2025-09-29

QRes - An Open-Source Dose–Response Data Platform for the R Community

Viera Kovacova
Viera Kovacova, Institute for Biological Physics, University of Cologne, Cologne, Germany

UNDER CONSTRUCTION

This repository is currently being revisited. An update will be released in early September 2025.

Executive Summary

Dose–response data are fundamental to research in microbiology, oncology, pharmacology, biophysics, and toxicology; yet, raw datasets remain fragmented, inconsistently formatted, and often lack metadata, making them challenging to integrate into reproducible R workflows. QRes platform is a web-based platform hosted at the University of Cologne that collects, processes, and analyzes dose–response datasets. This proposal seeks R Consortium support to expand QRes into a fully R-integrated, open-source resource for the global R community.

The project aims to deliver a public API and an R package designed for standardized data access, database storage, and data analysis. Additionally, it will provide an enhanced R Shiny interface that allows users to both upload and download raw data along with its associated metadata, as well as query and analyze this data. This interface will include an automated submission form for validating and storing new datasets, as well as tools for interactive visualization and analysis of dose–response measurements, all accessible directly through the browser.

QRes will integrate predictive tools developed in Python and Julia by Prof. Ville Mustonen (FiMAR, Helsinki) and Dr. Fernanda Pinheiro (Human Technopol, Milan), enabling cross-language interoperability. In addition, the DGrowthR package (developed by Medina Feldl at Helmholtz, Munich) will be incorporated as an analysis option for bulk dose–response data. Key outputs include an R package released on CRAN, standardized data formats, reproducible workflows, and community guidelines establishing best practices and gold standards for dose–response measurements. By bridging R with Python and Julia tools, Qres will provide a foundation for AI-driven analysis

and long-term reproducible science. This project directly addresses the R Consortium's mission by strengthening technical infrastructure, fostering interoperability, and serving a broad international user base across biology, biophysics, and medicine.

The Qres platform is an open-access dose–response database and web service designed to transform how raw microbial and cancer cell growth curve data are collected, curated, and analyzed. Measuring dose–response curves is a fundamental experiment in microbiology, oncology, biophysics, and pharmacology; yet, access to raw time-resolved data is limited, often buried in supplementary files and lacking standardized formats. This fragmentation hampers reproducibility, inhibits cross-study comparisons, and slows progress in understanding resistance dynamics and cellular responses under perturbation.

Qres addresses these challenges by providing a centralized, open, and R-integrated platform that: * Collects raw, time-resolved growth curves under antibiotic, chemical, genetic, or environmental perturbations. * Offers an automated Shiny-based submission form for uploading and validating raw measurements with curated metadata. * Implements standardized pipelines for extracting quantitative resistance traits and automating the fitting and benchmarking of growth curve models. * Provides a public API and R package to programmatically query datasets, returning either complete raw data or derived parameters. * Integrates curve fitting and predictive modeling tools developed in Python and Julia by Prof. Ville Mustonen (FiMAR, Helsinki) and Dr. Fernanda Pinheiro (Human Technopol, Milan), reimplemented or wrapped for use in R. * Connects with the DGrowthR package (Medina Feldl, Helmholtz, Munich) for analysis of microbial growth data.

Expected outcomes and deliverables include: * An R package released on CRAN with functions for standardized data access, preprocessing, and model fitting. * A robust R Shiny interface enabling data upload/download, metadata validation, and interactive visualization. * A curated, FAIR-compliant database with persistent identifiers for citation and cross-study reuse. * Benchmarked integration of cross-language predictive tools within R workflows. * Published best-practice guidelines and gold standards for microbial dose–response experiments.

Budget and Team: Funding is requested to support two PhD-level staff members for 12 months — one focusing on database/software development, and the other on R, Shiny, and API development. The PI (Viera Kováčová, University of Cologne) will supervise the project, but does not request salary support. Collaborations with FiMAR (Finland), Human Technopol (Italy), and Helmholtz (Germany) provide complementary expertise in predictive modeling and growth analysis.

By creating an open, standardized, and interactive infrastructure, Qres will accelerate reproducible research, establish community standards, and strengthen R's role as the primary platform for dose–response data analysis, serving a broad international user base across microbiology, oncology, pharmacology, and beyond.

