Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
Università degli Studi di Napoli Federico II

FRAMEWORK TECHNICAL -SCIENTIFIC REPORT OF THE IMPLEMENTING ENTITY

PRIN PNRR 2022

“ASsessing glaucoma rAte of progression using artificial intelligence on metabolomics and extracellular VEsicles phenotYping in tEars: new insight to prevent visual disability and blindness”

RESPONSABILE: PROF. M.D. TORO

Codice identificativo: P2022ES8ME; CUP: E53D23015400001

DIPARTIMENTO DI SANITÀ PUBBLICA

UNIVERSITA' DEGLI STUDI DI NAPOLI FEDERICO II

NATIONAL RECOVERY AND RESILIENCE PLAN (NRRP) – MISSION 4 COMPONENT 2 INVESTMENT 1.1 – "Fund for the National Research Program and for Projects of National Interest (NRP)"

INTRODUCTION

The Technical-Scientific Report is produced by the *Principal Investigator*, on the basis of a discussion with each local manager for the purposes of collecting data and documenting the progress of each operational unit. It is then submitted according to deadlines envisaged in the related notification.

The Report is composed of the following sections:

SECTION 1 – GENERAL TRENDS OF THE PROJECT, where information on the general trends of the projects are included by making reference to the achievement of intermediate and long-term goals, to the compliance with the project activities, to the *DNSH* and *Open Access* principles as per the approved project.

SECTION 2 – PROGRESS OF ACTIVITIES, which entails the description of carried out activities based on the details provided by the operational units involved in the implementation of the project during the collaboration period. It also includes the description of future activities, potential publications and challenges encountered as well as actions for improvement. This, in order to guarantee the achievement of goals that have been established when the financed project was presented.

SECTION 3 – COMMON INDICATORS, where EU common indicators connected to the specific investment should be enhanced.

SECTION 4 – PREDICTIVE ANALYSIS AND FINAL COMMENTS, regards the forecast scenario on the development of the project and final comments.

SECTION 1 – GENERAL TRENDS OF THE PROJECT

With regard to the specific timeframe (bi-monthly/end of project activities), it is below provided:

a) ***a brief summary of the project***

Glaucoma is a chronic, progressive neurodegenerative disease of retinal ganglion cells and remains one of the leading causes of irreversible blindness worldwide. Despite advances in diagnostic technologies and clinical management, the molecular determinants underlying inter-individual variability in disease progression remain poorly understood, particularly in patients exhibiting a fast rate of progression (RoP).
In this context, the present project was designed to investigate novel molecular mechanisms associated with glaucoma progression by adopting an innovative and multidisciplinary approach based on the analysis of tear fluid as a non-invasive biological matrix. The project integrated three complementary domains: cytofluorimetric phenotyping of extracellular vesicles (EVs), untargeted metabolomics profiling through high-resolution mass spectrometry, and artificial intelligence-based data analysis.
The main objectives of the project were: (i) to identify specific biological signatures in tear samples of glaucoma patients with different rates of functional progression, and (ii) to develop predictive models capable of estimating disease progression and supporting personalized therapeutic strategies.
This approach enabled the generation of a multidimensional dataset combining clinical, molecular, and computational information, representing a novel framework for precision medicine in glaucoma.

b) ***names of the operational units involved in the implementation of the project.***

The project was conducted through the collaboration of two operational units:

- Unit 1: University of Studies “G. d’Annunzio” of Chieti-Pescara (coordinating unit)
- Unit 2: University of Studies “Federico II” of Naples

The coordinating unit was responsible for the overall management of the project, including clinical recruitment, experimental activities, and integration of data, while the collaborating unit contributed to patient recruitment, clinical characterization, and data sharing.

c) ***description of the achievement of the objectives connected to the project and related outcomes***

Throughout the duration of the project, all planned activities were carried out in accordance with the predefined timeline and milestones, without significant deviations from the approved work plan. A progressive recruitment strategy allowed the enrolment of a clinically well-characterized cohort of glaucoma patients, reaching a total of approximately 200 subjects at the end of the project. Patients were classified according to their rate of functional progression, enabling comparative analyses between slow and fast progressors.
In parallel with clinical recruitment, the cytofluorimetric characterization of extracellular vesicles in tear samples was progressively optimized. Specific marker panels were implemented to identify vesicles of different cellular origins, and gating strategies were refined to ensure reproducibility and robustness of the analysis.
The metabolomics component of the project underwent several phases of development and optimization. Sample preparation protocols, chromatographic separation, and mass spectrometry

acquisition parameters were carefully refined to ensure high analytical reproducibility and coverage of the tear metabolome. This approach enabled the identification of a large number of molecular features, leading to the annotation of 630 metabolites, of which approximately 260 were recognized by bioinformatic databases and included in downstream analyses.

