

## Principal Investigator

**First Name:** Johann

**Last Name:** Proeve

**Degree:** PhD

**Primary Affiliation:** Cyntegrity

**E-mail:** [ting-yu.ho@cyntegrity.com](mailto:ting-yu.ho@cyntegrity.com)

**State or Province:** Hessen

**Country:** Germany

## General Information

### Key Personnel (other than PI):

**First Name:** Anastasia

**Last name:** Shapovalova

**Degree:** Optic - Electronic Engineer (Medical devices)

**Primary Affiliation:** Cyntegrity Germany GmbH

**SCOPUS ID:**

**Requires Data Access?** Yes

**First Name:** Ting-Yu

**Last name:** Ho

**Degree:** MA/MS/MSc

**Primary Affiliation:** Cyntegrity Germany GmbH

**SCOPUS ID:**

**Requires Data Access?** No

**First Name:** Kateryna

**Last name:** Bolshakova

**Degree:** MD

**Primary Affiliation:** Cyntegrity Germany GmbH

**SCOPUS ID:**

**Requires Data Access?** No

**First Name:** Stefan

**Last name:** Hiess

**Degree:** Diplom Informatik

**Primary Affiliation:** Cyntegrity Germany GmbH

**SCOPUS ID:**

**Requires Data Access?** No

**First Name:** Boris

**Last name:** Kaganov

**Degree:** Specialist

**Primary Affiliation:** Cyntegrity Germany GmbH

**SCOPUS ID:**

**Requires Data Access?** No

**First Name:** Artem

**Last name:** Andrianov

**Degree:** PhD

**Primary Affiliation:** Cyntegrity Germany GmbH

**SCOPUS ID:**

**Requires Data Access?** No

**Are external grants or funds being used to support this research?:** No external grants or funds are being used to support this research.

**How did you learn about the YODA Project?:** Data Holder (Company)

## Conflict of Interest

<https://yoda.yale.edu/wp-content/uploads/2024/02/YODA-TingYu.pdf>

[https://yoda.yale.edu/wp-content/uploads/2024/01/YODA\\_Stefan-Hiess.pdf](https://yoda.yale.edu/wp-content/uploads/2024/01/YODA_Stefan-Hiess.pdf)

[https://yoda.yale.edu/wp-content/uploads/2024/02/YODA\\_Ana.pdf](https://yoda.yale.edu/wp-content/uploads/2024/02/YODA_Ana.pdf)

[https://yoda.yale.edu/wp-content/uploads/2024/01/SV\\_57KskaKADT3U9Aq-R\\_8aXmE8Vuza7iiMR.pdf](https://yoda.yale.edu/wp-content/uploads/2024/01/SV_57KskaKADT3U9Aq-R_8aXmE8Vuza7iiMR.pdf)

<https://yoda.yale.edu/wp-content/uploads/2024/01/1.pdf>

[https://yoda.yale.edu/wp-content/uploads/2024/01/SV\\_57KskaKADT3U9Aq-R\\_6mcuFb9E1LMJPOB.pdf](https://yoda.yale.edu/wp-content/uploads/2024/01/SV_57KskaKADT3U9Aq-R_6mcuFb9E1LMJPOB.pdf)

[https://yoda.yale.edu/wp-content/uploads/2024/02/YODA\\_Artem.pdf](https://yoda.yale.edu/wp-content/uploads/2024/02/YODA_Artem.pdf)

## Certification

**Certification:** All information is complete; I (PI) am responsible for the research; data will not be used to support litigious/commercial aims.

