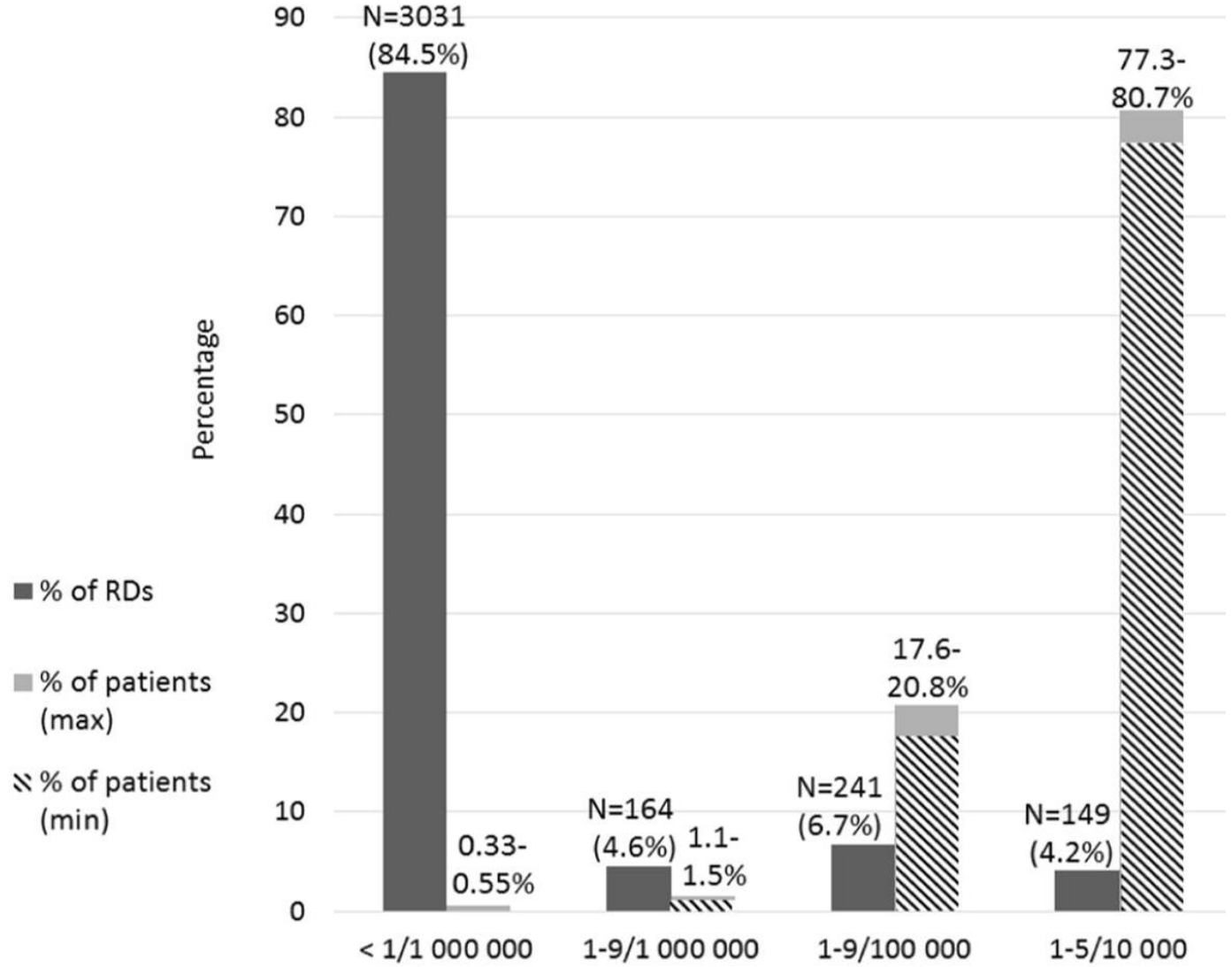
The PLUTO Project

research in forgotten diseases

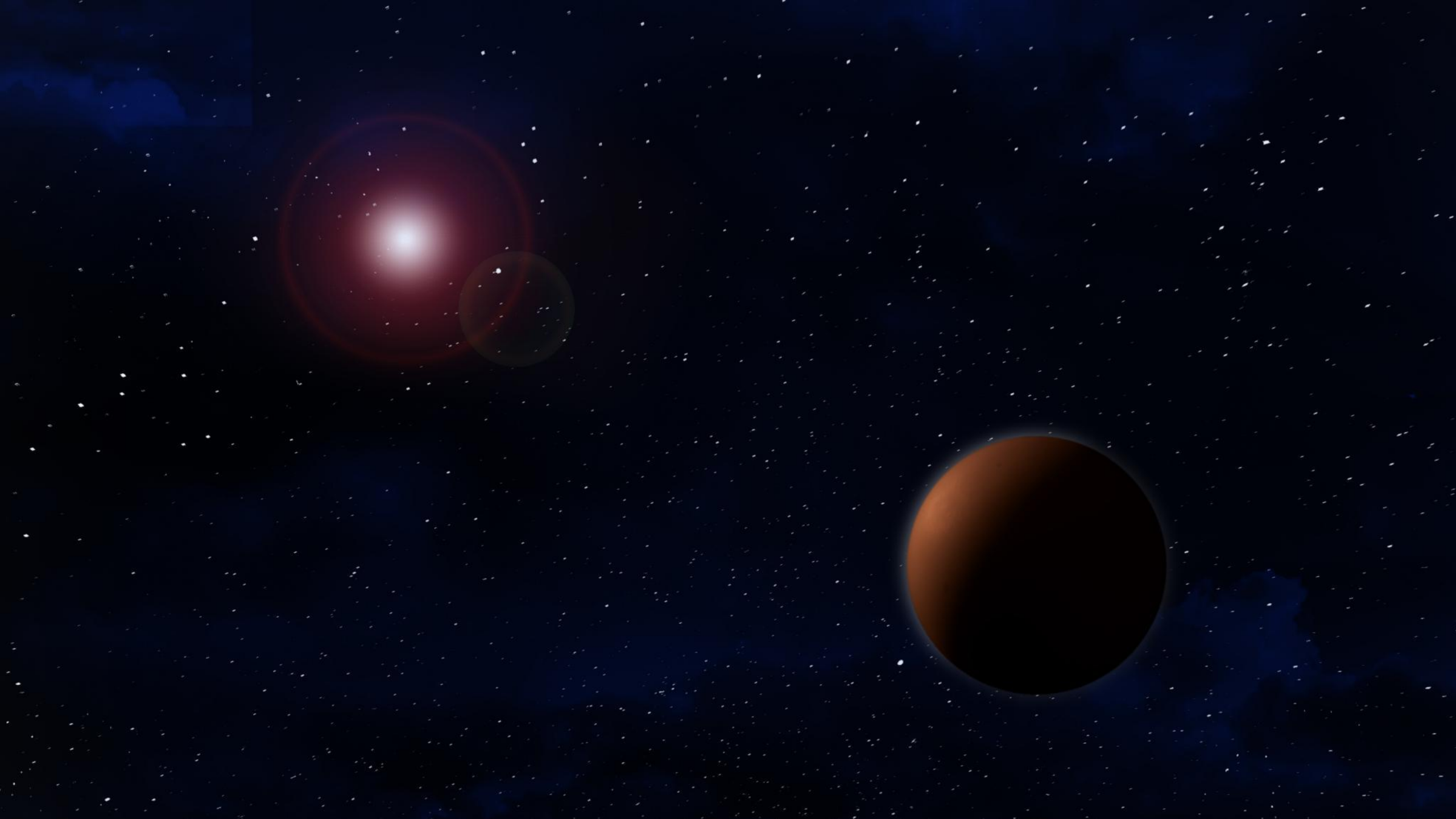
Diego Ardigò, MD PhD

ERICA ERN Research Conference
11- 13 December 2024 (Udine, Italy)