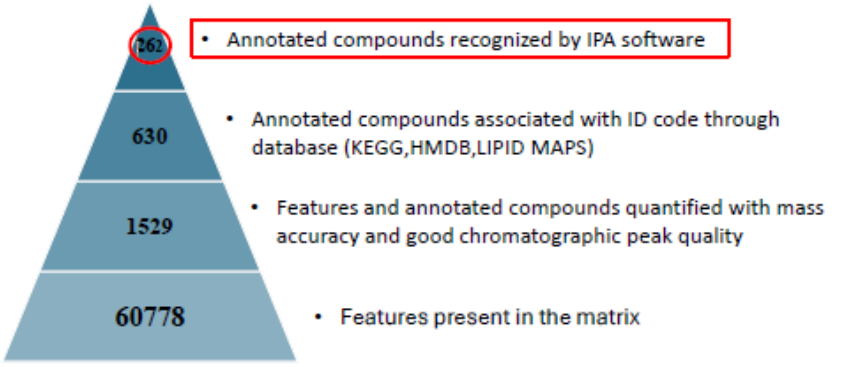

Figure 1. Annotated metabolites in the KEGG, HMDB and LIPID MAPS databases, but only 260 were recognized by the bioinformatic software IPA.

Subsequent statistical analyses allowed the identification of significantly modulated metabolites between patient groups. Differential expression analysis revealed the presence of upregulated metabolites in fast progressors, suggesting the activation of specific biological pathways. Functional enrichment analysis highlighted a prominent involvement of lipid metabolism, particularly arachidonic acid and prostaglandin pathways, suggesting the activation of compensatory and protective mechanisms associated with disease progression.
In addition to metabolomic findings, the analysis of extracellular vesicles provided further insights into the biological mechanisms associated with glaucoma progression. The cytofluorimetric characterization of EVs revealed distinct patterns related to different rates of progression, suggesting their potential role as mediators of intercellular communication and disease modulation.
The integration of EVs phenotyping with metabolomic data supports the hypothesis that patients with fast or dramatic progression exhibit the activation of compensatory and protective biological responses. These findings reinforce the concept that tear fluid represents a complex and dynamic biological environment, reflecting both pathogenic and adaptive processes occurring during glaucoma progression.

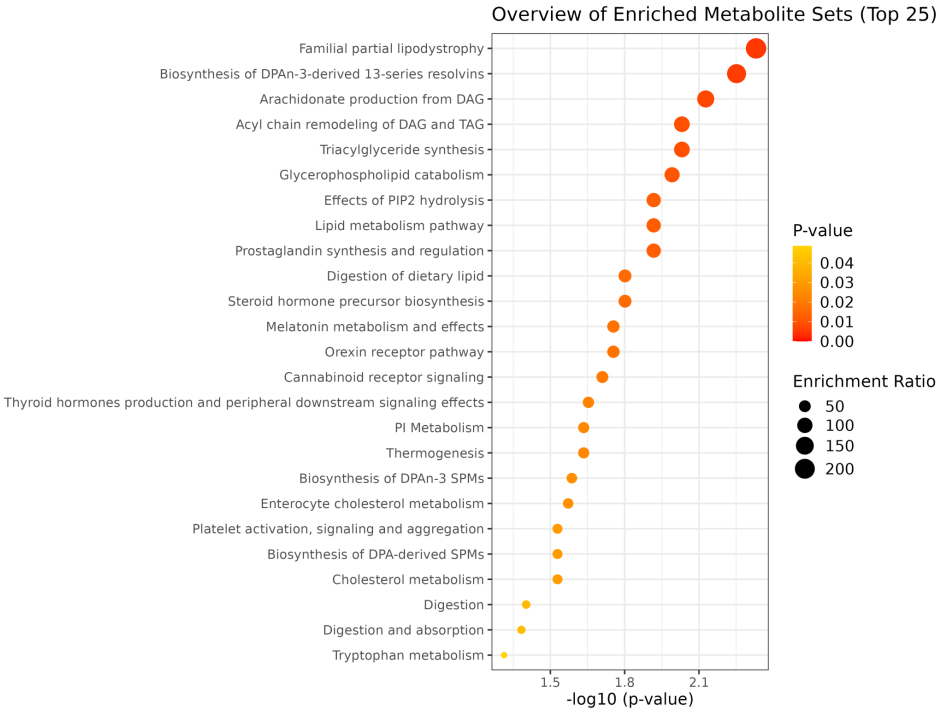

Figure 2. Bubble plot highlights the pathways involving metabolites quantified.

