maternal infection and fever during pregnancy have been implicated in the etiology of autism spectrum disorder (ASD); however, studies have not been able to separate the effects of fever itself from the impact of a specific infectious organism on the developing brain. We utilized data from the Study to Explore Early Development (SEED), a case-control study among 2- to 5-year-old children born between 2003 and 2006 in the United States, to explore a possible association between maternal infection and fever during pregnancy and risk of ASD and other developmental disorders (DDs). Three groups of children were included: children with ASD (N = 606) and children with DDs (N = 856), ascertained from clinical and educational sources, and children from the general population (N = 796), randomly sampled from state birth records. Information about infection and fever during pregnancy was obtained from a telephone interview with the mother shortly after study enrollment and maternal prenatal and labor/delivery medical records. ASD and DD status was determined by an in-person standardized developmental assessment of the child at 3-5 years of age. After adjustment for covariates, maternal infection anytime during pregnancy was not associated with ASD or DDs. However, second trimester infection accompanied by fever elevated risk for ASD approximately twofold ( $aOR = 2.19$ , 95% confidence interval 1.14-4.23). These findings of an association between maternal infection with fever in the second trimester and increased risk of ASD in the offspring suggest that the inflammatory response to the infectious agent may be etiologically relevant. Autism Res 2019, 12: 1551-1561. © 2019

International Society for Autism Research, Wiley Periodicals, Inc. LAY SUMMARY: Using data from a large multisite study in the United States—the Study to Explore Early Development—we found that women who had an infection during the second trimester of pregnancy accompanied by a fever are more likely to have children with ASD. These findings suggest the possibility that only more severe infections accompanied by a robust inflammatory response increase the risk of ASD.

PubMed ID: 31742491

Title: Prenatal Stress and Maternal Immune Dysregulation in Autism Spectrum Disorders: Potential Points for Intervention

Abstract: Background:

Genetics is a major etiological contributor to autism spectrum disorder (ASD). Environmental factors, however, also appear to contribute. ASD pathophysiology due to gene x environment is also beginning to be explored. One reason to focus on environmental factors is that they may allow opportunities for intervention or prevention.

Methods and results:

Herein, we review two such factors that have been associated with a significant proportion of ASD risk, prenatal stress exposure and maternal immune dysregulation. Maternal stress susceptibility appears to interact with prenatal stress exposure to affect offspring neurodevelopment. We also explore how maternal stress may interact with the microbiome in the neurodevelopmental setting. Additionally, understanding of the impact of maternal immune dysfunction on ASD has recently been advanced by recognition of specific fetal brain proteins targeted by maternal autoantibodies, and identification of unique mid-gestational maternal immune profiles. This might also be interrelated with maternal stress exposure. Animal models have been developed to explore pathophysiology targeting each of these factors.

Conclusion:

We are beginning to understand the behavioral, pharmacopathological, and epigenetic effects related to these interactions, and we are beginning to explore potential mitigating factors. Continued growth in understanding of these mechanisms may ultimately allow for the identification of multiple potential targets for prevention or intervention for this subset of environmental-associated ASD cases.

PubMed ID: 30523646

Title: A systematic review of gut-immune-brain mechanisms in Autism Spectrum Disorder

Abstract: Despite decades of research, the etiological origins of Autism Spectrum Disorder (ASD) remain elusive. Recently, the mechanisms of ASD have encompassed emerging theories involving the gastrointestinal, immune, and nervous systems. While each of these perspectives presents its own set of supporting evidence, the field requires an integration of these modular concepts and an overarching view of how these subsystems intersect. In this systematic review, we have synthesized relevant evidences from the existing literature, evaluating them in an interdependent manner and in doing so, outlining their possible connections. Specifically, we first discussed gastrointestinal and immuno-inflammation pathways in-depth, exploring the relationships between microbial composition, bacterial metabolites, gut mucosa, and immune system constituents. Accounting for temporal differences in the mechanisms involved in neurodevelopment, prenatal and postnatal phases were further elucidated, where the former focused on maternal immune activation (MIA) and fetal development, while the latter addressed the role of immune dysregulation in contributing to atypical neurodevelopment. As autism remains, foremost, a neurodevelopmental disorder, this

review presents an integration of disparate modules into a "Gut-Immune-Brain" paradigm. Existing gaps in the literature have been highlighted, and possible avenues for future research with an integrated physiological perspective underlying ASD have also been suggested.

