DATHARSIS

DATA SPEAKS, WE INTERPRET

# DATA ANALYSIS FOR HEALTH

Transforming complex data into efficient and personalized clinical solutions


www.datharsis.com

+34 657 28 85 85

info@datharsis .com

# TABLE OF CONTENTS

DATHARSIS
DATA SPEAKS, WE INTERPRET

DATHARSIS
DATA SPEAKS, WE INTERPRET

# 01 PRESENTATION

The generation of data is growing exponentially. However, the ability to integrate, clean, and rigorously interpret this information remains limited.

This deficit has **critical consequences** in research and clinical practice.

**Insufficient sample size** in clinical studies

**Multiple data sources** (records, imaging, omics) are difficult to integrate

**Poor data analysis practices** lead to miscalculated results

**Complex clinical studies** with multiple factors, covariates, comorbidities, missing values, outliers, and **inconclusive analyses**

Hundreds, thousands, or millions of potential **biomarkers** with complex and **difficult-to-determine** effects

In this context, Datharsis positions itself as a strategic ally, combining state-of-the-art **biostatistics** and **data science**, supported by cutting-edge international research, **to convert complex data into reliable, reproducible evidence with real impact** on clinical practice and pharmaceutical research.

# VALUE PROPOSITION

02

DATHARSIS
DATA SPEAKS, WE INTERPRET

**We turn complex data into solid evidence** so that researchers, hospitals, and pharmaceutical companies can make better decisions.

Our difference is based on three fundamental pillars:

**INTERPRETATION AND CAUSALITY**

We don't just tell you "what is happening," but we explain "why." This generates solid conclusions that can be applied directly to improve results and optimize processes.

**SCIENTIFIC RIGOR + INNOVATION**

We combine the rigor of biostatistics with cutting-edge data science techniques. This allows us to generate reliable and scalable explanatory and predictive models, guaranteeing the validity of each analysis.

**TOTAL SCALABILITY**

We work with a wide variety of data and cohorts of any size. We use optimized techniques to extract the maximum amount of information, which allows us to minimize the risk of publication rejection, accelerate R&D, and improve accuracy.

Many Artificial Intelligence solutions are based on massive and generic data, offering conclusions without explaining the "why". At Datharsis, **transparency** is our promise. We turn your own data into a **competitive advantage**, giving you findings with **clinical and strategic value**, supported by a methodology that allows you to understand the reason behind each result, **without the need to resort to artificial data**.

DATHARSIS
DATA SPEAKS, WE INTERPRET

# 03 TYPES OF DATA

At Datharsis, we understand that each project is unique, and the first step to success is understanding the available data.

We work with a wide range of information, and we are experts in **data fusion**. We combine multiple types of data to create a holistic and causal vision.

The types of data we work with include:

## Multi-omics data

Transcriptomics, metagenomics, metabolomics, proteomics, ...

To obtain a holistic and deeper vision of biological systems, identifying cause-and-effect relationships that could not be seen with a single type of data.

## Clinical and demographic data

Medical history, laboratory results, age, sex, ...

To contextualize molecular findings, correct models, and obtain the most complete possible vision of the patient's health status.

## Other types of data

Behavior, public health, medical images, ...

To go beyond molecular and demographic information and understand how the patient's habits, environment, and physical characteristics influence their health.

# AREAS OF APPLICATION

04

DATHARSIS
DATA SPEAKS, WE INTERPRET

At Datharsis, we offer solutions that combine **advanced biostatistics and data science with your clinical experience**. Our goal is to transform heterogeneous datasets into **applicable and reproducible evidence** that supports research, clinical practice, and pharmaceutical innovation.

## PRECISION MEDICINE

Patient stratification, biomarker discovery, and predictive models for diagnosis, treatment, and prevention

## CLINICAL STUDIES

Design and analysis of experiments to ensure scientific validity and guide health policies

## MULTI-OMICS

Analysis and integration of large multi-omics databases for research and clinical applications

## PHARMA/MEDTECH R&D

Optimization of drug and medical devices design

# AREAS OF APPLICATION

04.1

DATHARSIS
DATA SPEAKS, WE INTERPRET

## PRECISION MEDICINE

We support you in understanding the molecular and clinical particularities of each patient. Our consultancy focuses on **patient stratification** and the **identification of biomarkers** from clinical and multi-omics data. This allows us to generate predictive and interpretable models for diagnosis, treatment, and prevention.