**Data Use Agreement Training:** As the Principal Investigator of this study, I certify that I have completed the YODA Project Data Use Agreement Training

1. [NCT02493868 - A Randomized, Double-blind, Multicenter, Active-Controlled Study of Intranasal Esketamine Plus an Oral Antidepressant for Relapse Prevention in Treatment-resistant Depression](#)
2. [NCT00095134 - A Double-Blind Study Comparing Adjunctive Risperidone Versus Placebo in Major Depressive Disorder That Is not Responding to Standard Therapy](#)
3. [NCT02133001 - A Double-blind, Randomized, Placebo Controlled Study to Evaluate the Efficacy and Safety of Intranasal Esketamine for the Rapid Reduction of the Symptoms of Major Depressive Disorder, Including Suicidal Ideation, in Subjects Who Are Assessed to be at Imminent Risk for Suicide](#)
4. [NCT02264574 - A Randomized, Multi-center, Open-label, Phase 3 Study of the Bruton's Tyrosine Kinase Inhibitor Ibrutinib in Combination With Obinutuzumab Versus Chlorambucil in Combination With Obinutuzumab in Subjects With Treatment-naïve Chronic Lymphocytic Leukemia or Small Lymphocytic Lymphoma](#)
5. [NCT01578707 - A Randomized, Multicenter, Open-label, Phase 3 Study of the Bruton's Tyrosine Kinase \(BTK\) Inhibitor Ibrutinib \(PCI-32765\) Versus Ofatumumab in Patients With Relapsed or Refractory Chronic Lymphocytic Leukemia/Small Lymphocytic Lymphoma](#)
6. [NCT02489318 - A Phase 3 Randomized, Placebo-controlled, Double-blind Study of Apalutamide Plus Androgen Deprivation Therapy \(ADT\) Versus ADT in Subjects With Metastatic Hormone-sensitive Prostate Cancer \(mHSPC\)](#)
7. [NCT01946204 - A Multicenter, Randomized, Double-Blind, Placebo-Controlled, Phase III Study of ARN-509 in Men With Non-Metastatic \(M0\) Castration-Resistant Prostate Cancer](#)
8. [NCT01106014 - A Multicenter, Double-blind, Placebo-controlled Phase 3 Study Assessing the Safety and Efficacy of Selexipag on Morbidity and Mortality in Patients With Pulmonary Arterial Hypertension](#)
9. [NCT02207231 - Phase 3, Multicenter, Randomized, Double-blind, Placebo and Active Comparator-controlled Study Evaluating the Efficacy and Safety of Guselkumab in the Treatment of Subjects With Moderate to Severe Plaque-type Psoriasis](#)
10. [91-113 - Efficacy and pharmacokinetic/pharmacodynamic profile of ibuprofen chewable tablets versus ibuprofen suspension in febrile children. CSR 167S. Protocol 91-113, Unpublished Report 293A and Unpublished Report 1475.](#)
11. [NCT00034736 - A Multicenter, Randomized, Open-Label, Comparative Study to Compare the](#)