PubMed ID: 27217154

**Title:** Autism with intellectual disability is associated with increased levels of maternal cytokines and chemokines during gestation  
**Abstract:** Immune abnormalities have been described in some individuals with autism spectrum disorders (ASDs) as well as their family members. However, few studies have directly investigated the role of prenatal cytokine and chemokine profiles on neurodevelopmental outcomes in humans. In the current study, we characterized mid-gestational serum profiles of 22 cytokines and chemokines in mothers of children with ASD (N=415), developmental delay (DD) without ASD (N=188), and general population (GP) controls (N =428) using a bead-based multiplex technology. The ASD group was further divided into those with intellectual disabilities (developmental/cognitive and adaptive composite score <70) (ASD+ID, N=184) and those without (composite  $\Delta$ score70) (ASD-noID, N=201). Levels of cytokines and chemokines were compared between groups using multivariate logistic regression analyses, adjusting for maternal age, ethnicity, birth country and weight, as well as infant gender, birth year and birth month. Mothers of children with ASD+ID had significantly elevated mid-gestational levels of numerous cytokines and chemokines, such as granulocyte macrophage colony-stimulating factor, interferon  $\gamma$ , interleukin  $\alpha$ -1 (IL $\alpha$ -1) and IL-6, compared with mothers of children with either ASD-noID, those with DD, or GP controls. Conversely, mothers of children with either ASD-noID or with DD had significantly lower levels of the chemokines IL-8 and monocyte chemotactic protein-1 compared with mothers of GP controls. This observed immunologic distinction between mothers of children with ASD+ID from mothers of children with ASD-noID or DD suggests that the intellectual disability associated with ASD might be etiologically distinct from DD without ASD. These findings contribute to the ongoing efforts toward identification of early biological markers specific to subphenotypes of ASD.

PubMed ID: 27287966

**Title:** Maternal infection during pregnancy and risk of autism spectrum disorders: A systematic review and meta-analysis  
**Abstract:** Conflicting evidence exists with regard to the relationship between maternal infection during pregnancy and the risk of autism spectrum disorder (ASD) in offspring. The aim of this meta-analysis was to systematically assess this relationship. To identify relevant studies, we conducted systematic searches in PubMed and Embase of scientific articles published through March 2016. Random-effects models were adopted to estimate overall relative risk. A total of 15 studies (2 cohort and 13 case-control studies) involving more than 40,000 ASD cases were included in our meta-analysis. Our results showed that maternal infection during pregnancy was associated with an increased risk of ASD in offspring ( $OR=1.13$ , 95% confidence interval (CI): 1.03–1.23), particularly among those requiring hospitalization ( $OR=1.30$ , 95% CI: 1.14–1.50). Subgroup analyses suggested that risk may be modulated by the type of infectious agent, time of infectious exposure, and site of infection. These findings indicate that maternal infection during pregnancy increases the risk of ASD in offspring. Possible mechanisms may include direct effects of pathogens and, more indirectly, the effects of inflammatory responses on the developing brain.

PubMed ID: 24366406

**Title:** Maternal Infection During Pregnancy and Autism Spectrum Disorders  
**Abstract:** We conducted a nested case-control study including 407 cases and 2,075 frequency matched controls to investigate the association between maternal infections during pregnancy and risk of autism spectrum disorders (ASD). Cases, controls, and maternal infections were ascertained from Kaiser Permanente Northern California clinical databases. No overall association between diagnoses of any maternal infection during pregnancy and ASD was observed [adjusted odds ratio (ORadj) = 1.15, 95 % confidence interval (CI) 0.92–1.43]. However, women with infections diagnosed during a hospital admission (ORadj = 1.48, 95 % CI 1.07–2.04), particularly bacterial infections (ORadj = 1.58, 95 % CI 1.06–2.37), were at increased risk of delivering a child with ASD. Multiple infections during pregnancy were associated with ASD (ORadj = 1.36, 95 % CI 1.05–1.78).

PubMed ID: 25218900

Title: Maternal hospitalization with infection during pregnancy and risk of autism spectrum disorders

Abstract: Animal models indicate that maternal infection during pregnancy can result in behavioral abnormalities and neuropathologies in offspring. We examined the association between maternal inpatient diagnosis with infection during pregnancy and risk of ASD in a Swedish nationwide register-based birth cohort born 1984–2007 with follow-up through 2011. In total, the sample consisted of 2,371,403 persons with 24,414 ASD cases. Infection during pregnancy was defined from ICD codes. In the sample, 903 mothers of ASD cases (3.7%) had an inpatient diagnosis of infection during pregnancy. Logistic regression models adjusted for a number of covariates yielded odds ratios indicating approximately a 30% increase in ASD risk associated with any inpatient diagnosis of infection. Timing of infection did not appear to influence risk in the total Swedish population, since elevated risk of ASD was associated with infection in all trimesters. In a subsample analysis, infections were associated with greater risk of ASD with intellectual disability than for ASD without intellectual disability. The present study adds to the growing body of evidence, encompassing both animal and human studies, that supports possible immune-mediated mechanisms underlying the etiology of ASD.

