ARON TRIAL – STATISTICAL ANALYSIS PLAN

Impact of a diagnostic algorithm including clinically guided point-of-care C-reactive protein testing and safety netting advice on antibiotic prescribing rate and further management of acutely ill children presenting to ambulatory care: multicentre, cluster-randomized, parallel group pragmatic trial (the ARON project)

This statistical analysis plan provides guidelines for the final presentation and analysis for the ARON trial. This plan, along with all other documents relating to the analysis of this trial, will be stored in the Trial Master File electronically and/or in hard signed copy formats.

Prof Dr Jan Y Verbakel jan.verbakel@kuleuven.be **Department of Public Health and Primary Care, KU Leuven**	Chief Investigator
Dr Tine De Burghgraeve Tine.deburghgraeve@kuleuven.be **Department of Public Health and Primary Care, KU Leuven**	Project manager
Mrs. Annouschka Laenen annouschka.laenen@kuleuven.be **Leuven Biostatistics and Statistical Bioinformatics Centre, KU Leuven**	Trial statistician with final responsibility on trial analysis. She will validate the analysis for the primary outcome.
Ruben Burvenich Ruben.burvenich@kuleuven.be **Department of Public Health and Primary Care, KU Leuven**	Scientific collaborator who will perform the primary analysis as well as part of the secondary analysis.
Erinn D'hulster Erinn.dhulster@kuleuven.be **Department of Public Health and Primary Care, KU Leuven**	Scientific collaborator who will perform part of the secondary analysis, the full process evaluation as well as the health economic analysis.

Approval of statistical analysis plan

The undersigned confirm that the following SAP has been agreed and accepted:

Mrs. Annouschka Laenen
Date:
Signature:

Prof Dr Jan Y Verbakel
Date:
Signature:

Overall aim and research question

Background. The ARON trial will test the impact of a diagnostic algorithm including a standardized clinical assessment, a point-of-care (POC) C-reactive protein (CRP) test, and safety netting advice.

Question. Is this diagnostic algorithm able to safely reduce antibiotic prescribing in acutely ill children presenting to ambulatory care?

Aim. To establish the assumed superiority of a diagnostic algorithm including a standardized clinical assessment, a POC CRP test and safety netting advice over usual care to reduce antibiotic prescribing rates (both immediate and delayed prescribing).

Design of the trial

This study is a **multicentre, pragmatic cluster-randomized controlled superiority trial**:

- multicentre: six academic centres for primary care will be involved in the recruitment of 122 primary care practices
- pragmatic: criteria excluding patients from participation will be limited as well as procedures that do not reflect care as usual
- cluster: randomization occurs at the level of the primary care practice while patients are the unit of analysis
- randomized: block randomization per academic centre will be performed to assign primary care practices to either the intervention or the control arm
- controlled: the intervention will be compared with usual care
- superiority: the trial is designed to show that the intervention is superior to usual care

Randomization

In order to avoid contamination between physicians working in the same practice, randomization will happen at the level of the physician's practice.
General practice clinics and community paediatric centres will be randomized in one of the two study arms in a 1:1 ratio using a block randomization system stratified per recruiting academic centre in order to guarantee that allocation to either the usual care or the intervention arm is balanced within every region.
Stratified block randomization will be done using an electronic random numbers generator in blocks of four practices. Randomization and concealment will be centralized at the KU Leuven and conducted by a staff member not involved in data collection or delivering the intervention.

Sample size calculation

The original sample size calculation in the approved protocol version 2.0 was based on previous data from the ERNIE2 trial, assuming an overall antibiotic prescribing proportion of 26.5% in those children recruited for our trial.
The overall antibiotic proportion was found to be 18% in the first 4938 patients recruited in the ARON trial.

If we were to assume a reduction of 5.3% (proportionate to our original reduction) between the usual care group and the intervention group, using a 5% significance level (alpha 0.05), an intracluster correlation coefficient (ICC) of 0.063 (based on data from the ERNIE2 trial in the exact same population), and power of 90% (beta 0.1), this would require 63 clusters of 50 patients in both arms, resulting in 6300 children.

