

Bridging the gap of unmet need: using observational data to emulate clinical trials that address research priorities in cystic fibrosis

Gwyneth Davies, MBChB¹, Ruth H Keogh, DPhil²

¹ Population, Policy and Practice Research and Teaching Department, UCL Great Ormond Street Institute of Child Health, University College London, London, WC1N 1EH, United Kingdom

² Department of Medical Statistics, London School of Hygiene & Tropical Medicine, Keppel Street, London, WC1E 7HT, United Kingdom

ORCID iDs:

Gwyneth Davies 0000-0001-7937-2728

Ruth Keogh 0000-0001-6504-3253

Corresponding author:

Dr Gwyneth Davies

Email: gwyneth.davies@ucl.ac.uk

Address: PPP Research and Teaching Department, UCL Great Ormond Street Institute of Child Health, 30 Guilford Street, London WC1N 1EH

The field of cystic fibrosis (CF) research and clinical care is changing rapidly, yet there remain many unanswered questions. While randomised controlled trials (RCTs) are the gold standard for generating evidence on effects of treatment, they have limitations and are not always feasible. In populations where CFTR modulators are now established as standard care, anticipated effect sizes for other treatments may be smaller, requiring larger sample sizes to demonstrate efficacy. In those ineligible for CFTR modulators, the number of potential participants for traditional RCTs may fall below that required to achieve adequate power. Novel trial designs permitting smaller sample sizes, and increased opportunity for clinical trial participation, are highly welcomed.¹ Realistically, they will still not meet demand. Robust, complimentary, alternatives to the traditional RCT are required. An alternative approach is to ‘emulate’ an RCT using existing observational data.² Trial emulation using CF patient registry data offers the possibility of assessing a range of questions that are not feasible to address in RCTs. This is particularly topical in an era of a rapidly changing treatment landscape. The breadth and complexity of unanswered questions have been highlighted by the James Lind Alliance research priority setting exercises.^{3,4} Several priority questions pertain to the effects of treatments, including: *What are the long-term effects of medications (including CFTR modulators) in CF?* and *What are the effective ways of simplifying the treatment burden of people with CF?*

The trial emulation process, summarised in figure 1, begins with a research question. This is followed by specifying the protocol of the “target trial” one would like to conduct to answer that question. The protocol should include details about eligibility criteria; treatment strategies and assignment; follow-up period; outcomes; the causal effect of interest; and statistical analysis plan. The next step is to specify how each element of the target trial protocol is to be emulated using the available observational data, before proceeding to the final step of data setup and analysis. The key element of an RCT that cannot be emulated directly using observational data is the randomisation of participants to the treatment strategies being compared. In observational data there are likely to be imbalances in the characteristics of individuals who followed different treatment regimes, resulting in confounding. This is addressed at the analysis (and sometimes design) stage. The trial emulation framework also helps resolve less recognised sources of bias that can arise when using observational data to study treatment effects.⁵