PubMed ID: 24951035

Title: Neonatal cytokines and chemokines and risk of Autism Spectrum Disorder: the Early Markers for Autism (EMA) study: a case-control study

Abstract: Background:

Biologic markers of infection and inflammation have been associated with Autism Spectrum Disorders (ASD) but prior studies have largely relied on specimens taken after clinical diagnosis. Research on potential biologic markers early in neurodevelopment is required to evaluate possible causal pathways and screening profiles.

Objective:

To investigate levels of cytokines and chemokines in newborn blood specimens as possible early biologic markers for autism.

Methods:

We conducted a population-based case-control study nested within the cohort of infants born from July 2000 to September 2001 to women who participated in the prenatal screening program in Orange County, California, USA. The study population included children ascertained from the California Department of Developmental Services with Autism Spectrum Disorder (ASD, n = 84), or developmental delay but not ASD (DD, n = 49), and general population controls randomly sampled from the birth certificate files and frequency matched to ASD cases on sex, birth month and birth year (GP, n = 159). Cytokine and chemokine concentrations were measured in archived neonatal blood specimens collected for routine newborn screening.

Results:

Cytokines were not detected in the vast majority of newborn samples regardless of case or control status. However, the chemokine monocyte chemotactic protein-1 (MCP-1) was elevated and the chemokine Regulated upon Activation Normal T-Cell

Expressed and Secreted (RANTES) was decreased in ASD cases compared to GP controls. The chemokines macrophage inflammatory protein-1alpha (MIP $\alpha$ -1) and RANTES were decreased in children with DD compared to GP controls.

#### Conclusion:

Measurement of immune system function in the first few days of life may aid in the early identification of abnormal neurodevelopment and shed light on the biologic mechanisms underlying normal neurodevelopment.

PubMed ID: 23337946

Title: Elevated maternal C-reactive protein and autism in a national birth cohort

Abstract: Autism is a complex neuropsychiatric syndrome with a largely unknown etiology.

Inflammation during pregnancy may represent a common pathway by which infections and other insults increase risk for the disorder. Hence, we investigated the association between early gestational C-reactive protein (CRP), an established inflammatory biomarker, prospectively assayed in maternal sera, and childhood autism in a large national birth cohort with an extensive serum biobank. Other strengths of the cohort included nearly complete ascertainment of pregnancies in Finland (N=1.2 million) over the study period and national psychiatric registries consisting of virtually all treated autism cases in the population. Increasing maternal CRP levels, classified as a continuous variable, were significantly associated with autism in offspring. For maternal CRP levels in the highest quintile, compared with the lowest quintile, there was a significant, 43% elevated risk. This finding suggests that maternal inflammation may have a significant role in autism, with possible implications for identifying preventive strategies and pathogenic mechanisms in autism and other neurodevelopmental disorders.

PubMed ID: 38507750

Title: Risk of Autism after Prenatal Topiramate, Valproate, or Lamotrigine Exposure

Abstract: Background:

Maternal use of valproate during pregnancy has been associated with an increased risk of neurodevelopmental disorders in children. Although most studies of other antiseizure medications have not shown increased risks of these disorders, there are limited and conflicting data regarding the risk of autism spectrum disorder associated with maternal topiramate use.

#### Methods:

We identified a population-based cohort of pregnant women and their children within two health care utilization databases in the United States, with data from 2000 through 2020. Exposure to specific antiseizure medications was defined on the basis of prescription fills from gestational week 19 until delivery. Children who had been exposed to topiramate during the second half of pregnancy were compared with those unexposed to any antiseizure medication during pregnancy with respect to the risk of autism spectrum disorder. Valproate was used as a positive control, and lamotrigine was used as a negative control.

#### Results:

The estimated cumulative incidence of autism spectrum disorder at 8 years of age was 1.9% for the full population of children who had not been exposed to antiseizure medication (4,199,796 children). With restriction to children born to mothers

with epilepsy , the incidence was 4.2% with no exposure to antiseizure medication (8815 children) , 6.2% with exposure to topiramate (1030 children) , 10.5% with exposure to valproate (800 children) , and 4.1% with exposure to lamotrigine (4205 children) . Propensity score-adjusted hazard ratios in a comparison with no exposure to antiseizure medication were 0.96 (95% confidence interval [CI] , 0.56 to 1.65) for exposure to topiramate , 2.67 (95% CI , 1.69 to 4.20) for exposure to valproate , and 1.00 (95% CI , 0.69 to 1.46) for exposure to lamotrigine .

