BMJ Open

Long-term outcome of oesophageal atresia in adolescence (TransEAsome): a cohort study protocol

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2024-086303
Article Type:	Protocol
Date Submitted by the Author:	11-Mar-2024
Complete List of Authors:	Leroy, Mélanie; Lille Regional and University Hospital, Aumar, Madeleine; Lille Regional and University Hospital, Service de gastroentérologie pédiatrique Lille Duhamel, Marie; Université de Lille, Laboratoire Protéomique Réponse Inflammatoire Spectrométrie de Masse - PRISM, (INSERM U1192) Dauchet, Luc; Lille Regional and University Hospital, Departement of biostatistics Figeac, Martin; Université de Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, US 41 - UAR 2014 - PLBS, Gaillard, Segolene ; CHU Lyon, Hankard, Regis; Université de Poitiers, Labreuche, Julien; Lille Regional and University Hospital, Department of Biostatistics Marot, Guillemette; Université de Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, US 41 - UAR 2014 - PLBS; Université de Lille, CHU Lille, Inria, ULR 2694 - METRICS : Évaluation des technologies de santé et des pratiques médicales Reversat, Julie; Association Française de l'Atrésie de l'Œsophage Armand, Viviane; Association Française de l'Atrésie de l'Œsophage Salzet, Michel; Université de Lille, Laboratoire Protéomique Réponse Inflammatoire Spectrométrie de Masse - PRISM, (INSERM U1192) Sfeir, Rony; Lille Regional and University Hospital, Reference Center for Congenital Oesophageal Anomalies, CHU Lille, Infinite U1286 Vandel, Jimmy; Université de Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, US 41 - UAR 2014 - PLBS Gottrand, Frédéric; Lille Regional and University Hospital, Reference Center for Congenital Oesophageal Anomalies, CHU Lille, Infinite U1286
Keywords:	Adolescents < Adolescent, Oesophageal disease < GASTROENTEROLOGY, Patient Reported Outcome Measures, Quality of Life, Paediatric gastroenterology < GASTROENTEROLOGY, Cell biology < NATURAL SCIENCE DISCIPLINES

Long-term outcome of oesophageal atresia in adolescence (TransEAsome): a cohort study protocol

Mélanie Leroy[1], Madeleine Aumar[1], Marie Duhamel[2], Luc Dauchet[3], Martin Figeac[4], Ségolène Gaillard[5], Régis Hankard[6], Julien Labreuche[3], Guillemette Marot[4,8], Julie Reversat[9], Viviane Armand[9], Michel Salzet[2], Rony Sfeir[1], Jimmy Vandel[4], Frédéric Gottrand[1]

[1] Univ. Lille, Reference Center for Congenital Oesophageal Anomalies, CHU Lille, Infinite U1286, F-59000 Lille, France

[2] Univ. Lille, Inserm, CHU Lille, U-1192 - Laboratoire Protéomique Réponse Inflammatoire Spectrométrie de Masse - PRISM, F-59000, Lille, France

[3] CHU Lille, Department of Biostatistics, F59000-Lille, France

[4] Univ. Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, US 41 - UAR 2014 - PLBS, F-59000 Lille, France

[5] Hospices civils de Lyon, EPICIME-CIC 1407 de Lyon, Kids France, Inserm 1407, CHU-Lyon, 69677 Bron, France ; Université Lyon 1, CNRS UMR 5558, laboratoire de biométrie et biologie évolutive, 69622 Villeurbanne, France

[6] French Clinical Research Infrastructure Network (F-CRIN)-PEDSTART, INSERM, University of Tours, Tours, France

[8] Univ. Lille, CHU Lille, Inria, ULR 2694 - METRICS : Évaluation des technologies de santé et des pratiques médicales, F-59000 Lille, France

[9]Association Française de l'Atrésie de l'Œsophage, Maison Alfort, France

Corresponding author:

Pr Frédéric Gottrand

CENTRE HOSPITALIER UNIVERSITAIRE

Hôpital Jeanne de Flandre

Pôle Enfant

2 avenue Oscar Lambret - 59037 Lille Cedex, France

+33 3 20 44 61 26

Frederic.gottrand@chu-lille.fr

Abstract

Introduction

The *TransEAsome* project, funded by the Agence Nationale de la Recherche, aims to evaluate the long-term outcomes of oesophageal atresia (OA) patients between 13 and 14 years old and establish multi-omics profiles using data from the world biggest OA registry.

Methods and analysis

TransEAsome is a national multicentre population-based cohort study recruiting participants from all qualified French centre for oesophageal atresia surgery at birth. The primary objective is to assess the prevalence of gastro-oesophageal reflux disease in adolescence among OA patients, with several secondary objectives including the identification of risk factors and multi-omic profiles from oesophageal biopsies and blood samples collected between 13 and 14 years old, compared with a control group. This comprehensive characterisation of phenotype and omic profiles aims to enhance understanding of OA patient evolution and inform tailored care management strategies.

Ethics and dissemination

The study, co-constructed with input from patients, parents and research-expert adolescents, has obtained approval from the French ethics research committee. Findings will be disseminated to various target audiences, including the scientific community, research participants, patient community, the general public, regulatory authorities, and policymakers. Data will be made available in a FAIR format on the France Cohortes platform upon study completion.

Trial registration number

NCT05995171

Strengths and limitations of this study

Strengths:

- Multicentre, national, population-based cohort study conducted in all French centres qualified to perform oesophageal atresia surgery at birth,
- Longitudinal analysis facilitated by a nested cohort from the birth/1-year-old register (ReNaTo) and the 6-year-old cohort (COMAD6),
- Active involvement of patients and parents thanks to the users' representative committee, ensuring the study reflects patient perspectives,
- Comprehensive multi-omic analysis encompassing transcriptomic, epigenetic, proteomic, and metabolomic data, providing a holistic understanding of the disease.

Limitations:

- Recruitment bias may exist for control patients due to the requirement for oesophageal biopsies, potentially limiting generalizability,
- The study is exploratory and descriptive in nature, lacking predefined formal hypotheses, which may affect the interpretation of the results,
- Omic studies rely on oesophageal tissues samples collected for clinical surveillance purposes, which may not be systematic, and long-term storage could lead to RNA degradation over time, potentially impacting data quality.

Keywords

Oesophageal atresia, adolescence, gastro-oesophageal reflux, dysphagia, oesophageal cancer, eosinophilic esophagitis, outcome, epigenetic, proteogenomic metabolomics, quality of life

Introduction

Oesophageal atresia (OA) is a rare congenital anomaly affecting approximately 160 newborns annually in France, characterized by a discontinuity in the oesophagus. It is often concomitant with other malformations in over 50% of cases (1). Since the early 1950s, significant advancements in surgical techniques have vastly improved the management of this condition, resulting in a favourable prognosis. While the mortality rate has declined to less than 7% in developed nations like France (2), morbidity remains noteworthy (3,4).

Gastro-oesophageal reflux disease (GORD) is more prevalent among individuals with OA compared to the general population (5). Studies indicate that approximately 30% of OA patients experience GORD at least once in their lifetime (6). Complications of GORD, such as peptic esophagitis and Barrett's oesophagus, present heightened risks for oesophageal cancer (7), with documented cases of adenocarcinoma or squamous cell carcinoma in young adults with OA (8).

Dysphagia is a common concern among OA patients, with around 45% of five-year-olds experiencing swallowing difficulties attributed to anastomotic strictures or oesophageal dysmotility (9). Eosinophilic oesophagitis (EoE), a rising concern in oesophageal health, has been noted to occur more frequently in OA patients, contributing to dysphagia (10).

Challenges in eating often lead to undernutrition, affecting up to one-third of children with EA by the age of five (11). Respiratory complications are also prevalent (12), particularly in the early years, significantly impacting long-term quality of life (QoL) (13).

To better understand the long-term outcomes and predict complications, our objective is to investigate a nested population cohort of adolescents born with OA. Leveraging the existing national population-based registry (ReNaTo), and the nested clinical cohort, COMAD6, consisting of 382 children born between 2010 and 2012, we aim to

compile a new database comprising clinical data, blood samples, and oesophageal biopsies obtained during routine follow-up between 13 and 14 years old (Figure 1).

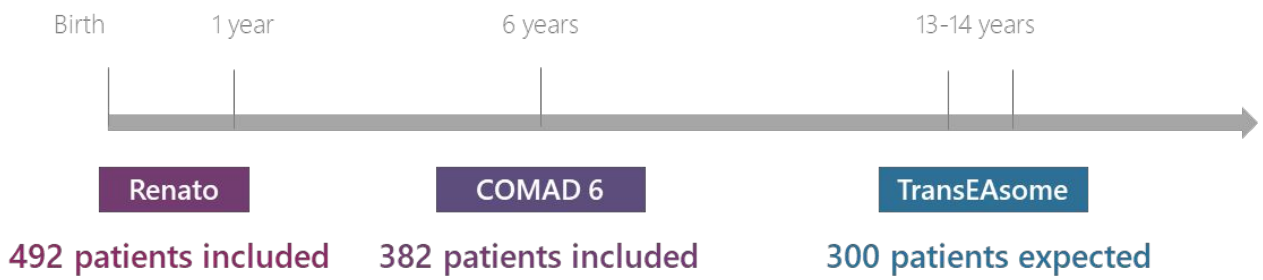


Figure 1: Follow-up of a nested cohort of OA patients born between 2010 and 2012

This nested cohort, within the world's largest registry of more than 2300 OA patients, offers a unique opportunity to answer questions about outcomes of OA in adolescence in terms of morbidity and health status, to improve care and follow-up and prevent long-term complications.

In addition to the phenotypic assessment, our investigations extend to exploring the biological mechanisms underpinning complication development through genomic, transcriptomic, proteomic, epigenetic, and metabolomic analyses of the oesophagus.

While upper gastrointestinal endoscopy (EGD) is standard in the care of symptomatic or not OA patients (5), its invasive nature necessitates hospitalization and general anaesthesia in children.

To date, there is only very few data on the epigenetic, genomic (14–18), or transcriptomic (19) profile of patients with OA. In a small study of 10 children with Barrett's oesophagus (some of them operated at birth for OA), the authors performed fluorescence in situ hybridization with probes on 4-micron sections taken from sequential paraffin-embedded biopsies and identified 4 probe sets reported to be associated to adult Barret adenocarcinoma (19). Genetic markers were also identified in adult Barrett's adenocarcinoma patients. This preliminary study shows that, even at an early age, Barrett's may show genetic changes associated with neoplastic

progression (20). Another recent study examined the relationship between eosinophilic esophagitis (EoE) and OA by profiling the transcriptional signature of EoE (21). Using an in-silico approach, they found 6 genes that were differentially expressed between the 2 entities (OA+ EoE+ and OA- EOE+). Two of them were associated with majored dysphagia, the development of strictures, and the need for dilatations in patients with OA+EoE+.

As EGD and oesophageal biopsies are invasive procedures that usually require hospitalization and general anaesthesia in children, using a more accessible biological sample as blood plasma will also be investigated to see if the markers of the oesophagus translate into blood.

The overarching goal of *TransEAsome* is to construct a vast longitudinal database including phenomenal and exposomal patient data amalgamating clinical, biological, environmental, and lifestyle data alongside multimodal omics analyses: RNA abundance, protein abundance, proteogenomic profiling, protein modification, metabolite abundance, methylated sites). This innovative approach holds promise for enhancing the prediction of adult OA outcomes, including the identification of risk factors for future health complications such as oesophageal cancer and EoE.

Objectives

TransEAsome will address the evolution of OA patients at the time of adolescence by establishing a unique and comprehensive database that integrates clinical and omics data in a structured and interoperable format to evaluate their health status and quality of life. This project is designed to achieve the following objectives:

1. Evaluate the prevalence of GORD during adolescence in the OA population.

2. Identify factors associated with GORD during adolescence in the OA population.

3. Assess the quality of life, nutritional status, and frequency of respiratory complications in the OA population.

4. Compare the single and multi-omic profiles derived from oesophageal biopsies in adolescents operated at birth for OA with those who did not, aiming to delineate alterations in biological pathways specific to OA.

5. Investigate changes in the expression of omics variables within the identified biological pathways over time in the OA population, utilizing data obtained from EGD and biopsies.

6. Evaluate the correlation between the expression levels of microRNAs and proteins in plasma compared to oesophageal biopsies in the OA population, utilizing data from individuals who underwent EGD, biopsies, and blood sampling at ages 13 and 14.

Methods and analysis

Recruitment sources

Patients for the *TransEAsome* study will be recruited from the French national registry for OA known as "ReNaTo," which involves all 35 OA competence centres across France. As of the conclusion of 2022, more than 2300 patients were included in ReNaTo. Among them, 492 are projected to reach 13 or 14 years of age during the recruitment period of the *TransEAsome* project. However, an anticipated loss to follow-up necessitates accounting for a pool of 300 patients. Leveraging the national register ReNaTo, the project management team will furnish each centre with a pseudo-anonymized patient roster, facilitating the identification of eligible candidates from their patient cohorts. Given the unique demographic characteristics of each participating centre, recruitment potentials vary, spanning from 1 to 34 patients per centre during the study period (Figure 2).

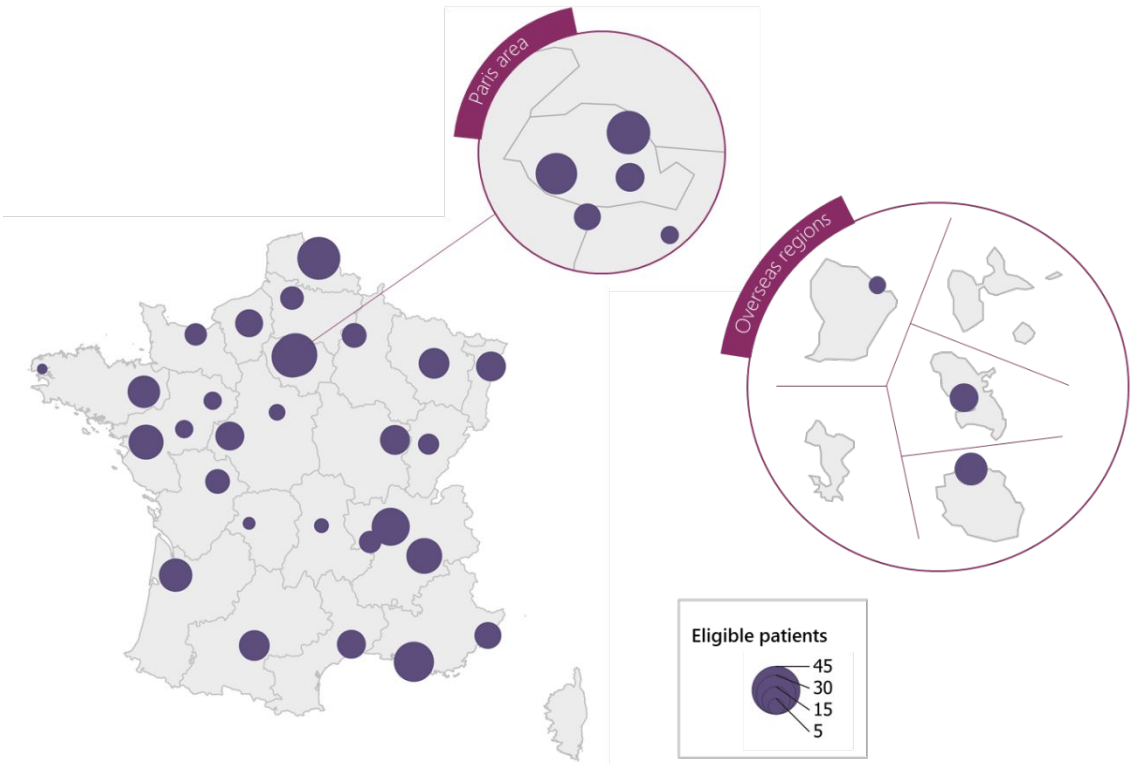


Figure 2: Localisation of recruitment in France

For the objectives involving multi-omic profiling, the control group will be drawn from six selected centres. Eligible patients will be identified through the extraction of cases meeting the study's inclusion criteria from the pathology department.

Study population

TransEAsome eligibility and exclusion criteria are listed in Table 1. The initial recruitment period will start in September 2023 and end in May 2025. Our goal will be to enrol 300 patients and 150 controls.

Table 1: Eligibility and exclusion criteria

Oesophageal atresia patients	
Eligibility criteria	• Born with OA in France. • Underwent oesophageal anastomosis at one of the participating centres

	• Age between 13 and 14 years during the inclusion period. • Willingness to participate. Blood sub-study: • Availability of oesophageal biopsy. • Consent for blood sampling.
Exclusion criteria	• Participation in an interventional trial (simultaneously or up to 3 months prior to inclusion. • Oesophageal replacement
Controls	
Eligibility criteria	• Age between 10 and 14 years during the inclusion period (born between 2010 and 2012). • Underwent esophagogastroduodenoscopy (EGD) as part of routine care for any digestive symptom to exclude organic aetiology (peptic esophagitis, gastritis, eosinophilic esophagitis, or ulcer). • EGD findings and histology are normal. • Absence of concurrent progressive chronic disease. • Provision of signed study consent.
Exclusion criteria	• Known underlying disease

The blood substudy as well as the inclusion of controls will only be performed in Lille, Paris Necker, Paris Robert Debré, Lyon, Grenoble, and Marseille for logistic, feasibility and cost saving reasons.

Patient and Public Involvement (PPI)

A young patient advisory committee and a parents/patient representativeness committee have been established to ensure the active involvement of young patients

and their parents throughout all phases of the *TransEAsome* project, from inception to dissemination of results. These committees have played a pivotal role in shaping the study protocol, including determining the data to be collected and reviewing informed assents and consents before regulatory submissions. Throughout the trial, these committees will contribute to the decision-making process of the steering committee and provide insights from a patient perspective on the conduct of the study.

The committees will remain actively engaged throughout the project, facilitating communication and dissemination efforts to the scientific community, patients, and the general public regarding key milestones, actions, and emerging knowledge. Upon project completion, the committees will review the analysis, discuss the findings, and aid in the dissemination of results.