Prevalence of rare diseases



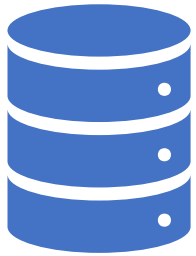
Nguengang Wakap S et al. European Journal of Human Genetics (2020) 28:165–173



Chiesi

METHODOLOGY OVERVIEW

Integration strategy

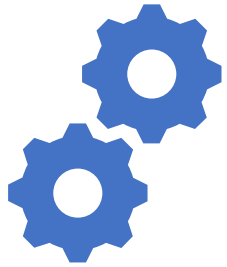


Automated Data Integration

Leverages various ontologies for semantic search

Identifies data related to rare diseases without manual curation

Ensures efficiency and consistency in data retrieval



Use of NLP Tool

Employs Linguamatics (IQVIA) for integration

Integrates multiple ontologies for comprehensive cross-referencing

PLUTO DATA CONNECTIONS

Pluto Project
Defining Disregarded Rare Diseases

DISEASE ONTOLOGIES and THESAURUS

SOURCE:

MeSH
MedDRA
DISEASE ONTOLOGY
ordo
mondo

EXCLUSIONS:
Neoplastic Disorders; Obsolete or NON-Rare; Other free terms like Cancers, melanoma, etc.

EPIDEMIOLOGY

SOURCE: orphadata

DATA TYPE:
- *Prevalence Class*
- *Epidemiology (descriptive)*

SCIENTIFIC KNOWLEDGE

SOURCE: PubMed.gov

DATA TYPE:
- *# of Publications*

GENETIC KNOWLEDGE

SOURCE: OMIM

DATA TYPE:
- *Gene details*
- *OMIM number*
- *OMIM group*

CLINICAL KNOWLEDGE

SOURCE: U.S. National Library of Medicine ClinicalTrials.gov; Trialtrove

DATA TYPE:
- *# of Clinical Trials*

SYMPTOMS

SOURCE: hpo

DATA TYPE:
- *# of Symptoms*
- *# of Organs affected*

SoC AND FUTURE THERAPIES

SOURCE: orphadata DRUGBANK

DATA TYPE:
- *# of Approved Drugs*
- *# of Orphan Drug Designation Granted (EU and US)*

RDs GENERAL KNOWLEDGE

SOURCE: orphadata

DATA TYPE:
- *Disease Definition*
- *Classification*
- *Clinical description*
- *Etiology*
- *Diagnostic methods*
- *Differential diagnosis*
- *Antenatal diagnosis*
- *Genetic counseling*
- *Management and treatment*
- *Prognosis*

PTS ASSOCIATIONS

SOURCE: orphadata

DATA TYPE:
- *# of PTS Associations*

EXPERT CENTERS

SOURCE: orphadata

DATA TYPE:
- *# of Expert Centers*

LABS & DIAGNOSTICS

SOURCE: orphadata

DATA TYPE:
- *# of Labs*
- *# of Diagnostics*

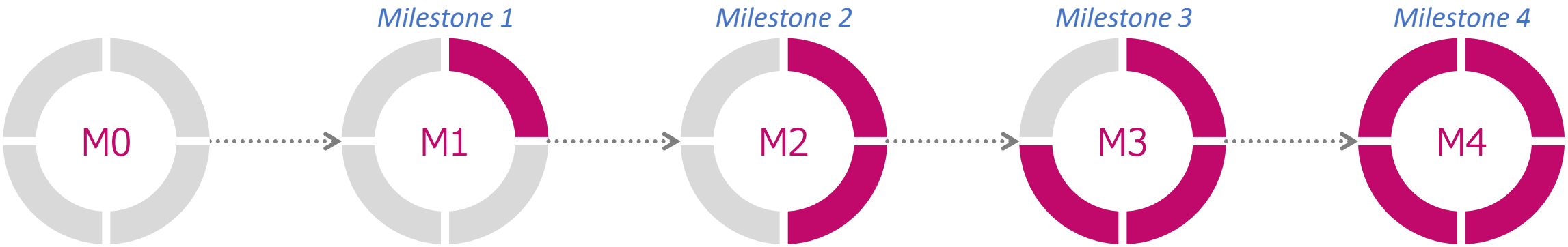
BIOBANKS & REGISTRIES

SOURCE: orphadata

DATA TYPE:
- *# of Biobanks*
- *# of Registries*

- *Not integrated into the final dataset*
- *Reliable data only for EU*

The journey of a rare disease to become treatable



Data Cleaning and Preprocessing

Initial Dataset Extraction

- Comprehensive dataset with 11,690 entries from Orphanet
- Dataset categorized into three groups: Groups of Disorders, Disorders, and Subtypes of Disorders

Data Cleaning Phase

- Exclusion of 2,153 entries classified under cancer or non-rare conditions
- Removal of 542 outliers in rare diseases

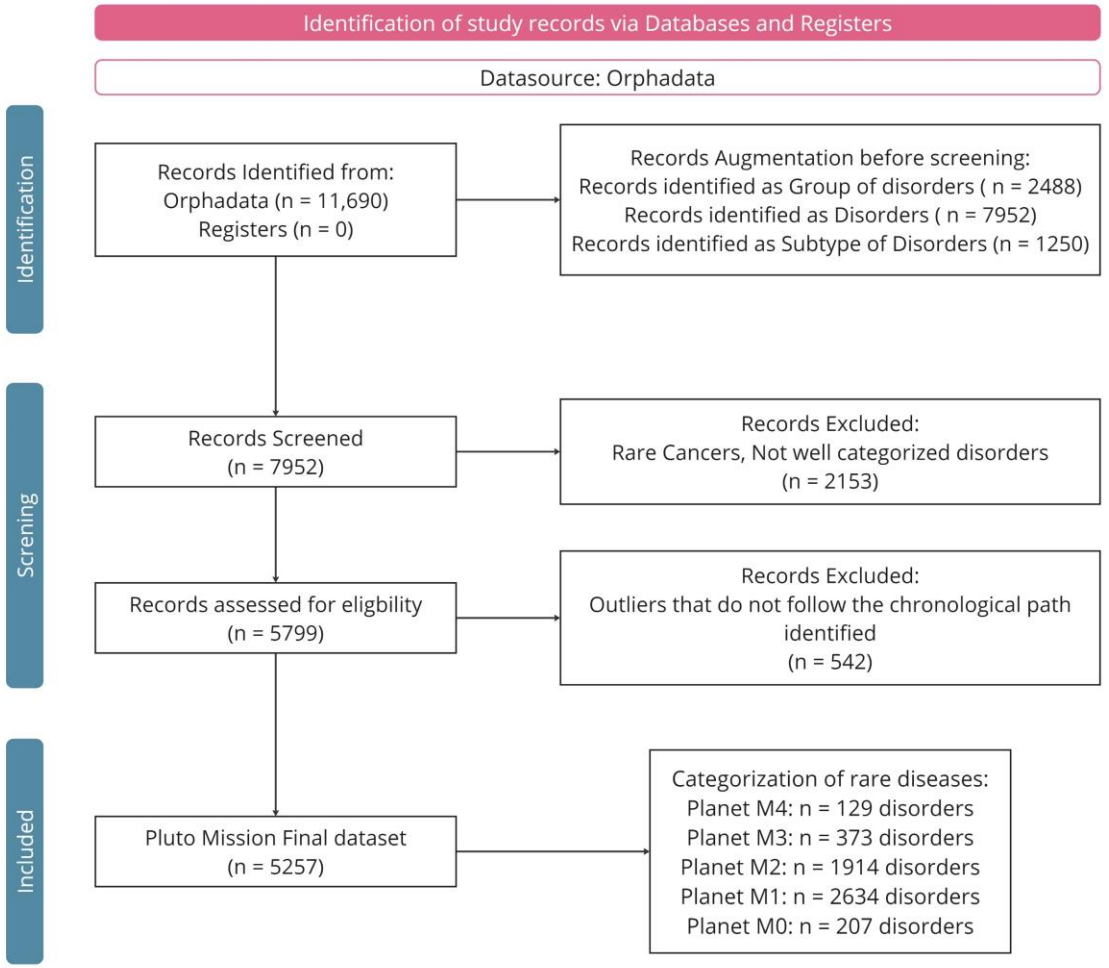
Final Streamlined Dataset

- Resulted in 5,257 unique entries

Additional Data Retrieval

- 19,464,268 publications
- 666,082 clinical trials
- 1,753 orphan drug designations (ODD)
- 273 marketing authorizations (MA)