#### Conclusions :

The incidence of autism spectrum disorder was higher among children prenatally exposed to the studied antiseizure medications than in the general population . However, after adjustment for indication and other confounders , the association was substantially attenuated for topiramate and lamotrigine , whereas an increased risk remained for valproate . (Funded by the National Institute of Mental Health .)

## V LLM Output.csv

```
chunk_id;node_1;entity_type_1;node_2;entity_type_2;edge
A1;autism spectrum disorder (asd);Disease;neurological disorder;Condition;Causes
A1;autism spectrum disorder (asd);Disease;early development;Stage of Development;Impacts
A1;maternal immune activation (mia);Event;fetal development;Stage of Development;Affects
A1;cytokines;Class of Molecules;maternal immune system;Organismal System;Activates
A1;interleukin -6 (il -6);Protein;fetal development;Stage of Development;Affects
A1;interleukin -6 (il -6);Protein;neuroinflammation;Condition;Causes
A1;interleukin -6 (il -6);Protein;neurodevelopmental processes;Stage of Development;Alters
A1;maternal immune activation (mia);Event;asd-like characteristics;Condition;Promotes
A1;interferon gamma (ifnγ-);Protein;fetal development;Stage of Development;Affects
A2;autism spectrum disorder (asd);Condition;neurological disorder;Concept;Is a type of
A2;maternal immune activation (mia);Condition;immune system activation;Concept;Causes
A2;interleukin -6 (il -6);Entity;cytokine;Concept;Is a type of
A2;interferon gamma (ifnγ-);Entity;cytokine;Concept;Is a type of
A2;placental barrier;Location;barrier;Concept;Allows transfer across
A2;fetus;Entity;developing organism;Concept;Is affected by
A2;neuroinflammation;Condition;inflammation;Concept;Is a type of
A2;gestational inflammation;Condition;inflammation;Concept;Is a type of
A3;autism spectrum disorder;Condition;neurological disorder;Condition;Is a type of
A3;autism spectrum disorder;Condition;maternal immune activation;Mechanism;Is associated with
A3;cytokines;Concept;fetal development;Process;Impact
A3;il -6;Concept;autism spectrum disorder;Condition;Is associated with
A3;il -6;Concept;neuroinflammation;Condition;May induce
A3;il -6;Concept;gestational inflammation;Condition;May promote
A3;mia;Mechanism;asd-like characteristics;Condition;May promote
A3;animal models;Concept;neurological effects;Condition;Support
A4;autism spectrum disorder (asd);Condition;neurological disorder;Condition;is_a
A4;maternal immune activation (mia);Event;asd;Condition;causes
A4;cytokines;Object;fetal development;Process;affects
A4;interleukin -6 (il -6);Object;asd;Condition;is associated with
A4;interferon gamma (ifnγ-);Object;asd;Condition;is associated with
A4;il -6;Object;placental barrier;Location;transfers across
A4;neuroinflammation;Condition;fetal development;Process;induces
A4;mia;Event;neurological effects;Condition;causes
A5;autism spectrum disorder (asd);Condition;neurological disorder;Condition;Is a type of
A5;maternal immune activation (mia);Event;immune system activation;Event;Is a type of
A5;interleukin -6 (il -6);Cytokine;proinflammatory cytokines;Category;Is a type of
A5;interferon gamma (ifnγ-);Cytokine;proinflammatory cytokines;Category;Is a type of
A5;neuroinflammation;Condition;inflammation;Category;Is a type of
```