**With Datharsis you will achieve**

**Well-defined patient subgroups**

This leads to more effective treatments and a notable reduction in therapeutic failures.

**Clear clinical evidence**

By generating robust and publishable results, we facilitate the financing and adoption of new solutions in the market.

**More robust diagnosis and prevention**

Our predictive models allow us to predict the response to treatments and the risk of diseases, improving clinical decision-making.

# AREAS OF APPLICATION

04.2

DATHARSIS
DATA SPEAKS, WE INTERPRET

We offer statistical consultancy to support the **design of clinical studies**, integrating your clinical and analytical experience with our expertise in maximizing the information in the data generated. In addition, we help you analyze the results to **evaluate the effectiveness of treatments and identify risk factors**, generating evidence for more robust and reliable decision-making.

CLINICAL STUDIES

## With Datharsis you will achieve:

### Optimized experimental design

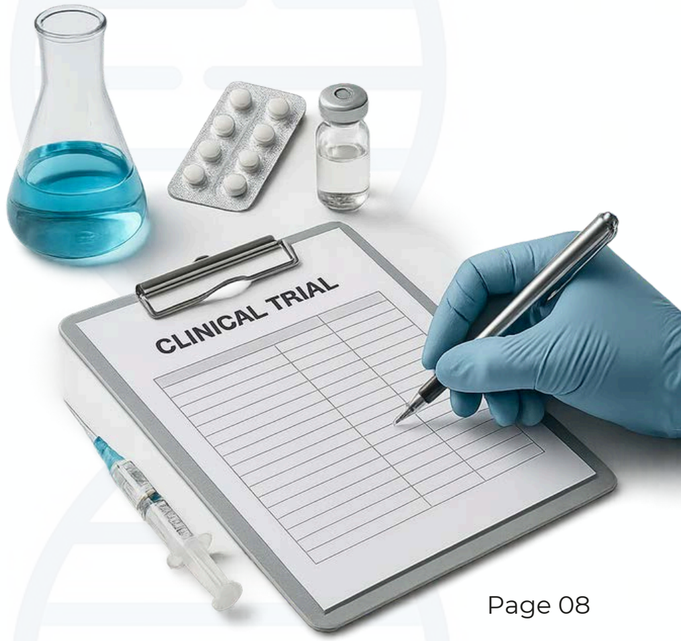
We ensure your study is robust from the start, reducing the risk of insufficient sample size and ensuring the integrity of the results.

### Rigorous data analysis

We transform complex trial data into clear clinical evidence, helping to determine the true effectiveness of treatments.

### Reliability in small cohorts

Our experience allows us to extract robust information even from cohorts with a small number of patients. This allows you to obtain reliable and solid conclusions, turning the sample limitation into a competitive advantage for your project.



# AREAS OF APPLICATION

04.3

DATHARSIS
DATA SPEAKS, WE INTERPRET

## MULTI-OMICS

We specialize in the analysis and integration of large multi-omics databases, such as genomics, proteomics, and metabolomics. We **identify molecular signatures associated with diseases**, providing key information for research and clinical applications.

**With Datharsis you will achieve:**

### Biomarker identification

Our biomarker analysis allows for more precise and personalized diagnoses, opening the door to new targeted therapies.

### Clinical applications with real impact

We convert multi-omics big data into findings that can be applied in clinical practice to improve patient care.

### Understanding of diseases

By integrating data at multiple levels, we reveal complex biological interactions, which allows you to better understand the pathology and accelerate the discovery of new therapies.

# AREAS OF APPLICATION

04.4

DATHARSIS
DATA SPEAKS, WE INTERPRET

Our consultancy **accelerates and optimizes the development cycle of drugs and medical devices**. We provide value from the identification of promising molecules and biomaterials to the design of clinical trials and validation studies using artificial intelligence.

**PHARMA/MEDTECH R&D**

**With Datharsis you will achieve:**

## Agility in research and development

By optimizing design and testing processes, we significantly reduce the time and resources needed to bring an innovation from the lab to the market (Time-to-Market), gaining agility and a competitive advantage.

