**A protocol for emulating four published randomised controlled trials using registry data: effects of antibiotics in cystic fibrosis**

Emily Granger1, Lorna Allen2, Susan C Charman2, Sarah L Clarke2, Gwyneth Davies3, Freddy Frost4, Hannah Plowman5, Laurie Tomlinson1, Ruth Keogh1

1 London School of Hygiene and Tropical Medicine, London, UK

2 UK Cystic Fibrosis Trust, London, UK

3 UCL Great Ormond Street Institute of Child Health, London UK

4 University of Liverpool, Liverpool, UK

5 Patient and Public Involvement representative

# Abstract

**Background.** Cystic fibrosis (CF) is an inherited disease affecting over 11,000 people in the UK. People with CF require multiple long-term treatments, and a key research priority in CF is to investigate long-term effects of treatments. Such questions are unlikely to be addressed in randomised controlled trials (RCTs) due to feasibility and cost. Real-world data provide an alternative resource for studying treatment effects but come with challenges. Target trial emulation is a framework for studying treatment effects using real-world data, which applies the study design principles of RCTs to help avoid common biases. This framework is increasingly used in many disease areas, but it’s use in CF is so far limited. The UK CF Registry is a potential resource for generating real-world evidence on treatment effects in CF, but it’s unclear how best to implement target trial emulation methodology using these data.

**Aims.** We aim this establish best practices for the application of target trial emulation methodology using UK CF Registry data to estimate effects of CF treatments.

**Methods and analysis.** We will determine optimal methods for using different components of the UK CF Registry data, including data on prescription dates and longitudinal data on clinical variables collected at annual review visits. We will conduct emulations of four published RCTs, which will allow us to compare the different methodological approaches, and help us to establish the best ways of using the registry to define the variables needed for target trial emulation and to illustrate its applicability in CF.

**Ethics and dissemination** We will seek ethical approval from the London School of Hygiene and Tropical Medicine Ethics Committee. The results of this study will be published in a peer-reviewed journal and presented at relevant scientific conferences.

**Keywords** cystic fibrosis; registries; observational study; target trial emulation.

# Introduction

A significant challenge faced by the cystic fibrosis (CF) community is the wide breadth of important clinical questions regarding treatment effectiveness that remain unanswered but are unlikely to be addressed in randomised controlled trials (RCTs), due to feasibility and cost. When RCTs are impractical, an alternative is to use real-world data. However, analyses based on real-world data are often criticized for their susceptibility to bias. Target trial emulation (TTE) [1], used alongside appropriate statistical analysis methods, offers a framework for using real-world data to study treatment effects, while helping to avoid common biases [2-3].

The TTE framework involves two steps. First, we describe the RCT we would like to conduct – the “target trial” – and secondly, we specify how each element of the target trial protocol will be emulated using real-world data. TTE is increasingly used across many disease areas [4-9] and is endorsed by the UK’s NICE Real-World Evidence Framework [10].

In CF, national patient registries are a key potential resource for generating evidence on treatment effects [11]. The UK CF Registry [12] collects longitudinal data on approximately 99% of all people with CF in the UK. Longitudinal data are collected at approximately annual review visits on variables including clinical measurements, health complications, treatment use, and from 2016, treatment prescription start and stop dates. The UK CF Registry has been used in recent TTEs, [13-15] demonstrating the feasibility of this approach within these data. However, these studies had some limitations, including assuming data were collected on a regular time grid and not making use of treatment prescription dates. This work aims to explore different ways in which the data can be used, including to help overcome these limitations.

One approach to assessing the reliability of results obtained using TTE is to emulate an existing RCT and compare the results between the trial emulation and the RCT – this is often referred to as “benchmarking” [16-17]. Benchmarking against an existing trial can provide evidence on whether observational analyses can be trusted to deliver valid estimates of treatment effects [16]. Benchmarking can also be used to compare different ways of implementing TTE methodology.

Our study aims to establish best practices for the application of TTE methodology using UK CF Registry data to estimate effects of CF treatments. Different methodological approaches will be compared in a series of trial emulations of published RCTs in CF. Our focus is on trials that study the effects of antibiotics, since these are commonly used in CF.

This protocol describes the target trials, the trial emulations using UK CF Registry data, and the details on different methodological approaches that will be compared in the trial emulations. Section 2 describes the target trials we plan to emulate. Section 3 provides details on the UK CF Registry data and Section 4 describes how the four target trials will be emulated using these data, as well as outlining the statistical methods that will be used. Section 5 discusses some of the reasons we may see differences in the results between the target trials and trial emulations.

# 2. Summary of selected trials

## 2.1 Selection of target trials

We selected four published RCTs which provide evidence on the effects of antibiotics in CF [18-21]. Trials were chosen on the basis of having a treatment strategy that is still relevant to the UK CF population, so that extensions of the trials (including long term effects of treatments) would be relevant for future trial emulations. Additionally, we selected trials with a range of different treatments, comparators (i.e., placebo-controlled trials and trials with active comparators), and outcome types, to assess how well these different design choices could be emulated using the UK CF Registry data. Table 1 summarises the primary research question addressed in each of the target trials.

|  |
| --- |
| *Table 1: Description of the primary research question addressed in each of the target trials*  |
|  | **Trial 1 [18]** | **Trial 2 [19]** | **Trial 3 [20]** | **Trial 4 [21]** |
| **Objective:** | Compare aztreonam for inhalation solution and tobramycin nebuliser solution for CF patients with airway *Pseudomonas aeruginosa* (PA) | Assess the efficacy and safety of inhaled colistimethate sodium in CF patients with chronic PA infection.  | Determine if azithromycin treatment improves lung function in paediatric CF patients uninfected with PA | Determine if an association between azithromycin use and pulmonary function exists in CF patients chronically infected with PA |
| **Population:** | CF patients, >6 years old, with PA-positive sputum culture  | CF patients, >6 years old, with chronic PA infection  | CF patients, 6-18 years old, with no culture positive for PA within 1 year | CF patients, >6 years old, within chronic PA infection  |
| **Treatment:** | Aztreonam for inhalation solution.  | Colobreathe dry powder for inhalation  | Azithromycin | Azithromycin |
| **Comparator:** | Tobramycin inhaled solution | Tobramycin inhaled solution | Placebo | Placebo |
| **Primary Outcome:** | Change in FEV1% | Change in FEV1% | Change in FEV1% | Change in FEV1% |
| **Secondary Outcomes:** | -Time to need for intravenous antibiotics-Time to first respiratory hospitalisation | -Time to first acute respiratory exacerbation-Nutritional status | -Additional pulmonary function end points-Changes in microbiology  | -Number of pulmonary exacerbations-Hospitalisation rates-Quality of life  |
| **Time:** | 24 weeks | 24 weeks | 24 weeks  | 24 weeks |
| **Setting:** | Recruitment from CF centres in Europe and the US, 2008-2010. | Recruitment from CF centres in Europe, 2003-2007.  | Recruitment from CF centres in US and Canada, 2007-2009 | Recruitment from CF centres in US, 2000-2002 |

## 2.2 Risk of bias in target trials

Target trials 1 and 2 were included in a Cochrane review on the use of long-term inhaled anti-pseudomonal antibiotics in cystic fibrosis [22] and target trials 3 and 4 were included in a Cochrane review on the use of macrolide antibiotics in cystic fibrosis [23]. The authors of these reviews assessed the risk of bias in each of the target trials. Their findings are summarised in Figure 1

Figure : Risk of bias summary: Cochrane review [22,23] authors' judgements about each risk of bias item for each included study



There was a high risk of performance bias (blinding or participants and clinicians) in target trials 1 and 2 [18-19], as they were described as open-label.