- [Efficacy and Safety of Levofloxacin and Standard of Care Therapy in the Treatment of Children With Community-Acquired Pneumonia in the Hospitalized or Outpatient Setting](#)
12. [NCT01474122 - Prospective, Randomized, Placebo-controlled, Double-blind, Multicenter, Parallel Group Study to Assess the Efficacy, Safety and Tolerability of Macitentan in Patients With Ischemic Digital Ulcers Associated With Systemic Sclerosis](#)
  13. [NCT02269917 - A Phase 3, Randomized, Active-controlled, Open-label Study to Evaluate the Efficacy, Safety and Tolerability of Switching to a Darunavir/ Cobicistat/ Emtricitabine/ Tenofovir Alafenamide \(D/C/F/TAF\) Once-daily Single-tablet Regimen Versus Continuing the Current Regimen Consisting of a Boosted Protease Inhibitor \(bPI\) Combined With Emtricitabine/Tenofovir Disoproxil Fumarate \(FTC/TDF\) in Virologically-suppressed, Human Immunodeficiency Virus Type 1 \(HIV-1\) Infected Subjects](#)
  14. [NCT02431247 - A Phase 3, Randomized, Active-controlled, Double-blind Study to Evaluate Efficacy and Safety of Darunavir/Cobicistat/Emtricitabine/Tenofovir Alafenamide \(D/C/F/TAF\) Once Daily Fixed Dose Combination Regimen Versus a Regimen Consisting of Darunavir/Cobicistat Fixed Dose Combination Coadministered With Emtricitabine/Tenofovir Disoproxil Fumarate Fixed Dose Combination in Antiretroviral Treatment-naive Human Immunodeficiency Virus Type 1 Infected Subjects](#)
  15. [NCT02269917 - A Phase 3, Randomized, Active-controlled, Open-label Study to Evaluate the Efficacy, Safety and Tolerability of Switching to a Darunavir/ Cobicistat/ Emtricitabine/ Tenofovir Alafenamide \(D/C/F/TAF\) Once-daily Single-tablet Regimen Versus Continuing the Current Regimen Consisting of a Boosted Protease Inhibitor \(bPI\) Combined With Emtricitabine/Tenofovir Disoproxil Fumarate \(FTC/TDF\) in Virologically-suppressed, Human Immunodeficiency Virus Type 1 \(HIV-1\) Infected Subjects](#)
  16. [NCT02431247 - A Phase 3, Randomized, Active-controlled, Double-blind Study to Evaluate Efficacy and Safety of Darunavir/Cobicistat/Emtricitabine/Tenofovir Alafenamide \(D/C/F/TAF\) Once Daily Fixed Dose Combination Regimen Versus a Regimen Consisting of Darunavir/Cobicistat Fixed Dose Combination Coadministered With Emtricitabine/Tenofovir Disoproxil Fumarate Fixed Dose Combination in Antiretroviral Treatment-naive Human Immunodeficiency Virus Type 1 Infected Subjects](#)
  17. [NCT01228383 - A Randomized, Single-blind, Active-controlled, Mono-center Phase II Study to Compare the Safety and Neutralizing Activity of Simulated Rabies Post-exposure Prophylaxis With CL184 in Combination With Purified Vero Cell Rabies Vaccine vs. Human Rabies Immune Globulin or Placebo in Combination With Purified Vero Cell Rabies Vaccine vs. CL184 or Placebo in Combination With Human Diploid Cell Rabies Vaccine in Healthy Adult Subjects](#)
  18. [NCT00656097 - A Randomized, Single-blind, Controlled, Monocentric Phase II Trial to Compare the Safety and Neutralizing Activity of Simulated Rabies Post-exposure Prophylaxis With CL184 in Combination With Rabies Vaccine vs. HRIG or Placebo in Combination With Rabies Vaccine in Healthy Adult Subjects](#)
  19. [NCT02019472 - A Multicenter, Randomized, Double-blind, Parallel Group Study of Sirukumab Monotherapy Compared With HUMIRA® Monotherapy Administered Subcutaneously, in Subjects With Active Rheumatoid Arthritis](#)
  20. [NCT02438787 - A Phase 3, Multicenter, Randomized, Double-blind, Placebo-controlled Study Evaluating the Efficacy and Safety of Ustekinumab in the Treatment of Anti-TNF\(Alpha\) Refractory Subjects With Active Radiographic Axial Spondyloarthritis](#)
  21. [NCT02437162 - A Phase 3, Multicenter, Randomized, Double-blind, Placebo-controlled Study Evaluating the Efficacy and Safety of Ustekinumab in the Treatment of Anti-TNF Alpha Naive Subjects With Active Radiographic Axial Spondyloarthritis](#)
  22. [NCT01004432 - Golimumab in Rheumatoid Arthritis Participants With an Inadequate Response to Etanercept \(ENBREL\) or Adalimumab \(HUMIRA\)](#)
  23. [NCT02034162 - A Double-Blind, Randomized, Multi-Center, Parallel-Group, Placebo-Controlled Study to Evaluate the Efficacy and Safety of a Single Dose of a 500-mg Chewable Tablet of Mebendazole in the Treatment of Soil-Transmitted Helminth Infections \(Ascaris Lumbricoides and Trichuris Trichiura\) in Pediatric Subjects](#)
  24. [NCT01173562 - An Open-Label, Single-Dose Study to Assess the Safety of 500-mg Mebendazole Chewable Formulation in Children 2 to 10 Years of Age, Inclusive](#)
  25. [NCT00207740 - A Phase 2, Multicenter, Randomized, Double-blind, Placebo-controlled, Parallel-group, Dose-ranging Study Evaluating the Efficacy and Safety of CNTO 148 Administered Subcutaneously in Symptomatic Subjects With Severe Persistent Asthma](#)