R code:
We used the n4props-function in the R package CRTSize, using the following command:
n4props(pc = 0.2065, pe = 0.1435, m = 50, ICC = 0.063, alpha = 0.05, power=0.9, AR=1, two.tailed=TRUE, digits=3).

Considering the pragmatic nature of this trial and in correspondence with the 10% of practices performing non-consecutive inclusion of patients (high risk of selection bias) during the ERNIE2 trial, we will perform sensitivity analyses only considering physicians who have recruited in a consecutive way. Taking into account the required sample size for this analysis of the primary study outcome, these assumptions result in a total sample size of 7000 patients.

Timing of the analysis

- **Primary and secondary outcomes**: April – June 2024.
- **Health economic study**: January - February 2025. This timeline is contingent upon possible delays in data availability from IMA and the efficiency of the data retrieval process.

Withdrawals

Withdrawal from the study can be initiated either by the child, their parent(s), or by the physician. As stated in the Informed Consent Form, children and their parents have the right to end their participation in the trial at any point in time and for any reason. In case a child or parent wishes to end their participation, they can do so by notifying their physician.
The treating physician can also consider a child for withdrawal from the study. However, the physician will have to discuss this option with the study coordinator and receive the consent of the latter before a final decision of withdrawal can be made.
In all instances, the reason for withdrawal will be asked for and recorded, but participants/parents are not required to share their reason for withdrawal.
Withdrawn children will not be replaced.

The patient and/or parent/legal guardian may withdraw consent at any time during the study. For the purposes of this trial, withdrawal is defined as:
The patient and/or parent/legal guardian would like to withdraw consent from study and is not willing to be followed up for the purposes of the trial at any further visits (i.e., only data collected prior to the withdrawal of consent can be used in the trial analysis).

The details of withdrawal should be clearly documented in the patient's health records and in the electronic case report form (REDCap).

Software used

REDCap version 13.7.25 (2024 Vanderbilt University).
QSR NVIVO software version 12 (QSR International Pty Ltd, Melbourne, Australia).
R software version 4.2.3 (1) within R Studio (2023.12.1 Build 402) (2) using the 'glmer' function of the 'lme4' package (3) and the 'n4props' function of the 'CRTSize' package (4).

(1) R Core Team (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

(2) Posit team (2023). RStudio: Integrated Development Environment for R. Posit Software, PBC, Boston, MA. URL http://www.posit.co/.

(3) Douglas Bates, Martin Maechler, Ben Bolker, Steve Walker (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1-48. doi:10.18637/jss.v067.i01.

(4) Rotondi MA (2023). _CRTSize: Sample Size Estimation Functions for Cluster Randomized Trials_. R package version 1.2, <https://CRAN.R-project.org/package=CRTSize>

Interim analyses

No interim analyses are planned.

Data Source

Our study utilized a prospective data collection approach, drawing from several sub-databases to create a comprehensive dataset for analysis:

1. **REDCap Database.** A secure web application for building and managing online surveys and databases. REDCap was utilized for its robust data capture capabilities, particularly for complex longitudinal data and multisite clinical research studies.

2. **ARON App Database.** A user-centred database collected through direct patient/parent contact via a mobile application over a 30-day follow-up period. These data encompass information on daily symptoms, healthcare consultations, hospital visits, and health-related quality of life (HRQoL).

3. **Administrative Database (healthcare utilization).** Data obtained from the Intermutualistische Agentschap (IMA) on potential hospitalizations, technical provisions, and reimbursed pharmaceutical usage up to 3 months post-index consultation. Access to this administrative data for ARON participants is enabled by a secure procedure developed to transmit national insurance numbers in compliance with GDPR regulations, facilitated by a trusted third party. This data is collected for all ARON trial participants with valid national insurance numbers.

4. **Interview Transcripts.** A process evaluation is nested within the ARON trial, aiming to gather insights into the utilization and perceptions of the intervention by physicians and (parents of the) patients. This qualitative process evaluation entails conducting individual semi-structured, in-depth interviews with GPs and parents, facilitated through video conferencing.