Kids France (Hospices Civils de Lyon, Pedstart) will oversee the patient and public involvement (PPI) activities in collaboration with the French OA patient support group (AFAO). Kids France will ensure that PPI activities are not only meaningful for the research but also beneficial for the patients themselves. Training related to PPI activities will be provided to the committees by Kids France. Reporting of activities to the steering committee, feedback on suggestions from patients and parents, and assessment of impact are integral components of the *TransEAsome* PPI process (22,23).

Outcomes and assessments

Outcomes and assessments are presented in with separate forms for parents and adolescent patients.

Table 2. Quality of life (QoL) questionnaires and blood sampling are the only assessments that are not part of routine practice. Both the generic paediatric QoL assessment (Pediatric Quality of Life Inventory™ (PedsQL) and the disease specific QoL questionnaire (EA QoL) (13) will be administred in this cohort, with separate forms for parents and adolescent patients.

Table 2: Outcomes and assessments between 13 and 14-year-old

Primary outcome	GORD will be determined based on the following criteria: • Positive pH-(impedance)metry results obtained within the previous year. • And/or presence of histological peptic esophagitis lesions at oesophageal biopsies collected during the previous year. • And/or history of anti-reflux surgery.
Secondary outcomes	Quality of life: • PedsQL total patient score • EA QoL total patient score • PedsQL total parent score • EA QoL total parent score Nutritional status: • Z-score Weight/Height • Z-score Height/Age • Type of feeding (oral, enteral, both) Digestive status: • Current proton pump inhibitor treatment

	• Dysphagia (defined by sensations such as blockage leading to vomiting or the need to drink to pass food, or slowness in eating) • GORD symptoms (regurgitation, vomiting, retrosternal pain, and heartburn) Respiratory status: • Frequency of cough • Asthma • Occurrence of exercise-induced symptoms (cough, dyspnoea) • Atopy • Wheezing • Stridor • Need for medications (corticosteroid, inhaled medications) History from 6 years to adolescence: • Associated malformations discovered after 6 years • Gastrostomy placement and/or use after 6 years old • Characteristics of eventual GORD surgery (type, complications, cardial dilatation, relapse) • Oesogastroduodenoscopy (number, histology) • Repermeabilization of the oesotracheal fistula (treatment, number, date) • Stenosis (date, dilatation, number, surgery) • Stricture (date, number and method of dilatation, surgery, use of corticosteroid) Psychological or psychiatric follow-up

	Neuro orthopaedic outcomes
	• Scoliosis
	• Kyphosis
	• Stature malposition
	
	Orthopedic treatment
	• Oral/stomatology
	• Oral disorders
	• Speech disorders
	• Dental conditions
	
	Education:
	• Physical/sport waiver
	• Adapted school rhythm
	• Specialized school
	• School absenteeism
	
	Omic profile:
	• In oesophageal biopsies:
	  ○ Protein abundances
	  ○ Metabolite abundances
	  ○ RNA profile
	  ▪ 3'RNA
	  ▪ miRNA
	  ○ Methylation status
	• In blood:
	  ○ Metabolite abundance
	  ○ Protein abundance
	  ○ miRNA profile

Other assessments	Parents' height Parents' socio-cultural level and highest diploma attained City of residence

Biological samples

Oesophageal tissue

EGD and biopsy sampling can be conducted at any point during the follow-up of OA patients, based on symptoms and routine evaluation (24). Patients who undergo these procedures will be requested to authorize the collection of four sections of formalin-fixed paraffin-embedded (FFPE) oesophageal tissue. These samples are stored in the pathology departments for a minimum of 10 years, in accordance with French regulations (décret N°88-280 du 24 mars 1988 de l'article L.761-11 du code de la santé publique) and are commonly stored for approximately 25 years.

FFPE storage is the established standard method for long-term preservation of biological tissue in routine care, universally practiced across all participating centres. Whenever feasible (subject to patient authorization and technical feasibility), previously collected oesophageal biopsies from included patients in routine care will also be analysed or pooled for analysis. FFPE sections sourced from all participating centres will be centralized at the Biological Resource Centre of Lille University Hospital (CRB du CIC 1403 du CHU de Lille, BRIF BB0033-00030) until procurement for analysis.

Comparative analysis of omic profiles measured at different time points for the same patient will facilitate longitudinal comparisons. The objective is to include biopsies from 150 patients in the analysis.

Plasma

Plasma collection will be conducted at selected centres following consultations for adolescents who have available oesophageal biopsies. Whole blood will be drawn into

a 5mL EDTA tube between the ages of 13 and 14 years and subsequently centrifuged. Plasma aliquots will then be stored at -80°C and centralized at the Biological Resource Centre of Lille University Hospital (CRB du CIC 1403 du CHU de Lille, BRIF BB0033-00030) until procurement for analysis.

Standard Operating Procedures (SOPs) for blood and biopsy collection, storage, and shipping will be uniform across all participating centres to ensure consistency and quality control.

Data sources, collection and monitoring

Clinical data collection

Clinical data will be recorded in both a traditional paper-based Case Report Form (CRF) and an electronic Case Report Form (eCRF) using Ennov Clinical platform. Monitoring of completion will be conducted, and participants will receive reminders to complete Quality of Life (QoL) questionnaires if necessary.

The study will adhere to the guidelines set forth by the International Conference on Harmonization (ICH) Good Clinical Practice (GCP) outlined in ICH GCP Topic E6 (1996), ensuring compliance with ethical and quality standards in clinical research.

Omics data generation

Omics data will be generated from both oesophageal mucosa and plasma samples.

For proteomics analysis, FFPE tissues and plasma extracellular vesicles (EVs) will undergo antigen retrieval prior to localized trypsin digestion of the regions of interest. The resulting peptides will be collected using a liquid micro junction prior to shotgun analyses. In parallel, total RNA extraction will be performed to prepare Next Generation Sequencing (NGS) 3'RNA-seq and small RNA-seq (sRNA-seq) libraries with molecular barcodes (Unique Molecular Identifiers - UMI). The incorporation of UMIs in the sequencing libraries will enable accurate quantification even in cases of low RNA yield and quality. The sRNA-seq libraries will facilitate the detection and quantification of

non-coding RNA such as microRNA (miRNA) and small nucleolar RNA (snoRNA). All libraries will be sequenced using a NovaSeq 6000 instrument.

For plasma proteomics analysis, EVs will initially be isolated using size exclusion chromatography. Fractions will be collected and assessed for size and concentration using Nanosight analysis. EV-positive fractions will then be pooled before enzymatic digestion of proteins and subsequent identification by nano-liquid chromatography-tandem mass spectrometry (nano-LC-MS/MS). RNA extraction from EVs in the pooled fractions will be performed using the Direct-zol RNA extraction kit, followed by analysis as described above.

Omics data preprocessing

The protein identification process will utilize MaxQuant software, comparing all MS/MS data with the human protein database from the Uniprot bank. Key parameters in MaxQuant will be defined, including trypsin as the digestion enzyme with a maximum of 2 missed cleavages, methionine oxidation and N-terminal protein acetylation as variable modifications, and carbamidomethylation of cysteines as a fixed modification. Label-free quantification (LFQ) will be performed using default settings. Initial mass tolerance will be set to 6 ppm for MS mode and 20 ppm for fragmentation data in MS/MS mode. Protein and peptide identification parameters will be configured with an FDR (false discovery rate) of less than 1%, requiring a minimum of 2 peptides per protein, including at least 1 unique peptide.

Pre-processing analysis will commence using Perseus software. LFQ intensity values of each sample will be imported from MaxQuant into Perseus, and the data matrix will undergo contaminant removal. Subsequently, the data will be log2 transformed, and a filter will be applied to retain proteins present in at least 70% of samples within at least one group.

The RNA-seq analysis pipeline will involve bclconverter for fastQ generation, fastp for trimming and UMI first-step processing, fastQC for quality control, STAR for alignment,

umi-tools for deduplication, qualimap for alignment quality control, and featureCounts for quantitative assessment of genes. The sRNA-seq analysis pipeline will follow a similar workflow to RNA-seq, utilizing bowtie for alignment with necessary adjustments.

For methylation analysis, the pipeline will include bclConverter, fastQC, trim_galore, additional python scripts, bismark (25), and bowtie2 for alignment and deduplication, with qualimap utilized for alignment QC.

Statistical analysis

Sample size

The rationale for the number of patients was determined based on the cohort of patients included in the oesophageal atresia (OA) registry who will be between 13 and 14 years old during the recruitment period (n=492). Accounting for potential dropouts, refusals to participate, and deaths, it is estimated that approximately 300 patients can be included, representing around 30% of those eligible. This estimation aligns with the observed rate of non-inclusion of 22% observed in the COMAD6 study.

With a cohort of 300 patients, the study will have the capacity to estimate the frequency of GORD in adolescents (primary objective) with a maximum absolute precision of 5.7%. This precision is represented by the half-width of the 95% confidence interval, particularly in scenarios where the observed frequency of GORD is 50%, which presents the greatest width of the confidence interval for an equivalent population size.

Regarding the control group without oesophageal atresia, the plan is to include 150 patients, taking into consideration logistical and financial constraints associated with various omic analyses.

For the blood sub-study, participating centres have a potential recruitment capacity of 176 patients with oesophageal atresia (ranging from 20 to 30 per centre). Anticipating

losses to follow-up and refusals, it is estimated that approximately 150 patients can be included in this sub-study.

Phenotyping

Frequency of GORD

The frequency of patients born with a diagnosis of oesophageal atresia and GORD in adolescence will be calculated along with its two-sided 95% confidence interval. The confidence interval will be estimated using the normal approximation method.

Here's the general formula to calculate the confidence interval for a proportion (in this case, the frequency of patients with OA and GORD):

$$Confidence\ interval = Frequency\ \pm Z\ \times \sqrt{\frac{Frequency \times (1 - Frequency)}{Sample\ size}}$$

Where:

- Frequency: Proportion of patients with OA and GORD

- Z: Z-score corresponding to the desired confidence level (for a 95% confidence interval, Z ≈ 1.96)

- Sample Size: Total number of patients in the study

Given the calculated proportion of patients with OA and GORD and the total sample size, we can plug these values into the formula to obtain the confidence interval.

Associated factors with GORD

To assess the independent factors associated with the presence of GORD in adolescence, the following methodology will be employed:

1. Bivariate Logistic Regression Models:

- Bivariate logistic regression models will be conducted to evaluate the odds ratios (ORs) of GORD and their corresponding 95% confidence intervals. These models will assess the relationship between each individual factor and the presence of GORD.

- For quantitative factors, the log-linearity assumption will be assessed using restricted cubic spline functions (26)

2. Selection of Factors for Multivariable Analysis:

- Factors associated with the presence of GORD with a significance level of less than 0.20 in bivariate analyses will be included in a multivariable backward-stepwise logistic regression model.

- If the number of events per candidate variable is insufficient (<10), a penalized method will be utilized.

3. Collinearity Assessment:

- Collinearity between the candidate factors for multivariate analysis will be examined by calculating the variance inflation factor (VIF). An alert threshold will be set at 2.5.

4. Development and Evaluation of Multivariable Model:

- Discrimination of the selected multivariable model, indicating its ability to differentiate between patients with and without GORD, will be assessed using C-statistics corrected for over-optimism via bootstrap resampling (26).

- Calibration, which measures the agreement between predicted and observed probabilities of GORD, will be evaluated using the Hosmer-Lemeshow goodness of fit test.

5. Handling Data:

- To prevent case deletion due to missing data in multivariable analysis, missing data will be imputed.

- Simple imputation will be used if the missing data rate is below 10%, while multiple imputations (27) will be employed otherwise.

- The number of imputations (m) will be determined based on the fraction of missing information (FMI).

- Imputations will be performed using the Multiple Imputation by Chained Equations (MICE) procedure, incorporating all variables included in the analyses (28). Missing data on quantitative variables will be imputed using the predictive mean matching method, while those on qualitative variables will be imputed using logistic regression models (binomial, ordinal or multinomial, depending on the number and order of modalities). In the case of multiple imputations, Rubin's rules will be applied to combine estimates obtained from each imputed dataset (29) in the case of multiple imputations.

Health status

The health status and quality of life assessment criteria in adolescence will be described using the following parameters:

1. Quality of Life Scores:

- Quality of life scores obtained from the PedsQL and the EA-QoL questionnaires will be reported.

2. Anthropometric Measures:

- Weight

- Height

- Weight/Height Z-score

- BMI Z-score

3. Frequency of Respiratory Complications

For each of these criteria, positional parameters, such as means or medians, will be calculated, and their two-sided 95% confidence intervals will be reported.

The confidence intervals will provide a range of plausible values for these parameters, allowing for an understanding of the precision of the estimates. These intervals will be calculated using appropriate statistical methods, such as the t-distribution for means or the bootstrap method for medians, depending on the distributional characteristics of the data.

Omic analysis

Omic profile between 13 and 14 years

For each type of omics data, differential analyses between the group with OA and the group without OA will be conducted using appropriate R packages. For example:

- Proteomic data will be analysed using the limma package (30).

- Sequencing data will be analysed using DESeq2 (31).

- Methylation levels extracted from bismark will be analysed using packages such as methylKit (32) and RnBeads (33).

Once differential analyses are performed, the impacted metabolic pathways will be identified through enrichment analyses. This can be accomplished using R packages like ClusterProfileR (34).

Further classification analysis will be conducted using R packages such as methylClass (35) to study the evolution of omic profiles over time.

For each identified metabolic pathway, differences in means of corresponding gene expression, methylation, or protein abundance between two time points will be tested using parametric or non-parametric tests, depending on the data distribution. Parametric tests will use similar approaches as for the initial differential analyses, while non-parametric tests such as Wilcoxon signed-rank tests may be employed.

To account for multiple testing, p-values of these tests will be corrected. The Bonferroni procedure may be used initially to control the Family-Wise Error Rate, followed by the Benjamini-Hochberg procedure to control the False Discovery Rate if necessary. These corrections will help ensure the reliability of the identified associations and reduce the likelihood of false positives.

Transposition from the biopsies to the blood

Scatterplots depicting microRNA expression and protein abundance in plasma will be generated, with levels in biopsies serving as the independent variable. These scatterplots will visually illustrate the relationships between the variables. Additionally, correlation coefficients will be calculated to quantify the strength and direction of these relationships, providing numerical measures of association.

Linear regression analysis will be conducted to further explore the relationships between microRNA expression, protein abundance in plasma, and levels in biopsies. This analysis will involve fitting a linear model to the data to determine the extent to which changes in one variable predict changes in another. The residuals from the

regression model will be examined to assess the adequacy of the model fit and identify any potential outliers or patterns in the data that may require further investigation.

Overall, these analyses will provide insights into the relationships between microRNA expression, protein abundance in plasma, and levels in biopsies, helping to uncover potential biomarkers or indicators of interest in the context of oesophageal atresia.

Ethics and dissemination

The study (protocol versions 1.1 and 2.0) have obtained approval from the French ethics research committee.

Informed consent

Patients and their parents will receive prior communication before the scheduled follow-up visit, providing them with information about the opportunity to participate in the study. The study protocol will be thoroughly explained, and an information note will be sent to them to allow ample time for consideration, with a minimum duration of one week.

On the day of the visit, any queries or concerns they may have will be addressed comprehensively. Subsequently, they will have the option to confirm or decline their participation in the study, based on their informed decision.

For the control group, eligible patients will be contacted to invite their participation in the study. A letter of non-opposition will be sent to obtain their agreement for the utilization of samples collected during routine care, which will be repurposed for research purposes.

Privacy

The TransEAsome team will prioritize the pseudo-anonymization of all clinical and omics data, ensuring that individual identities are solely identifiable by the patient inclusion number. Access to the data will be restricted, with each partner granted limited access based on their specific requirements for task performance.

The project will adhere to the guidelines established by the Commission Nationale de l'Informatique et des Libertés (CNIL), the French data protection authority, particularly following standard methodology 001 for clinical research involving the collection of patient consent. These measures aim to uphold patient confidentiality and privacy while facilitating valuable research endeavours.

Dissemination plan

Data

All data collected will undergo harmonization with standard terminologies encompassing pathologies, localization, treatments, units of measure, phenotypes, and health interventions. This harmonization process aims to enhance interoperability with other databases and enable seamless reuse by diverse research teams.

The data will be centralized within the France Cohortes platform, where each entry will be assigned a unique identifier. Additionally, metadata files will be generated in accordance with FAIR principles, ensuring that the data remains Findable, Accessible, Interoperable, and Reusable.

Results

The dissemination of findings from the TransEAsome project will target various audiences to ensure broad impact and relevance. These audiences include:

1. Scientific Community: Findings will be disseminated through peer-reviewed scientific publications and conference presentations, allowing researchers and clinicians to access and utilize the latest advancements in the field.

2. Research Participants: Participants in the study will be informed of the results through participant newsletters, ensuring that they remain informed about the outcomes of the research in which they contributed.

3. Patient Associations: Results will be shared with patient associations to empower patients and their families with knowledge about advancements in the understanding and management of oesophageal atresia.

4. General Public: Information about the project's findings will be communicated to the general public through social media platforms such as Twitter and LinkedIn, raising awareness and understanding of oesophageal atresia among the wider community.

5. Regulatory Authorities and Policy-Makers: Relevant findings will be shared with regulatory authorities and policy-makers to inform decision-making processes and contribute to the development of policies related to rare diseases and paediatric healthcare.

Communication efforts will adhere to the graphic charter of the project, ensuring consistency and professionalism in visual presentation. Additionally, all communications will acknowledge the funding provided by the Agence Nationale de la Recherche, as required. This comprehensive dissemination strategy aims to maximize the impact of the TransEAsome project across various stakeholders and promote the translation of research findings into tangible benefits for patients and society.

References

1. van Lennep M, Singendonk MMJ, Dall'Oglio L, Gottrand F, Krishnan U, Terheggen-Lagro SWJ, et al. Oesophageal atresia. Nat Rev Dis Primers. 2019 Apr 18;5(1):1–21.

2. Sfeir R, Michaud L, Sharma D, Richard F, Gottrand F. National Esophageal Atresia Register. Eur J Pediatr Surg. 2015 Dec;25(6):497–9.

3. Koivusalo AI, Sistonen SJ, Lindahl HG, Rintala RJ, Pakarinen MP. Long-term outcomes of oesophageal atresia without or with proximal tracheooesophageal fistula - Gross types A and B. J Pediatr Surg. 2017 Oct;52(10):1571–5.

