Team 0 - D&A
Project - Phase III

Aditya Nair (2023111029) Shreyas Deb (2023113024) Vishak Kashyap K (2023113012)

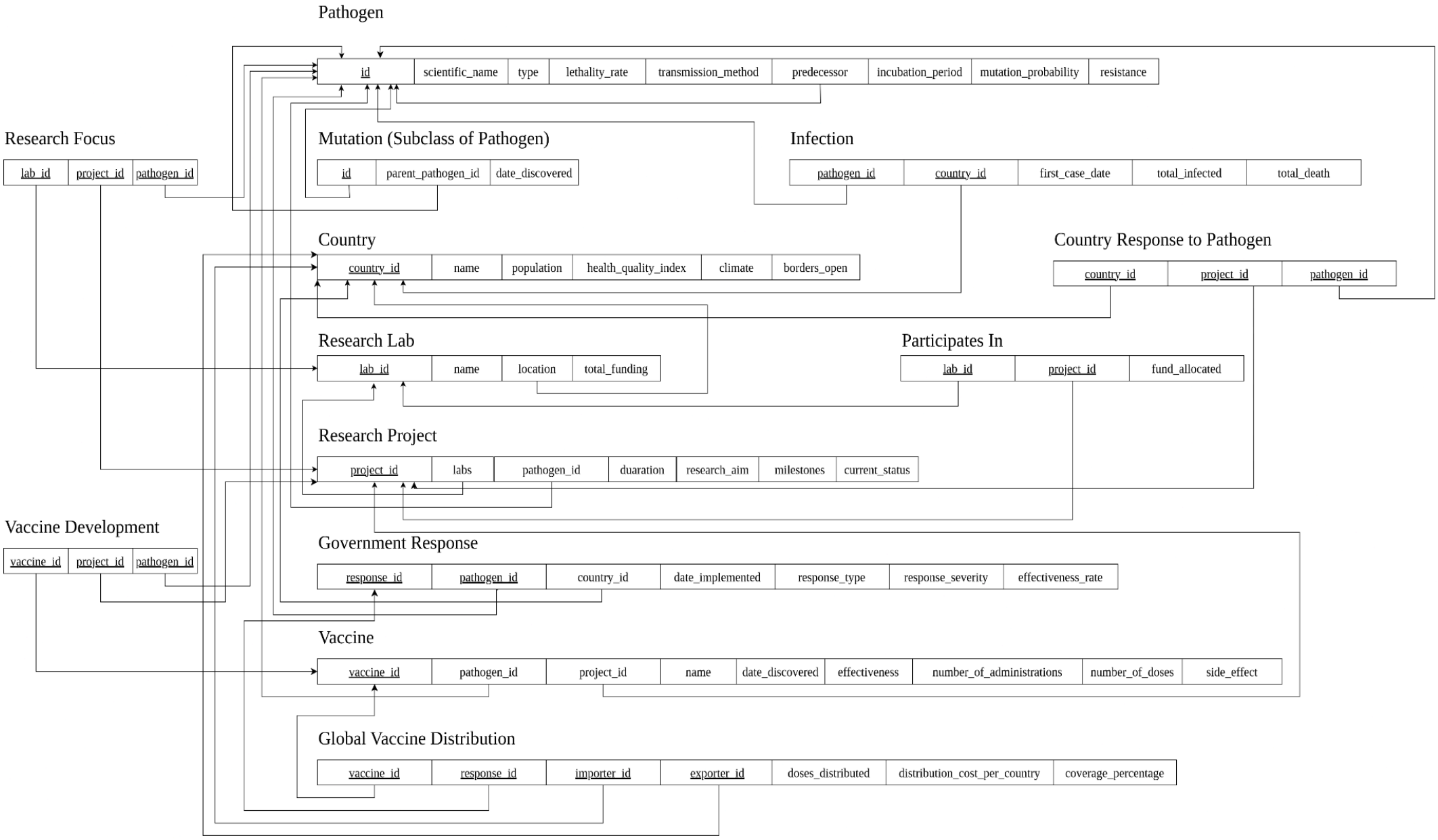
(total 8 pages)

Here is the draw.io link to our model diagrams - team 0 phase 3

1. **ERD to Relational Model**

   The changes made during this transition are as follows:

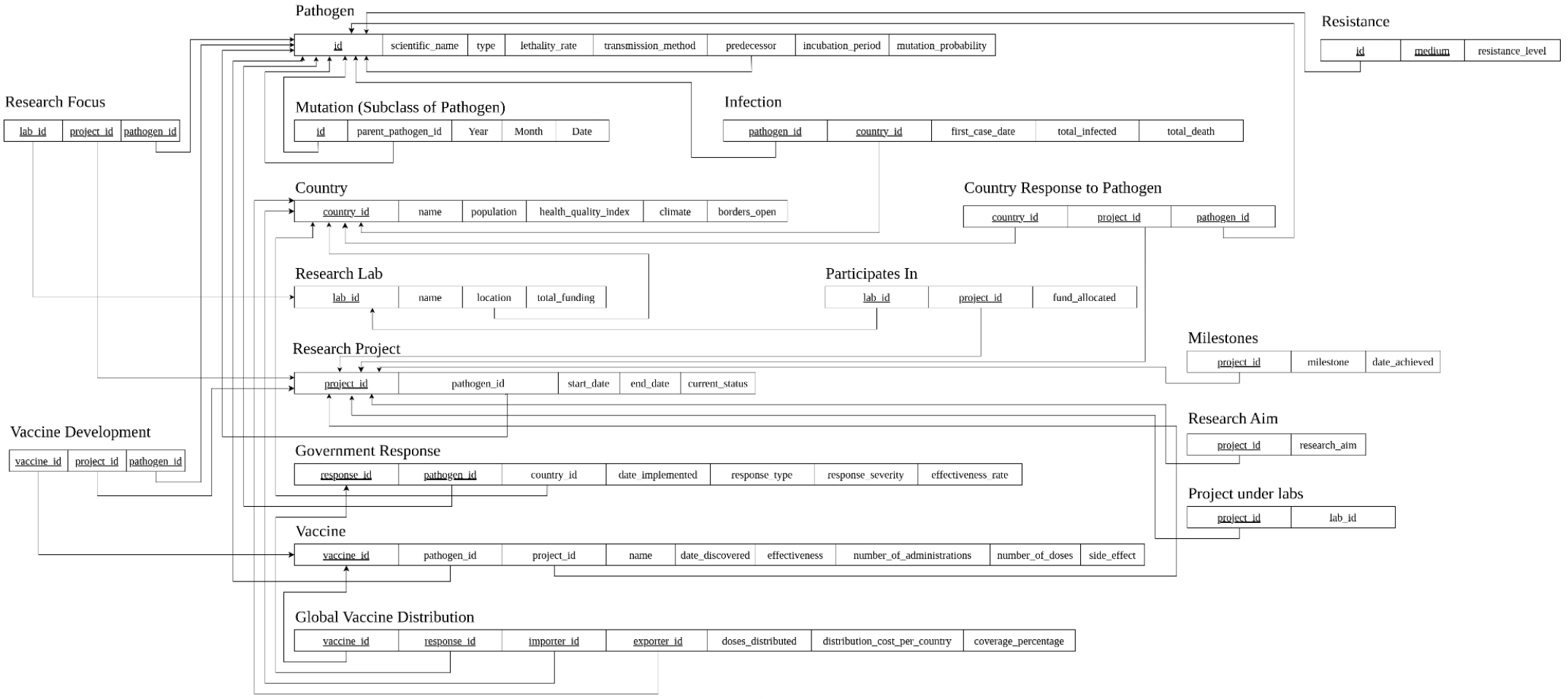
   - All the entity types are converted to relations. The arrow marks between the relations indicate the foreign key referencing.
   - We have used the option 8A (given in the textbook, pg no. 299) to represent the subclass-superclass relations.
   - All the composite and multivalued attributes have been kept as it is (We have assumed the traditional relational model, as the modern relational model is equivalent to 1NF).
   - All the 1:N relationship types have been represented using the foreign key approach (textbook, pg 294, step 4).
   - For the M:N binary relationships, the ternary and quaternary relationships, we have used the approach given in the textbook (step 5, pg 294 & pg 296).
   - New attributes have been created or have been made a foreign key to represent the two recursive relationships in our database design. One is in the pathogen relation, where the `predecessor` key is made a foreign key to the Pathogen relation itself. Another instance is in the Global Vaccine Distribution where `importer_id` and `exporter_id` attributes have been added which reference the country relation.


Pathogen
id
scientific_name
type
lethality_rate
transmission_method
predecessor
incubation_period
mutation_probability
resistance
Research Focus
lab_id
project_id
pathogen_id
Mutation (Subclass of Pathogen)
id
parent_pathogen_id
date_discovered
Infection
pathogen_id
country_id
first_case_date
total_infected
total_death
Country
country_id
name
population
health_quality_index
climate
borders_open
Country Response to Pathogen
country_id
project_id
pathogen_id
Research Lab
lab_id
name
location
total_funding
Participates In
lab_id
project_id
fund_allocated
Research Project
project_id
labs
pathogen_id
duaration
research_aim
milestones
current_status
Vaccine Development
vaccine_id
project_id
pathogen_id
Government Response
response_id
pathogen_id
country_id
date_implemented
response_type
response_severity
effectiveness_rate
Vaccine
vaccine_id
pathogen_id
project_id
name
date_discovered
effectiveness
number_of_administrations
number_of_doses
side_effect
Global Vaccine Distribution
vaccine_id
response_id
importer_id
exporter_id
doses_distributed
distribution_cost_per_country
coverage_percentage

2. Relational Model to 1NF

The following changes have been made during this transition:

- All the composite attributes have been represented in terms of their simple components.
- All the multivalued attributes have been represented in a new relation.
- There were two composite multivalued attributes, one is `resistance` in the Pathogen relation and the other is `milestones` in the Research Project relation.