Date of final version of dataset

1. **REDCap Data**

Estimated date of receipt 29/03/2024.

2. **ARON App Data**

The app database was finalized on 1 February 2024, 30 days after last patient visit.

3. **Administrative Data**

In the summer of 2023, a pilot retrieval of IMA data was conducted, which entailed the transmission of national insurance numbers for the first 495 patients recruited. The IMA data linked to the relevant clinical data became available for analysis on the IMA servers on 6 December 2023.

The second and final retrieval will encompass the remaining participants in the trial. Due to an estimated 9-month delay in data availability at IMA, the final retrieval is anticipated to be completed by the end of 2024.

4. **Interview Transcripts of process evaluation**

Interviews were conducted between February 2022 and February 2023, and analysis of the transcripts was completed in December 2023.

People responsible for data storage, data cleaning, and data analysis

Annouschka Laenen. Working at Leuven Biostatistics and Statistical Bioinformatics Centre (L-BioStat), she serves as the trial statistician, and all analyses will be conducted in close collaboration with her. She will validate the analysis for the primary endpoints.

Ruben Burvenich. He followed the KU Leuven Statistics: Online Course (with exam and certificate) in 2020, Introduction to R (with certificate) by FLAMES, Flanders' Training Network for Methodology and Statistics in 2020, and Open Online Introduction to R Course by Wolfgang Viechtbauer in 2020. He has a solid background in coding (R and SAS environments) and data analysis as a final year PhD candidate who is the first author and primary analyst of studies involving meta-analysis and meta-regression, moving average time series analysis, and network meta-analysis including network meta-regression. As such, he has collaborated and continues to work with biostatistician and Professor in Epidemiology at Maastricht University and KU Leuven, Professor Laure Wynants. Being a final-year PhD candidate with a primary focus on the clinical effectiveness of the ARON trial, he has read much of the national and international relevant literature and attended several international conferences on the topic. Over time, this has made him very familiar with this domain of research and the associated statistical methods.

Erinn D'hulster. She holds a degree in business engineering, where she completed courses on General Statistics, Statistical Modelling and Data Mining, Econometrics, Algorithms and Data Structures, Simulation Modelling and Analysis, and Programming in JavaScript. In 2021, she completed an online Biostatistics course at KU Leuven and obtained certification through examination. Furthermore, she attained a certificate in Introduction to R organized by the FLAMES training network. In her doctoral research, Erinn has undertaken several projects in health economics, resulting in first-author publications. Additionally, she conducted a budget impact analysis commissioned by the Flemish government, including a Python simulation model, which contributed to an additional allocation of 270 million euros to the Flemish childcare sector. As a final-year PhD candidate, Erinn has extensively engaged with relevant literature and has actively participated in various international conferences focusing on health economic analyses.

Inclusion and exclusion criteria

Practices' eligibility (and physicians within these practices) for inclusion in the study will be based on the following criteria:

- Being able to recruit acutely ill children, ideally consecutively.
- Agree to the terms of the clinical study agreement.

Practices will be **excluded** from study participation based on the following criteria:

- Current use of a POC CRP device as part of their routine care

Age, demographics, geographic region will not be used to exclude eligible practices. This will provide us with a real-life, representative subset of ambulatory care physicians.

Patients' eligibility for inclusion in the study will be based on the following criteria:

- Children aged 6 months to 12 years, provided informed consent can be obtained
- Presenting with an acute illness episode that started maximum 10 days before the index consultation

Patients will be **excluded** from study participation based on the following criteria:

- Children who were previously included in this trial
- Children with an underlying known chronic condition (e.g., asthma, immune deficiency)
- Clinically unstable warranting immediate care
- Immunosuppressant medication taken in the previous 30 days
- Trauma as the main presenting problem
- Antibiotics taken in the previous 7 days
- Unwillingness or inability to provide informed consent

Endpoints and covariates

Endpoints

The primary outcome is the proportion of subjects who were prescribed antibiotic treatment (both immediate and delayed) at the index consultation as recorded by the treating physician.