Further network-based analyses provided additional insights into the biological processes underlying glaucoma progression. In particular, the modulation of inflammatory mediators and lipid-derived signaling molecules was observed, supporting the hypothesis of a complex and dynamic biological response in fast-progressing patients.

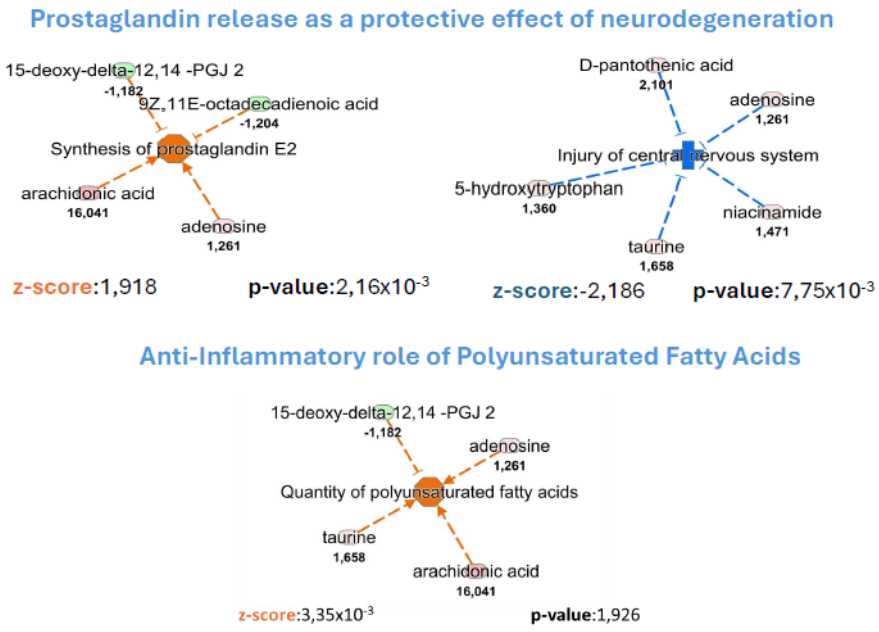

Figure 3. Networks showing a downstream effect in patient with ROP fast/dramatic compared

to patient with ROP slow

The biological interpretation of these findings suggests that the upregulation of lipid mediators, including prostaglandins and polyunsaturated fatty acids, may reflect an adaptive response aimed at counteracting neurodegenerative processes. In particular, the increased synthesis of prostaglandin E2 and the activation of anti-inflammatory pathways may contribute to intraocular pressure modulation and neuroprotection.
These results indicate that patients with fast progression may exhibit a potentially higher biological responsiveness to pharmacological treatment, opening new perspectives for personalized therapeutic strategies.
In the final phase of the project, all datasets were integrated into a unified computational framework. Clinical, cytometric and metabolomic variables were combined and analyzed using machine learning approaches. Feature selection techniques and classification models, including Random Forest algorithms, were applied to identify the most informative variables and to develop predictive models of disease progression.
Further analyses were conducted to assess the robustness and reproducibility of the machine learning approach. In particular, univariate feature selection (SelectKBest) was applied across different subsets of the dataset, confirming the stability of the most informative metabolomic features.
In addition, a preliminary hyperparameter tuning of the Random Forest classifier was performed, exploring variations in key parameters such as the number of estimators and maximum tree depth. These analyses resulted in a slight improvement in model stability, while performance metrics such as accuracy and AUC remained substantially unchanged.
Overall, these findings suggest that the predictive performance of the model is primarily influenced by the intrinsic biological variability of the samples rather than by methodological limitations, further supporting the validity of the proposed multi-omics approach.

Overall, the implemented models demonstrated moderate discriminative performance, reflecting both the presence of a biologically relevant signal and the intrinsic complexity of glaucoma pathophysiology.