4. Svoboda E, Fruithof J, Widenmann-Grolig A, Slater G, Armand F, Warner B, et al. A patient led, international study of long term outcomes of esophageal atresia: EAT 1. J Pediatr Surg. 2018 Apr;53(4):610–5.

5. Krishnan U, Mousa H, Dall'Oglio L, Homaira N, Rosen R, Faure C, et al. ESPGHAN-NASPGHAN Guidelines for the Evaluation and Treatment of Gastrointestinal and Nutritional Complications in Children With Esophageal Atresia-Tracheoesophageal Fistula. J Pediatr Gastroenterol Nutr. 2016 Nov;63(5):550–70.

6. Dimitrov G, Aumar M, Duhamel A, Wanneveich M, Gottrand F. Proton pump inhibitors in esophageal atresia: A systematic review and meta-analysis. J Pediatr Gastroenterol Nutr. 2024 Jan 23;

7. Schneider A, Gottrand F, Bellaiche M, Becmeur F, Lachaux A, Bridoux-Henno L, et al. Prevalence of Barrett Esophagus in Adolescents and Young Adults With Esophageal Atresia. Ann Surg. 2016 Dec;264(6):1004–8.

8. Vergouwe FW, Gottrand M, Wijnhoven BP, IJsselstijn H, Piessen G, Bruno MJ, et al. Four cancer cases after esophageal atresia repair: Time to start screening the upper gastrointestinal tract. World J Gastroenterol. 2018 Mar 7;24(9):1056–62.

9. Aumar M, Audrey N, Sfeir R, Seguy D, Gottrand F. Long term digestive outcome of œsophageal atresia. Best Practice & Research Clinical Gastroenterology. 2022 Feb 1;56–57:101771.

10. Lardenois E, Michaud L, Schneider A, Onea M, Rebeuh J, Gottrand-Aumar M, et al. Prevalence of Eosinophilic Esophagitis in Adolescents With Esophageal Atresia. J Pediatr Gastroenterol Nutr. 2019 Jul;69(1):52–6.

11. Traini I, Menzies J, Hughes J, Leach ST, Krishnan U. Oesophageal atresia: The growth gap. World J Gastroenterol. 2020 Mar 28;26(12):1262–72.

12. Dittrich R, Stock P, Rothe K, Degenhardt P. Pulmonary outcome of esophageal atresia patients and its potential causes in early childhood. J Pediatr Surg. 2017 Aug;52(8):1255–9.

13. Dellenmark-Blom M, Quitmann J, Dingemann C. Health-Related Quality of Life in Patients after Repair of Esophageal Atresia: A Review of Current Literature. Eur J Pediatr Surg. 2020 Jun;30(3):239–50.

14. Sy MR, Chauhan J, Prescott K, Imam A, Kraus A, Beleza A, et al. Exome sequencing efficacy and phenotypic expansions involving esophageal atresia/tracheoesophageal fistula plus. Am J Med Genet A. 2022 Sep 22;

15. Zhong G, Ahimaz P, Edwards NA, Hagen JJ, Faure C, Lu Q, et al. Identification and validation of candidate risk genes in endocytic vesicular trafficking associated with esophageal atresia and tracheoesophageal fistulas. HGG Adv. 2022 Jul 14;3(3):100107.

16. Thiem CE, Stegmann JD, Hilger AC, Waffenschmidt L, Bendixen C, Köllges R, et al. Re-sequencing of candidate genes FOXF1, HSPA6, HAAO, and KYNU in 522 individuals with VATER/VACTERL, VACTER/VACTERL-like association, and isolated anorectal malformation. Birth Defects Res. 2022 Jun;114(10):478–86.

17. Gehlen J, Giel AS, Köllges R, Haas SL, Zhang R, Trcka J, et al. First genome-wide association study of esophageal atresia identifies three genetic risk loci at CTNNA3, FOXF1/FOXC2/FOXL1, and HNF1B. HGG Adv. 2022 Apr 14;3(2):100093.

18. Brosens E, Brouwer RWW, Douben H, van Bever Y, Brooks AS, Wijnen RMH, et al. Heritability and De Novo Mutations in Oesophageal Atresia and Tracheoesophageal Fistula Aetiology. Genes (Basel). 2021 Oct 10;12(10):1595.

19. Ten Kate CA, de Klein A, de Graaf BM, Doukas M, Koivusalo A, Pakarinen MP, et al. Intrinsic Cellular Susceptibility to Barrett's Esophagus in Adults Born with Esophageal Atresia. Cancers (Basel). 2022 Jan 20;14(3):513.

20. Maltby EL, Dyson MJ, Wheeler MR, Thomson M, Sethuraman C, Cohen MC. Molecular abnormalities in pediatric barrett esophagus: can we test for potential of neoplastic progression? Pediatr Dev Pathol. 2010 Aug;13(4):310–7.

21. Krishnan U, Lijuan C, Andrew GJ, Rothenberg ME, Wen T. Analysis of eosinophilic esophagitis in children with repaired congenital esophageal atresia. J Allergy Clin Immunol. 2019 Apr;143(4):1455-1464.e2.

22. Gaillard S, Malik S, Preston J, Escalera BN, Dicks P, Touil N, et al. Involving children and young people in clinical research through the forum of a European Young Persons' Advisory Group: needs and challenges. Fundam Clin Pharmacol. 2018 Aug;32(4):357–62.

23. Preston J, Nafria B, Ohmer A, Gaillard S, Dicks P, West L, et al. Developing a More Tailored Approach to Patient and Public Involvement with Children and Families in Pediatric Clinical Research: Lessons Learned. Ther Innov Regul Sci. 2022 Nov;56(6):948–63.

24. Protocole national de diagnostic et de soins (PNDS) : atrésie de l'œsophage. Perfectionnement en Pédiatrie. 2019 Jun;2(2):98–115.

25. Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics. 2011 Jun 1;27(11):1571–2.

26. Harrell FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med. 1996 Feb 28;15(4):361–87.

27. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol. 1996 Dec;49(12):1373–9.

28. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med. 2011 Feb 20;30(4):377–99.

29. Rubin. Multiple Imputation for Nonresponse in Surveys. In: Multiple Imputation for Nonresponse in Surveys [Internet]. John Wiley & Sons, Ltd; 1987 [cited 2022 Dec 9]. p. i–xxix. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470316696.fmatter

30. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015 Apr 20;43(7):e47.

31. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology. 2014 Dec 5;15(12):550.

32. Wreczycka K, Gosdschan A, Yusuf D, Grüning B, Assenov Y, Akalin A. Strategies for analyzing bisulfite sequencing data. Journal of Biotechnology. 2017 Nov 10;261:105–15.

33. Müller F, Scherer M, Assenov Y, Lutsik P, Walter J, Lengauer T, et al. RnBeads 2.0: comprehensive analysis of DNA methylation data. Genome Biology. 2019 Mar 14;20(1):55.

34. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. The Innovation. 2021 Aug 28;2(3):100141.

35. methylClass: an R package to construct DNA methylation-based classification models | Briefings in Bioinformatics | Oxford Academic [Internet]. [cited 2024 Feb 1]. Available from: https://academic.oup.com/bib/article/25/1/bbad485/7513599

Acknowledgements

The authors express their gratitude to the scientific board of the *TransEAsome* project for their invaluable input in shaping its design. They also extend their appreciation to all investigation teams for their significant contributions to the establishment and execution of this study. Special thanks are extended to Katialine Groff and Dr. Rony Sfeir for their dedicated work on the oesophageal atresia national registry ReNaTo, which has been instrumental in the project's progress.

The authors also acknowledge the members of the user's committees for their insightful feedback, which has consistently enriched the project. Additionally, they express their gratitude to the paediatric clinical investigation centre at Lille Hospital for their technical support in clinical research, which has been indispensable to the authors.

This collaborative effort and support from various individuals and institutions have been essential in advancing the TransEAsome project and its objectives.

Collaborators

University hospital of Amiens, University hospital of Angers, University hospital of Besançon, University hospital of Bordeaux, University hospital of Brest, University hospital of Caen, Public hospital of Cayenne, University hospital of Clermont Ferrand, University hospital of Dijon, University hospital of Fort de France, University hospital of La Tronche, Public hospital of Le Mans, University hospital of Limoges, University hospital of Lyon, University hospital of Marseille, University hospital of Montpellier, University hospital of Nancy, University hospital of Nantes, University hospital of Nice, Public hospital of Orléans, University hospital of Paris Armand Trousseau, University hospital of Kremelin Bicêtre, University hospital of Créteil, University hospital of Paris Necker, University hospital of Paris Robert Debré, University hospital of Poitiers, University hospital of Pointe à Pitre, University hospital of Reims, University hospital of Rennes, University hospital of Rouen, University hospital of Saint Denis de la Réunion,

University hospital of Saint Etienne, University hospital of Strasbourg, University hospital of Toulouse and University hospital of Tours.

Authors' contributions

FG is the scientific coordinator. MF, RH, GM, MS are partner leaders in the project. MA is the Lille University Hospital's principal investigator, JV is the co-technical-head of the Billile platform, SG is head of the patient and public involvement task, JB is the head of the Lille University Hospital's biostatistics department, LD is a methodologist who revised the protocol submitted to ethics committee, MD is in charge of the omics analysis within the PRISM team, SF is a head of the ReNaTo registry, JR is a member of the French national OA patient association and ML is the project manager. All authors have contributed to the writing of the study protocol and have approved the final version of this manuscript.

Funding statement

This project has received funding from the ANR program France 2030 under the specific grant agreement ANR-21-PMRB-0011. As well as a funding from the Institution de prévoyance et Retraite Collective des Employés de Maison (IRCEM) fundation.

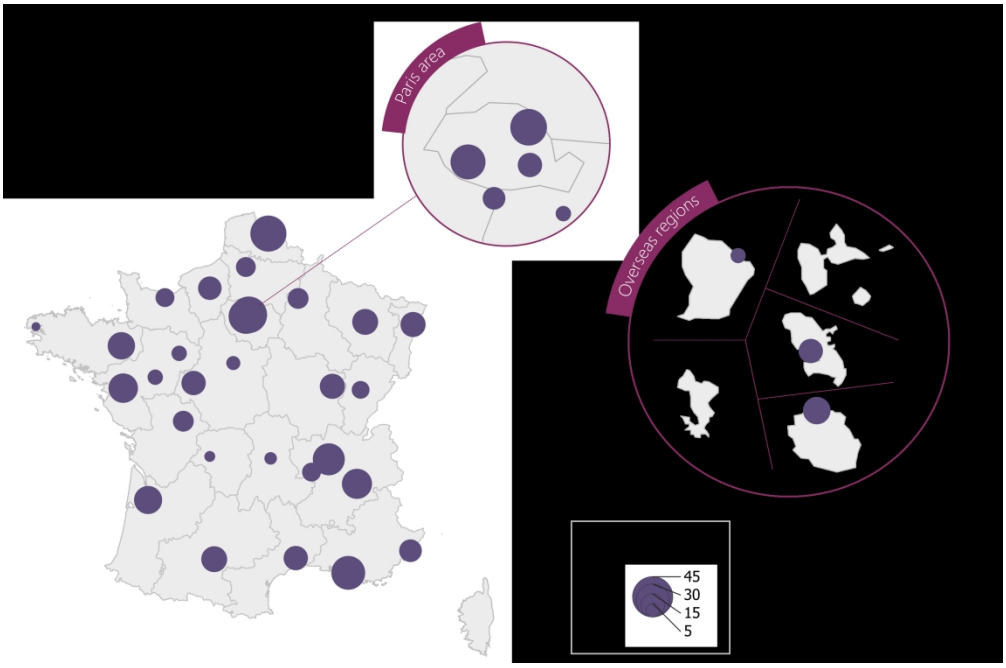
Competing interest statement

No competing interests.

Word count

4840 words

Open Access



Localisation of recruitment in France

619x408mm (130 x 130 DPI)

Follow-up of a nested cohort of OA patients born between 2010 and 2012

426x99mm (59 x 59 DPI)

BMJ Open

Long-term outcome of oesophageal atresia in adolescence (TransEAsome): a cohort study protocol

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2024-086303.R1
Article Type:	Protocol
Date Submitted by the Author:	22-Oct-2024
Complete List of Authors:	Leroy, Mélanie; Univ. Lille, Inserm, CHU Lille, U1286 - INFINITE - Institute for Translational Research in Inflammation, F-59000 Lille, France; CHU Lille, Reference Center for Congenital Oesophageal Anomalies, F-59000 Lille, France Aumar, Madeleine; Univ. Lille, Inserm, CHU Lille, U1286 - INFINITE - Institute for Translational Research in Inflammation, F-59000 Lille, France; Univ. Lille, CHU Lille, Infinite U1286, Reference Center for Congenital Oesophageal Anomalies , F-59000 Lille, France Duhamel, Marie; Univ. Lille, Inserm, CHU Lille, U1192 - Protéomique Réponse Inflammatoire et Spectrométrie de Masse - PRISM, F-59000 Lille, France Dauchet, Luc; CHU Lille, Department of Biostatistics, F59000-Lille, France Figeac, Martin; Univ. Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, US 41 - UAR 2014 - PLBS, F-59000 Lille, France Gaillard, Segolene ; Hospices civils de Lyon, EPICIME-CIC 1407 de Lyon, Kids France, Inserm 1407, CHU-Lyon, 69677 Bron, France ; Université Lyon 1, CNRS UMR 5558, laboratoire de biométrie et biologie évolutive, 69622 Villeurbanne, France Hankard, Regis; French Clinical Research Infrastructure Network (F-CRIN)-PEDSTART, INSERM, University of Tours, Tours, France Labreuche, Julien; CHU Lille, Department of Biostatistics, F59000-Lille, France Marot, Guillemette; Univ. Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, US 41 - UAR 2014 - PLBS, F-59000 Lille, France; Univ. Lille, CHU Lille, Inria, ULR 2694 - METRICS : Évaluation des technologies de santé et des pratiques médicales, F-59000 Lille, France Reversat, Julie; Association Française de l'Atrésie de l'Œsophage Armand, Viviane; Association Française de l'Atrésie de l'Œsophage Salzet, Michel; Univ. Lille, Inserm, CHU Lille, U1192 - Protéomique Réponse Inflammatoire et Spectrométrie de Masse - PRISM, F-59000 Lille, France Sfeir, Rony; Univ. Lille, CHU Lille, Infinite U1286, Reference Center for Congenital Oesophageal Anomalies , F-59000 Lille, France Vandel, Jimmy; Univ. Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, US 41 - UAR 2014 - PLBS, F-59000 Lille, France Gottrand, Frédéric; Univ. Lille, Inserm, CHU Lille, U1286 - INFINITE - Institute for Translational Research in Inflammation, F-59000 Lille, France; Univ. Lille, CHU Lille, Infinite U1286, Reference Center for Congenital Oesophageal Anomalies , F-59000 Lille, France

Primary Subject Heading:	Paediatrics
Secondary Subject Heading:	Gastroenterology and hepatology, Surgery
Keywords:	Adolescents < Adolescent, Oesophageal disease < GASTROENTEROLOGY, Patient Reported Outcome Measures, Quality of Life, Paediatric gastroenterology < GASTROENTEROLOGY

Long-term outcome of oesophageal atresia in adolescence (TransEAsome): a national French cohort study protocol

Mélanie Leroy[1,2], Madeleine Aumar[1,2], Marie Duhamel[3], Luc Dauchet[4], Martin Figeac[5], Ségolène Gaillard[6], Régis Hankard[7], Julien Labreuche[4], Guillemette Marot[5,8], Julie Reversat[9], Viviane Armand[9], Michel Salzet[3], Rony Sfeir[2], Jimmy Vandel[5], Frédéric Gottrand[1,2]

[1] Univ. Lille, Inserm, CHU Lille, U1286 - INFINITE - Institute for Translational Research in Inflammation, F-59000 Lille, France

[2] CHU Lille, Reference Center for Congenital Oesophageal Anomalies, F-59000 Lille, France

[3] Univ. Lille, Inserm, CHU Lille, U1192 - Protéomique Réponse Inflammatoire et Spectrométrie de Masse - PRISM, F-59000 Lille, France

[4] CHU Lille, Department of Biostatistics, F59000-Lille, France

[5] Univ. Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, US 41 - UAR 2014 - PLBS, F-59000 Lille, France

[6] Hospices civils de Lyon, EPICIME-CIC 1407 de Lyon, Kids France, Inserm 1407, CHU-Lyon, 69677 Bron, France ; Université Lyon 1, CNRS UMR 5558, laboratoire de biométrie et biologie évolutive, 69622 Villeurbanne, France

[7] French Clinical Research Infrastructure Network (F-CRIN)-PEDSTART, INSERM, University of Tours, Tours, France

[8] Univ. Lille, CHU Lille, Inria, ULR 2694 - METRICS : Évaluation des technologies de santé et des pratiques médicales, F-59000 Lille, France

[9]Association Française de l'Atrésie de l'Œsophage, Maison Alfort, France

Corresponding author:

Pr Frédéric Gottrand

CENTRE HOSPITALIER UNIVERSITAIRE

Hôpital Jeanne de Flandre

Pôle Enfant

2 avenue Oscar Lambret - 59037 Lille Cedex, France

+33 3 20 44 61 26

Frederic.gottrand@chu-lille.fr

Abstract

Introduction

The *TransEAsome* project, funded by the Agence Nationale de la Recherche, aims to evaluate the long-term outcomes of patients with oesophageal atresia (OA) between 13 and 14 years old and establish multi-omics profiles using data from the world biggest OA registry.

Methods and analysis

TransEAsome is a national multicentre population-based cohort study recruiting participants from all qualified French centre for oesophageal atresia surgery at birth. The primary objective is to assess the prevalence of gastro-oesophageal reflux disease in adolescence among patients with OA, with several secondary objectives including the identification of risk factors and multi-omic profiles from oesophageal biopsies and blood samples collected between 13 and 14 years old, compared with a control group. This comprehensive characterisation of phenotype and omic profiles aims to enhance understanding of the patient with OA evolution and inform tailored care management strategies.