Signatories

This section summarizes the individuals and groups who support, contribute to, or have been consulted on the Qres proposal. The project is already embedded in large research networks (CRC1310, CRC1678, FiMAR, Human Technopol, Helmholtz) and has broad community visibility.

Project team

- Viera Kováčová, PhD (viera.kovacova@uni-koeln.de) – Principal Investigator. Leads the development of Qres, supervises technical and scientific progress, and coordinates collaborations.

- Prof. Michael Lässig (lassig@uni-koeln.de) – Director of CRC1310. Provides strategic oversight and advice on the broader scientific vision.
- Stephan Kleinboltig (kleinboltig@uni-koeln.de) – CRC1310 project manager. Supports administrative coordination and R server infrastructure at the University of Cologne.

Two PhD Students (to be hired) – Dedicated to (1) software development and database/API implementation, and (2) R/Shiny/API development and user interface.

Contributors

- Prof. Ville Mustonen (FiMAR, Helsinki) – Provides predictive modeling expertise, helps shape the scientific scope, and engages international collaborators to contribute datasets.
- Dr. Fernanda Pinheiro (Human Technopol, Milan) – Leads development of the Kinbiont Julia package for bacterial growth modeling; collaborates on integration into Qres.
- Prof. Tobias Bollenbach (University of Cologne) – CRC1310 PI producing large-scale dose–response datasets; supports data contribution and benchmarking.

Consulted

University of Cologne colleagues, all contributing feedback and testing: - Theresa Finke - Leon Sieger - Gabriela Pertungaro - Rotem Gross

- C3RDM and ITCC departments, University of Cologne – Consulted on IT infrastructure, data management, and server support.
- Medina Feldl (Helmholtz, Munich) – Developer of the DGrowthR package; contributes to Qres integration for bacterial dose–response analysis.
- Prof. Andreas Beyer (andreas.beyer@uni-koeln.de) – Director of CRC1678. Advises on systems biology integration and data-driven approaches.
- Prof. Miroslav Baránek (Mendel University, Lednice, Czech Republic) – Provides perspective on dose–response measurements in plant pathogens.

Industry contacts: representatives from Agilent, Tecan, Molecular Devices, Promega, and BMGLabtech – provided test/sample files from microplate readers and laboratory robots to inform data ingestion pipelines.

Community Support

The project is known and endorsed within CRC1310 Area A (all PIs listed at crc1310.uni-koeln.de/project_a.html). This broad base of support ensures that Qres is well-positioned for adoption and sustainability, with immediate access to diverse microbial dose–response datasets.

The Problem

Dose–response measurements are among the most fundamental experiments in biology: they quantify how cells, microbes, and tissues respond to drugs, antibiotics, toxins, or environmental perturbations. These data are critical for understanding microbial resistance, cancer treatment response, and pharmacological efficacy. Despite their central role, the raw growth curve data underlying dose–response studies remain fragmented, inconsistently reported, and difficult to reuse.

Who it affects:

1. Experimental researchers face difficulties identifying optimal concentrations or conditions because raw data are buried in supplementary materials or inaccessible formats.
2. Modelers and computational biologists cannot easily benchmark or parameterize models due to lack of standardized, cross-study datasets.
3. The R community has strong modeling tools (e.g. drc, nlme, brms) but lacks a central resource of curated datasets with reproducible access pathways.

Why it is a problem:

1. Lack of accessible raw data prevents reproducibility and cross-study comparison.
2. Metadata are often incomplete, making it impossible to assess experimental context.
3. Absence of standards leads to inconsistent extraction of resistance traits and dose–response parameters.

What solving the problem enables:

1. Establishing Qres will provide the R community with a curated, FAIR-compliant infrastructure for dose–response data.
2. Researchers will be able to upload and share raw datasets with metadata, ensuring proper citation and recognition.
3. Modelers will have standardized, comparable inputs for parameter estimation, benchmarking, and AI-driven predictions.
4. R users will benefit from integrated pipelines linking raw data to statistical and mechanistic modeling.
5. For society, Qres will deepen our understanding of microbial resistance and its evolution, support better-informed strategies to combat antibiotic resistance, and provide benchmarks that improve the design of therapeutic interventions. By enabling systematic cross-study comparisons, Qres will help identify resistance patterns earlier, guide clinical and experimental decisions, and ultimately contribute to public health.