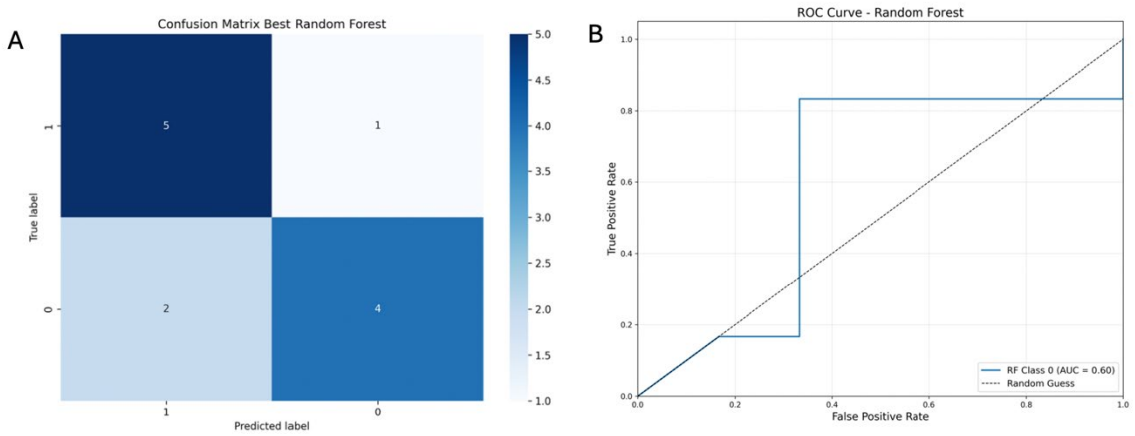

Figure 4. A: Confusion matrix of the best-performing Random Forest model after feature selection; B: ROC curve of the Random Forest model based on univariate feature selection.

Overall, the project achieved all planned objectives, generating a comprehensive dataset and providing novel insights into the molecular mechanisms underlying glaucoma progression.

d) ***description of the carried-out activities which are in compliance with the DNSH, Open Access principles as well as with gender, generational principles and with those of Equal opportunities***

All activities carried out within the project were conducted in full compliance with the "Do No Significant Harm" (DNSH) principle, in accordance with the requirements of the National Recovery and Resilience Plan (PNRR). The experimental procedures, including tear sampling, laboratory analyses, and computational processing, did not involve any environmental risk or impact. The project also adhered to Open Access principles, promoting the dissemination of scientific results through peer-reviewed publications and presentations at scientific meetings.
Furthermore, the project ensured compliance with gender equality, generational inclusion, and equal opportunity principles. Researchers at different career stages were actively involved in the project, fostering a collaborative and inclusive research environment.

e) ***description of the actions aimed at informing and disseminating knowledge***

During the course of the project, several activities aimed at disseminating scientific knowledge were undertaken. Preliminary results were presented at scientific meetings, congresses and workshops, contributing to the dissemination of the project outcomes within the scientific community. In addition, the data generated during the project led to the preparation of scientific manuscripts. In particular, two main lines of research are currently being developed: one focusing on metabolomics and extracellular vesicle profiling, and another dedicated to the development of predictive models based on artificial intelligence.
The results obtained provide a solid basis for future high-impact publications and for further research developments in the field of glaucoma and biomarker discovery.

Any relevant attached document which includes the abovementioned details

SECTION 2 – PROGRESS OF ACTIVITIES

With regard to the specific timeframe (bi-monthly/end of project activities), it is below provided:

a) ***detailed description of activities carried out by each operational unit with a focus on the timeframe for their implementation***