# Data Source

The four target trials outlined in Section 2 will be emulated using data from the UK CF Registry [12], which is a national database sponsored and hosted by the Cystic Fibrosis Trust, with NHS Research Ethics Committee approval. Data are collected on time-invariant variables (e.g. sex, genotype data, diagnosis data) alongside longitudinal data at approximately annual reviews clinic visits (e.g. clinical measurements, hospital admissions, health complications, chronic medications, microbiology, transplants and death). The registry was established in 1996 and in 2007, a new web-based data collection system was introduced which helped improve data quality. In 2016, the Registry moved to a new software platform with more advanced tools for data entry and tracking and started collecting data on treatment prescription start and stop dates. In this study we restrict to data from 2007 onwards, to make use of the higher quality data obtained after the web-based collection system was introduced. Data are entered onto the UK CF Registry by staff at CF centres in the UK using a secure web portal. The annual review proforma used provides details on the data available in the Registry [24].

# Research Methods

In this section we outline how the four target trials will be emulated using data from the UK CF Registry. In section 4.1. we describe aspects that are of relevance to all four trial emulations, including selecting the calendar time-period from which data will be used, defining “time zero”, and defining follow-up. In sections 4.2-4.5 we provide the full details of protocols for the target trials and the corresponding trial emulations. Section 4.6. provides an overview of the statistical analysis plans.

## 4.1 Setting

### 4.1.1Calendar time-periods

For each trial emulation, we consider using data from three calendar time-periods of three year’s duration. The first time-period corresponds to that used in the associated target trial. Using data from the same years as the target trial will ensure similar clinical settings between the target and emulated trials. However, performing a trial emulation using data from this time-period is not possible if: (1) the target trial was conducted prior to 2007 (as we restrict to data from 2007 onwards) or (2) the treatments under investigation in the target trial were not yet approved for use in CF in the UK (this would prevent us from finding sufficient numbers of people following the desired treatment strategies). Table 2 summarises the time-periods for the target trial and whether the original time-periods are possible to emulate.

|  |
| --- |
| *Table 2: Target trial time-periods and the feasibility of using these time-periods in trial emulations* |
| **Target trial**  | **Target trial time-period** | **Possible to emulate?** |
| 1  | 2008-2010 | × Aztreonam was approved for use in the UK in 2015.  |
| 2 | 2003-2007 | × Trial was conducted prior to 2007. |
| 3  | 2007-2009 | ✓  |
| 4  | 2000-2002 | × Trial was conducted prior to 2007. |

The second time-period considered is 2016-2018. This uses more recent data but ends before disease modifying treatments called CFTR modulators became widespread in clinical practice. The third time-period uses the most recent data available from 2020-2023. This allows us to compare results from the emulated trials pre- and post- the introduction of CFTR modulators. Both time-periods (pre-modulator: 2016-2018 and post-modulator: 2020-2023) are feasible for all target trials.

For each time-period, the first two years are used as the “recruitment-period”, defined as the period during which individuals are considered for inclusion in the emulated trial. Individuals are included in the emulated trial data if they meet the inclusion and exclusion criteria in at least one of their visits during the recruitment-period. Details on eligibility criteria are provided in Tables 3,5,7 and 9.

### Definition of time zero (0) and the index visit

A key consideration is how to define the time of treatment initiation in the emulated trial – we refer to this as time zero (0), and it is analogous to the time of randomisation in a randomised trial. Time is denoted using $t$ and is measured in weeks after time 0. Previous trial emulations [13-15] using these data have defined time 0 as the date of the annual review for which individuals meet the eligibility criteria. However, this does not necessarily align with the true date of treatment initiation. In the UK CF Registry, treatment use is recorded as a yes/no indicator indicating whether treatment has been prescribed any time over the past year. Therefore, we extract information on treatment use from the annual review subsequent to the review at which eligibility criteria were met. Defining time 0 as the date of the annual review for which individuals meet the eligibility criteria assumes that individuals initiated treatment on the date of that review when in reality, they may have initiated treatment at any point during the following year. This misalignment between time 0 and true treatment initiation may result in time-related biases.

In 2016, treatment prescription dates were first captured in the registry and these data can be used to more accurately define time 0. We describe two different approaches to the definition of time 0: (1) using annual review data only and (2) using prescription dates data. Figure 2 summarises the differences between the two definitions. Approach (1) will be used to define time 0 in trial emulations using data pre-2016. Both approaches will be used, and compared, in trial emulations using data post-2016.

When defining treatment groups, the “treatment strategy” will refer to the active treatment under study in the trial. The “control strategy” will refer to the comparator treatment strategy. Note that the control strategies include emulations of treatment strategies from placebo-controlled trials and active comparator trials. Individuals whose data are not consistent with either the treatment or control strategy will not be included in the emulated trial. The strategies are defined in sections 4.2.2, 4.3.2, 4.4.2 and 4.5.2.

*Approach 1: Using annual review data only*

Within each time-period, the index visit for a given individual is the first annual review visit at which they meet the eligibility criteria in the 2-year recruitment period. The follow-up visit is the next visit after time 0, where the date of that visit falls within 72 weeks after time 0. On average, we expect follow-up visits to be approximately 1 year after the index visit. We extend our follow-up time to 72 weeks to include data from individuals who have longer periods of time between clinical review visits. Individuals are included in the treated group if their treatment data is consistent with the treatment strategy at the follow-up visit. Individuals whose data are consistent with the control strategy at the follow-up visit are included in the control group. For both treated and control individuals, time 0 is defined as the date of the index visit. We use treatment information at the follow-up visit because the information recorded on treatment use at the annual review refers to treatment use over the past year.

*Approach 2: Using annual review data and prescription dates data*

In this approach, the index visit is defined as above. For control strategies which involve no treatment use (e.g., placebo), individuals are included in the control group if they do not have a prescription start date for the active treatment in the time between the index date and the subsequent annual review. Time 0 and the follow-up visit are then defined as above. For treatment strategies and control strategies which involve active treatments, individuals are included in the relevant groups if the prescription start dates data are consistent with the relevant treatment strategy in the time between the index visit and the subsequent annual review. For example, if a control strategy is defined as taking tobramycin but not aztreonam, the data would be consistent with this strategy if there was a start date for tobramycin and no start date for aztreonam. Time 0 is defined as the first date post-index visit at which the active treatment is prescribed (in our example, time 0 would be the prescription start date for tobramycin). The follow-up visit is defined as the next annual review visit which falls within 72 weeks after time 0.

Figure 2: Illustration of the different ways of defining time 0. In this example, we assume the study uses data from 2016-2018 and 2016 is the first year in which the eligibility criteria are met, and time is measured in years since the index visit.



### Follow-up period

The follow-up period for all target trials was 168 days (i.e., 24 weeks). The follow-up times in the emulated trials (i.e., the time between time 0 and the follow-up visit) will differ between individuals and will range from 0 to 72.

## Target and emulated trial 1: Eligibility criteria, treatment strategies and outcomes

In this section, we provide details on how target trial 1 will be emulated using the UK CF Registry data. Section 4.2.1 provides details on the eligibility criteria; section 4.2.2 describes the treatment strategies and section 4.2.3 lists the primary and secondary outcomes.