Manual Dataset Verification



PLUTO database in numbers

Pluto Project
Defining Disregarded Rare Diseases

ONTOLOGIES
5

SOURCES OF INFORMATION
7

DATABASE
1

RARE DISEASES
5257

Exclusions:
- 2,153 entries classified under cancer or non-rare conditions
- 542 outliers in rare diseases

CLINICAL TRIALS
666 K

PUBLICATIONS
19 M

ORPHAN DRUG DESIGNATIONS (EU)
1753

APPROVED DRUGS (EU)
273

EXPERT CENTERS
21 K

PATIENT ASSOCIATIONS
6.5 K

Chiesi

KEY FINDINGS

When the 5,257 rare diseases were first described

Correlation between rare diseases and historical events

- 1840: First publication on rare diseases
- Twentieth century: World Wars halted discoveries
- Post-WWII: Economic boom and scientific discoveries enhanced research

Impact of medical diagnostic tools and techniques

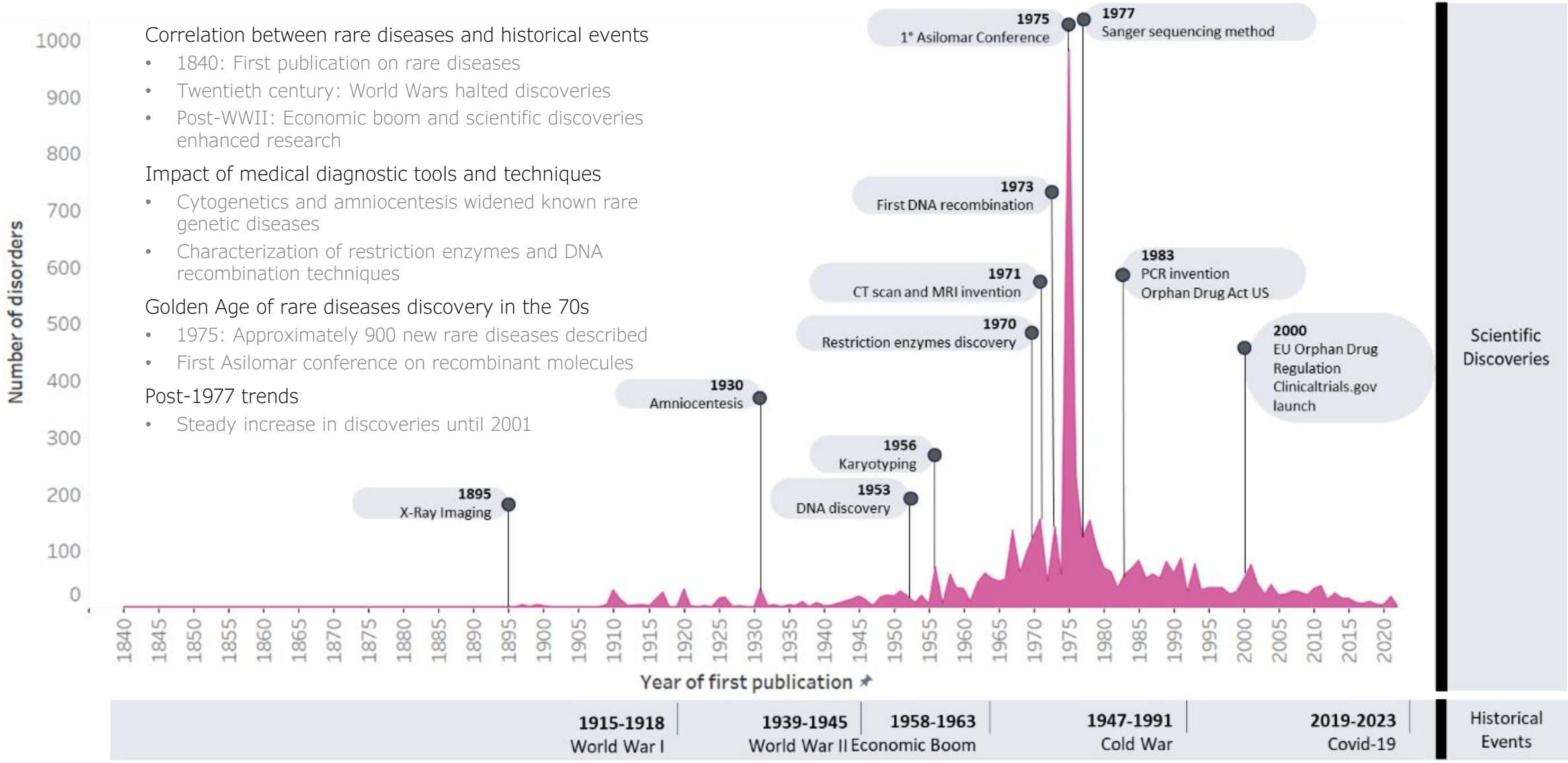
- Cytogenetics and amniocentesis widened known rare genetic diseases
- Characterization of restriction enzymes and DNA recombination techniques

Golden Age of rare diseases discovery in the 70s

- 1975: Approximately 900 new rare diseases described
- First Asilomar conference on recombinant molecules