Secondary endpoints:

- The duration (in days) until reaching full clinical recovery
- The proportion of subjects receiving additional testing (including, but not limited to (X-Ray, blood tests, urine tests) at index consultation (day 0) and/or during follow-up (day 1 to day 30)
- The proportion of subjects who re-consulted their physician during follow-up (day 0 to day 30)
- The proportion of subjects who were prescribed antibiotic treatment during follow-up (immediately after index consultation to day 30)

Exploratory endpoints:

- The proportion of subjects receiving additional testing (including, but not limited to (X-Ray, blood tests, urine tests) at index consultation (day 0)
- The proportion of subjects receiving additional testing (including, but not limited to (X-Ray, blood tests, urine tests) during follow-up (day 1 to day 30)
- The proportion of subjects referred to hospital at index consultation (day 0)
- The proportion of subjects referred to hospital during follow-up (day 1 to day 30)
- The proportion of subjects admitted to hospital at index consultation (day 0)
- The proportion of subjects admitted to hospital during follow-up (day 1 to day 30)
- The proportion of subjects who died at index consultation (day 0)
- The proportion of subjects who died during follow-up (day 1 to day 30)
- The proportion of subjects with full clinical recovery at day 7
- The proportion of subjects with full clinical recovery at day 30
- Patient's satisfaction (as part of the nested qualitative study)
- Parent's satisfaction (as part of the nested qualitative study)
- Physician's satisfaction (as part of the nested qualitative study)

- Cost-effectiveness of the intervention: healthcare expenditures in terms of hospitalization, consultations, pharmaceuticals (reimbursed and non-reimbursed), productivity, quality of life
- Adherence to the diagnostic algorithm
- The proportion of subjects who actually took antibiotics (from day 0 to Day 30)

Covariates (predictors in the model)

Important predictors were chosen a priori based on expert opinion by Prof Dr Jan Verbakel.

Predictor	Measurement scale	Df needed
Study arm	Binary: intervention versus control	1
Age	Continuous	1

Handling of missing values and other data conventions

1. **REDCap Data**

Drop-outs are unlikely to be a large problem since there will be only one study visit complemented by the collection of follow-up information.

We will analyse the whole population for the primary analysis and perform a sensitivity analysis excluding the low recruiting practices (less than 10 patients recruited for 12 months) to avoid our results to be susceptible to selection bias. In previous trials this resulted in approximately 10% of practices.

Practices that do not recruit children at all will be stopped after the first two monitoring visits and replaced by new practices to avoid reducing the total number of clusters available for analysis.

The percentage of missing data for the primary outcome of antibiotic prescribing proportion is expected to be low as this will be registered at the first contact consultation. In our previous ERNIE2 trial, we found 4% of missing data for the primary outcome of antibiotic prescribing proportion. Multiple imputation will be considered for the primary outcome measure as well as missing values for patient-reported outcomes (if there are any). Other outcomes will be collected from the patient health record of the physician.

Multiple imputation will be applied to deal with missing data. Imputation will be performed for the binary outcome variable and logistic regression will be used as imputation model. Predictors for the imputation model are baseline patient characteristics and intervention.

2. **ARON App Data**

A substantial drop-out rate in app completion is anticipated, with non-completion of the app not considered grounds for exclusion from the study. To address this, proactive measures were introduced to enhance participation rates, such as the provision of information about the smartphone app and instructions on installation to children and their parents during the baseline consultation. Additionally, the app was made available in three languages (Dutch, French, English).

Multiple imputation will be applied to deal with missing data. For imputation of continuous variables, a linear regression model will be employed, with baseline patient characteristics and intervention serving as predictors for the imputation model.

3. **Administrative Data**

Multiple imputation, following Rubin's rules, will be utilized to address missing data. Predictors for the imputation model will include baseline patient characteristics and the intervention.

4. Interview Transcripts

Imputation techniques are not applicable in this dataset.