### 4.2.1 Eligibility criteria

Table 3 details the inclusion and exclusion criteria used in target trial 1 and how we plan to emulate these criteria using the UK CF Registry data.

|  |
| --- |
| *Table 3: Eligibility criteria for target trial 1 and the associated emulated trial*  |
| **Target trial** | **Emulated Trial** |
| Inclusion criteria:1. Males or females aged 6 years and older
2. Subjects with diagnosed CF.
3. Documented *Pseudomonas aeruginosa* in an expectorated sputum or throat swab culture within 3 months prior to screening visit.
4. Subjects must be able to provide written informed consent as necessary prior to any study related procedure.
5. Subjects must have received previous treatment with aerosolised antibiotics without demonstration of drug intolerance.
6. FEV1%<75% predicted at screening visit
7. Ability to perform reproducible pulmonary function tests.
8. Chest radiograph at screening visit without significant acute findings.

Exclusion criteria:1. Current use of oral corticosteroids in doses exceeding the equivalent of 10mg prednisone a day or 20mg prednisone every other day.
2. History of sputum or throat swab culture yielding *Burkholderia cepacia* in the previous 2 years.
3. Current requirement for daily continuous oxygen supplementation or requirement for more than 2 L/minute at night.
4. Administration of any investigational drug or device within 28 days of screening visit or within 6 half-lives of the investigational drug (whichever is longer).
5. Known local or systemic hypersensitivity to monobactam antibiotics.
6. Known allergies/intolerance to tobramycin.
7. Inability to tolerate inhalation of a short acting beta agonist.
8. Changes in or initiation of chronic azithromycin treatment within 28 days prior to screening visit.
9. Administration of antipseudomonal antibiotics by inhalation, intravenous or oral routes within the 14 days prior to randomisation.
10. Changes in antimicrobial, bronchodilator, dornase alfa, or corticosteroid medications within 7 days prior to screening visit.
11. Changes in physiotherapy technique or schedule within 7 days prior to screening visit.
12. History of lung transplantation.
13. Abnormal renal or hepatic function or serum chemistry at screening visit, defined as aspartate aminotransferase (AST), alanine aminotransferase (ALT)>5 times upper limit of normal range (ULN) or creatinine > 2times ULN.
14. Positive pregnancy test at screening visit.
15. Female of childbearing potential who is lactating or is not (in the opinion of the investigator) practicing an acceptable method of birth control.
16. Any serious or active medical psychiatric illness, which in the opinion of the investigator, would interfere with patient treatment, assessment, or compliance with the protocol.
 | Inclusion criteria:1. Males or females aged 6 years and older.
2. NA - we assume people observed in the UK CF Registry have a CF diagnosis.
3. Record of *Pseudomonas aeruginosa* growth at index visit.
4. NA. Patient consent is given prior to entering the registry.
5. No record of intolerance to inhaled antibiotics and any previous record of inhaled antibiotics use.
6. FEV1% <75% predicted at index visit,
7. Any previous record of FEV1% (this indicates ability to perform these tests).
8. Unable to emulation this criterion.

Exclusion criteria:1. Current prescription of oral corticosteroids.
2. Record of *Burkholderia cepacia* bacterial growth at the index visit, or any annual review within 2 years prior to index visit.
3. Currently using continuous oxygen therapy.
4. Unable to emulate this criterion.
5. Unable to emulate this criterion.
6. Intolerant to tobramycin.
7. Unable to emulate this criterion.
8. Change in or initiation of chronic azithromycin within 28 days prior to index visit.
9. Prescription of antipseudomonal antibiotics (i.e., colistin, tobramycin, amikacin, aztreonam, IV antibiotics).
10. Changes in nebulised antibiotics, azithromycin, bronchodilator, dornase alfa, or corticosteroid medications within 7 days prior to index visit.
11. Unable to emulate this criterion as physiotherapy dates are not available.
12. History of organ transplantation.
13. Abnormal renal or hepatic function or serum chemistry, defined as AST or ALT > 5 x ULN or serum creatinine > 2 x ULN
14. Individual has record of undelivered pregnancy.
15. Unable to emulate this criterion.
16. Unable to emulate this criterion.
 |

### 4.2.2 Treatment strategies

Target trial 1 compared the following treatment strategies:

(1) Active: Aztreonam for inhalation solution (AZLI). 75mg is administered 3 times a day for 28 days for each treatment cycle via the PARI eFlow electronic nebulizer.

(2) Comparator: Tobramycin inhalation solution (300mg) administered 2 times a day for 28 days for each treatment cycle via the PARI LC Plus nebulizer with compressor or via another nebulizer compatible with country-specified labelling.

We will emulate these using data on treatment prescription (either using prescription dates or data from annual reviews). The active group will include people who are recorded as taking aztreonam and not tobramycin and the comparator group will include people who are recorded as taking tobramycin and not aztreonam.

### 4.2.3 Outcomes

Table 4 details the outcomes investigated in target trial 1 and how we plan to emulate these criteria using the UK CF Registry data.

|  |
| --- |
| *Table 4: Outcomes for target trial 1 and the associated emulated trial*  |
| **Target trial** | **Emulated trial** |
| Primary outcomes:1. Relative change from baseline in FEV1% at day 28 (calculated using Knudson equations)
2. Mean actual change from baseline in FEV1% across 3 treatment courses:
	1. Course 1 (Week 4)
	2. Course 2 (Week 12)
	3. Course 3 (Week 20)

Secondary outcomes:1. Relative change from baseline in FEV1% at day 28 in subjects who received inhaled tobramycin for >84 days in the 12 months prior to randomisation.
2. Mean actual change from baseline in FEV1% across 3 treatment courses in subjects who received inhaled tobramycin for >84 days in the 12 months prior to randomisation.
3. Time to need for intravenous antipseudomonal antibiotics for respiratory events.
4. Time to first respiratory hospitalisation.
 | Primary outcomes: 1. Same as target trial
2. Same as target trial

Secondary outcomes:1. Same as target trial
2. Same as target trial
3. Same as target trial
4. Time to first hospitalisation.
 |

## Target and emulated trial 2: Eligibility criteria, treatment strategies and outcomes

In this section, we provide details on how target trial 2 will be emulated using the UK CF Registry data. Section 4.3.1 provides details on the eligibility criteria; section 4.3.2 describes the treatment strategies and section 4.3.3 lists the primary and secondary outcomes.

### 4.3.1 Eligibility criteria

Table 5 details the inclusion and exclusion criteria used in target trial 2 and how we plan to emulate these criteria using the UK CF Registry data.

|  |
| --- |
| *Table 5: Eligibility criteria for target trial 2 and the associated emulated trial*  |
| **Target trial** | **Emulated trial** |
| Inclusion criteria:1. Males or female aged 6 years and above.
2. Patients who are receiving Tobi in a cycle of 28 days active therapy followed by 28 days of rest from treatment.
3. Patients who, on the first day of trial medication administration (Visit 1), will have received a minimum of 2 Tobi on/off cycles immediately prior to randomisation.
4. If the patient is female and post-menarche/pre-menopausal and heterosexually active, the patient must be using adequate effective contraceptive methods.
5. Patients are required to be non-smokers or a past smoker who has not smoked within the past 12 months prior to Visit 1.
6. Each patient or parent/guardian must be capable to reading and understanding informed consent and the clinical trial information leaflet.
7. Patients must have a documented diagnosis of CF from a specialist CF unit (genotype and/or positive sweat tests).
8. Current CF condition must be clinically stable in the investigator’s opinion i.e., there must be no evidence of a current acute respiratory exacerbation at Visit 1.
9. FEV1% must be at least 25% but no more than 75%.
10. Patients with *Pseudomonas aeruginosa* (PA) infection (including colonisation), defined as either:
	1. ≥ 50% samples (minimum of 3 samples: sputum samples or throat swabs) positive for PA over the previous 12 months prior to the first day of trial medication administration (Visit 1) OR
	2. two samples (sputum samples or throat swabs) positive for PA over the previous 6 months prior to the first day of trial medication administration (Visit 1).
11. Patients’ lung function must be clinically stable (investigators decision) after completing IV therapy (elective or treatment for exacerbation) at Visit 1 prior to randomisation.
12. Patients who, on the first day of trial medication administration, will have had at least 28 but no more than 35 days off TOBI.