## Safer, more effective, and scalable innovation

Our advanced models allow us to analyze mechanisms of action and device performance in digital environments (In-silico), selecting candidates with a higher probability of success and mitigating technical and clinical risks from the earliest phases.

## Optimization of clinical studies and validations

We design more efficient studies, reducing costs and execution times, while increasing data robustness to ensure regulatory approval and compliance.

# 05 SCOPE OF SERVICES

DATHARSIS
DATA SPEAKS, WE INTERPRET

At Datharsis, **we accompany you throughout the entire data lifecycle**, from experimental design, even before its generation, to its analysis, interpretation, and the delivery of results. This comprehensive approach allows us to guarantee rigor, precision, and quality in each phase, adapting to your specific needs.

| | PHASE | WHAT WE DO | WHAT YOU GET |
|---|---|---|---|
| 1 | **Experimental design** | Definition of key questions and sample size | Robust initial trials, with less risk of rejection and better use of resources |
| 2 | **Data curation** | Multi-source cleaning and integration | Reliable datasets, ready for analysis, and free from biases |
| 3 | **Exploratory analysis** | Visualization and descriptive statistics | Rapid understanding of patterns and anomalies, which guides advanced analysis |
| 4 | **Inference and causality** | Univariate and multivariate models | Solid and defensible conclusions |
| 5 | **Machine Learning and interpretable AI** | Predictive models and classification | Precise and transparent predictions, avoiding "black box" |
| 6 | **Delivery and publication** | Reports and writing support | Clear and publishable results that reinforce credibility and accelerate funding |

# 05.01 SCOPE OF SERVICES

DATHARSIS
DATA SPEAKS, WE INTERPRET

1 Experimental design

## Pre-clinical study experimental advisory

### The key to the success of your project

**Experimental design is the pillar of a successful data analysis.** We support you from the conception phase to define and validate the key questions of the analysis, ensure the validity of the results, and estimate the study's sample size with computational techniques and simulation.

- **Strategic planning:** We plan the study's strategy to guarantee the scientific rigor of the results, even in the earliest phases.
- **Question validation:** We define and validate that the key questions align correctly with the experimentation and analysis strategy to ensure you get the information you need.
- **Sample size estimation**: We use computational techniques to ensure an adequate sample size, optimizing resources and minimizing unnecessary costs.

# SCOPE OF SERVICES

DATHARSIS
DATA SPEAKS, WE INTERPRET

## Data cleaning and curation

### The foundation of a reliable analysis

2 Data curation

Once the data has been collected, cleaning and curation are crucial to **ensuring the quality and validity of the results**. Without a solid database, any subsequent analysis could lead to erroneous conclusions.

- **Data cleaning:** We are responsible for detecting and correcting errors, missing values, and duplicates, which are common in health sector data.
- **Format standardization:** We synchronize data from various sources to ensure they are consistent and compatible, facilitating their integration and analysis.
- **Data quality validation**: We verify the integrity and reliability of the information, giving you the peace of mind that your analysis is based on rigorous and high-quality data.

# SCOPE OF SERVICES

DATHARSIS
DATA SPEAKS, WE INTERPRET

3 Exploratory analysis

## Exploratory Data Analysis (EDA)

### The first step toward knowledge

EDA is the initial phase where **we discover the secrets hidden** in your data. It is a visual and statistical exploration that allows us to understand the structure, identify patterns, and detect anomalies before applying more complex analyses.

- **Data visualization:** We create personalized charts and tables that allow you to see and understand trends, relationships, and patterns intuitively.
- **Descriptive statistical analysis:** We summarize the main characteristics of your data, such as the mean, median, or standard deviation, to get a clear and concise overview.
- **Identification of relationships between variables:** We look for the interconnections between variables to understand how your data behaves. This serves as a roadmap for subsequent advanced analysis.

# SCOPE OF SERVICES

DATHARSIS
DATA SPEAKS, WE INTERPRET

## Statistical Inference

### The "why" behind the data

4 Inference and causality

Our specialty is going beyond correlation to establish **cause-and-effect relationships**. We provide you with solid conclusions that are applicable to your project through detailed inference adapted to the complexity of your data.