Ethics and dissemination

The study, co-constructed with input from patients, parents and research-expert adolescents, has obtained approval from the ethics research committee: Comité de protection des personnes Est II. Findings will be disseminated to various target audiences, including the scientific community, research participants, patient community, the general public, regulatory authorities, and policymakers. Data will be made available in a FAIR format on the France Cohortes platform upon study completion.

Trial registration number

NCT05995171

Strengths and limitations of this study

- Multicentre, national, population-based cohort study conducted in all French centres qualified to perform oesophageal atresia surgery at birth,
- Longitudinal analysis facilitated by a nested cohort from the birth/1-year-old register (ReNaTo) and the 6-year-old cohort (COMAD6),
- Active involvement of patients and parents thanks to the users' representative committee, ensuring the study reflects patient perspectives,
- Comprehensive multi-omic analysis encompassing transcriptomic, epigenetic, proteomic, and metabolomic data, providing a holistic understanding of the disease.
- Omic studies rely on oesophageal tissues samples collected for clinical surveillance purposes, which may not be systematic, and long-term storage could lead to RNA degradation over time, potentially impacting data quality.

Keywords

Oesophageal atresia, adolescence, gastro-oesophageal reflux, dysphagia, oesophageal cancer, eosinophilic esophagitis, outcome, epigenetic, proteogenomic metabolomics, quality of life

Introduction

Oesophageal atresia (OA) is a rare congenital anomaly affecting approximately 160 newborns annually in France, characterized by a discontinuity in the oesophagus. It is often concomitant with other malformations in over 50% of cases [1]. Since the early 1950s, significant advancements in surgical techniques have vastly improved the management of this condition, resulting in a favourable prognosis. While the mortality rate has declined to less than 7% in developed nations like France [2], morbidity remains noteworthy [3,4].

Gastro-oesophageal reflux disease (GORD) is more prevalent among individuals with OA compared to the general population [5]. Studies indicate that approximately 30% of patients with OA experience GORD at least once in their lifetime [6]. Complications of GORD, such as peptic esophagitis and Barrett's oesophagus, present heightened risks for oesophageal cancer [7], with documented cases of adenocarcinoma or squamous cell carcinoma in young adults with OA [8].

Dysphagia is a common concern among patients with OA, with around 45% of five-year-olds experiencing swallowing difficulties attributed to anastomotic strictures or oesophageal dysmotility [9]. Eosinophilic oesophagitis (EoE), a rising concern in oesophageal health, has been noted to occur more frequently in patients with OA, contributing to dysphagia [10].

Challenges in eating often lead to undernutrition, affecting up to one-third of children with EA by the age of five [11]. Respiratory complications are also prevalent [12], particularly in the early years, significantly impacting long-term quality of life (QoL) [13].

To better understand the long-term outcomes and predict complications, our objective is to investigate a nested population cohort of adolescents born with OA. Leveraging the existing national population-based registry (ReNaTo), and the nested clinical cohort, COMAD6, consisting of 382 children born between 2010 and 2012, we aim to

compile a new database comprising clinical data, blood samples, and oesophageal biopsies obtained during routine follow-up between 13 and 14 years old (Figure 1).

This nested cohort, within the world's largest registry of more than 2300 patients with OA, offers a unique opportunity to answer questions about outcomes of OA in adolescence in terms of morbidity and health status, to improve care and follow-up and prevent long-term complications.

In addition to the phenotypic assessment, our investigations extend to exploring the biological mechanisms underpinning complication development through genomic, transcriptomic, proteomic, epigenetic, and metabolomic analyses of the oesophagus.

While upper gastrointestinal endoscopy (EGD) is standard in the care of symptomatic or not patients with OA [5], its invasive nature necessitates hospitalization and general anaesthesia in children.

To date, there is only very few data on the epigenetic, genomic [14–18], or transcriptomic [19] profile of patients with OA. In a small study of 10 children with Barrett's oesophagus (some of them operated at birth for OA), the authors performed fluorescence in situ hybridization with probes on 4-micron sections taken from sequential paraffin-embedded biopsies and identified 4 probe sets reported to be associated to adult Barret adenocarcinoma [19]. Genetic markers were also identified in adult Barrett's adenocarcinoma patients. This preliminary study shows that, even at an early age, Barrett's may show genetic changes associated with neoplastic progression [20]. Another recent study examined the relationship between eosinophilic esophagitis (EoE) and OA by profiling the transcriptional signature of EoE [21]. Using an in-silico approach, they found 6 genes that were differentially expressed between the 2 entities (OA+ EoE+ and OA- EOE+). Two of them were associated with majored dysphagia, the development of strictures, and the need for dilatations in patients with OA+EoE+.

As EGD and oesophageal biopsies are invasive procedures that usually require hospitalization and general anaesthesia in children, using a more accessible biological sample as blood plasma will also be investigated to see if the markers of the oesophagus translate into blood.

The overarching goal of *TransEAsome* is to construct a vast longitudinal database including phenomenal and exposomal patient data amalgamating clinical, biological, environmental, and lifestyle data alongside multimodal omics analyses: RNA abundance, protein abundance, proteogenomic profiling, protein modification, metabolite abundance, methylated sites). This innovative approach holds promise for enhancing the prediction of adult OA outcomes, including the identification of risk factors for future health complications such as oesophageal cancer and EoE.

Objectives

TransEAsome will address the evolution of patients with OA at the time of adolescence by establishing a unique and comprehensive database that integrates clinical and omics data in a structured and interoperable format to evaluate their health status and quality of life. This project is designed to achieve the following objectives:

1. Evaluate the prevalence of GORD during adolescence in the OA population.

2. Identify factors associated with GORD during adolescence in the OA population.

3. Assess the quality of life, nutritional status, and frequency of respiratory complications in the OA population.

4. Compare the single and multi-omic profiles derived from oesophageal biopsies in adolescents operated at birth for OA with those who did not, aiming to delineate alterations in biological pathways specific to OA.

5. Investigate changes in the expression of omics variables within the identified biological pathways over time in the OA population, utilizing data obtained from EGD and biopsies.

6. Evaluate the correlation between the expression levels of microRNAs and proteins in plasma compared to oesophageal biopsies in the OA population, utilizing data from individuals who underwent EGD, biopsies, and blood sampling at ages 13 and 14.

Methods and analysis

Recruitment sources

Patients for the *TransEAsome* study will be recruited from the French national registry for OA known as "ReNaTo," which involves all 35 OA competence centres across France. As of the conclusion of 2022, more than 2300 patients were included in ReNaTo. Among them, 492 are projected to reach 13 or 14 years of age during the recruitment period of the *TransEAsome* project. However, an anticipated loss to follow-up necessitates accounting for a pool of 300 patients. Leveraging the national register ReNaTo, the project management team will furnish each centre with a pseudo-anonymized patient roster, facilitating the identification of eligible candidates from their patient cohorts. Given the unique demographic characteristics of each participating centre, recruitment potentials vary, spanning from 1 to 34 patients per centre during the study period (Figure 2).

For the objectives involving multi-omic profiling, the control group will be drawn from six selected centres. Eligible patients will be identified through the extraction of cases meeting the study's inclusion criteria from the pathology department.

Study population

TransEAsome eligibility and exclusion criteria are listed in Table 1. The initial recruitment period will start in September 2023 and end in May 2026. Our goal will be to enrol 300 patients and 150 controls.

Table 1: Eligibility and exclusion criteria

Patients with oesophageal atresia	
Eligibility criteria	• Born with OA in France (metropolitan or oversea) • Underwent oesophageal anastomosis at one of the participating centres • Included into the ReNaTo registry • Age between 13 and 14 years during the inclusion period. • Willingness to participate. • Patient with social security Blood sub-study: • Upper GI endoscopy performed as part of care between 13 and 14 years of age with oesophageal mucosal biopsy sampling • Consent for blood sampling.
Exclusion criteria	• Participation in an interventional trial (simultaneously or up to 3 months prior to inclusion. • Oesophageal replacement
Controls	
Eligibility criteria	• Upper GI endoscopy performed as part of care between 10 and 14 years of age with oesophageal mucosal biopsy sampling • Underwent esophagogastroduodenoscopy (EGD) as part of routine care for any digestive symptom to exclude organic aetiology (peptic esophagitis, gastritis, eosinophilic esophagitis, or ulcer). • EGD findings and histology are normal. • Absence of concurrent progressive chronic disease. • Provision of signed study consent.
Exclusion criteria	• Known underlying disease

The blood substudy as well as the inclusion of controls will only be performed in Lille, Paris Necker, Paris Robert Debré, Lyon, Grenoble, and Marseille for logistic, feasibility and cost saving reasons.

Patient and Public Involvement (PPI)

A young patient advisory committee and a parents/patient representativeness committee have been established to ensure the active involvement of young patients and their parents throughout all phases of the *TransEAsome* project, from inception to dissemination of results. These committees have played a pivotal role in shaping the study protocol, including determining the data to be collected and reviewing informed assents and consents before regulatory submissions. Throughout the trial, these committees will contribute to the decision-making process of the steering committee and provide insights from a patient perspective on the conduct of the study.

The committees will remain actively engaged throughout the project, facilitating communication and dissemination efforts to the scientific community, patients, and the general public regarding key milestones, actions, and emerging knowledge. Upon project completion, the committees will review the analysis, discuss the findings, and aid in the dissemination of results.

Kids France (Hospices Civils de Lyon, Pedstart) will oversee the patient and public involvement (PPI) activities in collaboration with the French patient with OA support group (AFAO). Kids France will ensure that PPI activities are not only meaningful for the research but also beneficial for the patients themselves. Training related to PPI activities will be provided to the committees by Kids France. Reporting of activities to the steering committee, feedback on suggestions from patients and parents, and assessment of impact are integral components of the *TransEAsome* PPI process [22,23].

Outcomes and assessments

Outcomes and assessments are presented in with separate forms for parents and adolescent patients.

Table 2. Quality of life (QoL) questionnaires and blood sampling are the only assessments that are not part of routine practice. Both the generic paediatric QoL assessment (Pediatric Quality of Life Inventory™ (PedsQL) and the disease specific QoL questionnaire (EA QoL) [13] will be administred in this cohort, with separate forms for parents and adolescent patients.

Table 2: Outcomes and assessments between 13 and 14-year-old

Primary outcome	GORD will be determined based on the following criteria: • Positive pH-(impedance)metry results obtained within the previous year. • And/or presence of histological peptic esophagitis lesions at oesophageal biopsies collected during the previous year. • And/or history of anti-reflux surgery.
Secondary outcomes	Quality of life: • PedsQL total patient and parent scores • EA QoL total patient and parent scores Nutritional status: • Z-score Weight/Height and Height/Age • Type of feeding (oral, enteral, both) (Food avoidance) Digestive status: • Current proton pump inhibitor treatment • Dysphagia (defined by sensations such as blockage leading to vomiting or the need to drink to pass food, or slowness in eating) • GORD symptoms (regurgitation, vomiting, retrosternal pain, and heartburn) Respiratory status: • Frequency of cough • Asthma • Occurrence of exercise-induced symptoms (cough, dyspnoea) • Atopy

	• Wheezing
	• Stridor
	• Need for medications (corticosteroid, inhaled medications)
	
	History from 6 years to adolescence:
	• Associated malformations discovered after 6 years
	• Gastrostomy placement and/or use after 6 years old
	• Characteristics of eventual GORD surgery (type, complications, cardial dilatation, relapse)
	• Oesogastroduodenoscopy (number, histology)
	• Repermeabilization of the oesotracheal fistula (treatment, number, date)
	• Stenosis (date, dilatation, number, surgery)
	• Stricture (date, number and method of dilatation, surgery, use of corticosteroid)
	
	Psychological or psychiatric follow-up
	
	Neuro orthopaedic outcomes:
	• Scoliosis
	• Kyphosis
	• Stature malposition
	
	Speech therapy:
	• Oral/stomatology
	• Oral disorders
	• Speech disorders
	• Dental conditions
	
	Education:
	• Physical/sport waiver
	• Adapted school rhythm
	• Specialized school
	• School absenteeism
	
	Omic profile:
	• In oesophageal biopsies:
	  ○ Protein abundances
	  ○ Metabolite abundances
	  ○ RNA profile
	  ▪ 3'RNA
	  ▪ miRNA
	  ○ Methylation status
	• In blood:

	○ Metabolite abundance ○ Protein abundance ○ miRNA profile
Other assessments	Parents' height Parents' socio-cultural level and highest diploma attained City of residence

Biological samples

Oesophageal tissue

EGD and biopsy sampling can be conducted at any point during the follow-up of patients with OA, based on symptoms and routine evaluation [24]. Patients who undergo these procedures will be requested to authorize the collection of four sections of formalin-fixed paraffin-embedded (FFPE) oesophageal tissue. These samples are stored in the pathology departments for a minimum of 10 years, in accordance with French regulations (décret N°88-280 du 24 mars 1988 de l'article L.761-11 du code de la santé publique) and are commonly stored for approximately 25 years.

FFPE storage is the established standard method for long-term preservation of biological tissue in routine care, universally practiced across all participating centres. Whenever feasible (subject to patient authorization and technical feasibility), previously collected oesophageal biopsies from included patients in routine care will also be analysed or pooled for analysis. FFPE sections sourced from all participating centres will be centralized at the Biological Resource Centre of Lille University Hospital (CRB du CIC 1403 du CHU de Lille, BRIF BB0033-00030) until procurement for analysis.

Comparative analysis of omic profiles measured at different time points for the same patient will facilitate longitudinal comparisons. The objective is to include biopsies from 150 patients in the analysis.

Plasma

Plasma collection will be conducted at selected centres following consultations for adolescents who have available oesophageal biopsies. Whole blood will be drawn into

a 5mL EDTA tube between the ages of 13 and 14 years and subsequently centrifuged. Plasma aliquots will then be stored at -80°C and centralized at the Biological Resource Centre of Lille University Hospital (CRB du CIC 1403 du CHU de Lille, BRIF BB0033-00030) until procurement for analysis.

Standard Operating Procedures (SOPs) for blood and biopsy collection, storage, and shipping will be uniform across all participating centres to ensure consistency and quality control.

Data sources, collection and monitoring

Clinical data collection

Clinical data will be recorded in both a traditional paper-based Case Report Form (CRF) and an electronic Case Report Form (eCRF) using Ennov Clinical platform. Monitoring of completion will be conducted, and participants will receive reminders to complete Quality of Life (QoL) questionnaires if necessary.

The study will adhere to the guidelines set forth by the International Conference on Harmonization (ICH) Good Clinical Practice (GCP) outlined in ICH GCP Topic E6 (1996), ensuring compliance with ethical and quality standards in clinical research.

Omics data generation

Omics data will be generated from both oesophageal mucosa and plasma samples.

For proteomics analysis, FFPE tissues and plasma extracellular vesicles (EVs) will undergo antigen retrieval prior to localized trypsin digestion of the regions of interest. The resulting peptides will be collected using a liquid micro junction prior to shotgun analyses. In parallel, total RNA extraction will be performed to prepare Next Generation Sequencing (NGS) 3'RNA-seq and small RNA-seq (sRNA-seq) libraries with molecular barcodes (Unique Molecular Identifiers - UMI). The incorporation of UMIs in the sequencing libraries will enable accurate quantification even in cases of low RNA yield and quality. The sRNA-seq libraries will facilitate the detection and quantification of

non-coding RNA such as microRNA (miRNA) and small nucleolar RNA (snoRNA). All libraries will be sequenced using a NovaSeq 6000 instrument.

For plasma proteomics analysis, EVs will initially be isolated using size exclusion chromatography. Fractions will be collected and assessed for size and concentration using Nanosight analysis. EV-positive fractions will then be pooled before enzymatic digestion of proteins and subsequent identification by nano-liquid chromatography-tandem mass spectrometry (nano-LC-MS/MS). RNA extraction from EVs in the pooled fractions will be performed using the Direct-zol RNA extraction kit, followed by analysis as described above.

Omics data preprocessing

The protein identification process will utilize MaxQuant software, comparing all MS/MS data with the human protein database from the Uniprot bank. Key parameters in MaxQuant will be defined, including trypsin as the digestion enzyme with a maximum of 2 missed cleavages, methionine oxidation and N-terminal protein acetylation as variable modifications, and carbamidomethylation of cysteines as a fixed modification. Label-free quantification (LFQ) will be performed using default settings. Initial mass tolerance will be set to 6 ppm for MS mode and 20 ppm for fragmentation data in MS/MS mode. Protein and peptide identification parameters will be configured with an FDR (false discovery rate) of less than 1%, requiring a minimum of 2 peptides per protein, including at least 1 unique peptide.

Pre-processing analysis will commence using Perseus software. LFQ intensity values of each sample will be imported from MaxQuant into Perseus, and the data matrix will undergo contaminant removal. Subsequently, the data will be log2 transformed, and a filter will be applied to retain proteins present in at least 70% of samples within at least one group.

The RNA-seq analysis pipeline will involve bclconverter for fastQ generation, fastp for trimming and UMI first-step processing, fastQC for quality control, STAR for alignment,

umi-tools for deduplication, qualimap for alignment quality control, and featureCounts for quantitative assessment of genes. The sRNA-seq analysis pipeline will follow a similar workflow to RNA-seq, utilizing bowtie for alignment with necessary adjustments.

For methylation analysis, the pipeline will include bclConverter, fastQC, trim_galore, additional python scripts, bismark [25], and bowtie2 for alignment and deduplication, with qualimap utilized for alignment QC.

Statistical analysis

Sample size

The rationale for the number of patients was determined based on the cohort of patients included in the oesophageal atresia (OA) registry who will be between 13 and 14 years old during the recruitment period (n=492). Accounting for potential dropouts, refusals to participate, and deaths, it is estimated that approximately 300 patients can be included, representing around 30% of those eligible. This estimation aligns with the observed rate of non-inclusion of 22% observed in the COMAD6 study.

With a cohort of 300 patients, the study will have the capacity to estimate the frequency of GORD in adolescents (primary objective) with a maximum absolute precision of 5.7%. This precision is represented by the half-width of the 95% confidence interval, particularly in scenarios where the observed frequency of GORD is 50%, which presents the greatest width of the confidence interval for an equivalent population size.

Regarding the control group without oesophageal atresia, the plan is to include 150 patients, taking into consideration logistical and financial constraints associated with various omic analyses.