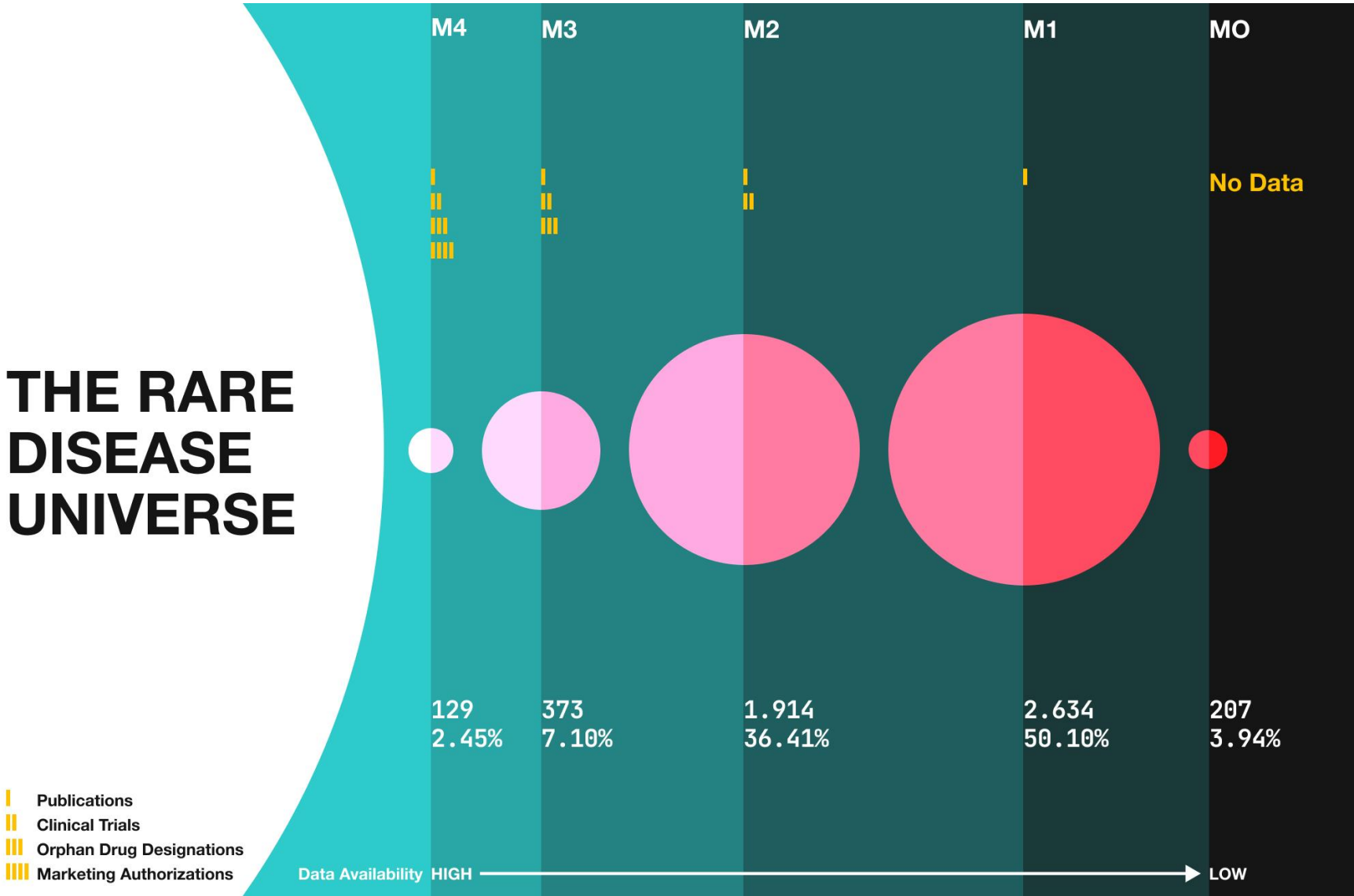
Post-1977 trends

- Steady increase in discoveries until 2001

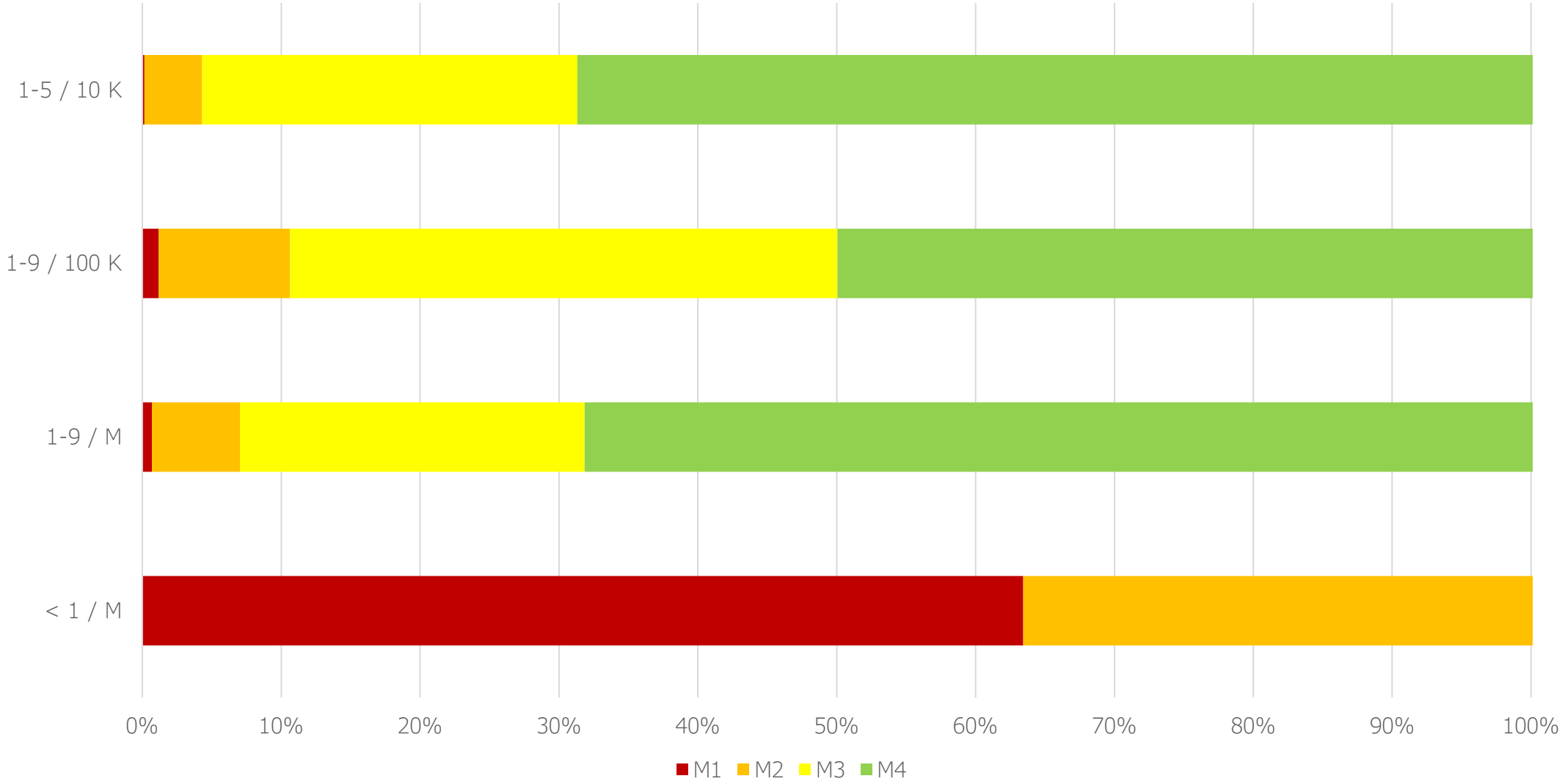


The rare diseases solar system

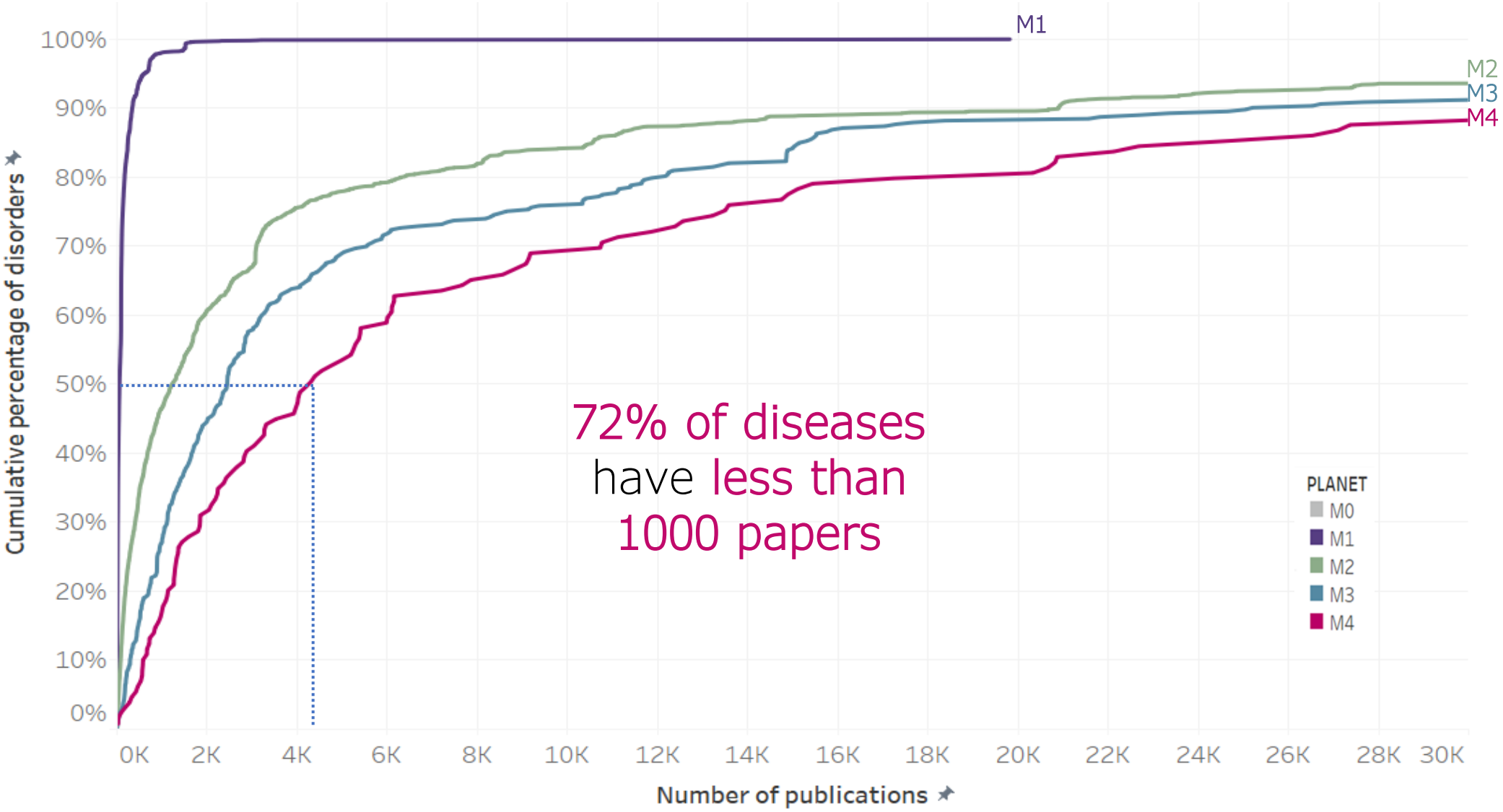
More than 50% of rare diseases have only literature publications, but there is no evidence of drug development



Distribution of RD «planets» across prevalence classes



The distribution of knowledge in rare diseases