A5; neurodevelopmental processes; Process; developmental processes; Category; Is a type of  
A5; maternal immune activation (mia); Event; autism spectrum disorder (asd); Condition; Causes  
A6; autism spectrum disorder (asd); Condition; neurological disorder; Condition; Is a type of  
A6; maternal immune activation (mia); Event; cytokines; Agent; Activate  
A6; interleukin -6 (il -6); Entity; cytokines; Agent; Is a type of  
A6; interferon gamma (ifn γ-); Entity; cytokines; Agent; Is a type of  
A6; placental barrier; Location; fetus; Entity; Separates  
A6; gestation; Event; asd diagnosis; Event; Leads to  
A6; maternal immune activation (mia); Event; neuroinflammation; Condition; Induces  
A6; mia; Event; asd-like characteristics; Condition; Promotes  
A7; autism spectrum disorder (asd); Condition; neurological disorder; Condition; Is a type of  
A7; cytokines; Concept; fetal development; Process; Affect  
A7; maternal immune activation (mia); Condition; cytokines; Concept; Cause  
A7; il -6; Concept; placental barrier; Object; Transfer across  
A7; il -6; Concept; fetus; Entity; Accumulates in  
A7; maternal immune activation (mia); Condition; neurodevelopmental processes; Process; Alters  
A7; il -6; Concept; asd; Condition; Is associated with  
A7; il -6; Concept; gestation; Process; Elevated levels during  
A7; mia; Condition; neurological effects; Process; Promotes the development of  
A8; autism spectrum disorder (asd); Condition; neurological disorder; Condition; Is a type of  
A8; maternal immune activation (mia); Event; neurodevelopmental processes; Process; Can alter  
A8; cytokines; Concept; fetal development; Process; Can impact  
A8; interleukin -6 (il -6); Concept; asd; Condition; Is associated with  
A8; interferon gamma (ifn γ-); Concept; asd; Condition; Is associated with  
A8; placental barrier; Object; cytokines; Concept; Can transfer  
A8; neuroinflammation; Condition; fetal development; Process; Can induce  
A8; maternal immune activation (mia); Event; neurological effects; Condition; Can promote  
A9; autism spectrum disorder (asd); Condition; neurological disorder; Condition; is\_a  
A9; maternal immune activation (mia); Event; cytokines; Object; activates  
A9; interleukin -6 (il -6); Object; cytokines; Object; is\_a  
A9; interferon gamma (ifn γ-); Object; cytokines; Object; is\_a  
A9; maternal immune activation (mia); Event; fetal development; Process; influences  
A9; neuroinflammation; Condition; fetal development; Process; can\_induce  
A9; autism spectrum disorder (asd); Condition; neurological disorder; Condition; can\_induce  
A10; autism spectrum disorder (asd); Condition; neurological disorder; Concept; manifestation  
A10; maternal immune activation (mia); Event; fetal development; Concept; impact  
A10; cytokines; Concept; fetal development; Process; influence  
A10; interleukin -6 (il -6); Concept; neuroinflammation; Condition; induction  
A10; il -6; Concept; asd; Condition; association  
A10; proinflammatory cytokines; Concept; asd; Condition; elevation  
A10; maternal immune activation (mia); Event; neurological effects; Condition; promotion  
A11; autism spectrum disorder (asd); Condition; neurological disorder; Condition; Is a type of  
A11; maternal immune activation (mia); Event; immune system activation; Event; Is a type of  
A11; cytokines; Concept; proinflammatory cytokines; Concept; Is a type of  
A11; interleukin -6 (il -6); Concept; cytokine; Concept; Is a type of  
A11; interferon gamma (ifn γ-); Concept; cytokine; Concept; Is a type of  
A11; fetal development; Event; developmental process; Event; Is affected by  
A11; neuroinflammation; Condition; inflammation; Condition; Is a type of  
A11; gestational inflammation; Event; inflammation; Event; Is a type of  
A11; maternal immune activation (mia); Event; autism spectrum disorder (asd); Condition; May promote development of  
A12; autism spectrum disorder (asd); Condition; neurological disorder; Condition; is a  
A12; maternal immune activation (mia); Event; fetal development; Event; affects  
A12; interleukin -6 (il -6); Condition; cytokine; Concept; is  
A12; interferon gamma (ifn γ-); Condition; cytokine; Concept; is  
A12; placental barrier; Location; fetus; Entity; allows transfer of  
A12; maternal immune system; Organization; immune system; Concept; has  
A12; gestation; Event; development; Event; during  
A12; neuroinflammation; Condition; neurodevelopmental processes; Event; alters  
A12; animal models; Entity; models; Concept; support  
A12; maternal immune activation (mia); Event; autism spectrum disorder (asd); Condition; promotes development of  
A13; autism spectrum disorder (asd); Condition; neurological disorder; Condition; has  
A13; asd; Condition; early development; Event; impacts