- **Hypothesis testing: univariate and multivariate**: This is the pillar of a reliable data analysis. We help you validate key findings and reduce the risk of erroneous conclusions.
  - **Univariate inference:** We use techniques like ANOVA based on permutation tests and with corrections like FDR and Q-value to compare groups and detect significant differences. This more traditional analysis is complementary to multivariate analysis..
  - **Multivariate inference:** It maximizes statistical power and allows for understanding complex biological relationships. We primarily use our experience in vASCA, our own technique with excellent statistical power and interpretability
- **Key visualization and interpretation techniques:** We provide you with clear results and intuitive graphics so you can interpret the findings. We rely on tools such as Manhattan Plot, Volcano Plot, and multivariate graphics.

We also offer other types of analysis, such as **linear regression** models to predict variables of interest, and **survival analysis** for studies that investigate the time to an event, such as disease recurrence.

# SCOPE OF SERVICES

DATHARSIS
DATA SPEAKS, WE INTERPRET

**5** Machine Learning and interpretable AI

## An accurate and reliable AI

### Decisions based on understanding

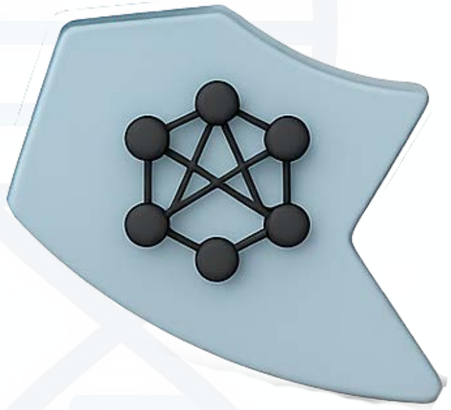
As an additional service, we can implement **Machine Learning and interpretable AI** models to complement your analyses. These models help you make safe decisions based on the data, since they not only give you a prediction but also explain the "why" behind each result so you can understand and trust the conclusions.

**Prediction:** Creation of models to predict the risk of disease or response to treatments.

**Classification:** Patient segmentation into subgroups for more personalized treatment.

**Anomaly detection:** Identification of unexpected patterns or atypical data points for in-depth analysis.

**Optimization:** Improvement of processes and operational efficiency based on data.



# SCOPE OF SERVICES

DATHARSIS
DATA SPEAKS, WE INTERPRET

## Delivery of results and publication support

6 Delivery and publication

We ensure that each project culminates in a clear and effective communication of the findings. We offer these services to complement the analysis or to help you in the final phase of your project.

We deliver a document with an executive summary, the key findings of the analysis, and personalized visualizations for a total understanding

**Drafting of the scientific-technical report**

We help you prepare the statistical analysis section in the materials and methods section of the article, create charts and result tables, and argue the methodological decisions based on scientific literature

**Support in writing the scientific article**

**Presentation meeting**

We meet with you to present the report, explain the results, and resolve any questions you may have

**Revisions and post-delivery advice**

If you need additional analysis to respond to revisions, we offer a package of hours for corrections and adjustments

# FLOWCHART

We have designed a clear, transparent, and results-oriented consultancy process. This guarantees that at each stage you know what to expect and what value you will obtain.

1. **Initial interview**: We meet to understand your needs, objectives, and the context of your project
2. **Project analysis and definition**: We prepare a report with an analysis proposal and define the key questions
3. **Proof of concept and data request**: If necessary, we validate our approach with a small analysis and request the necessary data
4. **Analysis execution**: We perform the advanced analysis of your data
5. **Report drafting and presentation**: We prepare a scientific-technical report and present the results to you and resolve all your doubts
6. **Revisions and post-delivery advice**: We accompany the revision phase with additional analysis and expert advice when you need it


07

# NEXT STEPS

In research and health, time and data quality are decisive. Every month lost in inefficient analysis or shaky conclusions delays innovation and increases costs.

Therefore, if you are looking for a partner that combines **data science, advanced biostatistics, and clinical experience**, let's talk now..

**Let's transform the future of health together**

## CONTACT US

DATHARSIS

DATA SPEAKS, WE INTERPRET

**www.datharsis.com**

**+34 657 28 85 85**

**info@datharsis .com**