PubMed Manuscripts for Selected Diseases

RARE DISEASES (median)

- **M1 = 54**
- **M2 = 1,221**
- **M3 = 2,445**
- **M4 = 4,308**

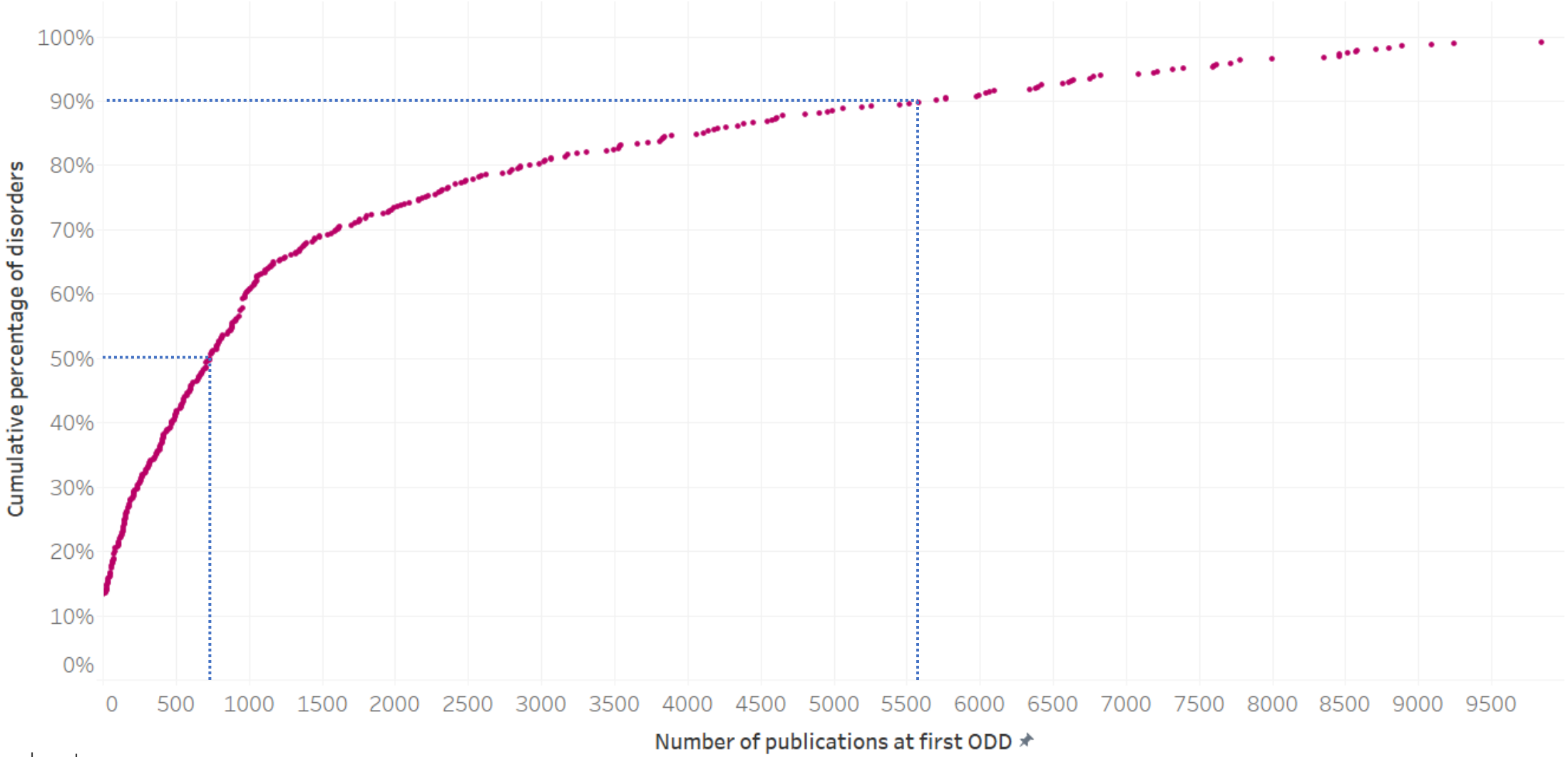
- Cystic fibrosis = ~ 320,000
- Tuberculosis = ~ 310,000
- Rare variants of common diseases

COMMON DISEASES

- Type 2 diabetes = 200,000 +
- COPD = ~110,000
- Chronic heart failure = 65,000 +
- Alzheimer's disease = ~210,000.
- Parkinson's disease = 200,000 +
- Liver cirrhosis = ~140,000