A13;asd;Condition;communicating;Concept;characterizes  
A13;asd;Condition;social behaviors;Concept;characterizes  
A13;asd;Condition;day-to-day activities;Concept;impacts  
A13;maternal immune activation (mia);Event;fetal development;Event;influences  
A13;cytokines;Concept;asd;Condition;related to  
A13;interleukin -6 (il -6);Concept;cytokines;Concept;part of  
A13;interferon gamma (ifn  $\gamma$ -);Concept;cytokines;Concept;part of  
A13;il -6;Concept;proinflammatory cytokines;Concept;part of  
A13;ifn  $\gamma$ -;Concept;proinflammatory cytokines;Concept;part of  
A13;il -6;Concept;fetus;Entity;transfers across  
A13;il -6;Concept;neuroinflammation;Event;induces  
A13;mia;Event;neurodevelopmental processes;Event;alters  
A13;individuals with asd;Entity;il -6;Concept;have elevated levels of  
A14;autism spectrum disorder (asd);Condition;neurological disorder;Condition;Is\_a  
A14;maternal immune activation (mia);Event;immune system activation;Event;Is\_part\_of  
A14;interleukin -6 (il -6);Cytokine;proinflammatory cytokines;Set of Cytokines;Is\_member\_of  
A14;interferon gamma (ifn  $\gamma$ -);Cytokine;proinflammatory cytokines;Set of Cytokines;  
Is\_member\_of  
A14;placental barrier;Anatomy;fetus;Organism;Contains  
A14;gestation;Event;development;Process;Is\_part\_of  
A14;neuroinflammation;Condition;developmental processes;Process;Affects  
A14;maternal immune activation (mia);Event;development of asd-like characteristics;Process;  
Causes  
A15;autism spectrum disorder (asd);Condition;neurological disorder;Concept;is a type of  
A15;maternal immune activation (mia);Concept;cytokine activation;Mechanism;precipitates  
A15;interleukin -6 (il -6);Cytokine;proinflammatory cytokine;Concept;is a type of  
A15;interferon gamma (ifn  $\gamma$ -);Cytokine;proinflammatory cytokine;Concept;is a type of  
A15;placental barrier;Anatomical structure;barrier;Concept;is a part of  
A15;neuroinflammation;Condition;inflammation;Concept;is a type of  
A15;neurodevelopmental processes;Process;developmental process;Concept;is a type of  
A16;autism spectrum disorder;Condition;neurological disorder;Condition;is a type of  
A16;maternal immune activation;Event;cytokine levels;Measurement;is associated with  
A16;interleukin -6;Entity;cytokine levels;Measurement;is a type of  
A16;interferon gamma;Entity;cytokine levels;Measurement;is a type of  
A16;proinflammatory cytokines;Entity;cytokine levels;Measurement;is a type of  
A16;maternal immune activation;Event;neuroinflammation;Condition;can induce  
A16;cytokines;Entity;fetal development;Event;impact  
A16;maternal immune activation;Event;autism spectrum disorder;Condition;is associated with  
A17;autism spectrum disorder;Condition;neurological disorder;Condition;is a type of  
A17;maternal immune activation;Event;cytokines;Object;activates  
A17;interleukin -6 (il -6);Condition;cytokines;Object;is a type of  
A17;interferon gamma (ifn  $\gamma$ -);Condition;cytokines;Object;is a type of  
A17;maternal immune activation;Event;fetal development;Event;impacts  
A17;autism spectrum disorder;Condition;cytokines;Object;is associated with  
A17;il -6;Condition;proinflammatory cytokines;Object;is a type of  
A17;ifn  $\gamma$ -;Condition;proinflammatory cytokines;Object;is a type of  
A17;il -6;Condition;placental barrier;Location;crosses  
A17;maternal immune activation;Event;neuroinflammation;Condition;induces  
A17;maternal immune activation;Event;neurodevelopmental processes;Event;alters  
A18;autism spectrum disorder (asd);Condition;neurological disorder;Condition;is a type of  
A18;maternal immune activation (mia);Condition;fetal development;Event;impacts  
A18;cytokines;Object;fetal development;Event;influence  
A18;interleukin -6 (il -6);Object;neurodevelopmental processes;Event;alters  
A18;maternal immune activation (mia);Condition;asd-like characteristics;Condition;promotes  
A18;il -6;Object;asd-like characteristics;Condition;is associated with  
A19;autism spectrum disorder (asd);Condition;neurological disorder;Condition;is\_a  
A19;cytokines;Concept;maternal immune activation (mia);Condition;affects  
A19;interleukin -6 (il -6);Concept;cytokines;Concept;is\_a  
A19;interferon gamma (ifn  $\gamma$ -);Concept;cytokines;Concept;is\_a  
A19;placental barrier;Location;fetus;Entity;contains  
A19;neuroinflammation;Condition;neurodevelopmental processes;Concept;alters  
A19;maternal immune activation (mia);Condition;autism-like characteristics;Condition;  
promotes  
A20;autism spectrum disorder (asd);Disease;neurological disorder;Condition;Is\_a type of

A20; autism spectrum disorder (asd); Disease; manifests during early development; Event; Is associated with

A20; maternal immune activation (mia); Event; cytokines; Substance; Have impact on

A20; interleukin -6 (il -6); Substance; proinflammatory cytokines; Category; Is a type of

A20; interleukin -6 (il -6); Substance; cytokines; Category; Is a type of

A20; interleukin -6 (il -6); Substance; maternal immune activation (mia); Event; Has impact on

A20; interferon gamma (ifn $\gamma$ -); Substance; proinflammatory cytokines; Category; Is a type of

A20; interferon gamma (ifn $\gamma$ -); Substance; cytokines; Category; Is a type of