26. [NCT02407236 - A Phase 3, Randomized, Double-blind, Placebo-controlled, Parallel-group, Multicenter Protocol to Evaluate the Safety and Efficacy of Ustekinumab Induction and Maintenance Therapy in Subjects With Moderately to Severely Active Ulcerative Colitis](#)
27. [NCT00488631 - A Phase 3 Multicenter, Randomized, Placebo-controlled, Double-blind Study to Evaluate the Safety and Efficacy of Golimumab Maintenance Therapy, Administered Subcutaneously, in Subjects With Moderately to Severely Active Ulcerative Colitis](#)
28. [NCT00487539 - A Phase 2/3 Multicenter, Randomized, Placebo-controlled, Double blind Study to Evaluate the Safety and Efficacy of Golimumab Induction Therapy, Administered Subcutaneously, in Subjects with Moderately to Severely Active Ulcerative Colitis](#)
29. [NCT00236665 - A Multicenter, Randomized, Double-Blind, Placebo-Controlled, Parallel Group Study of the Efficacy and Safety of Topiramate in the Treatment of Obese Patients With Mild to Moderate Essential Hypertension](#)
30. [NCT00816166 - Phase III Study of Pharos Vitesse Neurovascular Stent System Compared to Best Medical Therapy for the Treatment of Ischemic Disease](#)
31. [NCT02065791 - A Randomized, Double-blind, Event-driven, Placebo-controlled, Multicenter Study of the Effects of Canagliflozin on Renal and Cardiovascular Outcomes in Subjects With Type 2 Diabetes Mellitus and Diabetic Nephropathy](#)
32. [NCT01032629 - A Randomized, Multicenter, Double-Blind, Parallel, Placebo-Controlled Study of the Effects of JNJ-28431754 on Cardiovascular Outcomes in Adult Subjects With Type 2 Diabetes Mellitus](#)

**What type of data are you looking for?:** Individual Participant-Level Data, which includes Full CSR and all supporting documentation

## Research Proposal

### Project Title

Quality tolerance limit and duplicated patient investigation in clinical trials

#### Narrative Summary:

The upcoming research project at Cyntegrity focuses on improving the quality control of clinical trials, which are essential for developing new medical treatments. We want to make sure that the data collected in these trials is reliable and accurate. Our study has three main goals:

1. Quality Tolerance Limits (following : ICH E6 R2 5.0.4. Risk control , 5.0.7 Risk reporting; ICH E6 R3 (draft) 3.10.1.6. Risk Reporting ) Threshold (Mean, Upper, Lower limits) Benchmarking: We will carefully examine and compare quality tolerance limits (indicators crucial for assessment of clinical studies risks and/or execution) across different therapeutic areas, phases of clinical trials, and types of trials. By understanding these limits better, we aim to improve the methods used to control and ensure the quality of clinical trial data.

Cyntegrity is going to investigate the QTLs below: please check Table 1 in the attached document.

2. Data Behavior and Statistical Distribution: We will analyze the patterns in clinical trial data to understand how it behaves statistically. This will help us develop better quality control methods that are specific to the unique aspects of clinical trials.

3. Duplication Probability in Patient Data: We will investigate the likelihood of having duplicate patient records in the data collected during clinical trials. Using advanced statistical techniques, we aim to identify and address issues related to data integrity, participant safety, and data management processes. Additionally, we will explore and evaluate different statistical algorithms to efficiently detect duplicate patient data, further improving data quality assurance measures.

Our overall goal is to revolutionize how clinical trial data is managed, providing valuable insights and solutions to researchers, clinicians, health authorities, and other stakeholders. By doing so, we hope

to enhance the quality and reliability of outcomes in clinical trials, ultimately contributing to advancements in medical research and public health.