Existing work: 1. R packages like drc (Ritz et al. 2015), nlme (Pinheiro and Bates 2023), and brms (Bürkner 2017) support model fitting, but they assume users already have curated datasets. 2. DGrowthR (Helmholtz) (Feldl and Olayo Alarcon 2025) provides functions for analyzing microbial growth, but it does not provide access to standardized raw data repositories.

To date, there is no centralized, open-access platform in R that combines data storage, metadata standards, reproducible access, and interactive visualization. This gap limits reproducibility, slows the pace of discovery, and prevents the R community from fully supporting the next generation of data-driven microbiology and pharmacology.

The proposal

Approach and Work Plan

We will address the problem of fragmented and inaccessible dose–response datasets by developing Qres into an open, standardized, R-integrated platform with the following concrete actions: ** API and Database Development (Months 1–4) * Implement a REST API with standardized JSON schema. * Expand the database schema to store raw growth curves and curated metadata. * Establish

FAIR-compliant pipelines for data ingestion. ** R Package (Months 5–6) * Develop an open-source R package (QresR) to query the database and API. * Provide functions for data preprocessing, extraction of resistance traits, and integration with existing modeling packages (drc, nlme, brms). * Begin integration of DGrowthR for bulk growth curve analysis. ** Shiny Interface and Submission Form (Months 7–8) * Extend the existing Shiny prototype into a full interface supporting upload and download of raw growth data with metadata. * Implement automated submission validation and error reporting. * Add interactive visualization and curve-fitting modules. ** Cross-Language Integration (Months 9–10) * Reimplement or wrap selected Python/Julia predictive tools developed by collaborators (Mustonen, Pinheiro). * Benchmark results across platforms to ensure reproducibility. ** Testing, Standards, and Dissemination (Months 11–12) * Pilot test data submissions with internal and external collaborators. * Optimize submission form and analysis workflows. * Publish the R package on CRAN, complete documentation, tutorials, and case studies. * Release community guidelines establishing gold standards for dose–response experiments.

Risk Management

1.) Risk: Data heterogeneity -> Mitigation: enforce a standardized metadata schema and automated validation in the Shiny submission form. 2.) Risk: Integration complexity of Python/Julia tools -> Mitigation: begin with a small subset of methods and provide R wrappers before attempting full reimplementations. 3.) Risk: Limited adoption -> Mitigation: leverage CRC1310, FiMAR, and Human Technopol networks to ensure early use and dataset contributions; provide persistent identifiers for proper citation.

Timeline

Months 1–4: Database + API development. Months 5–6: R package prototype. Months 7–8: Shiny submission system. Months 9–10: Cross-language integration. Month 11: Testing and optimization. Month 12: Release, dissemination, and publication of guidelines.

Overview

The Qres platform is an open-access dose–response database and web service that collects, organizes, and analyzes raw, time-resolved microbial growth curves under antibiotic, genetic, and environmental perturbations. Currently, raw dose–response data are fragmented, inconsistently reported, and often hidden in supplementary files, making it difficult for researchers to reproduce analyses, perform cross-study comparisons, or build robust models of microbial resistance and cancer evolution.

Qres directly addresses this problem by providing the R community with a standardized, open-source infrastructure for dose–response data. The platform combines a robust database and public API, an R package (QresR) for reproducible workflows, and an extended Shiny interface that enables both upload and download of raw data with curated metadata, along with interactive visualization and analysis.

The project is developed in collaboration with international partners: Predictive tools from FiMAR (Helsinki) and Human Technopol (Milan) will be reimplemented or wrapped for integration into R. The DGrowthR package (Helmholtz, Munich) will be incorporated for bacterial growth curve analysis.

Benefits to the R community: * Access to curated, FAIR-compliant raw dose–response datasets. * Reproducible workflows for microbial growth, cancel cell growth and resistance modeling. *

Integration of cross-language predictive tools into R. * Community guidelines establishing gold standards for dose–response data.

By solving the accessibility and reproducibility gap in dose–response research, the Qres will accelerate discovery, improve experimental design, and expand R's role as the central platform for quantitative biology (for example, in the field of antibiotic resistance).