Exclusion criteria:1. Known sensitivity, or previous intolerance, to colistimethate sodium or beta2 agnoists.
2. Administration of any investigational drug within 28 days prior to first trial medication administration.
3. Patients who have received treatment which has permanently reduced *Pseudomonas aeruginosa* infection status will not be included (e.g., effective anti-pseudomonal vaccination and gene therapy).
4. Existence of any pre-study medical conditions which, in the judgement of the investigator, warrants exclusion from the study.
5. Patients who are pregnant or breast-feeding, or who plan to become pregnant during the study period.
6. Inability to communicate or co-operate with the investigator due to language problems, poor mental development or impaired cerebral function.
7. Objection by the patient’s usual CF caregiver to their participation in the study.
8. Inability to comply with any of the study procedures or the study regimen (including inability to use study devices).
9. Laboratory parameters failing outside the expected normal ranges for CF (investigator decision).
10. Children who in the opinion of the investigator would not be reliable in handling the devices.
11. Patients who, at Visit 1, will have had less than 28 days off inhaled colistin.
12. Patients who, at Visit 1, are receiving anti-pseudomonal agents specifically for the treatment of an exacerbation or pre-scheduled prophylactic courses of oral antibiotics. Long-term use of antibiotics for prophylaxis or anti-inflammatories is permitted, provided the regimen has been used for at least 28 days and is planned to continue throughout the study.
13. Patients who need to receive other anti-pseudomonal agent as part of their standard care during the off-cycle of Tobi.
14. Patients who are infected/colonised with *Burkholderia cepacia*.
15. Patients who are complicated by symptoms related to ABPA.
16. Patients who are awaiting heart-lung or lung transplantation, where the transplant is likely to take place within 6 months of the first day of trial medication administration.
 | Inclusion criteria:1. Males or female aged 6 years and above.
2. People who are prescribed tobramycin at index visit.
3. People who have been prescribed. tobramycin for a minimum of 2 on/off cycles (i.e., at least 4 months of tobramycin use).
4. Unable to emulate this criterion.
5. Non-smokers.
6. NA – consent to be in the Registry has been obtained.
7. NA – assume people in the Registry have a documented diagnosis of CF.
8. Unable to emulate this criterion. We assume clinical stability at index visit.
9. FEV1% must be at least 25% but no more than 75%.
10. Record of *Pseudomonas aeruginosa* bacterial growth at index visit.
11. Unable to emulate this criterion
12. Individuals who have had at least 28, but no more than 35, days off tobramycin.

Exclusion criteria: 1. Intolerant to colistin.
2. Unable to emulate this criterion
3. People who have gone from chronic/intermittent *Pseudomonas aeruginosa* to negative *Pseudomonas aeruginosa* status.
4. Unable to emulate this criterion
5. Individuals with an undelivered pregnancy.
6. NA
7. NA
8. NA
9. Unable to emulate this criterion
10. NA
11. People who have had less than 28 days off inhaled colistin.
12. Unable to emulate this criterion. Reasons for treatment prescription are not recorded in the Registry.
13. Unable to emulate this criterion.
14. Record of *Burkholderia cepacia* bacterial growth.
15. Record of ABPA.
16. Unable to emulate criterion
 |

### 4.3.2 Treatment strategies

Target trial 2 compared the following treatment strategies:

(1) Active: Colobreathe dry powder for inhalation (CDPI, one capsule containing colistimethate sodium 1 662 500 IU, twice daily)

(2) Comparator: Three 28-day cycles with twice-daily 300mg/5ml tobramycin inhaler solution (TIS)

We will emulate these using data on treatment prescription (either using prescription dates or data from annual reviews). The active group will include people who are recorded as taking colistin and not tobramycin and the comparator group will include people who are recorded as taking tobramycin and not colistin.

### 4.3.3 Outcomes

Table 6 details the outcomes investigated in target trial 2 and how we plan to emulate these criteria using the UK CF Registry data.

|  |
| --- |
| *Table 6: Outcomes for target trial 2 and the associated emulated trial*  |
| **Target trial** | **Emulated trial** |
| Primary outcomes:1. Change in mean FEV1% from baseline at week 24. (Knudson equations used).

Secondary outcomes:1. Changes in microbiological sensitivity of *Pseudomonas aeruginosa* in each treatment group as measured by in-vitro MIC.
2. Quality of life.
3. Nutritional status (weight and BMI)
4. Time to first acute respiratory exacerbation.
5. Time to first use (and duration of use) of additional anti-pseudomonal intravenous or oral antibiotics administered for the management of acute respiratory exacerbation.
 | Primary outcome: 1. Same as target trial.

Secondary outcomes:1. Unable to emulate.
2. Unable to emulate.
3. Weight and BMI z-score.
4. Unable to emulate.
5. Time to first use of IV antibiotics home or hospital.
 |

## Target and emulated trial 3: Eligibility criteria, treatment strategies and outcomes

In this section, we provide details on how target trial 3 will be emulated using the UK CF Registry data. Section 4.4.1 provides details on the eligibility criteria; section 4.4.2 describes the treatment strategies and section 4.4.3 lists the primary and secondary outcomes.

### 4.4.1 Eligibility criteria

Table 7 details the inclusion and exclusion criteria used in target trial 3 and how we plan to emulate these criteria using the UK CF Registry data.

|  |
| --- |
| *Table 7: Eligibility criteria for target trial 3 and the associated emulated trial*  |
| **Target trial** | **Emulated trial** |
| Inclusion criteria:1. Male or female, aged 6-18 years.
2. Confirmed diagnosis of CF.
3. Written informed consent.
4. Clinically stable at enrolment as assessed by the site investigator.
5. FEV1%>50.
6. Ability to comply with medication use, study visits, and study procedures.
7. Ability to swallow a 250mg tablet.