For the blood sub-study, participating centres have a potential recruitment capacity of 176 patients with oesophageal atresia (ranging from 20 to 30 per centre). Anticipating

losses to follow-up and refusals, it is estimated that approximately 150 patients can be included in this sub-study.

Phenotyping

Frequency of GORD

The frequency of patients born with a diagnosis of oesophageal atresia and GORD in adolescence will be calculated along with its two-sided 95% confidence interval. The confidence interval will be estimated using the normal approximation method.

Here's the general formula to calculate the confidence interval for a proportion (in this case, the frequency of patients with OA and GORD):

$$Confidence\ interval = Frequency\ \pm Z\ \times \sqrt{\frac{Frequency \times (1 - Frequency)}{Sample\ size}}$$

Where:

- Frequency: Proportion of patients with OA and GORD

- Z: Z-score corresponding to the desired confidence level (for a 95% confidence interval, Z ≈ 1.96)

- Sample Size: Total number of patients in the study

Given the calculated proportion of patients with OA and GORD and the total sample size, we can plug these values into the formula to obtain the confidence interval.

Associated factors with GORD

To assess the independent factors associated with the presence of GORD in adolescence, the following methodology will be employed:

1. Bivariate Logistic Regression Models:

- Bivariate logistic regression models will be conducted to evaluate the odds ratios (ORs) of GORD and their corresponding 95% confidence intervals. These models will assess the relationship between each individual factor and the presence of GORD.

- For quantitative factors, the log-linearity assumption will be assessed using restricted cubic spline functions [26]

2. Selection of Factors for Multivariable Analysis:

- Factors associated with the presence of GORD with a significance level of less than 0.20 in bivariate analyses will be included in a multivariable backward-stepwise logistic regression model.

- If the number of events per candidate variable is insufficient (<10), a penalized method will be utilized.

3. Collinearity Assessment:

- Collinearity between the candidate factors for multivariate analysis will be examined by calculating the variance inflation factor (VIF). An alert threshold will be set at 2.5.

4. Development and Evaluation of Multivariable Model:

- Discrimination of the selected multivariable model, indicating its ability to differentiate between patients with and without GORD, will be assessed using C-statistics corrected for over-optimism via bootstrap resampling [26].

- Calibration, which measures the agreement between predicted and observed probabilities of GORD, will be evaluated using the Hosmer-Lemeshow goodness of fit test.

5. Handling Data:

- To prevent case deletion due to missing data in multivariable analysis, missing data will be imputed.

- Simple imputation will be used if the missing data rate is below 10%, while multiple imputations [27] will be employed otherwise.

- The number of imputations (m) will be determined based on the fraction of missing information (FMI).

- Imputations will be performed using the Multiple Imputation by Chained Equations (MICE) procedure, incorporating all variables included in the analyses [28]. Missing data on quantitative variables will be imputed using the predictive mean matching method, while those on qualitative variables will be imputed using logistic regression models (binomial, ordinal or multinomial, depending on the number and order of modalities). In the case of multiple imputations, Rubin's rules will be applied to combine estimates obtained from each imputed dataset [29] in the case of multiple imputations.

Health status

The health status and quality of life assessment criteria in adolescence will be described using the following parameters:

1. Quality of Life Scores:

- Quality of life scores obtained from the PedsQL and the EA-QoL questionnaires will be reported.

2. Anthropometric Measures:

- Weight

- Height

- Weight/Height Z-score

- BMI Z-score

3. Frequency of Respiratory Complications

For each of these criteria, positional parameters, such as means or medians, will be calculated, and their two-sided 95% confidence intervals will be reported.

The confidence intervals will provide a range of plausible values for these parameters, allowing for an understanding of the precision of the estimates. These intervals will be calculated using appropriate statistical methods, such as the t-distribution for means or the bootstrap method for medians, depending on the distributional characteristics of the data.

Omic analysis

Omic profile between 13 and 14 years

For each type of omics data, differential analyses between the group with OA and the group without OA will be conducted using appropriate R packages. For example:

- Proteomic data will be analysed using the limma package [30].

- Sequencing data will be analysed using DESeq2 [31].

- Methylation levels extracted from bismark will be analysed using packages such as methylKit [32] and RnBeads [33].

Once differential analyses are performed, the impacted metabolic pathways will be identified through enrichment analyses. This can be accomplished using R packages like ClusterProfileR [34].

Further classification analysis will be conducted using R packages such as methylClass [35] to study the evolution of omic profiles over time.

For each identified metabolic pathway, differences in means of corresponding gene expression, methylation, or protein abundance between two time points will be tested using parametric or non-parametric tests, depending on the data distribution. There will be a particular focus on the biological pathways of eosinophilic esophagitis and barrett's esophagus. Parametric tests will use similar approaches as for the initial differential analyses, while non-parametric tests such as Wilcoxon signed-rank tests may be employed.

To account for multiple testing, p-values of these tests will be corrected. The Bonferroni procedure may be used initially to control the Family-Wise Error Rate, followed by the Benjamini-Hochberg procedure to control the False Discovery Rate if necessary. These corrections will help ensure the reliability of the identified associations and reduce the likelihood of false positives.

Transposition from the biopsies to the blood

Scatterplots depicting microRNA expression and protein abundance in plasma will be generated, with levels in biopsies serving as the independent variable. These scatterplots will visually illustrate the relationships between the variables. Additionally, correlation coefficients will be calculated to quantify the strength and direction of these relationships, providing numerical measures of association.

Linear regression analysis will be conducted to further explore the relationships between microRNA expression, protein abundance in plasma, and levels in biopsies. This analysis will involve fitting a linear model to the data to determine the extent to

which changes in one variable predict changes in another. The residuals from the regression model will be examined to assess the adequacy of the model fit and identify any potential outliers or patterns in the data that may require further investigation.

Overall, these analyses will provide insights into the relationships between microRNA expression, protein abundance in plasma, and levels in biopsies, helping to uncover potential biomarkers or indicators of interest in the context of oesophageal atresia.

Ethics and dissemination

The study (protocol versions 1.1, 2.0, 3.0 and 4.0) has obtained approval from the French ethics research committee (Comité de protection des personnes Est II).

Informed consent

Patients and their parents will receive prior communication before the scheduled follow-up visit, providing them with information about the opportunity to participate in the study. The study protocol will be thoroughly explained, and an information note will be sent to them to allow ample time for consideration, with a minimum duration of one week.

On the day of the visit, any queries or concerns they may have will be addressed comprehensively. Subsequently, they will have the option to confirm or decline their participation in the study, based on their informed decision.

For the control group, eligible patients will be contacted to invite their participation in the study. A letter of non-opposition will be sent to obtain their agreement for the utilization of samples collected during routine care, which will be repurposed for research purposes.

Privacy

The TransEAsome team will prioritize the pseudo-anonymization of all clinical and omics data, ensuring that individual identities are solely identifiable by the patient

inclusion number. Access to the data will be restricted, with each partner granted limited access based on their specific requirements for task performance.

The project will adhere to the guidelines established by the Commission Nationale de l'Informatique et des Libertés (CNIL), the French data protection authority, particularly following standard methodology 001 for clinical research involving the collection of patient consent. These measures aim to uphold patient confidentiality and privacy while facilitating valuable research endeavours.

Dissemination plan

Data

All data collected will undergo harmonization with standard terminologies encompassing pathologies, localization, treatments, units of measure, phenotypes, and health interventions. This harmonization process aims to enhance interoperability with other databases and enable seamless reuse by diverse research teams.

The data will be centralized within the France Cohortes platform, where each entry will be assigned a unique identifier. Additionally, metadata files will be generated in accordance with FAIR principles, ensuring that the data remains Findable, Accessible, Interoperable, and Reusable.

Results

The dissemination of findings from the TransEAsome project will target various audiences to ensure broad impact and relevance. These audiences include:

1. Scientific Community: Findings will be disseminated through peer-reviewed scientific publications and conference presentations, allowing researchers and clinicians to access and utilize the latest advancements in the field.

2. Research Participants: Participants in the study will be informed of the results through participant newsletters, ensuring that they remain informed about the outcomes of the research in which they contributed.

3. Patient Associations: Results will be shared with patient associations to empower patients and their families with knowledge about advancements in the understanding and management of oesophageal atresia.

4. General Public: Information about the project's findings will be communicated to the general public through social media platforms such as Twitter and LinkedIn, raising awareness and understanding of oesophageal atresia among the wider community.

5. Regulatory Authorities and Policy-Makers: Relevant findings will be shared with regulatory authorities and policy-makers to inform decision-making processes and contribute to the development of policies related to rare diseases and paediatric healthcare.

Communication efforts will adhere to the graphic charter of the project, ensuring consistency and professionalism in visual presentation. Additionally, all communications will acknowledge the funding provided by the Agence Nationale de la Recherche, as required. This comprehensive dissemination strategy aims to maximize the impact of the TransEAsome project across various stakeholders and promote the translation of research findings into tangible benefits for patients and society.

References

1. van Lennep M, Singendonk MMJ, Dall'Oglio L, Gottrand F, Krishnan U, Terheggen-Lagro SWJ, et al. Oesophageal atresia. Nat Rev Dis Primers [Internet]. 2019 Apr 18 [cited 2022 Feb 4];5(1):1–21. Available from: https://www.nature.com/articles/s41572-019-0077-0

2. Sfeir R, Michaud L, Sharma D, Richard F, Gottrand F. National Esophageal Atresia Register. Eur J Pediatr Surg [Internet]. 2015 Dec;25(6):497–9. Available from: https://pubmed.ncbi.nlm.nih.gov/26642387/

3. Koivusalo AI, Sistonen SJ, Lindahl HG, Rintala RJ, Pakarinen MP. Long-term outcomes of oesophageal atresia without or with proximal tracheooesophageal fistula - Gross types A and B. J Pediatr Surg [Internet]. 2017 Oct;52(10):1571–5. Available from: https://pubmed.ncbi.nlm.nih.gov/28499713/

4. Svoboda E, Fruithof J, Widenmann-Grolig A, Slater G, Armand F, Warner B, et al. A patient led, international study of long term outcomes of esophageal atresia: EAT 1. J Pediatr Surg [Internet]. 2018 Apr;53(4):610–5. Available from: https://pubmed.ncbi.nlm.nih.gov/28622972/

5. Krishnan U, Mousa H, Dall'Oglio L, Homaira N, Rosen R, Faure C, et al. ESPGHAN-NASPGHAN Guidelines for the Evaluation and Treatment of Gastrointestinal and Nutritional Complications in Children With Esophageal Atresia-Tracheoesophageal Fistula. J Pediatr Gastroenterol Nutr [Internet]. 2016 Nov;63(5):550–70. Available from: https://pubmed.ncbi.nlm.nih.gov/27579697/

6. Dimitrov G, Aumar M, Duhamel A, Wanneveich M, Gottrand F. Proton pump inhibitors in esophageal atresia: A systematic review and meta-analysis. J Pediatr Gastroenterol Nutr. 2024 Jan 23;

7. Schneider A, Gottrand F, Bellaiche M, Becmeur F, Lachaux A, Bridoux-Henno L, et al. Prevalence of Barrett Esophagus in Adolescents and Young Adults With Esophageal Atresia. Ann Surg [Internet]. 2016 Dec;264(6):1004–8. Available from: https://pubmed.ncbi.nlm.nih.gov/26720426/

8. Vergouwe FW, Gottrand M, Wijnhoven BP, IJsselstijn H, Piessen G, Bruno MJ, et al. Four cancer cases after esophageal atresia repair: Time to start screening the upper gastrointestinal tract. World J Gastroenterol [Internet]. 2018 Mar 7 [cited 2022 Mar 11];24(9):1056–62. Available from: https://pubmed.ncbi.nlm.nih.gov/29531469/

9. Aumar M, Audrey N, Sfeir R, Seguy D, Gottrand F. Long term digestive outcome of œsophageal atresia. Best Practice & Research Clinical Gastroenterology

[Internet]. 2022 Feb 1;56–57:101771. Available from: https://pubmed.ncbi.nlm.nih.gov/35331402/

10. Lardenois E, Michaud L, Schneider A, Onea M, Rebeuh J, Gottrand-Aumar M, et al. Prevalence of Eosinophilic Esophagitis in Adolescents With Esophageal Atresia. J Pediatr Gastroenterol Nutr [Internet]. 2019 Jul;69(1):52–6. Available from: https://pubmed.ncbi.nlm.nih.gov/30614952/

11. Traini I, Menzies J, Hughes J, Leach ST, Krishnan U. Oesophageal atresia: The growth gap. World J Gastroenterol [Internet]. 2020 Mar 28 [cited 2022 Mar 11];26(12):1262–72. Available from: https://pubmed.ncbi.nlm.nih.gov/32256015/

12. Dittrich R, Stock P, Rothe K, Degenhardt P. Pulmonary outcome of esophageal atresia patients and its potential causes in early childhood. J Pediatr Surg [Internet]. 2017 Aug;52(8):1255–9. Available from: https://pubmed.ncbi.nlm.nih.gov/28094013/

13. Dellenmark-Blom M, Quitmann J, Dingemann C. Health-Related Quality of Life in Patients after Repair of Esophageal Atresia: A Review of Current Literature. Eur J Pediatr Surg [Internet]. 2020 Jun;30(3):239–50. Available from: https://pubmed.ncbi.nlm.nih.gov/32483796/

14. Sy MR, Chauhan J, Prescott K, Imam A, Kraus A, Beleza A, et al. Exome sequencing efficacy and phenotypic expansions involving esophageal atresia/tracheoesophageal fistula plus. Am J Med Genet A. 2022 Sep 22;

15. Zhong G, Ahimaz P, Edwards NA, Hagen JJ, Faure C, Lu Q, et al. Identification and validation of candidate risk genes in endocytic vesicular trafficking associated with esophageal atresia and tracheoesophageal fistulas. HGG Adv [Internet]. 2022 Jul 14;3(3):100107. Available from: https://pubmed.ncbi.nlm.nih.gov/35519826/

16. Thiem CE, Stegmann JD, Hilger AC, Waffenschmidt L, Bendixen C, Köllges R, et al. Re-sequencing of candidate genes FOXF1, HSPA6, HAAO, and KYNU in 522 individuals with VATER/VACTERL, VACTER/VACTERL-like association, and isolated anorectal malformation. Birth Defects Res. 2022 Jun;114(10):478–86.

17. Gehlen J, Giel AS, Köllges R, Haas SL, Zhang R, Trcka J, et al. First genome-wide association study of esophageal atresia identifies three genetic risk loci at CTNNA3, FOXF1/FOXC2/FOXL1, and HNF1B. HGG Adv [Internet]. 2022 Apr 14;3(2):100093. Available from: https://pubmed.ncbi.nlm.nih.gov/35199045/

18. Brosens E, Brouwer RWW, Douben H, van Bever Y, Brooks AS, Wijnen RMH, et al. Heritability and De Novo Mutations in Oesophageal Atresia and Tracheoesophageal Fistula Aetiology. Genes (Basel). 2021 Oct 10;12(10):1595.

19. Ten Kate CA, de Klein A, de Graaf BM, Doukas M, Koivusalo A, Pakarinen MP, et al. Intrinsic Cellular Susceptibility to Barrett's Esophagus in Adults Born with Esophageal Atresia. Cancers (Basel). 2022 Jan 20;14(3):513.

20. Maltby EL, Dyson MJ, Wheeler MR, Thomson M, Sethuraman C, Cohen MC. Molecular abnormalities in pediatric barrett esophagus: can we test for potential of neoplastic progression? Pediatr Dev Pathol. 2010 Aug;13(4):310–7.

21. Krishnan U, Lijuan C, Andrew GJ, Rothenberg ME, Wen T. Analysis of eosinophilic esophagitis in children with repaired congenital esophageal atresia. J Allergy Clin Immunol [Internet]. 2019 Apr;143(4):1455-1464.e2. Available from: https://pubmed.ncbi.nlm.nih.gov/30527929/

22. Gaillard S, Malik S, Preston J, Escalera BN, Dicks P, Touil N, et al. Involving children and young people in clinical research through the forum of a European Young Persons' Advisory Group: needs and challenges. Fundam Clin Pharmacol. 2018 Aug;32(4):357–62.

23. Preston J, Nafria B, Ohmer A, Gaillard S, Dicks P, West L, et al. Developing a More Tailored Approach to Patient and Public Involvement with Children and Families in Pediatric Clinical Research: Lessons Learned. Ther Innov Regul Sci. 2022 Nov;56(6):948–63.

24. Protocole national de diagnostic et de soins (PNDS) : atrésie de l'œsophage. Perfectionnement en Pédiatrie [Internet]. 2019 Jun [cited 2022 Mar 11];2(2):98–115. Available from: https://linkinghub.elsevier.com/retrieve/pii/S2588932X19300452

25. Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics [Internet]. 2011 Jun 1 [cited 2024 Feb 1];27(11):1571–2. Available from: https://doi.org/10.1093/bioinformatics/btr167

26. Harrell FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med. 1996 Feb 28;15(4):361–87.

27. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol. 1996 Dec;49(12):1373–9.

28. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med. 2011 Feb 20;30(4):377–99.

29. Rubin. Multiple Imputation for Nonresponse in Surveys. In: Multiple Imputation for Nonresponse in Surveys [Internet]. John Wiley & Sons, Ltd; 1987 [cited 2022

Dec 9]. p. i–xxix. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470316696.fmatter

30. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015 Apr 20;43(7):e47.

31. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology [Internet]. 2014 Dec 5 [cited 2023 Sep 20];15(12):550. Available from: https://doi.org/10.1186/s13059-014-0550-8

32. Wreczycka K, Gosdschan A, Yusuf D, Grüning B, Assenov Y, Akalin A. Strategies for analyzing bisulfite sequencing data. Journal of Biotechnology [Internet]. 2017 Nov 10 [cited 2024 Feb 1];261:105–15. Available from: https://www.sciencedirect.com/science/article/pii/S0168165617315936

33. Müller F, Scherer M, Assenov Y, Lutsik P, Walter J, Lengauer T, et al. RnBeads 2.0: comprehensive analysis of DNA methylation data. Genome Biology [Internet]. 2019 Mar 14 [cited 2024 Feb 1];20(1):55. Available from: https://doi.org/10.1186/s13059-019-1664-9

34. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. The Innovation [Internet]. 2021 Aug 28 [cited 2023 Sep 20];2(3):100141. Available from: https://www.sciencedirect.com/science/article/pii/S2666675821000667

35. methylClass: an R package to construct DNA methylation-based classification models | Briefings in Bioinformatics | Oxford Academic [Internet]. [cited 2024 Feb 1]. Available from: https://academic.oup.com/bib/article/25/1/bbad485/7513599

Acknowledgements

The authors express their gratitude to the scientific board of the *TransEAsome* project for their invaluable input in shaping its design. They also extend their appreciation to all investigation teams for their significant contributions to the establishment and execution of this study. Special thanks are extended to Katialine Groff and Dr. Rony Sfeir for their dedicated work on the oesophageal atresia national registry ReNaTo, which has been instrumental in the project's progress.