Detail

Minimum Viable Product

The smallest version of Qres that delivers value to the R community will include: 1. A database + REST API for storing and retrieving standardized dose–response datasets. 2. An R package (QresR) providing functions to query the API and return data in tidy formats for downstream analysis with packages like drc, nlme, and brms. 3. A basic Shiny interface that allows users to upload raw growth curve data with metadata, validates submissions, and enables download of standardized datasets.

This MVP ensures that from the beginning, R users can both contribute new datasets and reuse existing data in reproducible workflows.

Architecture

At a high level, the architecture will consist of: 1. Database Layer * Stores raw growth measurements, curated metadata, and derived resistance traits. * Implements FAIR-compliant schemas with persistent identifiers for citation. 2. API Layer * RESTful API (JSON-based) for standardized programmatic access. * Enables both raw data retrieval and derived parameter queries. 3. R Package (QresR) * Provides user-facing functions to query the API, clean and preprocess data, and connect to modeling packages. * Interfaces with external analysis tools such as DGrowthR. 4. Web Interface (Shiny) * Supports upload of raw data + metadata via automated submission form. * Performs validation, error reporting, and visualization of fitted curves. * Allows users to explore and download datasets interactively. 5. External Tool Integration * Wrappers or reimplementations for Python/Julia tools (Mustonen, Pinheiro). * Optional DGrowthR workflows for bulk analysis.

Assumptions

The project assumes that: 1. Community willingness: Researchers will contribute raw dose–response datasets when given easy submission tools and proper citation credit (via persistent identifiers). 2. Technical feasibility: Core predictive tools from Python/Julia can be either wrapped or reimplemented in R without loss of fidelity. 3. Resource stability: Hosting on University of Cologne servers remains available for the duration of the project, with minimal cost. 4. Sufficient adoption: By leveraging CRC1310, FiMAR, and Human Technopol networks, Qres will gain early adopters who validate and stress-test the system.

If any of these assumptions were false (e.g. very low data submission, or inability to wrap external tools), Qres would still deliver value as a curated database + API + R package, but the broader interoperability and uptake could be delayed.

External dependencies

Qres builds on several external components: 1. R ecosystem: Packages including drc, nlme, brms, and shiny for modeling and visualization. 2. DGrowthR (Helmholtz, Munich): External R package for analysis, integrated into Qres workflows. 3. Python/Julia predictive tools: Developed by collaborators at FiMAR (Helsinki) and Human Technopol (Milan); require wrapping or partial reimplementation in R. 4. Server infrastructure: Hosted on University of Cologne's R Shiny server,

with additional database backend. 5. Community datasets: Contributions from CRC1310, FiMAR, Human Technopol, and beyond are essential to seed the database.

Project plan

Start-up phase

Collaboration platform: We will host all code on GitHub under an open repository (qres-platform), including the API, Shiny interface, and R package. Issues and pull requests will be used to manage contributions from collaborators and the community.

License decisions: All code will be released under the MIT license; data schemas and metadata standards will be released under CC-BY 4.0 to maximize reuse.

Reporting framework: We will provide quarterly progress updates to the ISC, including blog posts summarizing milestones. The GitHub repository will include a project board to track deliverables.

Technical delivery

Months 1–2: Finalize schema design, API blueprint, GitHub repo setup, and automated CI/CD pipeline for the R package.

Months 3–4: Implement database backend and API endpoints; first internal prototype of data upload form.

Months 5–6: Release alpha version of QresR package (GitHub); API documentation available online.

Months 7–8: Extend Shiny interface with validated data submission and visualization modules; first external test submissions.

Months 9–10: Integrate predictive tools (Python/Julia) and DGrowthR; provide R wrappers.

Months 11–12: Community testing with CRC1310/FiMAR/Helmholtz datasets; finalize CRAN release of QresR; publish white paper/guidelines.

Other aspects

Licensing: Code under MIT; data standards and metadata under CC-BY 4.0.

Hosting: Public GitHub repository with issue tracking and contribution guidelines; Shiny app hosted on University of Cologne's R Shiny server.

Dissemination: 1. Announcement blog post on R Consortium blog at project start. 2. Quarterly blog updates on progress. 3. Delivery blog post on release of QresR (CRAN). 4. Presentations at UseR! 2026, ISC meetings, and relevant microbiology conferences. 5. Social media announcements (via CRC1310, FiMAR, and partner institutes).