Exclusion criteria:1. Weigh less than 18.0kg.
2. Respiratory culture positive for *Pseudomonas aeruginosa*, NTM, or *Burkholderia cepacia* complex within 1 year or at screening, or AFB positive at screening.
3. Allergy to macrolide antibiotics.
4. Use of macrolide antibiotics within 60 days of screening.
5. Use of systemic corticosteroids or intravenous or oral antibiotics within 14 days of screening.
6. Initiation of high dose ibuprofen, dornase alfa, hypertonic saline or aerosolized antibiotics within 30 days of screening.
7. Chronic therapy with drugs known to have rare but serious interactions with azithromycin: amiodarone, digoxin, disopyramide, lovastatin, pimozide, rifabutin and nelfinavir.
8. Investigational drug use within 30 days of screening.
9. Laboratory abnormalities (creatinine, liver function, or neutropenia) at screening and confirmed at follow-up testing prior to randomisation.
10. History of biliary cirrhosis, portal hypertension, or splenomegaly, or splenomegaly on physical exam.
11. History of ventricular arrhythmia.
12. Other major organ dysfunction, excluding pancreatic dysfunction.
13. History of lung transplantation or currently on lung transplant list.
14. Relative decrease FEV1% >20 between the screening and enrolment visit.
15. Positive serum pregnancy test at screening.
16. Pregnant, breastfeeding, or if post-menarche female, unwilling to practice birth control during participation in the study.
17. History of alcohol, illicit drug or medication abuse within 1 year of screening in the judgement of the site investigator.
18. Presence of a condition or abnormality that in the opinion of the site investigator would compromise the safety of the subject or the quality of data.
 | Inclusion criteria:1. Males or females aged 6-18 years.
2. NA – assume confirmed diagnosis of CF if they are in the Registry.
3. NA – written consent to be in Registry has been obtained
4. NA – assume clinical stability at annual review visits.
5. FEV1%>50.
6. NA – assume ability to comply with medication use.
7. NA – assume ability to swally tablet.

Exclusion criteria:1. Weigh less than 18.0kg.
2. Positive growth cultures for *Pseudomonas aeruginosa*, NTM or *Burkholderia cepacia* at index visit or any annual review within 1 year prior to index visit.
3. Intolerance to macrolides.
4. Prescription of macrolides within 14 days prior to index visit.
5. Prescription of systemic corticosteroids or oral antibiotics, use of IV antibiotics.
6. Prescription of high dose ibuprofen, dornase alfa, hypertonic saline or inhaled antibiotics.
7. Unable to emulate this criterion.
8. Unable to emulate this criterion.
9. Abnormal renal or hepatic function or serum chemistry, defined as AST or ALT > 5 x ULN or serum creatinine > 2 x ULN.
10. History of portal hypertension.
11. History of arrhythmia (bradycardia or tachyarrhythmia).
12. Kidney/renal complications (including hypertension).
13. History of lung transplantation.
14. Unable to emulate this criterion
15. Unable to emulate this criterion
16. Individuals with an undelivered pregnancy.
17. Unable to emulate this criterion.
18. Unable to emulate this criterion.
 |

### 4.4.2 Treatment strategies

Target trial 3 compared the following treatment strategies:

(1) Active: One azithromycin 250mg tablet three times weekly for patients who weigh 40-79 lbs. Two azithromycin 250mg tablets three times weekly for patients who weigh greater than or equal to 80 lbs.

(2) Comparator: One placebo pill three times weekly for patients who weigh 40-79 lbs. Two placebo pills three times weekly for patients who weigh greater than or equal to 80lbs.

We will emulate these using data on treatment prescription (either using prescription dates or data from annual reviews). The active group will include people who are recorded as taking azithromycin and the comparator group will include people who are not recorded as taking azithromycin.

### 4.4.3 Outcomes and follow-up period

Table 8 details the outcomes investigated in target trial 3 and how we plan to emulate these criteria using the UK CF Registry data.

|  |
| --- |
| *Table 8: Outcomes for target trial 3 and the associated emulated trial*  |
| **Target trial** | **Emulated trial** |
| Primary outcome:1. Change in FEV1% from baseline to end of treatment at day 168.

Secondary outcomes:1. Additional pulmonary function end points
2. Pulmonary exacerbations
3. Changes in weight and height
4. Use of new antibiotics
5. Hospitalisations
6. Changes in microbiology
 | Primary outcome: 1. Same as in target trial.

Secondary outcomes:1. FVC1%.
2. Unable to emulate.
3. Same as target trial.
4. Same as target trial.
5. Same as target trial.
6. Same as target trial.
 |

## Target and emulated trial 4: Eligibility criteria, treatment strategies and outcomes

In this section, we provide details on how target trial 4 will be emulated using the UK CF Registry data. Section 4.5.1 provides details on the eligibility criteria; section 4.5.2 describes the treatment strategies and section 4.5.3 lists the primary and secondary outcomes.

### 4.5.1 Eligibility criteria

Table 9 details the inclusion and exclusion criteria used in target trial 4 and how we plan to emulate these criteria using the UK CF Registry data.

|  |
| --- |
| *Table 9: Eligibility criteria for target trial 4 and the associated emulated trial*  |
| **Target trial** | **Emulated trial** |
| Inclusion criteria:1. Males or female aged 6 years and above.
2. A documented diagnosis of CF.
3. Weight of 25kg or more.
4. Chronic infection with *Pseudomonas aeruginosa*, defined by a positive respiratory tract culture 1 year or more before screening and at screening.
5. FEV1%>30.

Exclusion criteria:1. *Burkholderia cepacia* complex isolated from the respiratory tract at screening or within 2 years or screening.
2. Nontuberculous mycobacteria within 2 years of screening or acid-fast bacillus smear positive at screening.
3. History of biliary cirrhosis or portal hypertension, splenomegaly on examination, or liver function results of 2 or more times the upper limit of normal.
4. Use of intravenous antibiotics, quinolones, or other oral antibiotics within 14 days of screening.
5. Use of systemic corticosteroids within 30 days of screening.
6. Initiation of tobramycin solution for inhalation, recombinant human dornase alfa solution, or high-dose ibuprofen within 60 days of screening. Long term use of these medication was permissible.
 | Inclusion criteria:1. Males or female aged 6 years and above.
2. NA – Assume people in the Registry have a documented diagnosis of CF.
3. Weight of 25kg or more.
4. Recorded *Pseudomonas aeruginosa* infection at index visit or any annual review within 1 year prior to index visit.
5. FEV1%>30.

Exclusion criteria: 1. *Burkholderia cepacia* recorded at index visit or within 2 years prior to index visit.
2. Nontuberculous mycobacteria within 2 years prior to index visit.
3. History of biliary cirrhosis or portal hypertension, or liver function results of 2 or more times the upper limit of normal.
4. Use of intravenous antibiotics or other oral antibiotics within 14 days of index visit.
5. Use of oral corticosteroids within 30 days of the index visit.
6. Initiation of tobramycin solution for inhalation, dornase alfa or high-dose ibuprofen within 60 days of index visit.
 |

### 4.5.2 Treatment strategies

Target trial 4 compared the following treatment strategies:

(1) Active: Azithromycin as 250-mg tablets. Participants who weighed less than 40kg take 1 tablet 3 days a week and participants who weight 40kg or more take 2 tables 3 days a week.

(2) Comparator: Placebo tablets. Participants who weighed less than 40kg take 1 tablet 3 days a week and participants who weight 40kg or more take 2 tables 3 days a week.

We will emulate these using data on treatment prescription. The active group will include people who are recorded as taking azithromycin and the comparator group will include people who are not recorded as taking azithromycin.

### 4.5.3 Outcomes and follow-up period

Table 10 details the outcomes investigated in target trial 4 and how we plan to emulate these criteria using the UK CF Registry data.

|  |
| --- |
| *Table 10: Outcomes for target trial 4 and the associated emulated trial*  |
| **Target trial** | **Emulated trial** |
| Primary outcomes:1. Change from baseline in FEV1% at day 168 (calculated using Knudson equations)
2. Safety, as determined by adverse events including:
	1. Self-reported symptoms
	2. Physical findings
	3. Laboratory tests
	4. Audiology
	5. Microbiology

Secondary outcomes:1. Change in FVC1%
2. Change in body weight
3. Time until first pulmonary exacerbation.
4. Number of pulmonary exacerbations.
5. Hospitalisation rates.
6. Use of non-quinolone oral antibiotics.
7. Changes in inflammatory markers.
8. Quality of life.