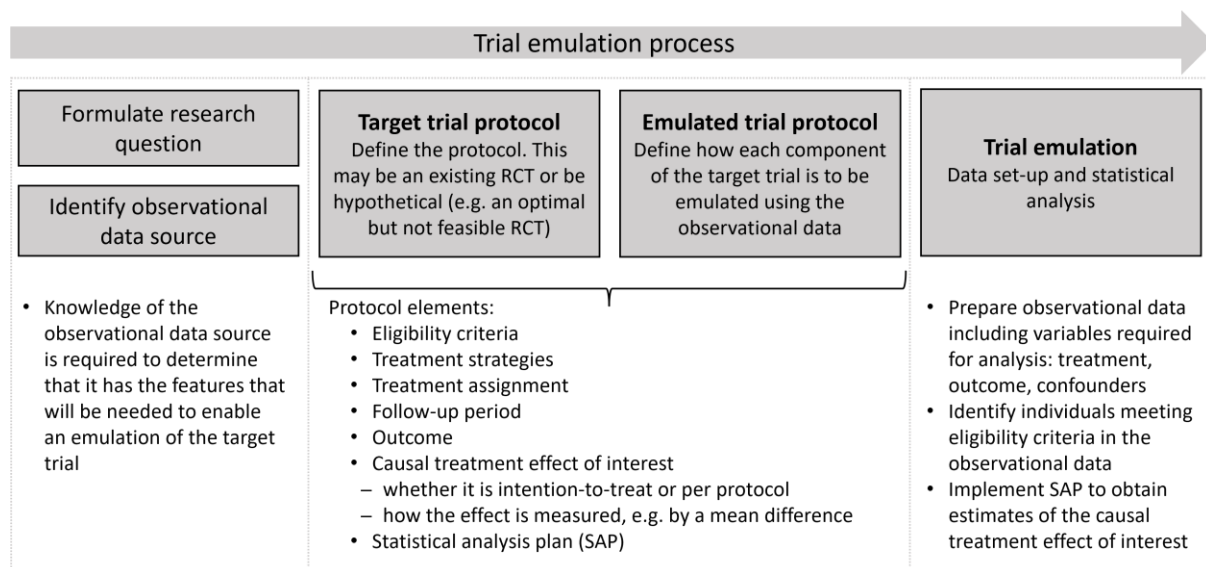
Applications of trial emulation are increasing across a number of disease areas.^{6–9} In CF, data from patient registries are ripe for enabling use of this approach. Key questions that could be assessed include those concerning generalisability of treatment effects across diverse CF populations, treatment effect heterogeneity, long-term effects, and effects of treatments used in combination. National CF patient registries are already a key resource for understanding the characteristics of CF populations and tracking changes over time. The largest are the US Cystic Fibrosis Foundation Patient Registry (~32,000 individuals)¹⁰ and the UK CF Registry (over 10,500 individuals)¹¹, which captures nearly all of the UK CF population. There is not yet widespread awareness of the trial emulation framework within the CF community or the wider respiratory disease community, although examples demonstrating its feasibility using CF registry data are emerging.^{12,13} The UK CF Registry was highlighted in the NICE Real World Evidence Framework as an example data source for generating real-world evidence on effects of interventions.¹⁴

A crucial step in establishing the potential of target trial emulation in CF will be to use registry data to emulate existing RCTs, to demonstrate the extent to which RCT findings can be replicated and identify sources of difference. Emulation of an existing trial has been referred to as ‘benchmarking’⁸ and in this

case the target trial protocol is given by that of the index RCT. Collaborative groups have already been established with the aim of emulating existing trials. These include the RCT DUPLICATE Initiative⁶ (<https://www.rctduplicate.org/>), which has published findings from emulations of 32 existing trials using US insurance claims data, and the Observational Patient Evidence for Regulatory Approval Science and Understanding Disease (OPERAND) project, which aims to examine whether real-world data can be used to inform regulatory decision making¹⁵. Initiatives that focus on the CF context are required, and this motivated the newly established CF Trial Emulation Network (CF-TEN). CF-TEN aims to lead international efforts to build the evidence base for target trial emulation in CF using national patient registries. The network is multidisciplinary and collaborative, involving clinicians, statisticians, data experts, and the patient community. CF-TEN activities include emulating existing RCTs using registry data from the pre-CFTR modulator era, to demonstrate whether the findings can be replicated for populations similar to those used in the original trial, and if not, to explore reasons for differences observed. If findings from existing trials can be replicated, this will provide confidence that emulation of target trials that are hypothetical can provide reliable evidence about treatment effects. Subsequent CF-TEN activities will include 1) repeating emulations of existing trials but in CFTR modulator-exposed populations; 2) extending emulations of existing trials to wider patient groups than those considered in the original RCTs, and to other outcomes and longer-term follow-up; and 3) emulating hypothetical trials to estimate effects of treatment strategies for which there is little evidence from RCTs, such as treatments used in combination and treatment withdrawal.

Trial emulation will not be a panacea or a replacement for RCTs. Challenges include the ability to adequately control for confounding and that certain aspects of a target trial may not be possible to emulate exactly, such as specification of treatment dose, and adherence. Trial emulation is also restricted to treatments that have been used within the populations of interest. However, trial emulation in CF using data from national patient registries presents an exciting opportunity to contribute to the evidence base for treatment effects for the benefit of the CF community.