### **Scientific Abstract:**

#### **Background:**

Clinical trials represent a cornerstone in advancing medical knowledge and healthcare interventions. As the volume and complexity of clinical trial data continue to grow, there is a critical need to understand and optimize the quality, behavior, and integrity of this data. The proposed research project by Cyntegrity aims to comprehensively investigate quality tolerance limits, data behavior, and duplication probability in clinical trials across diverse therapeutic areas, phases, and trial types. The background of this project stems from the increasing challenges associated with managing and interpreting vast datasets within the context of clinical trial quality. By leveraging advanced data science techniques and statistical methodologies, the research seeks to unravel patterns, benchmarks, and probabilities that significantly impact the reliability and precision of clinical trial outcomes. The significance of this project is multiform. It addresses critical needs in current practices and contributes to the advancement of clinical trial research and data management in the following ways.

#### **Objective:**

Firstly, the clinical trial design will be optimized. Understanding benchmark quality tolerance limits will enable the optimization of clinical trial designs, leading to more efficient studies. Secondly, it enhanced the data quality and reliability. Insights into data behavior and statistical distributions will enhance the accuracy and reliability of predictive models, fostering more dependable clinical trial data. Thirdly, the investigation into duplication probability directly impacts patient data integrity, minimizing risks associated with duplicate records and improving overall data management practices. Overall, by disseminating findings through publications and collaboration with the scientific community, the project aims to contribute to the advancement of industry standards, fostering continuous improvement and innovation in quality control of clinical trial research.

#### **Study Design:**

The study is composed of data management and statistical analysis. Study data will be categorized according to phases and therapeutical areas. Based on the requirements of quality tolerance limit such as rate of adverse event, lost to follow-up, missing endpoint, etc., data from various studies is converted into a standardized format which is used in calculating the matrices on site, country, and study level. To understand the best model for study in different stages and therapeutical areas, multiple distributions are tested on matrices and summarize the best parameter set for each group of study type.

#### **Participant:**

In the context of the two projects, we are going to investigate the clinical trial data quality and integrity. Therefore, the project is not limited to a specific patient group. To develop a solid solution, the project would like to include patients in different ages, ethnicities, races, and medical history backgrounds, so that the developed algorithm is applicable to different types of studies.

#### **Primary and Secondary Outcome Measure(s):**

The primary outcome of the project is to determine the best distribution that either the site or center monitor can follow during clinical trials. Comparing the selected distribution with the study data can assess whether a study is under control or not. Regarding patient duplication detection the primary outcome is to assess the similarity between the given patient and all the patients in the database. For both topics, we will not take the secondary endpoint into account because the purpose of the research is to develop a new method of data management. We are not going to investigate/compare etc. product efficacy/safety etc.

#### **Statistical Analysis:**

The goodness of fit and patient similarity is evaluated by numeric analysis approach, Anderson-Darling statistic, and significant test. By patient similarity, the statistic test defines the threshold of

identity. Regarding the quality tolerance limit, some matrices depend on the number of study days. The numeric analysis approach can identify the pattern of the time series variable and then use statistical tests again to evaluate the goodness of fit. Moreover, considering the hypothesis of the best-fit distribution, the researchers assess if the quality tolerance limits are under the same scenario as the best-fit distribution.

### **Brief Project Background and Statement of Project Significance:**

Clinical trials represent a cornerstone in advancing medical knowledge and healthcare interventions. As the volume and complexity of clinical trial data continue to grow, there is a critical need to understand and optimize the quality, behaviour, and integrity of this data. The proposed research project by Cyntegrity aims to comprehensively investigate quality tolerance limits, data behaviour, and duplication probability in clinical trials across diverse therapeutic areas, phases, and trial types. The background of this project stems from the increasing challenges associated with managing and interpreting vast datasets within the context of clinical trials quality. By leveraging advanced data science techniques and statistical methodologies, the research seeks to unravel patterns, benchmarks, and probabilities that significantly impact the reliability and precision of clinical trial outcomes.

#### **Significance of the Project:**

The significance of this project is multiform. It addresses critical needs in current practices and contributes to the advancement of clinical trial research and data management in the following ways:

#### **Optimizing Clinical Trial Design:**

Understanding benchmark quality tolerance limits will enable the optimization of clinical trial designs, leading to more efficient studies.

#### **Enhancing Data Quality and Reliability:**

Insights into data behaviour and statistical distributions will enhance the accuracy and reliability of predictive models, fostering more dependable clinical trial data.