Statistical Methodology

Statistical Procedures

Baseline characteristics & Clinical features

Presentation of baseline characteristics of the study population and comparability of the two arms will be based on the following variables:

- Age (median and 25-75 percentiles)
- Gender (percentage)

Baseline characteristics and clinical features will be reported using frequencies and percentages, means and standard deviations or medians and interquartile ranges, and minimum / maximum, as appropriate.

Primary endpoint

For the analysis of the primary outcome, we will use a mixed-effects logistic regression analysis to account for the clustering at practice level. The child's age and study arm will be included as predefined covariates in the model. We will do this using R software using the 'glmer' function of the 'lme4' package. Results will be reported by an odds ratio (OR) with 95% confidence intervals.

Multiple imputation will be applied to deal with missing data. Ten complete datasets will be constructed and analysed. The 10 results will be combined into a final result following Rubin's rule.

We will analyse the primary endpoint according to the intention-to-treat (ITT) approach, given the pragmatic nature of our trial.

The participating physicians were asked to recruit children with an acute illness consecutively. GPs were likely to have breached the assumption of consecutive inclusions if they:

- recruited fewer than 10 patients per year;
- performed a POC CRP test on nearly all (>90%) of the included children; and/or
- included nearly exclusively children (>90%) based on a negative result on all three items of the clinical decision tree that do not need antibiotics and exclude those that might need antibiotic treatment.

Apart from the intention-to-treat analysis, a per protocol sensitivity analysis will be performed only considering physicians who have recruited children consecutively.

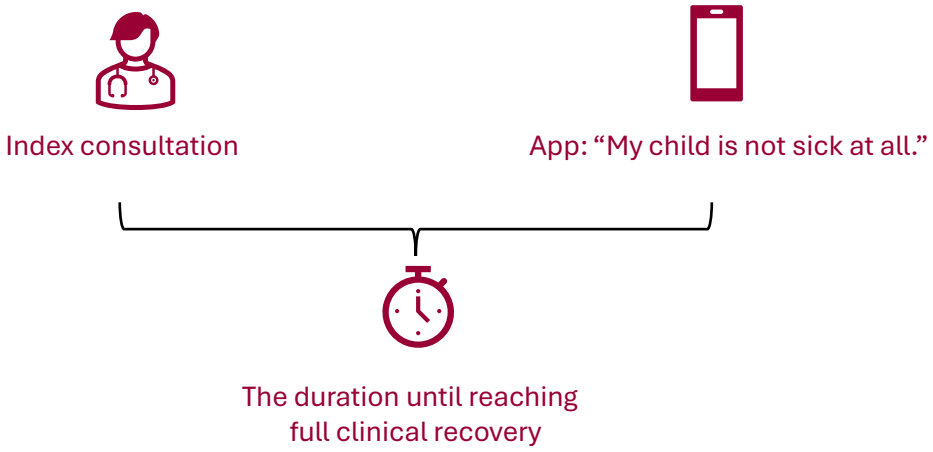
Secondary endpoint – Proportional measures

We will provide a descriptive analysis of the proportion of subjects with the outcome of interest for each randomization group.

A mixed-effects logistic regression model will be employed to address clustering at the practice level. The child's age and study arm will be included as predefined covariates in the model. Odds ratios, along with their corresponding 95% confidence intervals will be reported.

Secondary endpoint – Duration until reaching full clinical recovery

The duration (in days) until reaching full clinical recovery:



Null hypothesis. The diagnostic algorithm including a standardized clinical assessment, a POC CRP test and safety netting advice is inferior to usual care in terms of the number of days until subjects reach full clinical recovery.

Analysis. We will conduct a descriptive analysis to determine the median time to recovery for each randomization group. IQR will also be provided for each group.

For graphical representation, Kaplan-Meier curves will be generated. These curves will be used in combination with a log-rank test to compare the survival curves of the randomization groups.

Additionally, will employ a mixed-effects Cox regression model to address clustering at the practice level. The child’s age and study arm will be included as predefined covariates in the model. Hazard ratios, along with their corresponding 95% confidence intervals, will be reported.