Figure 1. Key steps in trial emulation



Acknowledgements

On behalf of the Cystic Fibrosis Trial Emulation Network (CF-TEN). GD and RHK are co-leads of CF-TEN.

Funding

RHK is funded by UK Research and Innovation (Future Leaders Fellowship MR/S017968/1). GD is funded by UK Research and Innovation (Future Leaders Fellowship MR/T041285/1).

Conflict of Interest statement

GD reports institutional fees for clinical trial leadership and advisory board roles (Vertex Pharmaceuticals) and speaker honoraria from Vertex Pharmaceuticals and Chiesi Ltd, outside the submitted work. GD is a member of the UK CF Registry steering committee. RHK reports speaker honoraria from Vertex Pharmaceuticals, outside the submitted work.

References

1. Mayer-Hamblett N, Clancy JP, Jain R, Donaldson SH, Fajac I, Goss CH, et al. New roadmaps to advance a pipeline of cystic fibrosis clinical trials: A global trial network perspective.
2. 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.
3. Rowbotham NJ, Smith S, Leighton PA, Rayner OC, Gathercole K, Elliott ZC, et al. The top 10 research priorities in cystic fibrosis developed by a partnership between people with CF and healthcare providers. *Thorax*. 2018 Apr;73(4):388–90.
4. Rowbotham NJ, Smith S, Elliott ZC, Cupid B, Allen LJ, Cowan K, et al. A refresh of the top 10 research priorities in cystic fibrosis. *Thorax* [Internet]. 2023 Jun 7 [cited 2023 Jun 11]; Available from: <https://thorax.bmj.com/content/early/2023/06/07/thorax-2023-220100>
5. 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 Nov;79:70–5.
6. 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.
7. Xie Y, Bowe B, Xian H, Loux T, McGill JB, Al-Aly Z. Comparative effectiveness of SGLT2 inhibitors, GLP-1 receptor agonists, DPP-4 inhibitors, and sulfonylureas on risk of major adverse cardiovascular events: emulation of a randomised target trial using electronic health records. *Lancet Diabetes Endocrinol* [Internet]. 2023 Jul 24 [cited 2023 Aug 9];0(0). Available from: [https://www.thelancet.com/journals/landia/article/PIIS2213-8587\(23\)00171-7/fulltext](https://www.thelancet.com/journals/landia/article/PIIS2213-8587(23)00171-7/fulltext)
8. Matthews AA, Dahabreh IJ, Fröbert O, Lindahl B, James S, Feychting M, et al. 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):1652–65.

9. Fu EL, Evans M, Clase CM, Tomlinson LA, van Diepen M, Dekker FW, et al. Stopping Renin-Angiotensin System Inhibitors in Patients with Advanced CKD and Risk of Adverse Outcomes: A Nationwide Study. *J Am Soc Nephrol JASN*. 2021 Feb;32(2):424–35.
10. Cystic Fibrosis Foundation Patient Registry. Patient Registry Annual Data Report 2021 [Internet]. Available from: <https://www.cff.org/medical-professionals/patient-registry>
11. Cystic Fibrosis Trust. Annual Registry Report 2021 [Internet]. Available from: <https://www.cysticfibrosis.org.uk/the-work-we-do/uk-cf-registry/reporting-and-resources>
12. 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 Oct;8(4):00170–2022.
13. 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* [Internet]. 2023 Jul 13 [cited 2023 Jul 19]; Available from: <https://thorax.bmj.com/content/early/2023/07/13/thorax-2023-220031>
14. Introduction to real-world evidence in NICE decision making | NICE real-world evidence framework | Guidance | NICE [Internet]. NICE; 2022 [cited 2023 Apr 10]. Available from: <https://www.nice.org.uk/corporate/ecd9/chapter/introduction-to-real-world-evidence-in-nice-decision-making>
15. Crown W, Dahabreh IJ, Li X, Toh S, Bierer B. Can Observational Analyses of Routinely Collected Data Emulate Randomized Trials? Design and Feasibility of the Observational Patient Evidence for Regulatory Approval Science and Understanding Disease Project. *Value Health*. 2023 Feb 1;26(2):176–84.