The authors also acknowledge the members of the user's committees for their insightful feedback, which has consistently enriched the project. Additionally, they express their gratitude to the paediatric clinical investigation centre at Lille Hospital for their technical support in clinical research, which has been indispensable to the authors.

This collaborative effort and support from various individuals and institutions have been essential in advancing the TransEAsome project and its objectives.

Collaborators

University hospital of Amiens, University hospital of Angers, University hospital of Besançon, University hospital of Bordeaux, University hospital of Brest, University hospital of Brest, University hospital of Caen, Public hospital of Cayenne, University hospital of Clermont Ferrand, University hospital of Dijon, University hospital of Fort de France, University hospital of La Tronche, Public hospital of Le Mans, University hospital of Limoges, University hospital of Lyon, University hospital of Marseille, University hospital of Montpellier, University hospital of Nancy, University hospital of Nantes, University hospital of Nice, Public hospital of Orléans, University hospital of Paris Armand Trousseau, University hospital of Kremelin Bicêtre, University hospital of Créteil, University hospital of Paris Necker, University hospital of Paris Robert Debré, University hospital of Poitiers, University hospital of Pointe à Pitre, University hospital of Reims, University hospital of Rennes, University hospital of Rouen, University hospital

of Saint Denis de la Réunion, University hospital of Saint Etienne, University hospital of Strasbourg, University hospital of Toulouse and University hospital of Tours.

Authors' contributions

FG is the scientific coordinator and guarantor. MF, RH, GM, MS are partner leaders in the project. MA is the Lille University Hospital's principal investigator, JV is the co-technical-head of the Billile platform, SG is head of the patient and public involvement task, JL is the head of the Lille University Hospital's biostatistics department, LD is a methodologist who revised the protocol submitted to ethics committee, MD is in charge of the omics analysis within the PRISM team, SF is a head of the ReNaTo registry, JR and VA are a members of the French national patient with OA association and ML is the project manager. All authors have contributed to the writing of the study protocol and have approved the final version of this manuscript.

Funding statement

This project has received funding from the ANR program France 2030 under the specific grant agreement ANR-21-PMRB-0011. As well as a funding from the Institution de prévoyance et Retraite Collective des Employés de Maison (IRCEM) fundation.

Competing interest statement

No competing interests.

Word count

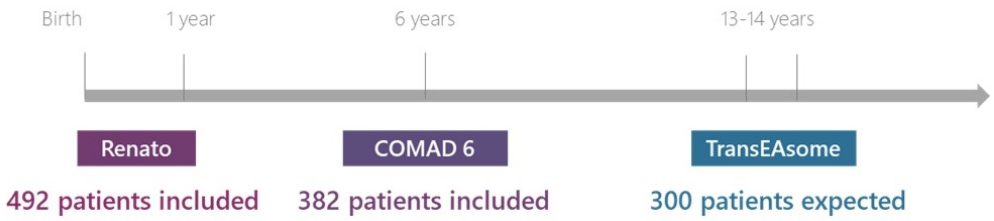
4840 words

Open Access

Figure legends

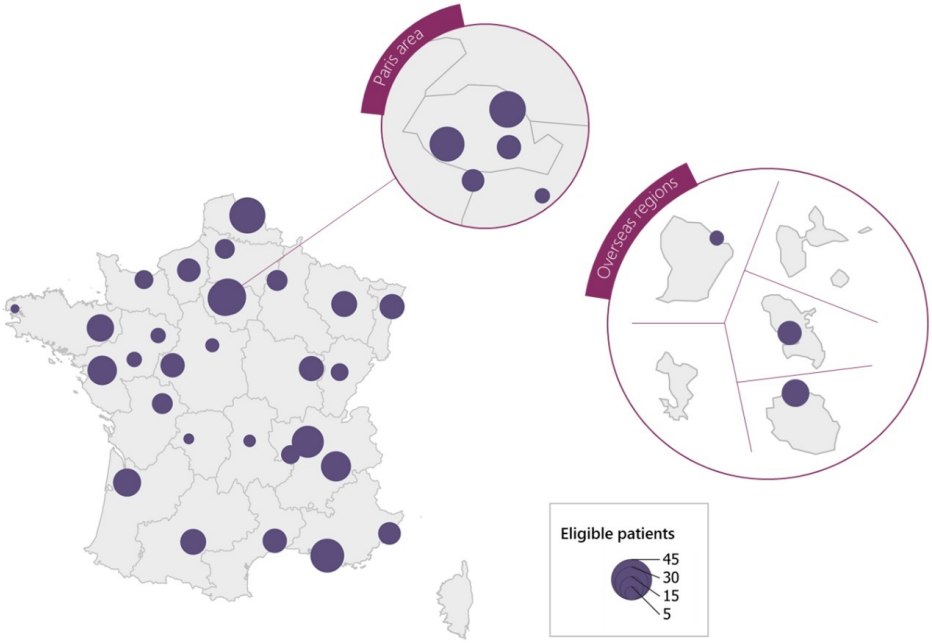
Figure 1: Follow-up of a nested cohort of OA patients born between 2010 and 2012

Figure 2: Localisation of recruitment in France



Follow-up of a nested cohort of OA patients born between 2010 and 2012

83x19mm (300 x 300 DPI)



Localisation of recruitment in France

115x76mm (300 x 300 DPI)

BMJ Open

Long-term outcome of oesophageal atresia in adolescence (TransEAsome): a cohort study protocol

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2024-086303.R2
Article Type:	Protocol
Date Submitted by the Author:	02-Dec-2024
Complete List of Authors:	Leroy, Mélanie; Univ. Lille, Inserm, CHU Lille, U1286 - INFINITE - Institute for Translational Research in Inflammation, F-59000 Lille, France; CHU Lille, Reference Center for Congenital Oesophageal Anomalies, F-59000 Lille, France Aumar, Madeleine; Univ. Lille, Inserm, CHU Lille, U1286 - INFINITE - Institute for Translational Research in Inflammation, F-59000 Lille, France; Univ. Lille, CHU Lille, Infinite U1286, Reference Center for Congenital Oesophageal Anomalies , F-59000 Lille, France Duhamel, Marie; Univ. Lille, Inserm, CHU Lille, U1192 - Protéomique Réponse Inflammatoire et Spectrométrie de Masse - PRISM, F-59000 Lille, France Dauchet, Luc; CHU Lille, Department of Biostatistics, F59000-Lille, France Figeac, Martin; Univ. Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, US 41 - UAR 2014 - PLBS, F-59000 Lille, France Gaillard, Segolene ; Hospices civils de Lyon, EPICIME-CIC 1407 de Lyon, Kids France, Inserm 1407, CHU-Lyon, 69677 Bron, France ; Université Lyon 1, CNRS UMR 5558, laboratoire de biométrie et biologie évolutive, 69622 Villeurbanne, France Hankard, Regis; French Clinical Research Infrastructure Network (F-CRIN)-PEDSTART, INSERM, University of Tours, Tours, France Labreuche, Julien; CHU Lille, Department of Biostatistics, F59000-Lille, France Marot, Guillemette; Univ. Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, US 41 - UAR 2014 - PLBS, F-59000 Lille, France; Univ. Lille, CHU Lille, Inria, ULR 2694 - METRICS : Évaluation des technologies de santé et des pratiques médicales, F-59000 Lille, France Reversat, Julie; Association Française de l'Atrésie de l'Œsophage Armand, Viviane; Association Française de l'Atrésie de l'Œsophage Salzet, Michel; Univ. Lille, Inserm, CHU Lille, U1192 - Protéomique Réponse Inflammatoire et Spectrométrie de Masse - PRISM, F-59000 Lille, France Sfeir, Rony; Univ. Lille, CHU Lille, Infinite U1286, Reference Center for Congenital Oesophageal Anomalies , F-59000 Lille, France Vandel, Jimmy; Univ. Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, US 41 - UAR 2014 - PLBS, F-59000 Lille, France Gottrand, Frédéric; Univ. Lille, Inserm, CHU Lille, U1286 - INFINITE - Institute for Translational Research in Inflammation, F-59000 Lille, France; Univ. Lille, CHU Lille, Infinite U1286, Reference Center for Congenital Oesophageal Anomalies , F-59000 Lille, France

Primary Subject Heading:	Paediatrics
Secondary Subject Heading:	Gastroenterology and hepatology, Surgery
Keywords:	Adolescents < Adolescent, Oesophageal disease < GASTROENTEROLOGY, Patient Reported Outcome Measures, Quality of Life, Paediatric gastroenterology < GASTROENTEROLOGY

Long-term outcome of oesophageal atresia in adolescence (TransEAsome): a national French cohort study protocol

Mélanie Leroy[1,2], Madeleine Aumar[1,2], Marie Duhamel[3], Luc Dauchet[4], Martin Figeac[5], Ségolène Gaillard[6], Régis Hankard[7], Julien Labreuche[4], Guillemette Marot[5,8], Julie Reversat[9], Viviane Armand[9], Michel Salzet[3], Rony Sfeir[2], Jimmy Vandel[5], Frédéric Gottrand[1,2]

[1] Univ. Lille, Inserm, CHU Lille, U1286 - INFINITE - Institute for Translational Research in Inflammation, F-59000 Lille, France

[2] CHU Lille, Reference Center for Congenital Oesophageal Anomalies, F-59000 Lille, France

[3] Univ. Lille, Inserm, CHU Lille, U1192 - Protéomique Réponse Inflammatoire et Spectrométrie de Masse - PRISM, F-59000 Lille, France

[4] CHU Lille, Department of Biostatistics, F59000-Lille, France

[5] Univ. Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, US 41 - UAR 2014 - PLBS, F-59000 Lille, France

[6] Hospices civils de Lyon, EPICIME-CIC 1407 de Lyon, Kids France, Inserm 1407, CHU-Lyon, 69677 Bron, France ; Université Lyon 1, CNRS UMR 5558, laboratoire de biométrie et biologie évolutive, 69622 Villeurbanne, France

[7] French Clinical Research Infrastructure Network (F-CRIN)-PEDSTART, INSERM, University of Tours, Tours, France

[8] Univ. Lille, CHU Lille, Inria, ULR 2694 - METRICS : Évaluation des technologies de santé et des pratiques médicales, F-59000 Lille, France

[9]Association Française de l'Atrésie de l'Œsophage, Maison Alfort, France

Corresponding author:

Pr Frédéric Gottrand

CENTRE HOSPITALIER UNIVERSITAIRE

Hôpital Jeanne de Flandre

Pôle Enfant

2 avenue Oscar Lambret - 59037 Lille Cedex, France

+33 3 20 44 61 26

Frederic.gottrand@chu-lille.fr

Abstract

Introduction

The *TransEAsome* project, funded by the Agence Nationale de la Recherche, aims to evaluate the long-term outcomes of patients with oesophageal atresia (OA) between 13 and 14 years old and establish multi-omics profiles using data from the world's biggest OA registry.

Methods and analysis

TransEAsome is a national multicentre population-based cohort study recruiting participants from all qualified French centres for oesophageal atresia surgery at birth. The primary objective is to assess the prevalence of gastro-oesophageal reflux disease in adolescence among patients with OA, with several secondary objectives including the identification of risk factors and multi-omic profiles from oesophageal biopsies and blood samples collected between 13 and 14 years old, compared with a control group. This comprehensive characterisation of phenotype and omic profiles aims to enhance the understanding of disease evolution in patients with OA and inform tailored care management strategies.

Ethics and dissemination

The study, co-constructed with input from patients, parents and research-expert adolescents, has obtained approval from the ethics research committee: Comité de protection des personnes Est II. Findings will be disseminated to various target audiences, including the scientific community, research participants, the patient community, the general public, regulatory authorities, and policymakers. Data will be made available in a Findable, Accessible, Interoperable, Reusable (FAIR) format on the France Cohortes platform upon study completion.

Trial registration number

NCT05995171

Strengths and limitations of this study

- Multicentre, national, population-based cohort study conducted in all French centres qualified to perform oesophageal atresia surgery at birth,
- Longitudinal analysis facilitated by a nested cohort design. The birth and 1-year-old visits are in the national register of oesophageal atresia (Registre National de l'ATresie de l'Oesophage (ReNATO)), the 6-year-old visits into the nested cohort COMAD6, while *TransEAsome* focuses on the 13-14-year-old visits,
- Active involvement of patients and parents thanks to the users' representative committee, ensuring the study reflects patient perspectives,
- Comprehensive multi-omic analysis encompassing transcriptomic, epigenetic, proteomic, and metabolomic data, providing a holistic understanding of the disease.
- Omic studies rely on oesophageal tissue samples collected for clinical surveillance purposes, which may not be systematic, and long-term storage could lead to RNA degradation over time, potentially impacting data quality.

Keywords

Oesophageal atresia, adolescence, gastro-oesophageal reflux, dysphagia, oesophageal cancer, eosinophilic esophagitis, outcome, epigenetic, proteogenomic metabolomics, quality of life

Introduction

Oesophageal atresia (OA) is a rare congenital anomaly affecting approximately 160 newborns annually in France, characterised by a discontinuity in the oesophagus. It is often associated with other malformations in over 50% of cases [1]. Since the early 1950s, significant advancements in surgical techniques have vastly improved the management of this condition, resulting in a favourable prognosis. While the mortality rate has declined to less than 7% in developed nations like France [2], morbidity remains noteworthy [3,4].

Gastro-oesophageal reflux disease (GORD) is more prevalent among individuals with OA compared to the general population [5]. Studies indicate that approximately 30% of patients with OA experience GORD at least once in their lifetime [6]. Complications of GORD, such as peptic esophagitis and Barrett's oesophagus, present heightened risks for oesophageal cancer [7], with documented cases of adenocarcinoma or squamous cell carcinoma in young adults with OA [8].

Dysphagia is a common concern among patients with OA, with around 45% of five-year-olds experiencing swallowing difficulties attributed to anastomotic strictures or oesophageal dysmotility [9]. Eosinophilic oesophagitis (EoE), a rising concern in oesophageal health, has been noted to occur more frequently in patients with OA, contributing to dysphagia [10].

Challenges in eating often lead to undernutrition, affecting up to one-third of children with EA by the age of five [11]. Respiratory complications are also prevalent [12], particularly in the early years, significantly impacting long-term quality of life (QoL) [13].

To better understand the long-term outcomes and predict complications, our objective is to investigate a nested population cohort of adolescents born with OA. Leveraging the existing national population-based registry (ReNaTo), and the nested clinical cohort, COMAD6, consisting of 382 children born between 2010 and 2012, we aim to

compile a new database comprising clinical data, blood samples, and oesophageal biopsies obtained during routine follow-up between 13 and 14 years old (Figure 1).

This nested cohort, within the world's largest registry of more than 2300 patients with OA, offers a unique opportunity to answer questions about outcomes of OA at adolescence in terms of morbidity and health status, to improve care and follow-up and prevent long-term complications.

In addition to the phenotypic assessment, our investigations extend to exploring the biological mechanisms underpinning complication development through genomic, transcriptomic, proteomic, epigenetic, and metabolomic analyses of the oesophagus.

While upper gastrointestinal endoscopy (EGD) is standard in the care of symptomatic or not patients with OA [5], its invasive nature necessitates hospitalization and general anaesthesia in children.

To date, there is little data on the epigenetic, genomic [14–18], or transcriptomic [19] profile of patients with OA. In a small study of 10 children with Barrett's oesophagus (some of them operated at birth for OA), the authors performed fluorescence in situ hybridization with probes on 4-micron sections taken from sequential paraffin-embedded biopsies and identified 4 probe sets reported to be associated to adult Barret adenocarcinoma [19]. Genetic markers were also identified in adult Barrett's adenocarcinoma patients. This preliminary study shows that, even at an early age, Barrett's may show genetic changes associated with neoplastic progression [20]. Another recent study examined the relationship between eosinophilic esophagitis (EoE) and OA by profiling the transcriptional signature of EoE [21]. Using an in-silico approach, they found 6 genes differentially expressed between the 2 entities (OA+ EoE+ and OA- EOE+). Two of them were associated with major dysphagia, the development of strictures, and the need for dilatations in patients with OA+EoE+.

As EGD and oesophageal biopsies are invasive procedures that usually require hospitalization and general anaesthesia in children, using a more accessible biological sample as blood plasma will also be investigated to see if the markers of the oesophagus translate into blood.

The overarching goal of *TransEAsome* is to construct a vast longitudinal database including phenomenal and exposomal patient data amalgamating clinical, biological, environmental, and lifestyle data alongside multimodal omics analyses: RNA abundance, protein abundance, proteogenomic profiling, protein modification, metabolite abundance, methylated sites). This innovative approach holds promise for enhancing the prediction of adult OA outcomes, including the identification of risk factors for future health complications such as oesophageal cancer and EoE.

Objectives

TransEAsome will address the evolution of patients with OA at the time of adolescence by establishing a unique and comprehensive database that integrates clinical and omics data in a structured and interoperable format to evaluate their health status and quality of life. This project is designed to achieve the following objectives:

1. Evaluate the prevalence of GORD during adolescence in the OA population.

2. Identify factors associated with GORD during adolescence in the OA population.

3. Assess the quality of life, nutritional status, and frequency of respiratory complications in the OA population.

4. Compare the single and multi-omic profiles derived from oesophageal biopsies in adolescents operated at birth for OA with those who did not, aiming to delineate alterations in biological pathways specific to OA.

5. Investigate changes in the expression of omics variables within the identified biological pathways over time in the OA population, utilizing data obtained from EGD and biopsies.