Exacerbations were defined as the use of intravenous antipseudomonal antibiotics or oral quinolones for 7 or more days.  | Primary outcomes: 1. Same as target trial.
2. Safety, as determined by hearing loss. The remaining safety outcomes in the target trial are not recorded in the UK CF Registry.

Secondary outcomes:1. Same as target trial.
2. Same as target trial.
3. Time until first use of intravenous antibiotics (for 7 days or more).
4. Unable to emulate.
5. Unable to emulate.
6. Unable to emulate.
7. Unable to emulate.
8. Unable to emulate.
 |

## 4.6 Statistical Analysis Plan

### 4.6.1 Notation

Let $A$ denote an indicator variable for treatment strategy ($A=1$ indicates active treatment and$A=0$indicates comparator). Let $Y\_{t}\left(0\right)$ denote the potential outcome under treatment $A=0$ at time $t,$ where $t$ denotes weeks after $t=0. $ Similarly, let $Y\_{t}\left(1\right)$ denote the potential outcome under treatment $A=1$ at time $t.$ Finally, $C$denotes the minimally sufficient set of confounding variables.

### 4.6.2 Causal estimands of interest

In all trial emulations, the primary outcome is FEV1%, which is a continuous outcome. Secondary outcomes involve variety of variable types including continuous, binary and time-to-event. Here we define the causal estimand of interest for continuous, binary and time-to-event outcomes, using the potential outcomes framework.

For continuous outcomes, the causal estimand of interest is the difference in expected means:

|  |  |
| --- | --- |
| $$E\left\{Y\_{24}\left(1\right)\right\}-E\left\{Y\_{24}\left(0\right)\right\}$$ | (1) |

For FEV1%, this will be interpreted as the expected difference in FEV1% between two hypothetical worlds: one in which everyone follows the active treatment strategy and one in which everyone follows the comparison treatment strategy.

For binary outcomes, the causal estimands of interest will include odds ratios:

|  |  |
| --- | --- |
| $$\frac{{P(Y\_{24}\left(1\right) =1)}/{P(Y\_{24}\left(1\right) =0)}}{{P(Y\_{24}\left(0\right) =1)}/{P(Y\_{24}\left(0\right) =0)}}$$ | (2) |

and risk ratios:

|  |  |
| --- | --- |
| $$\frac{P(Y\_{24}\left(1\right) =1)}{P(Y\_{24}\left(0\right) =1)}$$ | (3) |
|  |  |

For time-to-event outcomes, the cause specific hazard function will be of interest:

|  |  |
| --- | --- |
| Hazard function: $h(t)=\lim\_{∆t\to 0}\frac{1}{∆t}P\left(t<T<t+∆t, D=1 \right|T>t)$  |  |

where T is the event time and $D=1$ denotes the event of interest (as opposed to competing events such as death or organ transplant). The estimand of interest will be the cause specific hazard ratio:

|  |  |
| --- | --- |
| $$\frac{h^{1}(t) }{h^{0}(t) }$$ | (4) |

Where $h^{0}\left(t\right)$ and $h^{1}(t)$ are the potential cause specific hazards at time $t$ under treatment $A=0$ and $A=1$, respectively$.$ We assume this ration is constant over time.

These estimands are average treatment effects (ATE). The expectations and probabilities refer to the population of individuals meeting the trial eligibility criteria.

### 4.6.3 Main analysis

In our analyses, we assume positivity, no interference, consistency and conditional exchangeability (conditional on $C$) [25]. For each trial emulation, directed acyclic graphs will be developed to help inform which variables must be included in the adjustment set, $C. $We will then use augmented inverse-probability-of-treatment weighting (AIPTW) to control for potential confounding by $C$. AIPTW involves defining a treatment model (i.e., the propensity score model) and an outcome model. The AIPTW estimator for $E[Y\_{t}(a)], (a \in \{0,1\})$ is then given by:

|  |  |
| --- | --- |
| $$\hat{E}[Y\_{t}(a)]=\frac{1}{n}\sum\_{i=1}^{n}\left\{\frac{I(A\_{i}=a)Y\_{ti}}{\hat{P}(A\_{i}=a|C\_{i})}-\frac{I(A\_{i}=a\_{i})-\hat{P}(A\_{i}=a|C\_{i})}{\hat{P}(A\_{i}=a|C\_{i})}\hat{E}(Y\_{ti}|A\_{i}=a,C\_{i})\right\}$$ | (5) |

The probability of treatment, $P(A\_{i}=a|C\_{i})$ will be estimating using logistic regression with treatment indicator as the outcome and including$C $as linear predictor terms. The conditional expectations $E\left(A\_{i}=a,C\_{i}\right)$ will be estimated using multivariable linear regression models conditional on treatment, covariates, time and an interaction between treatment and time:

|  |  |
| --- | --- |
| $$Y\_{t}= β\_{0}+β\_{A}A+β\_{C}C+β\_{t}t+β\_{At}At+ε$$ | (6) |

Including time in the model allows different individuals to have different follow-up times relative to time 0. Estimates of $E\left(A\_{i}=a,C\_{i}\right)$ will be obtained using predictions from model (6) by plugging in $t=24.$ The estimator in equation 5 can be used to estimate $E\left[Y\_{24}\left(0\right)\right] $and $E[Y\_{24}(1)],$ and the difference between these two expectations is an estimate of the average treatment effect in the population. Standard errors can be obtained based on the efficient influence function [26].

The distribution of weights will be assessed using summary statistics and plots. Standardised mean differences will be used to compare the balance in the distribution of confounders between treatment and control groups in the original and weighted samples.

### 4.6.4 Secondary outcomes

The secondary outcomes include continuous variables (FEV1%, FVC%, BMI z-score and weight), binary outcomes (use of new antibiotics and changes in microbiology) and time-to-event outcomes (time to first need of IV antibiotics and time to first hospitalisation). Binary outcomes will be analysed using AIPTW but with logistic regression (if the interest is in odds ratios) or Poisson regression (if the interest is in risk ratios) for the outcome model. In time-to-event analyses, combining multivariable Cox models and propensity score weighting does not form a doubly robust estimator [27]. Versions of AIPW for time-to-event outcomes have been proposed, but these target survival probabilities rather than a hazard ratio [27]. Our estimand of interest is the cause-specific hazard ratio and these will be estimated using a weighted Cox regression. Censoring will occur at 72 weeks, or prior in the event of a competing event (death or organ transplant).

### 4.6.5 Subgroup analyses

In each emulated trial, the main analysis will be conducted on the entire cohort of individuals who meet the eligibility criteria. However, in the post-modulator time period (2020-2023), it will be of interest to assess whether our findings differ between modulator users and non-users. We will repeat the analysis in subgroups of the cohort, defined by modulator use. We define three subgroups: (1) non-users, (2) ivacaftor-users and (3) elexacaftor/tezacaftor/ivacaftor -users. Dual-therapy users will be excluded from the subgroup analyses.

### 4.6.6 Sensitivity analyses

A key assumption in our analysis is that there are no unmeasured confounders. Unfortunately, there may exist some factors that are associated with both treatment prescription and the outcome, which are not captured in the registry. Sensitivity to unmeasured confounders will be summarised using E-values.