The activities carried out during the project followed a progressive and structured development, in line with the work plan and timeline defined in the original proposal. The project was based on a strong integration between clinical, experimental and computational components, coordinated by the leading unit and supported by the collaborating institution.
Patient recruitment represented a core activity throughout the entire duration of the project. In the initial phases, specific efforts were devoted to defining inclusion criteria and establishing a standardized classification system based on the rate of functional progression. Patients were stratified into subgroups according to their clinical evolution, initially including slow, fast and dramatic progressors. As the project progressed, and based on the observed distribution of cases, a refinement of this classification was introduced, with the merging of the dramatic subgroup into the fast progressors, allowing a more robust statistical framework.
The recruitment process was continuous across all project phases and led to the enrolment of approximately 200 glaucoma patients at the end of the study. Tear samples were systematically collected from all participants, ensuring consistency and reproducibility across the two operational units.
The cytofluorimetric analysis of extracellular vesicles was progressively implemented and optimized. In the early stages of the project, specific marker panels were developed to identify EVs derived from different cellular populations, including immune, neural, endothelial and mesenchymal cells. Gating strategies were refined over time to improve accuracy and reproducibility. In later phases, computational approaches were introduced to support automated data processing, including dimensionality reduction and clustering techniques, which enabled the identification of distinct EV subpopulations and their association with clinical phenotypes.
In parallel, the metabolomics component underwent several stages of development. Initial efforts focused on the optimization of sample preparation and extraction protocols, ensuring adequate metabolite coverage and minimizing variability. Chromatographic separation and mass spectrometry acquisition parameters were progressively refined to achieve high resolution and reproducibility. Quality control procedures confirmed the robustness of the analytical platform.
As the project advanced, untargeted metabolomics analysis was applied to the collected tear samples, leading to the identification and annotation of a large number of metabolites. Subsequent statistical analyses allowed the identification of significantly modulated compounds between patient groups. These results were further explored through bioinformatic tools, including pathway enrichment and network analysis, providing insights into the biological processes associated with disease progression.
In the final phase of the project, all datasets were integrated into a unified analytical framework. Clinical, cytofluorimetric and metabolomic variables were combined and analyzed using artificial intelligence approaches. Machine learning techniques, including feature selection methods and classification algorithms such as Random Forest, were applied to identify the most informative variables and to develop predictive models of glaucoma progression. These models demonstrated moderate but consistent performance, supporting the feasibility of multi-omics integration for disease stratification.

Overall, the activities carried out by both operational units were complementary and synergistic, contributing to the successful achievement of all planned objectives.
Further analyses were conducted to assess the robustness and reproducibility of the machine learning approach. In particular, univariate feature selection (SelectKBest) was applied across different subsets of the dataset, confirming the stability of the most informative metabolomic features.
In addition, a preliminary hyperparameter tuning of the Random Forest classifier was performed, exploring variations in key parameters such as the number of estimators and maximum tree depth. These analyses resulted in a slight improvement in model stability, while performance metrics such as accuracy and AUC remained substantially unchanged.
Overall, these findings suggest that the predictive performance of the model is primarily influenced by the intrinsic biological variability of the samples rather than by methodological limitations, further supporting the validity of the proposed multi-omics approach.

b) ***description of potential changes to what has been originally approved mentioning the impacts on the aim of the intervention, on the achievement of intermediate and long- term goals, on the proposed actions for improvement***

No substantial changes were introduced with respect to the originally approved research plan, and all activities were conducted in accordance with the predefined objectives and timeline.
A minor methodological adjustment was implemented during the course of the project concerning the classification of patients based on the rate of progression. Due to the relatively low number of cases presenting an extreme ("dramatic") progression pattern, this subgroup was merged with the fast progressors. This adjustment did not affect the overall aims of the project, nor did it compromise the achievement of intermediate and long-term goals. On the contrary, it allowed a more robust statistical analysis and improved the interpretability of the results.
No modifications were required in terms of experimental design, analytical methodologies, or overall project strategy. Similarly, no significant economic variations or deviations from the planned budget were observed.

c) ***description of potential challenges encountered and of the proposed actions for improvement***

The main challenge encountered during the implementation of the project was related to the recruitment of patients with very rapid disease progression. Such cases are relatively rare in clinical practice, and their limited availability represented a constraint for subgroup analyses. This limitation was addressed through a combination of strategies. Recruitment efforts were intensified across the clinical facilities involved in the project, and patient stratification criteria were adapted, as previously described, to ensure adequate group sizes for statistical analysis. Despite the slightly lower-than-expected number of extreme cases, the overall cohort remained sufficiently large and well characterized to support all planned analyses.
Another challenge was represented by the intrinsic complexity of integrating heterogeneous datasets, including clinical, cytometric and metabolomic data. This required iterative optimization of computational workflows and data processing pipelines. The introduction of machine learning approaches and feature selection techniques allowed effective management of this complexity, enabling the identification of meaningful patterns and predictive features.
Overall, the challenges encountered did not significantly impact the execution of the project or the validity of the results, and were effectively managed through adaptive methodological strategies.

d) ***brief description of potential publications.***

The results obtained during the project have already led to the preparation and dissemination of scientific outputs. Preliminary findings have been presented at national and international scientific meetings, contributing to the dissemination of the project outcomes within the research community.
In addition, the data generated during the project are currently being organized into scientific manuscripts. Two main lines of research are under development. The first focuses on the characterization of tear extracellular vesicles and metabolomics profiles in relation to glaucoma progression. The second is dedicated to the development of predictive models based on artificial intelligence and multi-omics data integration.
These studies are expected to result in peer-reviewed publications and represent a solid basis for future research developments and funding opportunities.