6. Evaluate the correlation between the expression levels of microRNAs and proteins in plasma compared to oesophageal biopsies in the OA population, utilizing data from individuals who underwent EGD, biopsies, and blood sampling at ages 13 and 14.

Methods and analysis

Recruitment sources

Patients for the *TransEAsome* study will be recruited from the French national registry for OA known as "ReNaTo," which involves all 35 OA competence centres across France. As of the conclusion of 2022, more than 2300 patients were included in ReNaTo. Among them, 492 are projected to reach 13 or 14 years of age during the recruitment period of the *TransEAsome* project. However, an anticipated loss to follow-up necessitates accounting for a pool of 300 patients. Leveraging the national register ReNaTo, the project management team will furnish each centre with a pseudo-anonymized patient roster, facilitating the identification of eligible candidates from their patient cohorts. Given the unique demographic characteristics of each participating centre, recruitment potentials vary, spanning from 1 to 34 patients per centre during the study period (Figure 2).

For the objectives involving multi-omic profiling, the control group will be drawn from six selected centres. Eligible patients will be identified through the extraction of cases meeting the study's inclusion criteria from the pathology department.

Study population

TransEAsome eligibility and exclusion criteria are listed in Table 1. The initial recruitment period will start in September 2023 and end in May 2026. Our goal will be to enrol 300 patients and 150 controls.

Table 1: Eligibility and exclusion criteria

Patients with oesophageal atresia	
Eligibility criteria	• Born with OA in France (metropolitan or overseas) • Underwent oesophageal anastomosis at one of the participating centres • Included in the ReNaTo registry • Age between 13 and 14 years during the inclusion period. • Willingness to participate. • Patient with social security Blood sub-study: • Upper GI endoscopy performed as part of care between 13 and 14 years of age with oesophageal mucosal biopsy sampling • Consent for blood sampling.
Exclusion criteria	• Participation in an interventional trial (simultaneously or up to 3 months before inclusion. • Oesophageal replacement
Controls	
Eligibility criteria	• Upper GI endoscopy performed as part of care between 10 and 14 years of age with oesophageal mucosal biopsy sampling • Underwent esophagogastroduodenoscopy (EGD) as part of routine care for any digestive symptom to exclude organic aetiology (peptic esophagitis, gastritis, eosinophilic esophagitis, or ulcer). • EGD findings and histology are normal. • Absence of concurrent progressive chronic disease. • Provision of signed study consent.
Exclusion criteria	• Known underlying disease

The blood sub study as well as the inclusion of controls will only be performed in Lille, Paris Necker, Paris Robert Debré, Lyon, Grenoble, and Marseille for logistic, feasibility and cost-saving reasons.

Patient and Public Involvement (PPI)

A young patient advisory committee and a parents/patient representativeness committee have been established to ensure the active involvement of young patients and their parents throughout all phases of the *TransEAsome* project, from inception to dissemination of results. These committees have played a pivotal role in shaping the study protocol, including determining the data to be collected and reviewing informed assents and consents before regulatory submissions. Throughout the trial, these committees will contribute to the decision-making process of the steering committee and provide insights from a patient perspective on the conduct of the study.

The committees will remain actively engaged throughout the project, facilitating communication and dissemination efforts to the scientific community, patients, and the public regarding key milestones, actions, and emerging knowledge. Upon project completion, the committees will review the analysis, discuss the findings, and aid in disseminating results.

Kids France (Hospices Civils de Lyon, Pedstart) will oversee the patient and public involvement (PPI) activities in collaboration with the French patient with OA support group (AFAO). Kids France will ensure that PPI activities are not only meaningful for the research but also beneficial for the patients themselves. Training related to PPI activities will be provided to the committees by Kids France. Reporting activities to the steering committee, feedback on suggestions from patients and parents, and assessment of impact are integral components of the *TransEAsome* PPI process [22,23].

Outcomes and assessments

Outcomes and assessments are presented in with separate forms for parents and adolescent patients.

Table 2. Quality of life (QoL) questionnaires and blood sampling are the only assessments that are not part of routine practice. Both the generic paediatric QoL assessment (Pediatric Quality of Life Inventory™ (PedsQL) and the disease-specific QoL questionnaire (EA QoL) [13] will be administered in this cohort, with separate forms for parents and adolescent patients.

Table 2: Outcomes and assessments between 13 and 14-year-old

Primary outcome	GORD will be determined based on the following criteria: • Positive pH-(impedance)metry results obtained within the previous year. • And/or presence of histological peptic esophagitis lesions at oesophageal biopsies collected during the previous year. • And/or history of anti-reflux surgery.
Secondary outcomes	Quality of life: • PedsQL total patient and parent scores • EA QoL total patient and parent scores Nutritional status: • Z-score Weight/Height and Height/Age • Type of feeding (oral, enteral, both) (Food avoidance) Digestive status: • Current proton pump inhibitor treatment • Dysphagia (defined by sensations such as blockage leading to vomiting or the need to drink to pass food, or slowness in eating) • GORD symptoms (regurgitation, vomiting, retrosternal pain, and heartburn) Respiratory status: • Frequency of cough • Asthma • Occurrence of exercise-induced symptoms (cough, dyspnoea) • Atopy

	• Wheezing
	• Stridor
	• Need for medications (corticosteroid, inhaled medications)
	
	History from 6 years to adolescence:
	• Associated malformations discovered after 6 years
	• Gastrostomy placement and/or use after 6 years old
	• Characteristics of eventual GORD surgery (type, complications, cardial dilatation, relapse)
	• Oesogastroduodenoscopy (number, histology)
	• Repermeabilization of the oesotracheal fistula (treatment, number, date)
	• Stenosis (date, dilatation, number, surgery)
	• Stricture (date, number and method of dilatation, surgery, use of corticosteroid)
	
	Psychological or psychiatric follow-up
	
	Neuro orthopaedic outcomes:
	• Scoliosis
	• Kyphosis
	• Stature malposition
	
	Speech therapy:
	• Oral/stomatology
	• Oral disorders
	• Speech disorders
	• Dental conditions
	
	Education:
	• Physical/sport waiver
	• Adapted school rhythm
	• Specialized school
	• School absenteeism
	
	Omic profile:
	• In oesophageal biopsies:
	  ○ Protein abundances
	  ○ Metabolite abundances
	  ○ RNA profile
	  ▪ 3'RNA
	  ▪ miRNA
	  ○ Methylation status
	• In blood:

	○ Metabolite abundance ○ Protein abundance ○ miRNA profile
Other assessments	Parents' height Parents' socio-cultural level and highest diploma attained City of residence

Biological samples

Oesophageal tissue

EGD and biopsy sampling can be conducted at any point during the follow-up of patients with OA, based on symptoms and routine evaluation [24]. Patients who undergo these procedures will be requested to authorize the collection of four sections of formalin-fixed paraffin-embedded (FFPE) oesophageal tissue. These samples are stored in the pathology departments for a minimum of 10 years, per French regulations (décret N°88-280 du 24 mars 1988 de l'article L.761-11 du code de la santé publique) and are commonly stored for approximately 25 years.

FFPE storage is the established standard method for the long-term preservation of biological tissue in routine care, universally practised across all participating centres. Whenever feasible (subject to patient authorization and technical feasibility), previously collected oesophageal biopsies from included patients in routine care will also be analysed or pooled for analysis. FFPE sections sourced from all participating centres will be centralized at the Biological Resource Centre of Lille University Hospital (CRB du CIC 1403 du CHU de Lille, BRIF BB0033-00030) until procurement for analysis.

Comparative analysis of omic profiles measured at different time points for the same patient will facilitate longitudinal comparisons. The objective is to include biopsies from 150 patients in the analysis.

Plasma

Plasma collection will be conducted at selected centres following consultations for adolescents who have available oesophageal biopsies. Whole blood will be drawn into

a 5mL EDTA tube between the ages of 13 and 14 years and subsequently centrifuged. Plasma aliquots will then be stored at -80°C and centralized at the Biological Resource Centre of Lille University Hospital (CRB du CIC 1403 du CHU de Lille, BRIF BB0033-00030) until procurement for analysis.

Standard Operating Procedures (SOPs) for blood and biopsy collection, storage, and shipping will be uniform across all participating centres to ensure consistency and quality control.

Data sources, collection and monitoring

Clinical data collection

Clinical data will be recorded in both a traditional paper-based Case Report Form (CRF) and an electronic Case Report Form (eCRF) using the Ennov Clinical platform. Monitoring of completion will be conducted, and participants will receive reminders to complete Quality of Life (QoL) questionnaires if necessary.

The study will adhere to the guidelines set forth by the International Conference on Harmonization (ICH) Good Clinical Practice (GCP) outlined in ICH GCP Topic E6 (1996), ensuring compliance with ethical and quality standards in clinical research.

Omics data generation

Omics data will be generated from both oesophageal mucosa and plasma samples.

For proteomics analysis, FFPE tissues and plasma extracellular vesicles (EVs) will undergo antigen retrieval before localized trypsin digestion of the regions of interest. The resulting peptides will be collected using a liquid micro junction before shotgun analyses. In parallel, total RNA extraction will be performed to prepare Next Generation Sequencing (NGS) 3'RNA-seq and small RNA-seq (sRNA-seq) libraries with molecular barcodes (Unique Molecular Identifiers - UMI). The incorporation of UMIs in the sequencing libraries will enable accurate quantification even in cases of low RNA yield and quality. The sRNA-seq libraries will facilitate the detection and quantification of

non-coding RNA such as microRNA (miRNA) and small nucleolar RNA (snoRNA). All libraries will be sequenced using a NovaSeq 6000 instrument.

For plasma proteomics analysis, EVs will initially be isolated using size exclusion chromatography. Fractions will be collected and assessed for size and concentration using Nanosight analysis. EV-positive fractions will then be pooled before enzymatic digestion of proteins and subsequent identification by nano-liquid chromatography-tandem mass spectrometry (nano-LC-MS/MS). RNA extraction from EVs in the pooled fractions will be performed using the Direct-zol RNA extraction kit, followed by analysis as described above.

Omics data preprocessing

The protein identification process will utilize MaxQuant software, comparing all MS/MS data with the human protein database from the Uniprot bank. Key parameters in MaxQuant will be defined, including trypsin as the digestion enzyme with a maximum of 2 missed cleavages, methionine oxidation and N-terminal protein acetylation as variable modifications, and carbamidomethylation of cysteines as a fixed modification. Label-free quantification (LFQ) will be performed using default settings. Initial mass tolerance will be set to 6 ppm for MS mode and 20 ppm for fragmentation data in MS/MS mode. Protein and peptide identification parameters will be configured with an FDR (false discovery rate) of less than 1%, requiring a minimum of 2 peptides per protein, including at least 1 unique peptide.

Pre-processing analysis will commence using Perseus software. LFQ intensity values of each sample will be imported from MaxQuant into Perseus, and the data matrix will undergo contaminant removal. Subsequently, the data will be log2 transformed, and a filter will be applied to retain proteins present in at least 70% of samples within at least one group.

The RNA-seq analysis pipeline will involve bclconverter for fastQ generation, fastp for trimming and UMI first-step processing, fastQC for quality control, STAR for alignment,

umi-tools for deduplication, qualimap for alignment quality control, and featureCounts for quantitative assessment of genes. The sRNA-seq analysis pipeline will follow a similar workflow to RNA-seq, utilizing bowtie for alignment with necessary adjustments.

For methylation analysis, the pipeline will include bclConverter, fastQC, trim_galore, additional python scripts, bismark [25], and bowtie2 for alignment and deduplication, with qualimap utilized for alignment QC.

Statistical analysis

Sample size

The rationale for the number of patients was determined based on the cohort of patients included in the oesophageal atresia (OA) registry who will be between 13 and 14 years old during the recruitment period (n=492). Accounting for potential dropouts, refusals to participate, and deaths, it is estimated that approximately 300 patients can be included, representing around 30% of those eligible. This estimation aligns with the observed rate of non-inclusion of 22% observed in the COMAD6 study.

With a cohort of 300 patients, the study will have the capacity to estimate the frequency of GORD in adolescents (primary objective) with a maximum absolute precision of 5.7%. This precision is represented by the half-width of the 95% confidence interval, particularly in scenarios where the observed frequency of GORD is 50%, which presents the greatest width of the confidence interval for an equivalent population size.

Regarding the control group without oesophageal atresia, the plan is to include 150 patients, taking into consideration logistical and financial constraints associated with various omics analyses.

For the blood sub-study, participating centres have a potential recruitment capacity of 176 patients with oesophageal atresia (ranging from 20 to 30 per centre). Anticipating

losses to follow-up and refusals, it is estimated that approximately 150 patients can be included in this sub-study.

Phenotyping

Frequency of GORD

The frequency of patients born with a diagnosis of oesophageal atresia and GORD in adolescence will be calculated along with its two-sided 95% confidence interval. The confidence interval will be estimated using the normal approximation method.

Here's the general formula to calculate the confidence interval for a proportion (in this case, the frequency of patients with OA and GORD):

$$Confidence\ interval = Frequency\ \pm Z\ \times \sqrt{\frac{Frequency \times (1 - Frequency)}{Sample\ size}}$$

Where:

- Frequency: Proportion of patients with OA and GORD

- Z: Z-score corresponding to the desired confidence level (for a 95% confidence interval, Z ≈ 1.96)

- Sample Size: Total number of patients in the study

Given the calculated proportion of patients with OA and GORD and the total sample size, we can plug these values into the formula to obtain the confidence interval.

Associated factors with GORD

To assess the independent factors associated with the presence of GORD in adolescence, the following methodology will be employed:

1. Bivariate Logistic Regression Models:

- Bivariate logistic regression models will be conducted to evaluate the odds ratios (ORs) of GORD and their corresponding 95% confidence intervals. These models will assess the relationship between each individual factor and the presence of GORD.

- For quantitative factors, the log-linearity assumption will be assessed using restricted cubic spline functions [26]

2. Selection of Factors for Multivariable Analysis:

- Factors associated with the presence of GORD with a significance level of less than 0.20 in bivariate analyses will be included in a multivariable backward-stepwise logistic regression model.

- If the number of events per candidate variable is insufficient (<10), a penalized method will be utilized.

3. Collinearity Assessment:

- Collinearity between the candidate factors for multivariate analysis will be examined by calculating the variance inflation factor (VIF). An alert threshold will be set at 2.5.

4. Development and Evaluation of Multivariable Model:

- Discrimination of the selected multivariable model, indicating its ability to differentiate between patients with and without GORD, will be assessed using C-statistics corrected for over-optimism via bootstrap resampling [26].

- Calibration, which measures the agreement between predicted and observed probabilities of GORD, will be evaluated using the Hosmer-Lemeshow goodness of fit test.

5. Handling Data:

- To prevent case deletion due to missing data in multivariable analysis, missing data will be imputed.

- Simple imputation will be used if the missing data rate is below 10%, while multiple imputations [27] will be employed otherwise.

- The number of imputations (m) will be determined based on the fraction of missing information (FMI).

- Imputations will be performed using the Multiple Imputation by Chained Equations (MICE) procedure, incorporating all variables included in the analyses [28]. Missing data on quantitative variables will be imputed using the predictive mean matching method, while those on qualitative variables will be imputed using logistic regression models (binomial, ordinal or multinomial, depending on the number and order of modalities). In the case of multiple imputations, Rubin's rules will be applied to combine estimates obtained from each imputed dataset [29] in the case of multiple imputations.

Health status

The health status and quality of life assessment criteria in adolescence will be described using the following parameters:

1. Quality of Life Scores:

- Quality of life scores obtained from the PedsQL and the EA-QoL questionnaires will be reported.

2. Anthropometric Measures:

- Weight

- Height

- Weight/Height Z-score

- BMI Z-score

3. Frequency of Respiratory Complications

For each of these criteria, positional parameters, such as means or medians, will be calculated, and their two-sided 95% confidence intervals will be reported.

The confidence intervals will provide a range of plausible values for these parameters, allowing for an understanding of the precision of the estimates. These intervals will be calculated using appropriate statistical methods, such as the t-distribution for means or the bootstrap method for medians, depending on the distributional characteristics of the data.

Omic analysis

Omic profile between 13 and 14 years

For each type of omics data, differential analyses between the group with OA and the group without OA will be conducted using appropriate R packages. For example:

- Proteomic data will be analysed using the limma package [30].

- Sequencing data will be analysed using DESeq2 [31].

- Methylation levels extracted from bismark will be analysed using packages such as methylKit [32] and RnBeads [33].

Once differential analyses are performed, the impacted metabolic pathways will be identified through enrichment analyses. This can be accomplished using R packages like ClusterProfileR [34].

Further classification analysis will be conducted using R packages such as methylClass [35] to study the evolution of omic profiles over time.

For each identified metabolic pathway, differences in means of corresponding gene expression, methylation, or protein abundance between two-time points will be tested using parametric or non-parametric tests, depending on the data distribution. There will be a particular focus on the biological pathways of eosinophilic esophagitis and Barrett's oesophagus. Parametric tests will use similar approaches as for the initial differential analyses, while non-parametric tests such as Wilcoxon signed-rank tests may be employed.

To account for multiple testing, the p-values of these tests will be corrected. The Bonferroni procedure may be used initially to control the Family-Wise Error Rate, followed by the Benjamini-Hochberg procedure to control the False Discovery Rate if necessary. These corrections will help ensure the reliability of the identified associations and reduce the likelihood of false positives.

Transposition from the biopsies to the blood

Scatterplots depicting microRNA expression and protein abundance in plasma will be generated, with levels in biopsies serving as the independent variable. These scatterplots will visually illustrate the relationships between the variables. Additionally, correlation coefficients will be calculated to quantify the strength and direction of these relationships, providing numerical measures of association.

Linear regression analysis will be conducted to further explore the relationships between microRNA expression, protein abundance in plasma, and levels in biopsies. This analysis will involve fitting a linear model to the data to determine the extent to

which changes in one variable predict changes in another. The residuals from the regression model will be examined to assess the adequacy of the model fit and identify any potential outliers or patterns in the data that may require further investigation.

Overall, these analyses will provide insights into the relationships between microRNA expression, protein abundance in plasma, and levels in biopsies, helping to uncover potential biomarkers or indicators of interest in the context of oesophageal atresia.

Ethics and dissemination

The study (protocol versions 1.1, 2.0, 3.0 and 4.0) has obtained approval from the French ethics research committee (Comité de protection des personnes Est II).