Exploratory endpoints – Proportional measures

We will provide a descriptive analysis of the proportion of subjects with the outcome of interest for each randomization group.

A mixed-effects logistic regression model will be employed to address clustering at the practice level. The child’s age and study arm will be included as predefined covariates in the model. Odds ratios, along with their corresponding 95% confidence intervals will be reported.

Exploratory endpoints – Satisfactory measures

Statistical analysis not applicable due to the qualitative nature of the data.

Upon transcription of the interviews verbatim, an inductive thematic analysis approach is employed to analyse the data. The entire data analysis process was facilitated using qualitative data analysis software QSR NVIVO software version 12 (QSR International Pty Ltd, Melbourne, Australia).

Findings will be centred on main themes and subthemes closely aligned with the research inquiries and reflective of the interview content. In addition, illustrative quotes will be included.

Exploratory endpoints – Cost-effectiveness of the intervention

Different types of economic evaluations will be conducted to compare participants in the intervention group with those in the control group:

- **Costing study** to comprehend the overall cost structure, identify significant cost drivers, and compare the cost impact.
    - Calculation of costs: In both study groups, the subcategory costs and total costs per patient will be calculated. These costs include the direct medical costs of healthcare use, the cost of the intervention itself, and the indirect costs for patients.
        - Between-group differences: Differences between the intervention and control groups will be calculated for each subcategory and total costs.
            - Descriptive Statistics: Initial assessment of cost difference will involve descriptive statistics to provide unadjusted information regarding the magnitude and distribution of the costs. T-tests and Mann-Whitney U tests will be employed to assess differences between the randomization groups.
            - Generalized Linear Modelling: To evaluate total cost differences between both groups while adjusting for independent factors such as patient characteristics (age and gender).
- **Cost-consequences analysis**:
    - Compares costs and consequences, encompassing hospitalizations, consultations, pharmaceuticals, and productivity between intervention and control group.
    - Considers perspectives from healthcare payers and society.
    - Utilizes decision-analytic modelling, employing a combination of decision trees and Markov models to track hypothetical cohorts over time.
    - The results of this analysis will be presented as incremental cost-effectiveness ratios, calculated based on natural effects. In case of a dominant or dominated intervention, net benefits will be computed and plotted on a net benefit plot.
- **Cost-utility analysis:**
    - Compares costs and ‘utilities’ (calculated using HRQoL data collected through questionnaires in ARON app and the Belgian value set for the EQ5D-Y questionnaire) between intervention and control group.
    - Considers perspectives from healthcare payers and society.
    - Utilizes decision-analytic modelling, employing a combination of decision trees and Markov models to track hypothetical cohorts over time.
    - The results of this analysis will be presented as incremental cost-utility ratios, calculated based on Quality-Adjusted Life Years (QALYs). In case of a dominant or dominated intervention, net benefits will be computed and plotted on a net benefit plot.

We will perform one-way, multi-way, and probabilistic sensitivity analyses, including Monte Carlo simulation and non-parametric bootstrapping, to account for various sources of uncertainty. The outcomes of these analyses will be depicted using cost-effectiveness acceptability curves.

Subgroup Analysis

Subgroup analysis will be performed in order to investigate how the primary outcome behaves in function of:

- Age categories: 0-1 years old, 2-6 years old, 7-12 years old
- Gender: male, female, prefer not to say

Presentation of Results

Figure 1. Flowchart of inclusions/exclusions.

Table 1. Demographic background data, potential predictors, and outcomes (n=x)

Characteristic	Mean +-SD; median [IQR]; event(yes)/n (percentage)		
	Total	Control	Intervention

Table 2. Odds ratios with 95% CI and p-value of the treatment effect (n=x)

Outcome	OR (95 %CI)	P-value

Supplementary Tables. Subgroup and sensitivity analyses

Outcome	OR (95 %CI)	P-value

Rationale for any deviation from pre-specified analysis plan

Not applicable.