#### **Improving Patient Data Management:**

The investigation into duplication probability directly impacts patient data integrity, minimizing risks associated with duplicate records and improving overall data management practices.

#### **Contributing to Best Practices:**

The evaluation and recommendation of statistical algorithms for duplicate detection will contribute to establishing best practices in data management, guiding researchers and organizations in adopting effective strategies.

#### **Advancing Industry Standards:**

By disseminating findings through publications and collaboration with the scientific community, the project aims to contribute to the advancement of industry standards, fostering continuous improvement and innovation in quality control of clinical trial research.

#### **Use of Information for Generalizable Scientific and/or Medical Knowledge:**

The insights gained from this research will be disseminated through peer-reviewed publications, conference presentations, and collaborative engagement with the scientific community. By sharing methodologies, findings, and recommendations, the information obtained will materially enhance knowledge and experience in clinical trial quality control. This knowledge, in turn, will inform future research endeavours, influence data management practices, and contribute to the collective understanding of how to optimize clinical trials for the betterment of science and public health.

#### **References:**

ICH E6 (R3) Guideline on good clinical practice (GCP)\_Step 2b (europa.eu)

ICH: E6 (R2): Guideline for good clinical practice - Step 5 (europa.eu)

ICH guideline E8(R1) Step 2b on general considerations for clinical studies (europa.eu)

## Specific Aims of the Project:

### Topic 1 (QTL):

Examine prevalent or unique subject related data and benchmarking for subject-related quality tolerance limits employed in clinical trials. Gain insights into the types of distributions, distribution parameters (with recommendations), and assess whether there are specific patterns related to therapeutic areas or study phases.

### Topic 2 (Duplicated records):

Identify a set of subject characteristics that can facilitate the detection of Duplicated Subject.

## Study Design:

Meta-analysis (analysis of multiple trials together)

### What is the purpose of the analysis being proposed? Please select all that apply.

Summary-level data meta-analysis

Meta-analysis using data from the YODA Project and other data sources

Develop or refine statistical methods

Research on clinical prediction or risk prediction

## Research Methods

### Data Source and Inclusion/Exclusion Criteria to be used to define the patient sample for your study:

The analysis will be based on three data sources which are Cyntegrity owned data, open-source clinical trials data and Data from YODA. The study data from phase 1 and 4 will be excluded. During data cleaning stage, the records with unclear date or missing information will not be considered.

### Primary and Secondary Outcome Measure(s) and how they will be categorized/defined for your study:

#### Topic 1.

The primary outcome of the project is to determine the best guidance that either the site or center monitor can follow during clinical trials. The guidance is distributions which can explain the quality tolerance limit and statistically significant represent the distribution of metrics calculated by a study from each subgroup. For instance, the center monitor can know from the guidance that their study is overreporting or underreporting adverse events. Furthermore, according to the guidance, the center monitor can use it as a model to adjust the parameters that could better fit their study requirement for example, a higher adverse event report rate at an early stage is expected.

#### Topic 2.

The primary outcome of the patient duplication detection is to assess the similarity between the given patient and all the patients in the database. Therefore, the optimal outcome is a score of similarity taking the vital sign, geolocation, and medical history into consideration. In addition, another probability score can indicate the rarity of getting this similarity score.

For both projects, we will not take the secondary endpoint into account because the purpose of the research is to develop a new method of data management. We are not going to investigate/compare etc. product efficacy /safety etc.

### Main Predictor/Independent Variable and how it will be categorized/defined for your

## study:

Topic 1. The analysis will be based on three data sources which are Cyntegrity owned data, open-source clinical trials data and Data from YODA. Study data will all be anonymized and categorized by the therapeutic area and study phase.

Topic 2. Patient data will be anonymized. The predictor variables are vital sign, site location, medical history and concomitant medication, which will not be categorized but one-hot encoding and convert into multi-dimensional vector for the similarity calculation algorithms.