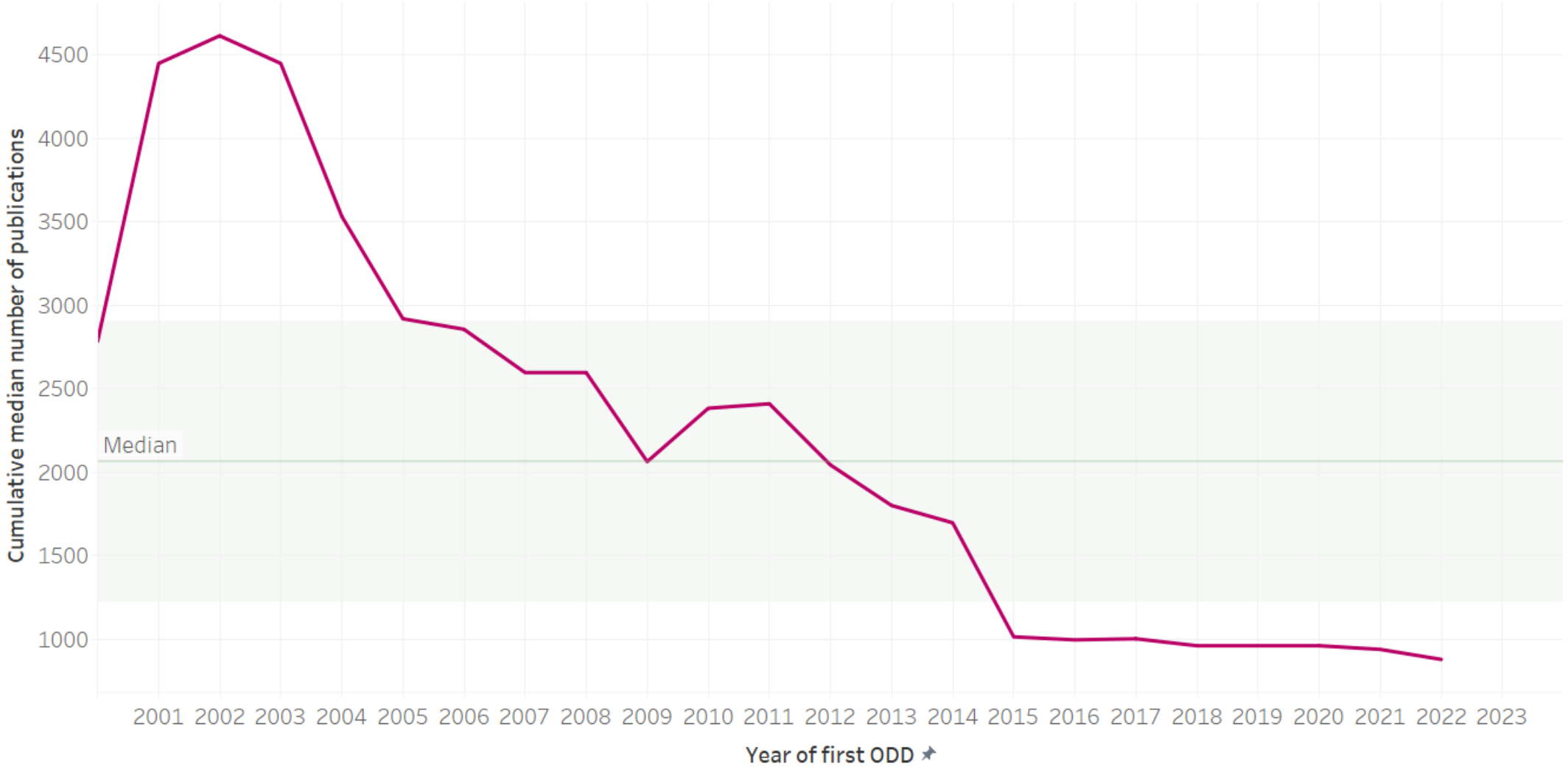
Number of published manuscripts in PubMed as of index date (December 31st, 2023).

Body of knowledge at the time of the first ODD

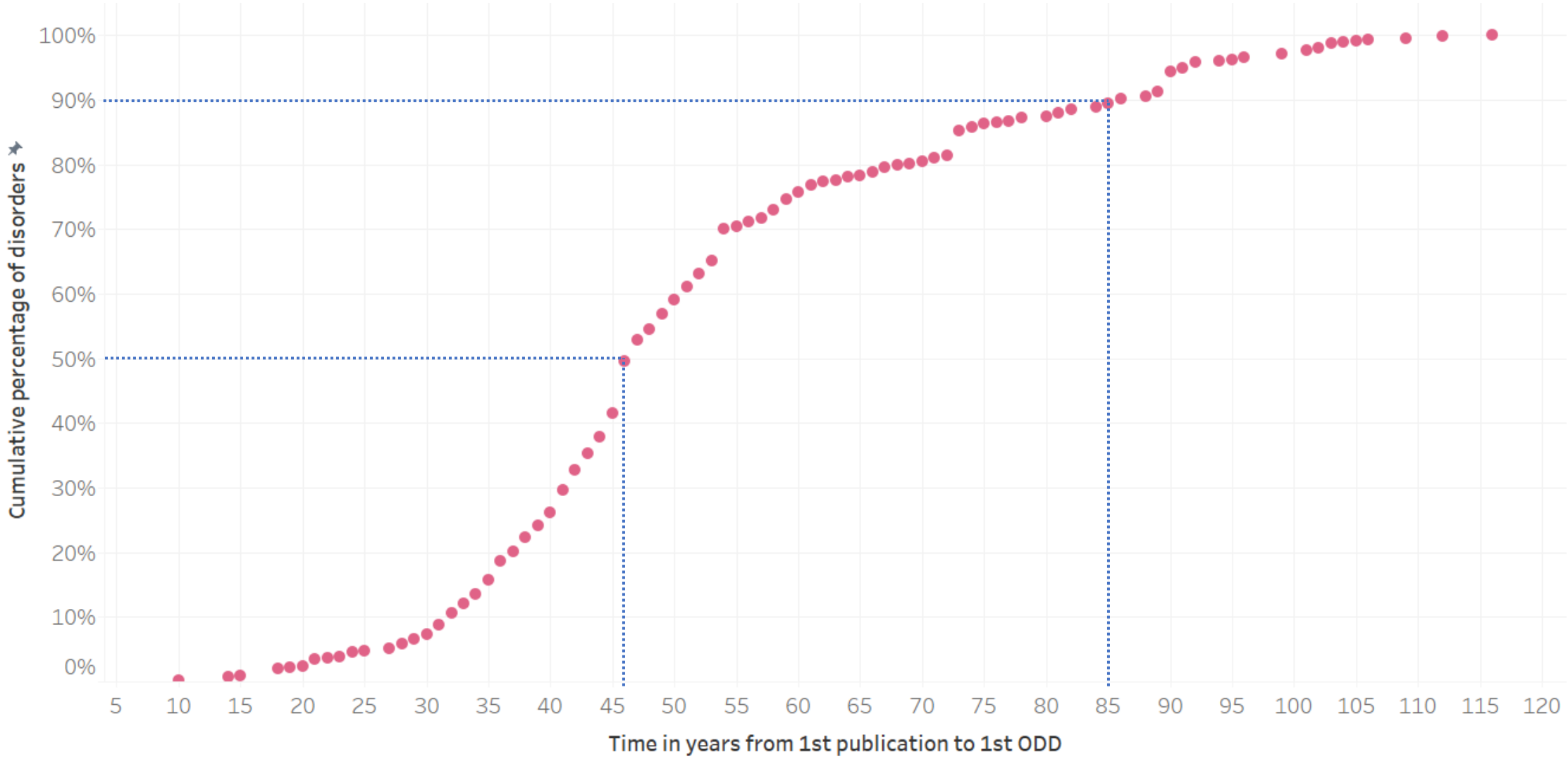


M3 and M4 «planets»