A20; maternal immune activation (mia); Event; fetal development; Process; Has impact on

A20; neuroinflammation; Condition; fetal development; Process; Is a result of

A20; neuroinflammation; Condition; autism spectrum disorder (asd); Disease; Is associated with

A21; autism spectrum disorder (asd); Condition; neurological disorder; Concept; has type

A21; autism spectrum disorder (asd); Condition; maternal immune activation (mia); Concept; related to

A21; interleukin -6 (il -6); Concept; cytokine; Concept; is a type of

A21; interleukin -6 (il -6); Concept; proinflammatory cytokine; Concept; is a type of

A21; interferon gamma (ifn $\gamma$ -); Concept; cytokine; Concept; is a type of

A21; interferon gamma (ifn $\gamma$ -); Concept; proinflammatory cytokine; Concept; is a type of

A21; maternal immune activation (mia); Concept; gestational inflammation; Concept; is related to

A22; autism spectrum disorder; Condition; neurological disorder; Condition; has type

A22; maternal immune activation; Event; fetal development; Process; influences

A22; interleukin -6 (il -6); Entity; cytokines; Entity; is a type of

A22; interferon gamma (ifn $\gamma$ -); Entity; cytokines; Entity; is a type of

A22; maternal immune activation; Event; autism spectrum disorder; Condition; causes

A22; neuroinflammation; Process; fetal development; Process; is related to

A23; autism spectrum disorder; Condition; neurological disorder; Condition; has\_part

A23; maternal immune activation; Event; fetal development; Process; affects

A23; cytokines; Object; maternal immune system; Organization; produced\_by

A23; interleukin -6 (il -6); Object; proinflammatory cytokines; Object; is\_a

A23; interferon gamma (ifn $\gamma$ -); Object; proinflammatory cytokines; Object; is\_a

A23; placental barrier; Object; fetus; Person; transports

A23; maternal immune activation; Event; neuroinflammation; Condition; causes

A23; il -6; Object; asd-like characteristics; Condition; contributes\_to

A24; autism spectrum disorder (asd); Condition; neurological disorder; Condition; is\_a

A24; maternal immune activation (mia); Event; neuroinflammation; Condition; may\_induce

A24; interleukin -6 (il -6); Cytokine; proinflammatory cytokines; Concept; is\_a

A24; interferon gamma (ifn $\gamma$ -); Cytokine; proinflammatory cytokines; Concept; is\_a

A24; maternal immune activation (mia); Event; autism spectrum disorder (asd); Condition; may\_promote

A25; autism spectrum disorder (asd); Condition; neurological disorder; Condition; has type

A25; fetal development; Process; maternal immune activation (mia); Event; is affected by

A25; cytokines; Chemical; fetal development; Process; affects

A25; interleukin -6 (il -6); Chemical; elevated levels; Event; causes

A25; maternal immune activation (mia); Event; neuroinflammation; Condition; can induce

A25; il -6; Chemical; placental barrier; Object; can transfer across

A25; il -6; Chemical; fetus; Object; can accumulate in

A25; maternal immune activation (mia); Event; asd-like characteristics; Condition; can promote development of

A26; autism spectrum disorder (asd); Condition; neurological disorder; Concept; has\_subtype

A26; maternal immune activation (mia); Condition; immune system activation; Concept; causes

A26; cytokines; Substance; interleukin -6 (il -6); Specific substance; has\_component

A26; cytokines; Substance; interferon gamma (ifn $\gamma$ -); Specific substance; has\_component

A26; fetal development; Process; neuroinflammation; Condition; may\_induce

A26; maternal immune activation (mia); Condition; autism spectrum disorder (asd); Condition; may\_promote

A26; animal models; Concept; social interaction impairment; Condition; may\_induce

A26; animal models; Concept; repetitive behavior increase; Condition; may\_induce

A27; autism spectrum disorder; Condition; neurological disorder; Condition; is a type of