## Other Variables of Interest that will be used in your analysis and how they will be categorized/defined for your study:

NA

## Statistical Analysis Plan:

Topic 1:

The quality tolerance limit monitors clinical trials data through numeric analysis. The numeric analysis is conducted on derived metrics such as percentage of dropout, rate of adverse event etc. Metrics will be converted into density-based distribution and will be observe on site level and exam through the 108 types of continuous distribution's probability density function to figure out the best match parameter set which is with less deviation of curve area between real-world data and simulation. Moreover, we quantify the goodness of fit by using Anderson-Darling statistic on continuous data. The selected study will be labelled by their therapeutical area and phases and classify. The hypothesis is that studies from different therapeutic area and phase can show different patterns on the quality tolerance limit.

Topic 2:

Duplicated patients' detection is conducted within and across studies and subgrouping patient features such as medical history, vital sign, geolocation and other physical examinations. The features from patient will be converted into multi-dimensional matrix. Similarity is evaluated by the dot product of matrices.

## Software Used:

Python

## Project Timeline:

1. Data Cleaning Std-on (1-2 month):

- Clean and standardize the collected data to ensure consistency.
- Develop a standardized form for presenting the data.
- Verify data soundness and completeness.

2. Data Analysis (2-3 months):

- Perform statistical analyses to extract meaningful insights.
- Identify patterns, trends, and correlations within the data.
- Validate results and conduct sensitivity analyses.

3. Conclusions and Discussion (1 month):

- Summarize key findings and draw conclusions from the analysis.
- Discuss the implications of the results in the context of existing literature.
- Identify any limitations and propose avenues for future research.

4. Manuscript Writing (1 month):

- Outline the structure of the article (Introduction, Methods, Results, Discussion, Conclusion).
- Begin drafting each section, ensuring clarity and coherence.
- Review and revise the manuscript iteratively.

5. Review and Feedback (0.5 month):

- Seek input from colleagues, mentors, or collaborators.
- Address feedback and make necessary revisions.
- Ensure compliance with journal guidelines and standards.



6. Submission (1 month):

- Finalize the manuscript and prepare supplementary materials.
- Submit the article to the selected journal.
- Monitor the peer-review process and respond to reviewer comments.

7. Revision and Publication (0.5 month):

- Revise the manuscript based on peer-reviewer comments.
- Prepare the final version for publication.
- Await the journal's decision and publication.

**Dissemination Plan:**

Publications/Presentations: We will publish results in Applied Clinical Trials, present results in different conferences such as DIA, Phuse, Scope and SCDM

Auditory: Study Risk Management, Centralized Monitoring Management

The thresholds (medium and high) and distributions of Quality Tolerance limits are of paramount significance in the realm of clinical study data quality control. These metrics (provided they are derived from meticulously evaluated and mathematically validated data related to similar trials) are indispensable for the formulation of a Sponsor's quality strategy and for the management of ongoing clinical trials. The development of a quality strategy typically falls within the purview of Study Risk Management, with the ongoing oversight of studies from a risk management perspective being the responsibility of the Centralized Monitoring management team. In addition to mitigating and managing study risks, Centralized Monitoring teams are typically entrusted with overseeing study data fraud detection, which means that patient data duplication serving as a pivotal and challenging checkpoint in this regard. The findings of Cyntegrity's research endeavors are slated for dissemination and presentation through relevant publications or platforms (Society for Clinical Trials ([sctweb.org](http://sctweb.org)), Journal of Clinical Monitoring and Computing ([springer.com](http://springer.com))), and conferences (DIA, Phuse, Scope and SCDM) aimed at the aforementioned target audiences.

**Bibliography:**

[ICH E6 \(R3\) Guideline on good clinical practice \(GCP\)\\_Step 2b \(europa.eu\)](#)

[ICH: E 6 \(R2\): Guideline for good clinical practice – Step 5 \(europa.eu\)](#)

[ICH guideline E8\(R1\) Step 2b on general considerations for clinical studies \(europa.eu\)](#)

**Supplementary Material:**

<https://yoda.yale.edu/wp-content/uploads/2024/02/CT-Quality-Management-for-YODA.pdf>

<https://yoda.yale.edu/wp-content/uploads/2024/02/YODA-Application.pdf>