### 4.6.7 Missing data

We will summarise the amount of missing data in each variable by treatment group. If there is missingness in variables that are usually static for long time periods, we will use a simple imputation approach. For missing visits where the value of the variable recorded at the prior visit and subsequent visit are equal, we will assume the missing value is also equal and impute accordingly. This approach will be used for variables denoting infections or treatment use which can be assumed to be chronic. Missingness patterns in the remaining missing data will be explored. If there are missing outcomes that are missing at random conditional on $C, $then a complete case analysis is appropriate. If a complete case analysis is not appropriate,more complex missing data methods such multiple imputation by chained equations may be considered.

### 4.6.8 Comparison of results against the target trial

Results from the trial emulations will be compared to those from the associated trials to determine whether they are compatible. We will use the following criteria, as were used in the RCT DUPLICATE Project [28]:

(1) do the estimated ATEs from the emulated trials replicate the direction and statistical significance of the estimated ATE in the target trial?

(2) do the estimated ATEs from the emulated trials lie within the 95% confidence intervals for the ATE estimates reported in the target trial?

(3) is there evidence against the null hypothesis of no difference between the ATE estimates from the emulated trials and that from target trial? To assess this, we calculate the standardised mean difference between the effect estimate obtained in the target trial and that obtained in the emulated trial. Evidence against the null hypothesis at the 5% level is indicated by a standardised mean difference greater than 1.96.

# 5. Potential sources of biases and limitations of the UK CF Registry data

Results from the emulated trials will be compared to the target trials, with the assumption that the target trial results represent “the truth”. Target trials 3 and 4 were rated as low risk of bias and so it can be reasonably expected that the results presented in these trials are similar to the true underlying effects that are being estimated. However, target trials 1 and 2 were rated as having high risk of performance bias. These trials were described as open-label and so participants and clinicians were not blinded to treatment allocation. When interpreting our findings, we must take into consideration the risk of bias in the results of the target trials.

There are also several limitations of the CF Registry data which may account for any differences in the results. Such limitations may lead to bias, or to differences in the study populations between the target and emulated trials. The details of these are discussed below.

#### Limitations in emulating the target trials eligibility criteria

Some criteria are impossible to emulate as the relevant variables are not collected in the registry, for example chest radiograph results. Other criteria may be impossible to emulate as they depend on the opinion of the investigator. Some of the eligibility criteria specify a timespan that is infeasible to emulate using the Registry data. For example, target trial 1 excluded participants who had changes in physiotherapy technique within 7 days prior to screening visit. Data on physiotherapy dates are not available in the CF Registry.

Some of the target trials specify eligibility criteria based on laboratory results. While these are collected in the registry, we believe that they contain a large amount of missingness. If there is insufficient data on laboratory results in the Registry, proxy variables could be used. For example, the criterion “exclude individuals with liver function tests more than twice the laboratory upper limit” can be approximated by excluding individuals with recorded acute liver failure.

Clinical stability at the time of treatment initiation is a common criterion for the target trials. We assume clinical stability at annual review visits, but when time 0 is defined for a date other than an annual clinical review visit, clinical stability cannot be assumed.

#### Limitations in emulating the target trials treatment strategies

Treatment strategies in the target trials are more precisely defined with respect to dosage. The Registry does not provide reliable data on treatment doses, and it is possible that individuals in the registry will take different doses to those given in the target trials. Moreover, there is no data on adherence and our emulated trials rely on data on treatment prescription, which may differ from actual treatment use.

#### Limitations in emulating the target trials outcomes of interest

The target trials calculate the primary outcome, FEV1%, using the Knudson equations [29]. We will use the Global Lung Initiative (GLI) equations [30] in the emulated trials as these are now more commonly used. Previous research suggests that results will be minimally affected by choice of reference equations [31].

A secondary outcome of interest is the time to first respiratory hospitalisation. Reliable data on reason for hospitalisation is not available in the Registry and we will use time to first hospitalisation as a proxy, assuming individuals are hospitalised for an exacerbation. Another secondary outcome of interest is the time to first IV antibiotics administered for respiratory exacerbation. Reasons for treatment prescription are not recorded in the Registry and so time to first prescription of IV antibiotics is used as a proxy variable with the assumption that the IV antibiotics are administered for respiratory exacerbation.

#### Confounding bias

Our analyses rely on the assumption that all confounding of the treatment-outcome association is accounted for in the analysis. It is possible that there are some factors associated with both treatment and the outcome, that are not collected in the registry and may lead to residual confounding bias. We plan a sensitivity analysis to assess how sensitive our results are to unmeasured confounders.

# 6 Ethics

This project will use anonymised data from the UK CF Registry, which has Research Ethics Approval (ref: 24/EE/0012). We will seek ethical approval from the London School of Hygiene and Tropical Medicine Ethics Committee.

# Acknowledgements

The authors would like to thank the Cystic Fibrosis Trust for the use of patient registry data to conduct this study. The authors also thank the patients, care providers and clinic coordinators at CF centres throughout the UK for their contributions to the patient registries.

# Data Statement

UK Registry Data are available following application to the UK CF Registry Research Committee. https://www.cysticfibrosis.org.uk/the-work-we-do/uk-cf-registry/apply-for-data-from-the-uk-cf-registry.

# Conflicts of Interest

GD reports speaker honoraria from Vertex Pharmaceuticals and Chiesi Ltd, and advisory board and clinical trial leadership roles with Vertex, unrelated to the current manuscript. RHK reports a speaker honorarium from Vertex Pharmaceuticals. FF reports speaker honoraria from Chiesi Ltd, Vertex Pharmaceuticals. SLC and SCC have no personal conflicts. However, their institution (Cystic Fibrosis Services Ltd (CFSL)) has received funding from Vertex Pharmaceuticals Inc. and Chiesi Farmaceutici S.p.A.to provide data capture, statistical analysis, and reporting through the UK Cystic Fibrosis Registry. CFSL previously received funding from Teva Pharmaceuticals B.V. to conduct the long-term safety study of colistimethate sodium (Colobreathe®) through the UK Cystic Fibrosis Registry.

# Funding

This work is supported by an NIHR RfPB grant [NIHR206915]. GD is supported by a UKRI Future Leaders Fellowship [MR/T041285/1]. RHK is supported by a UKRI Future Leaders fellowship [MR/S017968/1] awarded to RHK. LAT is funded by an NIHR Research Professorship [NIHR302405].

# References

[1] Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. Am J Epidemiol. 2016 Apr 15;183(8):758-64.

[2] Hernán MA, Sauer BC, Hernández-Díaz S, Platt R, Shrier I. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. J Clin Epidemiol 2016;79:70-75. DOI:10.1016/j.jclinepi.2016.04.014 9.

[3] Dickerman BA, García-Albéniz X, Logan RW, Denaxas S, Hernán MA. Avoidable flaws in observational analyses: an application to statins and cancer. Nat Med. 2019;25(10):1601-1606. DOI:10.1038/s41591-019-0597-x

[4] Wing K, Williamson E, Carpenter JR, et al. Real world effects of COPD medications: a cohort study with validation against results from randomised controlled trials. [Eur Respir J.](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8176192/) 2021 Mar; 57(3): 2001586. doi: [10.1183/13993003.01586-2020](https://doi.org/10.1183/13993003.01586-2020)

[5] Fu EL, Evans M, Clase CM, et al. Stopping Renin-Angiotensin System Inhibitors in Patients with Advanced CKD and Risk of Adverse Outcomes: A Nationwide Study. J Am Soc Nephrol 2021 Feb;32(2):424-435

[6] Cain SE, Saag MS,Petersen M, et al. Using observational data to emulate a randomized trial of dynamic treatment-switching strategies: an application to antiretroviral therapy. Int J Epidemiol. 2016 Dec 1;45(6):2038-2049

[7] García-Albéniz X, Hsu J, Hernán MA. The value of explicitly emulating a target trial when using real world evidence: an application to colorectal cancer screening. *Eur J Epidemiol.* 2017;32:495–500.