Evolution of knowledge at the time of the first ODD

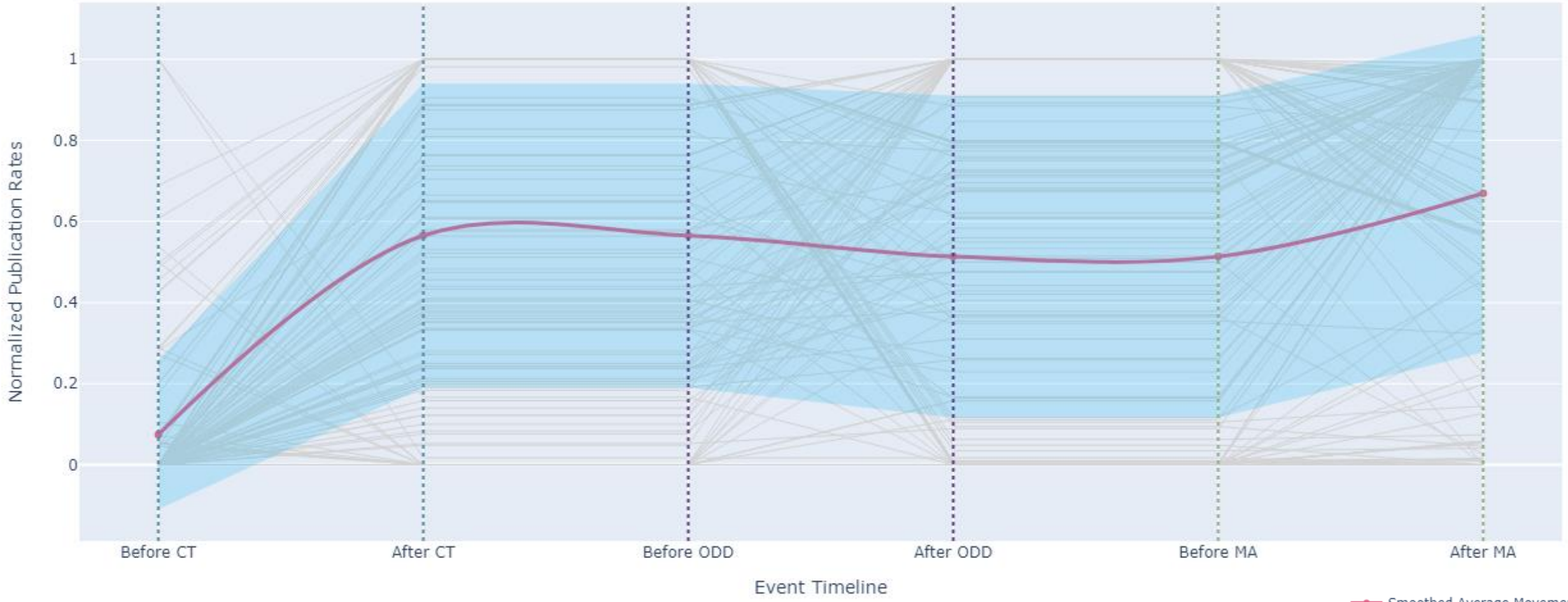


Time from disease discovery to the first ODD

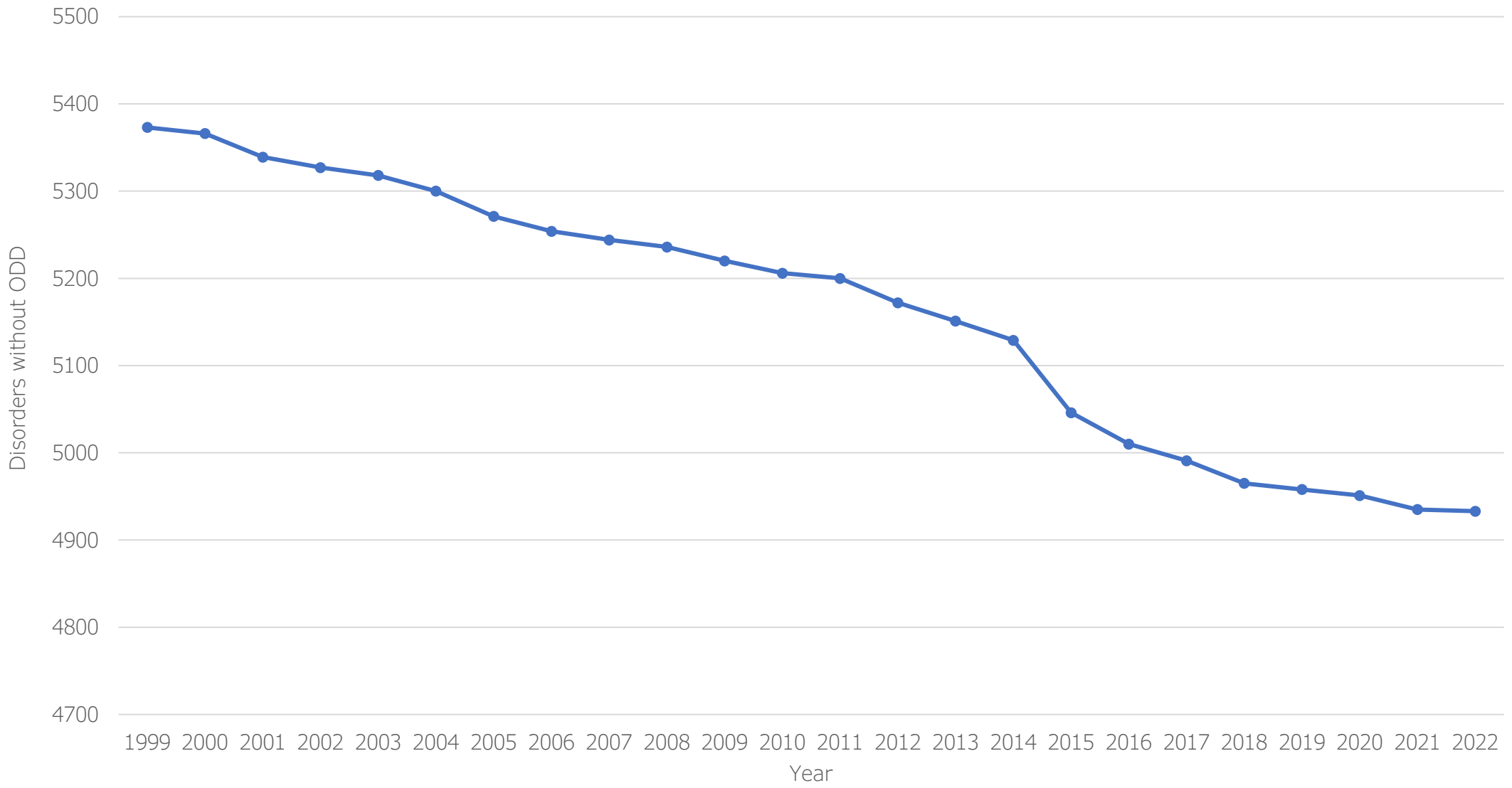


M3 and M4 «planets»