Informed consent

Patients and their parents will receive prior communication before the scheduled follow-up visit, providing them with information about the opportunity to participate in the study. The study protocol will be thoroughly explained, and an information note will be sent to them to allow ample time for consideration, with a minimum duration of one week.

On the day of the visit, any queries or concerns they may have will be addressed comprehensively. Subsequently, they will have the option to confirm or decline their participation in the study, based on their informed decision.

For the control group, eligible patients will be contacted to invite their participation in the study. A letter of non-opposition will be sent to obtain their agreement for the utilization of samples collected during routine care, which will be repurposed for research purposes.

Privacy

The TransEAsome team will prioritize the pseudo-anonymization of all clinical and omics data, ensuring that individual identities are solely identifiable by the patient

inclusion number. Access to the data will be restricted, with each partner granted limited access based on their specific requirements for task performance.

The project will adhere to the guidelines established by the Commission Nationale de l'Informatique et des Libertés (CNIL), the French data protection authority, particularly following standard methodology 001 for clinical research involving the collection of patient consent. These measures aim to uphold patient confidentiality and privacy while facilitating valuable research endeavours.

Dissemination plan

Data

All data collected will undergo harmonization with standard terminologies encompassing pathologies, localization, treatments, units of measure, phenotypes, and health interventions. This harmonization process aims to enhance interoperability with other databases and enable seamless reuse by diverse research teams.

The data will be centralized within the France Cohortes platform, where each entry will be assigned a unique identifier. Additionally, metadata files will be generated per FAIR principles, ensuring that the data remains Findable, Accessible, Interoperable, and Reusable.

Results

The dissemination of findings from the TransEAsome project will target various audiences to ensure broad impact and relevance. These audiences include:

1. Scientific Community: Findings will be disseminated through peer-reviewed scientific publications and conference presentations, allowing researchers and clinicians to access and utilize the latest advancements in the field.

2. Research Participants: Participants in the study will be informed of the results through participant newsletters, ensuring that they remain informed about the outcomes of the research to which they contributed.

3. Patient Associations: Results will be shared with patient associations to empower patients and their families with knowledge about advancements in the understanding and management of oesophageal atresia.

4. General Public: Information about the project's findings will be communicated to the general public through social media platforms such as Twitter and LinkedIn, raising awareness and understanding of oesophageal atresia among the wider community.

5. Regulatory Authorities and Policy-makers: Relevant findings will be shared with regulatory authorities and policy-makers to inform decision-making processes and contribute to the development of policies related to rare diseases and paediatric healthcare.

Communication efforts will adhere to the graphic charter of the project, ensuring consistency and professionalism in visual presentation. Additionally, all communications will acknowledge the funding provided by the Agence Nationale de la Recherche, as required. This comprehensive dissemination strategy aims to maximize the impact of the TransEAsome project across various stakeholders and promote the translation of research findings into tangible benefits for patients and society.

References

1. van Lennep M, Singendonk MMJ, Dall'Oglio L, Gottrand F, Krishnan U, Terheggen-Lagro SWJ, et al. Oesophageal atresia. Nat Rev Dis Primers [Internet]. 2019 Apr 18 [cited 2022 Feb 4];5(1):1–21. Available from: https://www.nature.com/articles/s41572-019-0077-0

2. Sfeir R, Michaud L, Sharma D, Richard F, Gottrand F. National Esophageal Atresia Register. Eur J Pediatr Surg [Internet]. 2015 Dec;25(6):497–9. Available from: https://pubmed.ncbi.nlm.nih.gov/26642387/

3. Koivusalo AI, Sistonen SJ, Lindahl HG, Rintala RJ, Pakarinen MP. Long-term outcomes of oesophageal atresia without or with proximal tracheooesophageal fistula - Gross types A and B. J Pediatr Surg [Internet]. 2017 Oct;52(10):1571–5. Available from: https://pubmed.ncbi.nlm.nih.gov/28499713/

4. Svoboda E, Fruithof J, Widenmann-Grolig A, Slater G, Armand F, Warner B, et al. A patient led, international study of long term outcomes of esophageal atresia: EAT 1. J Pediatr Surg [Internet]. 2018 Apr;53(4):610–5. Available from: https://pubmed.ncbi.nlm.nih.gov/28622972/

5. Krishnan U, Mousa H, Dall'Oglio L, Homaira N, Rosen R, Faure C, et al. ESPGHAN-NASPGHAN Guidelines for the Evaluation and Treatment of Gastrointestinal and Nutritional Complications in Children With Esophageal Atresia-Tracheoesophageal Fistula. J Pediatr Gastroenterol Nutr [Internet]. 2016 Nov;63(5):550–70. Available from: https://pubmed.ncbi.nlm.nih.gov/27579697/

6. Dimitrov G, Aumar M, Duhamel A, Wanneveich M, Gottrand F. Proton pump inhibitors in esophageal atresia: A systematic review and meta-analysis. J Pediatr Gastroenterol Nutr. 2024 Jan 23;

7. Schneider A, Gottrand F, Bellaiche M, Becmeur F, Lachaux A, Bridoux-Henno L, et al. Prevalence of Barrett Esophagus in Adolescents and Young Adults With Esophageal Atresia. Ann Surg [Internet]. 2016 Dec;264(6):1004–8. Available from: https://pubmed.ncbi.nlm.nih.gov/26720426/

8. Vergouwe FW, Gottrand M, Wijnhoven BP, IJsselstijn H, Piessen G, Bruno MJ, et al. Four cancer cases after esophageal atresia repair: Time to start screening the upper gastrointestinal tract. World J Gastroenterol [Internet]. 2018 Mar 7 [cited 2022 Mar 11];24(9):1056–62. Available from: https://pubmed.ncbi.nlm.nih.gov/29531469/

9. Aumar M, Audrey N, Sfeir R, Seguy D, Gottrand F. Long term digestive outcome of œsophageal atresia. Best Practice & Research Clinical Gastroenterology

[Internet]. 2022 Feb 1;56–57:101771. Available from: https://pubmed.ncbi.nlm.nih.gov/35331402/

10. Lardenois E, Michaud L, Schneider A, Onea M, Rebeuh J, Gottrand-Aumar M, et al. Prevalence of Eosinophilic Esophagitis in Adolescents With Esophageal Atresia. J Pediatr Gastroenterol Nutr [Internet]. 2019 Jul;69(1):52–6. Available from: https://pubmed.ncbi.nlm.nih.gov/30614952/

11. Traini I, Menzies J, Hughes J, Leach ST, Krishnan U. Oesophageal atresia: The growth gap. World J Gastroenterol [Internet]. 2020 Mar 28 [cited 2022 Mar 11];26(12):1262–72. Available from: https://pubmed.ncbi.nlm.nih.gov/32256015/

12. Dittrich R, Stock P, Rothe K, Degenhardt P. Pulmonary outcome of esophageal atresia patients and its potential causes in early childhood. J Pediatr Surg [Internet]. 2017 Aug;52(8):1255–9. Available from: https://pubmed.ncbi.nlm.nih.gov/28094013/

13. Dellenmark-Blom M, Quitmann J, Dingemann C. Health-Related Quality of Life in Patients after Repair of Esophageal Atresia: A Review of Current Literature. Eur J Pediatr Surg [Internet]. 2020 Jun;30(3):239–50. Available from: https://pubmed.ncbi.nlm.nih.gov/32483796/

14. Sy MR, Chauhan J, Prescott K, Imam A, Kraus A, Beleza A, et al. Exome sequencing efficacy and phenotypic expansions involving esophageal atresia/tracheoesophageal fistula plus. Am J Med Genet A. 2022 Sep 22;

15. Zhong G, Ahimaz P, Edwards NA, Hagen JJ, Faure C, Lu Q, et al. Identification and validation of candidate risk genes in endocytic vesicular trafficking associated with esophageal atresia and tracheoesophageal fistulas. HGG Adv [Internet]. 2022 Jul 14;3(3):100107. Available from: https://pubmed.ncbi.nlm.nih.gov/35519826/

16. Thiem CE, Stegmann JD, Hilger AC, Waffenschmidt L, Bendixen C, Köllges R, et al. Re-sequencing of candidate genes FOXF1, HSPA6, HAAO, and KYNU in 522 individuals with VATER/VACTERL, VACTER/VACTERL-like association, and isolated anorectal malformation. Birth Defects Res. 2022 Jun;114(10):478–86.

17. Gehlen J, Giel AS, Köllges R, Haas SL, Zhang R, Trcka J, et al. First genome-wide association study of esophageal atresia identifies three genetic risk loci at CTNNA3, FOXF1/FOXC2/FOXL1, and HNF1B. HGG Adv [Internet]. 2022 Apr 14;3(2):100093. Available from: https://pubmed.ncbi.nlm.nih.gov/35199045/

18. Brosens E, Brouwer RWW, Douben H, van Bever Y, Brooks AS, Wijnen RMH, et al. Heritability and De Novo Mutations in Oesophageal Atresia and Tracheoesophageal Fistula Aetiology. Genes (Basel). 2021 Oct 10;12(10):1595.

19. Ten Kate CA, de Klein A, de Graaf BM, Doukas M, Koivusalo A, Pakarinen MP, et al. Intrinsic Cellular Susceptibility to Barrett's Esophagus in Adults Born with Esophageal Atresia. Cancers (Basel). 2022 Jan 20;14(3):513.

20. Maltby EL, Dyson MJ, Wheeler MR, Thomson M, Sethuraman C, Cohen MC. Molecular abnormalities in pediatric barrett esophagus: can we test for potential of neoplastic progression? Pediatr Dev Pathol. 2010 Aug;13(4):310–7.

21. Krishnan U, Lijuan C, Andrew GJ, Rothenberg ME, Wen T. Analysis of eosinophilic esophagitis in children with repaired congenital esophageal atresia. J Allergy Clin Immunol [Internet]. 2019 Apr;143(4):1455-1464.e2. Available from: https://pubmed.ncbi.nlm.nih.gov/30527929/

22. Gaillard S, Malik S, Preston J, Escalera BN, Dicks P, Touil N, et al. Involving children and young people in clinical research through the forum of a European Young Persons' Advisory Group: needs and challenges. Fundam Clin Pharmacol. 2018 Aug;32(4):357–62.

23. Preston J, Nafria B, Ohmer A, Gaillard S, Dicks P, West L, et al. Developing a More Tailored Approach to Patient and Public Involvement with Children and Families in Pediatric Clinical Research: Lessons Learned. Ther Innov Regul Sci. 2022 Nov;56(6):948–63.

24. Protocole national de diagnostic et de soins (PNDS) : atrésie de l'œsophage. Perfectionnement en Pédiatrie [Internet]. 2019 Jun [cited 2022 Mar 11];2(2):98–115. Available from: https://linkinghub.elsevier.com/retrieve/pii/S2588932X19300452

25. Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics [Internet]. 2011 Jun 1 [cited 2024 Feb 1];27(11):1571–2. Available from: https://doi.org/10.1093/bioinformatics/btr167

26. Harrell FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med. 1996 Feb 28;15(4):361–87.

27. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol. 1996 Dec;49(12):1373–9.

28. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med. 2011 Feb 20;30(4):377–99.

29. Rubin. Multiple Imputation for Nonresponse in Surveys. In: Multiple Imputation for Nonresponse in Surveys [Internet]. John Wiley & Sons, Ltd; 1987 [cited 2022

Dec 9]. p. i–xxix. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470316696.fmatter

30. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015 Apr 20;43(7):e47.

31. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology [Internet]. 2014 Dec 5 [cited 2023 Sep 20];15(12):550. Available from: https://doi.org/10.1186/s13059-014-0550-8

32. Wreczycka K, Gosdschan A, Yusuf D, Grüning B, Assenov Y, Akalin A. Strategies for analyzing bisulfite sequencing data. Journal of Biotechnology [Internet]. 2017 Nov 10 [cited 2024 Feb 1];261:105–15. Available from: https://www.sciencedirect.com/science/article/pii/S0168165617315936

33. Müller F, Scherer M, Assenov Y, Lutsik P, Walter J, Lengauer T, et al. RnBeads 2.0: comprehensive analysis of DNA methylation data. Genome Biology [Internet]. 2019 Mar 14 [cited 2024 Feb 1];20(1):55. Available from: https://doi.org/10.1186/s13059-019-1664-9

34. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. The Innovation [Internet]. 2021 Aug 28 [cited 2023 Sep 20];2(3):100141. Available from: https://www.sciencedirect.com/science/article/pii/S2666675821000667

35. methylClass: an R package to construct DNA methylation-based classification models | Briefings in Bioinformatics | Oxford Academic [Internet]. [cited 2024 Feb 1]. Available from: https://academic.oup.com/bib/article/25/1/bbad485/7513599

Acknowledgements

The authors express their gratitude to the scientific board of the *TransEAsome* project for their invaluable input in shaping its design. They also extend their appreciation to all investigation teams for their significant contributions to the establishment and execution of this study. Special thanks are extended to Katialine Groff and Dr. Rony Sfeir for their dedicated work on the oesophageal atresia national registry ReNaTo, which has been instrumental in the project's progress.

The authors also acknowledge the members of the user's committees for their insightful feedback, which has consistently enriched the project. Additionally, they express their gratitude to the paediatric clinical investigation centre at Lille Hospital for their technical support in clinical research, which has been indispensable to the authors.

This collaborative effort and support from various individuals and institutions have been essential in advancing the TransEAsome project and its objectives.

Collaborators

University hospital of Amiens, University hospital of Angers, University hospital of Besançon, University hospital of Bordeaux, University hospital of Brest, University hospital of Brest, University hospital of Caen, Public hospital of Cayenne, University hospital of Clermont Ferrand, University hospital of Dijon, University hospital of Fort de France, University hospital of La Tronche, Public hospital of Le Mans, University hospital of Limoges, University hospital of Lyon, University hospital of Marseille, University hospital of Montpellier, University hospital of Nancy, University hospital of Nantes, University hospital of Nice, Public hospital of Orléans, University hospital of Paris Armand Trousseau, University hospital of Kremelin Bicêtre, University hospital of Créteil, University hospital of Paris Necker, University hospital of Paris Robert Debré, University hospital of Poitiers, University hospital of Pointe à Pitre, University hospital of Reims, University hospital of Rennes, University hospital of Rouen, University hospital

of Saint Denis de la Réunion, University hospital of Saint Etienne, University hospital of Strasbourg, University hospital of Toulouse and University hospital of Tours.

Authors' contributions

FG is the scientific coordinator and guarantor. MF, RH, GM, MS are partner leaders in the project. MA is the Lille University Hospital's principal investigator, JV is the co-technical-head of the Billile platform, SG is head of the patient and public involvement task, JL is the head of the Lille University Hospital's biostatistics department, LD is a methodologist who revised the protocol submitted to ethics committee, MD is in charge of the omics analysis within the PRISM team, SF is a head of the ReNaTo registry, JR and VA are a members of the French national patient with OA association and ML is the project manager. All authors have contributed to the writing of the study protocol and have approved the final version of this manuscript.

Funding statement

This project has received funding from the ANR program France 2030 under the specific grant agreement ANR-21-PMRB-0011. As well as a funding from the Institution de prévoyance et Retraite Collective des Employés de Maison (IRCEM) fundation.

Competing interest statement

No competing interests.

Word count

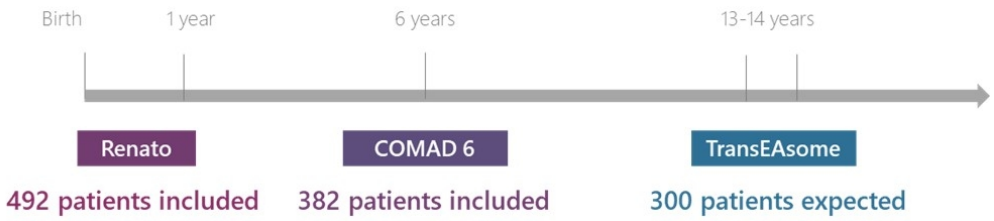
4840 words

Open Access

Figure legends

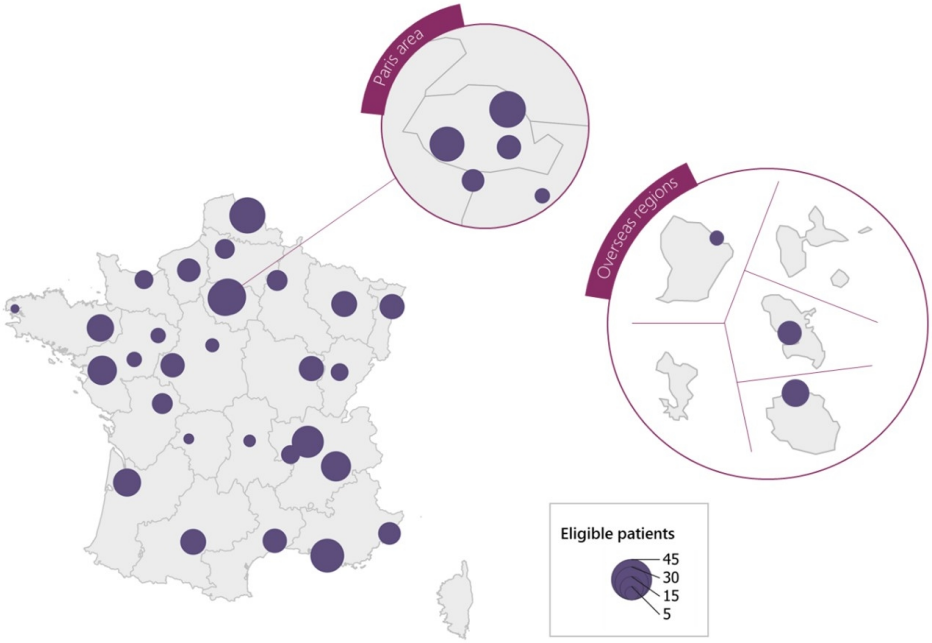
Figure 1: Follow-up of a nested cohort of OA patients born between 2010 and 2012

Figure 2: Localisation of recruitment in France



Follow-up of a nested cohort of OA patients born between 2010 and 2012

83x19mm (300 x 300 DPI)



Localisation of recruitment in France

115x76mm (300 x 300 DPI)