Any relevant attached document which includes the abovementioned details

SECTION 3 – COMMON INDICATORS

Below the updates on the indicator RRFCI 8 – "*Number of researchers who work in research centres which are recipients of financial support (women; men; non-binary)*" – as per the description in the guidelines included in the n.34 MEF notification from the 17th of October 2022.

Common indicators	Planned value	Implemented value
Researchers who work in research centers which are recipients of financial support (women)		
Researchers who work in research centers which are recipients of financial support (men)		
Researchers who work in research centers which are recipients of financial support (non-binary)		

SECTION 4 – PREDICTIVE ANALYSIS AND FINAL COMMENTS

Below it is provided a description of the forecast scenario on the development of the project, any potential change which is deemed necessary for the future as well as comments on the document.

1) ***Predictive analysis***

The results obtained during the project provide a solid foundation for future developments in both research and clinical applications. The integration of clinical data, extracellular vesicle phenotyping, and metabolomics profiles through artificial intelligence approaches has demonstrated the feasibility of a multi-omics framework for the analysis of glaucoma progression.
Future research activities will be primarily directed toward the validation of the identified molecular signatures in independent and larger patient cohorts, in order to confirm their robustness and generalizability. In particular, the metabolites and pathways identified in the present study, especially those related to lipid metabolism and inflammatory processes, represent promising candidates for further investigation as potential biomarkers of disease progression.
Another key direction will involve the refinement and optimization of the predictive models developed within the project. Although the current models demonstrated moderate discriminative performance, their accuracy can be improved by increasing sample size, enhancing feature selection strategies, and incorporating additional clinical and imaging variables. The integration of longitudinal data may further strengthen the predictive capability of these models.
From a translational perspective, the results of this project open the possibility of developing clinically applicable tools for early risk stratification of glaucoma patients. The identification of subjects at higher risk of rapid progression could enable a more personalized therapeutic approach, optimizing treatment strategies and improving long-term outcomes.
Furthermore, the use of tear fluid as a non-invasive source of molecular information represents a particularly promising avenue for future research. This approach may be extended to other ocular and systemic diseases, potentially contributing to the development of minimally invasive diagnostic platforms.
Overall, the predictive framework established in this project provides a valuable basis for future studies aimed at bridging the gap between molecular research and clinical practice in glaucoma. The integration of extracellular vesicles and metabolomic data represents a promising strategy to further improve predictive modelling in future studies.

2) ***Final comments***

The project was successfully completed in full accordance with the objectives, timeline, and methodological framework defined in the original proposal and in the Atto d'Obbligo. All planned activities were carried out without significant deviations, and the milestones established at the beginning of the project were achieved.
The integration of different methodological approaches, including metabolomics, extracellular vesicle analysis, and artificial intelligence, allowed the generation of a comprehensive and

multidimensional dataset. This integrative strategy provided novel insights into the biological mechanisms underlying glaucoma progression and highlighted the complexity of the disease.
The results obtained demonstrate that tear fluid represents a valuable and non-invasive source of biomarkers, capable of reflecting molecular alterations associated with disease progression. In particular, the identification of metabolomic signatures and the involvement of specific biological pathways, such as those related to lipid metabolism and inflammatory processes, support the hypothesis of an active and dynamic biological response in patients with different progression rates.
The development of predictive models further strengthens the relevance of the project, providing a proof-of-concept for the use of multi-omics data in patient stratification. Although further validation is required, these models represent a promising step toward the implementation of personalized medicine approaches in glaucoma.
In conclusion, the project has achieved its primary objectives and has generated results of significant scientific and potential clinical impact. The findings obtained lay the groundwork for future translational research and may contribute to the development of innovative diagnostic and prognostic tools aimed at improving the management of glaucoma patients.