[8] Trevisan M, Fu EL, Yang X, et al. Stopping mineralocorticoid receptor antagonists after hyperkalaemia: trial emulation in data from routine care. European Journal of Heart Failure 2021; 10: 1698-1707. doi.org/10.1002/ejhf.2287

[9] Ioannou GN, Locke ER, O'Hare AM, Bohnert ASB, Boyko EJ, Hynes DM, Berry K. COVID-19 Vaccination Effectiveness Against Infection or Death in a National U.S. Health Care System: A Target Trial Emulation Study. *Ann Intern Med*. 2022;175(3):352-361.

[10] National Institute for Health and Care Excellence (NICE). NICE real world evidence framework (ECD9) 2022. Available at: https://rb.gy/9vxl6 [Accessed 15:52 11/09/2023].

[11] Davies G, Keogh RH; Cystic Fibrosis Trial Emulation Network. Trial emulation with observational data in cystic fibrosis. Lancet Respir Med. 2023 Oct;11(10):859-861. doi: 10.1016/S2213-2600(23)00328-4.

[12] Taylor-Robinson D, Archangelidi O, Carr SB, Cosgriff R, Gunn E, Keogh RH, MacDougall A, Newsome S, Schlüter DK, Stanojevic S, Bilton D; CF-EpinNet collaboration. Data Resource Profile: The UK Cystic Fibrosis Registry. Int J Epidemiol. 2018 Feb 1;47(1):9-10e.

[13] Granger E, Keogh RH, Frost F. The long-term effects of insulin use in incident cystic fibrosis-related diabetes: a target trial emulated using longitudinal national registry data. ERJ Open Res. 2022 Nov 7;8(4):00170-2022.

[14] Granger E, Davies G, Keogh RH. Emulated trial investigating effects of multiple treatments: estimating combined effects of mucoactive nebulisers in cystic fibrosis using registry data. Thorax. 2023 Oct;78(10):1011-1018.

[15] Keogh RH, Gran JM, Seaman SR, Davies G, Vansteelandt S. Causal inference in survival analysis using longitudinal observational data: Sequential trials and marginal structural models. Stat Med. 2023 Jun 15;42(13):2191-2225. doi: 10.1002/sim.9718.

[16] Matthews AA, Dahabreh IJ, Fröbert O, Lindahl B, James S, Feychting M, Jernberg T, Berglund A, Hernán MA. Benchmarking Observational Analyses Before Using Them to Address Questions Trials Do Not Answer: An Application to Coronary Thrombus Aspiration. Am J Epidemiol. 2022 Aug 22;191(9):1652-1665. doi: 10.1093/aje/kwac098.

[17] Dahabreh IJ, Matthews A, Steingrimsson JA, Scharfstein DO, Stuart EA. Using Trial and Observational Data to Assess Effectiveness: Trial Emulation, Transportability, Benchmarking, and Joint Analysis. Epidemiol Rev. 2024 Dec 16;46(1):1-16. doi: 10.1093/epirev/mxac011.

[18] Assael BM et al; AZLI Active Comparator Study Group. Inhaled aztreonam lysine vs. inhaled tobramycin in cystic fibrosis: a comparative efficacy trial. J Cyst Fibros. 2013 Mar;12(2):130-40

[19] Schuster A et al; Freedom Study Group. Safety, efficacy and convenience of colistimethate sodium dry powder for inhalation (Colobreathe DPI) in patients with cystic fibrosis: a randomised study. Thorax. 2013 Apr;68(4):344-50.

[20] Saiman L et al; AZ0004 Azithromycin Study Group. Effect of azithromycin on pulmonary function in patients with cystic fibrosis uninfected with Pseudomonas aeruginosa: a randomized controlled trial. JAMA. 2010 May 5;303(17):1707-15.

[21] Saiman L, Marshall BC, Mayer-Hamblett N, Burns JL, Quittner AL, Cibene DA, Coquillette S, Fieberg AY, Accurso FJ, Campbell PW 3rd; Macrolide Study Group. Azithromycin in patients with cystic fibrosis chronically infected with Pseudomonas aeruginosa: a randomized controlled trial. JAMA. 2003 Oct 1;290(13):1749-56. doi: 10.1001/jama.290.13.1749.

[22] Smith S, Rowbotham NJ. Inhaled anti‐pseudomonal antibiotics for long‐term therapy in cystic fibrosis. Cochrane Database of Systematic Reviews 2022, Issue 11. Art. No.: CD001021. DOI: 10.1002/14651858.CD001021.pub4

[23] Southern KW, Solis-Moya A, Kurz D, Smith S. Macrolide antibiotics (including azithromycin) for cystic fibrosis. Cochrane Database of Systematic Reviews 2024, Issue 2. Art. No.: CD002203. DOI: 10.1002/14651858.CD002203.pub5.

[24] Cystic Fibrosis Trust. Cystic Fibrosis Trust UK CF Registry Annual review proforma. 2023. Available at: <https://www.cysticfibrosis.org.uk/sites/default/files/2023-08/Annual%20review%20proforma%20-%20complete%20-%20June%202023.pdf> Accessed: 17:09 19/03/2025.

[25] Hernán MA, Robins JM (2020). *Causal Inference: What If (Chapter 3)*. Boca Raton: Chapman & Hall/CRC

[26] Fisher, A., & Kennedy, E. H. (2020). Visually Communicating and Teaching Intuition for Influence Functions. *The American Statistician*, *75*(2), 162–172. <https://doi.org/10.1080/00031305.2020.1717620>

[27] Gabriel EE, Sachs MC, Waernbaum I, Goetghebeur E, Blanche PF, Vansteelandt S, Sjölander A, Scheike T. Propensity weighting plus adjustment in proportional hazards model is not doubly robust. Biometrics. 2024 Jul 1;80(3):ujae069. doi: 10.1093/biomtc/ujae069. PMID: 39036984.

[28] Wang SV, Schneeweiss S, RCT-DUPLICATE Initiative. Emulation of Randomized Clinical Trials With Nonrandomized Database Analyses: Results of 32 Clinical Trials. JAMA. 2023 Apr 25;329(16):1376–85.

[29] Knudson RJ, Lebowitz MD, Holberg CJ, Burrows B. Changes in the normal maximal expiratory flow-volume curve with growth and aging. *Am Rev Respir Dis*. 1983; **127**: 725– 734.

[30] Quanjer PH, Stanojevic S, Cole TJ, Baur X, Hall GL, Culver BH, Enright PL, Hankinson JL, Ip MS, Zheng J, Stocks J; ERS Global Lung Function Initiative. Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations. Eur Respir J. 2012 Dec;40(6):1324-43

[31] Konstan MW, Wagener JS, VanDevanter DR, Pasta DJ, Millar SJ, Morgan WJ; Scientific Advisory Group and the Investigators and Coordinators of the Epidemiologic Study of Cystic Fibrosis. Comparison of FEV1 reference equations for evaluating a cystic fibrosis therapeutic intervention. Pediatr Pulmonol. 2017 Aug;52(8):1013-1019.