A27; maternal immune activation; Event; fetal development; Event; impacts

A27; cytokines; Object; maternal immune system; Organization; are produced by

A27; interleukin -6; Object; proinflammatory cytokines; Object; is a type of

A27; interferon gamma; Object; proinflammatory cytokines; Object; is a type of

A27; il -6; Object; cytokine; Object; is an example of

A27; placental barrier; Object; fetus; Entity; transfers through

A27; il -6; Object ;neuroinflammation ;Event ;can induce  
A27; maternal immune activation ;Event ;asd-like characteristics ;Condition ;is associated with  
A28; autism spectrum disorder ;Condition ;neurological disorder ;Concept ;is a type of  
A28; cytokines ;Condition ;proinflammatory cytokines ;Concept ;are a type of  
A28; maternal immune activation ;Event ;gestational inflammation ;Condition ;is a type of  
A28; interleukin -6; Condition ;proinflammatory cytokine ;Concept ;is a type of  
A28; interferon gamma ;Condition ;proinflammatory cytokine ;Concept ;is a type of  
A28; placental barrier ;Location ;fetus ;Entity ;separates  
A28; neuroinflammation ;Condition ;neurodevelopmental processes ;Concept ;affects  
A28; autism-like characteristics ;Condition ;neurological effects ;Concept ;are associated with  
A28; maternal immune activation ;Event ;development of asd-like characteristics ;Condition ;may promote  
A29; autism spectrum disorder (asd) ;Condition ;neurological disorder ;Condition ;Is a type of  
A29; maternal immune activation (mia) ;Event ;fetal development ;Event ;Affects  
A29; cytokines ;Concept ;fetal development ;Event ;Influences  
A29; interleukin -6 (il -6) ;Concept ;neuroinflammation ;Condition ;Causes  
A29; interferon gamma (ifn γ-) ;Concept ;neuroinflammation ;Condition ;Causes  
A29; il -6; Concept ;asd ;Condition ;Is associated with  
A29; ifn γ-; Concept ;asd ;Condition ;Is associated with  
A29; maternal immune activation (mia) ;Event ;neurological effects ;Condition ;Causes  
A30; autism spectrum disorder (asd) ;Condition ;neurological disorder ;Condition ;Is a type of  
A30; cytokines ;Concept ;maternal immune activation (mia) ;Event ;Contribute to  
A30; interleukin -6 (il -6) ;Concept ;cytokines ;Concept ;Is a type of  
A30; interferon gamma (ifn γ-) ;Concept ;cytokines ;Concept ;Is a type of  
A30; il -6; Concept ;proinflammatory cytokines ;Concept ;Is a type of  
A30; mia ;Event ;fetal development ;Process ;Affects  
A30; il -6; Concept ;neuroinflammation ;Condition ;Induces  
A30; asd ;Condition ;neurological effects ;Condition ;Is associated with

S

---

## **Investigating Maternal Inflammation During Pregnancy and its Association with Autism Spectrum Disorder in Offspring**

### **Abstract:**

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder that affects social interaction, communication, interests, and behavior. The exact cause of ASD is unknown, but research suggests that maternal inflammation during pregnancy may increase the risk of developing ASD in offspring. In this project, we created a knowledge graph to investigate the association between maternal inflammation and ASD. We parsed medical article abstracts from PubMed, and performed Named Entity Recognition to extract entities with two distinct approaches. Initially with the SciSpaCy library and subsequently with the LLama3 model. The retrieved entities and relations were integrated into a structured table and deployed to Neo4j resulting in a strong knowledge graph. Our results show that maternal inflammation during pregnancy is associated with an increased risk of ASD in offspring, and that certain cytokines and chemokines are key players in this association. This project provides an organized and integrated picture of biomedical information, making it a useful tool for medical researchers, scholars, and practitioners.

**Keywords:** Autism Spectrum Disorder, maternal inflammation, knowledge graph, Disease Ontology, Neo4j, LLM

---

---

## **Investigation de l'inflammation maternelle pendant la grossesse et son association avec le trouble du spectre autistique chez les enfants**

### **Résumé:**

Le trouble du spectre autistique (TSA) est un trouble du neurodéveloppement qui affecte l'interaction sociale, la communication, les intérêts et le comportement. La cause exacte du TSA est inconnue, mais des recherches suggèrent que l'inflammation maternelle pendant la grossesse peut augmenter le risque de développer un TSA chez les enfants. Dans ce projet, nous avons créé un graphe de connaissances pour étudier l'association entre l'inflammation maternelle et le TSA. Nous avons analysé les résumés d'articles médicaux de PubMed et avons effectué une reconnaissance d'entités nommées pour extraire les entités avec deux approches distinctes. Initialement avec la bibliothèque SciSpacy, puis avec le modèle LLama3. Les entités et relations extraites ont été intégrées dans un tableau structuré et déployées dans Neo4j, ce qui a abouti à un graphe de connaissances solide. Nos résultats montrent que l'inflammation maternelle pendant la grossesse est associée à un risque accru de TSA chez les enfants, et que certains cytokines et chimiokines sont des acteurs clés dans cette association. Ce projet offre une image organisée et intégrée des informations biomédicales, en faisant un outil utile pour les chercheurs médicaux, les universitaires et les praticiens.

**Mots clés:** Trouble du spectre de l'autisme, Inflammation maternelle, Graph de connaissances, ontologie des maladies, Neo4j, LLM

---