Pathogen
id
scientific_name
type
lethality_rate
transmission_method
predecessor
incubation_period
mutation_probability
Resistance
id
medium
resistance_level
Research Focus
lab_id
project_id
pathogen_id
Mutation (Subclass of Pathogen)
id
parent_pathogen_id
Year
Month
Date
Infection
pathogen_id
country_id
first_case_date
total_infected
total_death
Country
country_id
name
population
health_quality_index
climate
borders_open
Country Response to Pathogen
country_id
project_id
pathogen_id
Research Lab
lab_id
name
location
total_funding
Participates In
lab_id
project_id
fund_allocated
Milestones
project_id
milestone
date_achieved
Research Project
project_id
pathogen_id
start_date
end_date
current_status
Vaccine Development
vaccine_id
project_id
pathogen_id
Research Aim
project_id
research_aim
Government Response
response_id
pathogen_id
country_id
date_implemented
response_type
response_severity
effectiveness_rate
Project under labs
project_id
lab_id
Vaccine
vaccine_id
pathogen_id
project_id
name
date_discovered
effectiveness
number_of_administrations
number_of_doses
side_effect
Global Vaccine Distribution
vaccine_id
response_id
importer_id
exporter_id
doses_distributed
distribution_cost_per_country
coverage_percentage

3. 1NF to 2NF

The following changes have been made during this transition:

- Government Response relation has `date_implemented` and `response_type` dependent only on the `response_id` and the others dependent on both `response_id` and `pathogen_id` both, and are broken off into a new relation Response Effect.
- Global Vaccine Distribution has `distribution_cost_per_country` depending only on the `vaccine_id` and `exporter_id` as it depends only on the exporting country for a vaccine, and refers to the customs and such costs excluding cost of actual vaccine units. Hence, we broke it off into a new relation Vaccine Distribution Cost.
- While creating a new relation for the multivalued attribute `labs` of Research Project, it turns out equivalent to an existing relationship type, based on which the relation Participates In was created. Hence we kept the relations Participates In and renamed it to `Project Under Lab` as well, to better reflect its purpose, and further remove redundancy.
- Also, to be noted that, the relations `Research Focus`, `Vaccine Development`, and `Country Response to Pathogen` can, and should, be implemented as Views in RDBMS implementation. However, we are retaining them in the relational model as they represent important relationship types.


Pathogen
id
scientific_name
type
lethality_rate
transmission_method
predecessor
incubation_period
mutation_probability
Resistance
id
medium
resistance_level
Research Focus
lab_id
project_id
pathogen_id
Mutation (Subclass of Pathogen)
id
parent_pathogen_id
Year
Month
Date
Infection
pathogen_id
country_id
first_case_date
total_infected
total_death
Country
country_id
name
population
health_quality_index
climate
borders_open
Country Response to Pathogen
country_id
project_id
pathogen_id
Research Lab
lab_id
name
location
total_funding
Project under Lab
lab_id
project_id
fund_allocated
Milestones
project_id
milestone
date_achieved
Research Project
project_id
pathogen_id
start_date
end_date
current_status
Vaccine Development
vaccine_id
project_id
pathogen_id
Research Aim
project_id
research_aim
Government Response
response_id
country_id
date_implemented
response_type
Response Effect
response_id
pathogen_id
response_severity
effectiveness_rate
Vaccine
vaccine_id
pathogen_id
project_id
name
date_discovered
effectiveness
number_of_administrations
number_of_doses
side_effect
Global Vaccine Distribution
vaccine_id
response_id
importer_id
exporter_id
doses_distributed
coverage_percentage
Vaccine Distribution Cost
vaccine_id
exporter_id
distribution_cost_per_country

4. 2NF to 3NF

The following (and only) change has been made during this transition:

- It is observed that `pathogen_id` in Vaccine relation can be inferred from the non-prime attribute `project_id` as each Project, as well as Vaccine, focuses on only one Pathogen in our universe/mini-world. Since `pathogen_id` is already present in the Project that a given Vaccine emerges out of, it was simply removed from Vaccine.


Pathogen
id
scientific_name
type
lethality_rate
transmission_method
predecessor
incubation_period
mutation_probability
Resistance
id
medium
resistance_level
Research Focus
lab_id
project_id
pathogen_id
Mutation (Subclass of Pathogen)
id
parent_pathogen_id
Year
Month
Date
Infection
pathogen_id
country_id
first_case_date
total_infected
total_death
Country
country_id
name
population
health_quality_index
climate
borders_open
Country Response to Pathogen
country_id
project_id
pathogen_id
Research Lab
lab_id
name
location
total_funding
Project under Lab
lab_id
project_id
fund_allocated
Milestones
project_id
milestone
date_achieved
Research Project
project_id
pathogen_id
start_date
end_date
current_status
Vaccine Development
vaccine_id
project_id
pathogen_id
Research Aim
project_id
research_aim
Government Response
response_id
country_id
date_implemented
response_type
Response Effect
response_id
pathogen_id
response_severity
effectiveness_rate
Vaccine
vaccine_id
project_id
name
date_discovered
effectiveness
number_of_administrations
number_of_doses
Global Vaccine Distribution
vaccine_id
response_id
importer_id
exporter_id
doses_distributed
coverage_percentage
Vaccine Distribution Cost
vaccine_id
exporter_id
distribution_cost_per_country