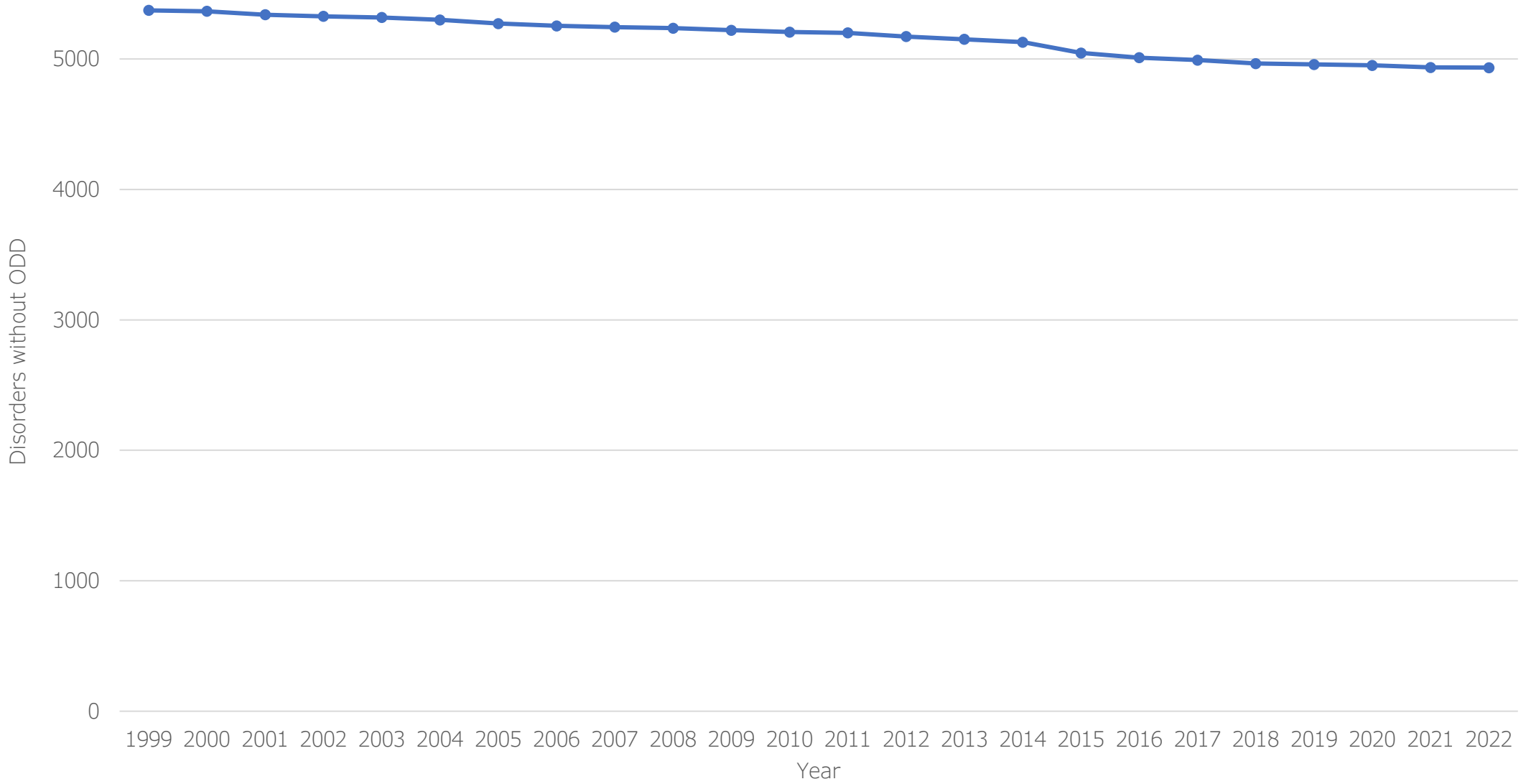
Evolution of publication productivity across «planets»



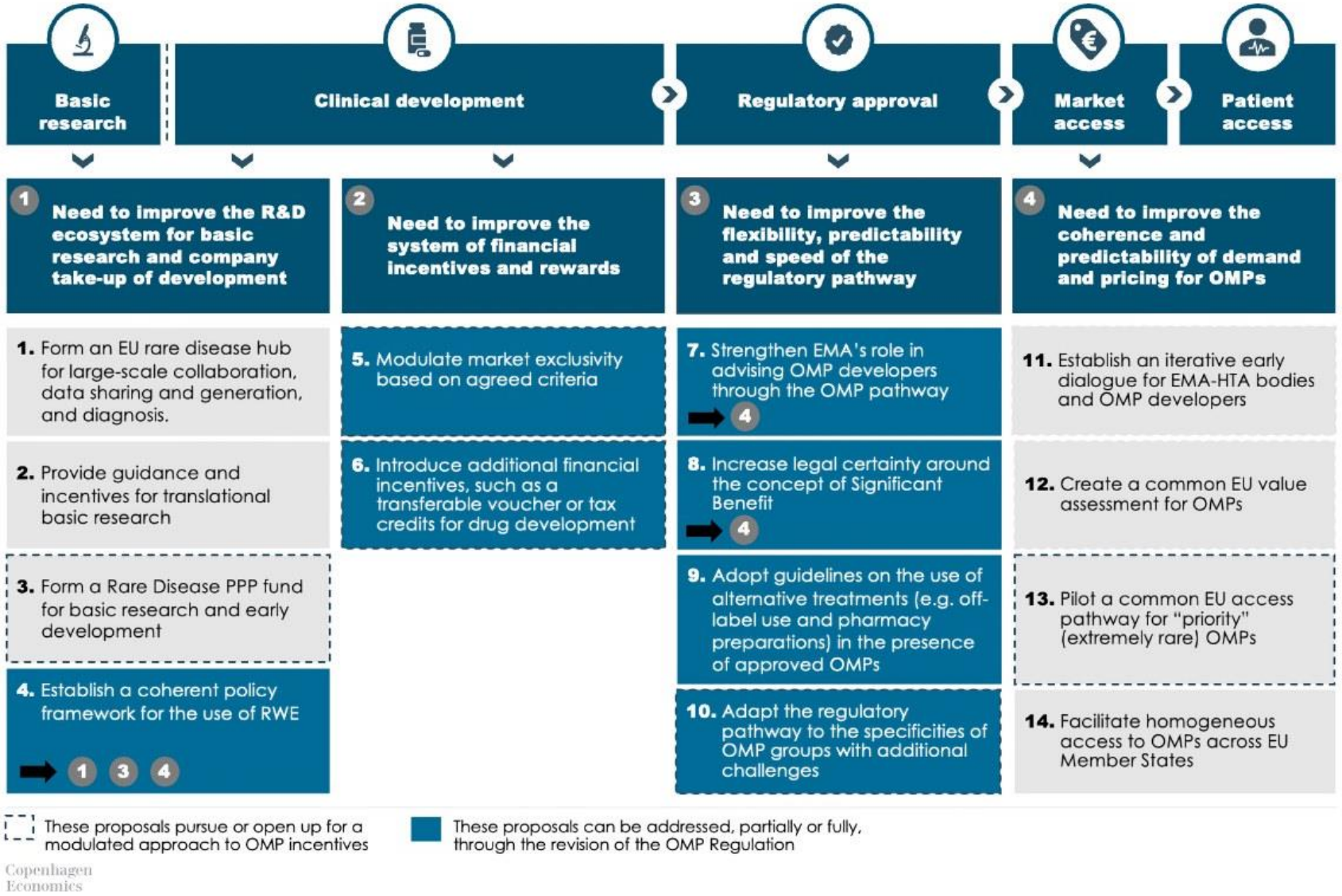
Diseases remaining without ODD since 1999



Diseases remaining without ODD since 1999 (in scale from 0)



OD expert group framework and recommendations



https://od-expertgroup.eu/

ODDG –
How do you start the development of your product?


MENU nature reviews drug discovery

Search E-alert Submit Login

COMMENT · 20 APRIL 2020

Boosting delivery of rare disease therapies: the IRDiRC Orphan Drug Development Guidebook

The International Rare Diseases Research Consortium (IRDiRC) has created a Guidebook to facilitate drug development for rare diseases by organizing available tools into a standardized framework.

Anneliene Hechtelt Jonker, Virginie Hivert, Michela Gabaldo, Liliana Batista, Daniel O'Connor, Annemieke Aartsma-Rus, Simon Day, Ken Sakushima & Diego Ardigo

S

T

A

R

T

STakeholders mapping

Available information on the disease

Financial **R**esources

Target Patient Value Profile