Budget & funding plan

Funding is requested primarily for labor costs: Two PhD students (12 months each) 1. PhD student 1 (Software Development & Data Engineering): Responsible for database, API, and storage pipelines. 2. PhD student 2 (R/Shiny/API Development): Responsible for R package, Shiny interface, and automated data submission. 3. The PI (Viera Kováčová, University of Cologne) will supervise the project but does not request salary support. 4. No funds are requested for travel, publication fees, or hardware. University of Cologne provides server infrastructure.

Role	FTE	Duration	Cost per year (€)	Total (€)
PhD Student 1 – Software Development & Data Engineering	1.0	12 months	76,500	76,500
PhD Student 2 – R/Shiny/API Development	1.0	12 months	76,500	76,500
Total Personnel	2.0	12 months	–	153,000

Milestones tied to funding release: Month 4: API + database schema complete → €30,600 Month 6: Alpha R package + API documentation → €30,600 Month 8: Shiny submission system + visualization → €30,600 Month 10: Integration with external tools + DGrowthR → €30,600 Month 12: CRAN release, guidelines, dissemination → €30,600

Budget Justification

The requested funding is dedicated entirely to two full-time PhD student positions for 12 months. One PhD student will focus on software development and data engineering (database design, storage pipelines, REST API), while the second will focus on R/Shiny and API development (R package, submission interface, visualization). This division of labor is essential to ensure rapid and parallel progress across the backend and user-facing components of Qres. The PI (Viera Kováčová) will supervise the project but does not request salary support, ensuring that 100% of ISC funds are invested directly into infrastructure development for the R community.

Success

Definition of done

The project will be considered complete and successful when: 1. The Qres database and REST API are fully operational, allowing storage and retrieval of dose–response datasets with curated metadata. 2. The QresR package is publicly released on CRAN and GitHub, enabling R users to programmatically query, preprocess, and analyze datasets. 3. The Shiny interface supports automated data submission (with validation) and interactive download/visualization of raw data and fitted models. 4. Integration with at least one external predictive tool (Python/Julia) and DGrowthR is demonstrated through vignettes and tutorials. 5. Community guidelines / white paper on best practices for dose–response experiments are published and disseminated.

Measuring success

Success will be measured using concrete, trackable outputs: 1. Technical outputs: * QresR package accepted on CRAN. * API documentation publicly available. * Shiny interface deployed with submission/download capability.

2. Community uptake:

- At least 5 external datasets submitted and validated through the automated submission system.

- Engagement of early adopters from CRC1310, FiMAR, Human Technopol, and Helmholtz.

3. Dissemination:

- Release of at least 2 vignettes/tutorials demonstrating workflows in R.
- Publication of guidelines for dose–response data management and modeling.

Future work

Once Qres reaches full functionality, we will prepare a high-impact publication describing its architecture, features, and scientific use cases. The manuscript will highlight: * The need for standardized, FAIR-compliant dose–response data. * Qres infrastructure (database, API, R package, Shiny interface). * Integration of predictive tools across R, Python, and Julia. * Demonstration datasets from microbial resistance and cancer evolution. * Benchmarks comparing automated analysis pipelines (e.g. DGrowthR, Kinbiont).

Target journals could include Nature Biotechnology, Nature Communications, Nucleic Acids Research (Database issue), or Bioinformatics.

This publication will provide a citable reference for Qres, ensuring that dataset contributors receive recognition and that the platform is embedded in the global research ecosystem.

Bürkner, Paul-Christian. 2017. "Brms: An r Package for Bayesian Multilevel Models Using Stan." *Journal of Statistical Software* 80 (1): 1–28.

Feldl, Medina, and Roberto Olayo Alarcon. 2025. *DGrowthR: Bacterial Growth Curve Analysis*. https://bio-datascience.github.io/DGrowthR/.

Pinheiro, José, and Douglas Bates. 2023. *Mixed-Effects Models in s and s-PLUS*. Springer.

Ritz, Christian, Florent Baty, Jens C Streibig, and Daniel Gerhard. 2015. "Dose-Response Analysis Using r." *PLOS ONE* 10 (12): e0146021.