Economic Evaluation of the Liverpool Civic Data Cooperative

December 2025

oxford insights

Lateral Economics

This report was produced by Oxford Insights with support from Lateral Economics.

Authors:

Emma Hankins Law
Jasmine Kendall

Economic advisors:

Gene Tunny
Sulamaan Rahim

Research support was provided by Joram Brouwer.

Contents

Executive summary

Our analysis demonstrates that **the CDC delivers benefits that substantially exceed its costs**, generating significant economic, social, and public health value for the region. Drawing on interviews with 13 stakeholders closely involved in the programme, workshops, and desk-based research, we identify four key findings.

1. **The CDC delivers a strong return on investment through the creation of high-value data assets.**

By enabling the development of major datasets, including CIPHA, M-RIC, and CGULL, we conservatively estimate that the CDC achieves a benefit–cost ratio (BCR) of between **3.96 and 8.87**. This indicates that **for every £1 invested, between £3.96 and £8.87 of economic value is generated**. Using market-based methods, which look at the value of comparable purchasable data assets, these datasets can be valued at up to £25 million. Meanwhile, a cost-based approach, which estimates the expense of recreating the datasets through alternative means, places the value of the CDC's contribution at £46.8 million.

2. **The CDC has played a critical role in securing competitive research funding for the region.**

Stakeholder evidence confirms that the CDC was instrumental in securing **£55 million in research funding** for the Liverpool City Region, with **£6.4 million in funding solely attributable to the CDC.** Interviewees consistently identified the CDC's technical expertise, public and patient involvement and engagement (PPIE) capability, and trusted regional brand as decisive factors in the success of these bids.

3. **The CDC has enabled tangible improvements in public health outcomes and service delivery.**

Case study analysis demonstrates how the CDC's work translates directly into measurable impacts. For example, an enhanced case-finding tool supported the expansion of a remote monitoring programme, avoiding an estimated **397 emergency admissions** and generating **£145,897 in A&E cost savings**. In addition, the use of real-time NHS data to inform Safe and Well home visits increased the identification of vulnerable households from **80% to 95%**, with a projected **BCR of 1.95** if implemented at scale across the region.

4. **The full value of the CDC is likely understated in quantitative estimates.**

Throughout the analysis, we have applied conservative assumptions to ensure results are robust and defensible. However, these estimates do not fully capture important intangible benefits, including **public trust, civic leadership, and the social licence to operate**. These factors are critical enablers of data use and value creation and suggest that the CDC's true contribution to the region is likely greater than the figures presented here.

Background

What is the Civic Data Cooperative?

The Civic Data Cooperative (CDC) is a data stewardship initiative established to enhance the use of data for public good in the Liverpool City Region (LCR). Funded by the Liverpool City Region Combined Authority and hosted by the Civic Health Innovation Labs (CHIL) at the University of Liverpool, the CDC aims to connect civic organisations, industry experts, and the community to mobilise data across the region to improve the lives of residents.

The CDC conducts a wide range of activities, which vary in type and intensity across work programmes and over time. Some of the CDC's primary activities include:

- Supporting data and technical architecture design for new projects and platforms;
- designing information governance protocols for new projects and platforms;
- performing patient and public involvement and engagement (PPIE) work;
- funding staff to perform project management and technical development on projects; and
- leveraging any of the above activities to contribute to funding bids for new projects.

Purpose of this work

In 2025, the CDC commissioned SQW to conduct a series of process and impact evaluations. While these provided valuable insight into the organisation's activities and outcomes, quantifying the economic benefits of the CDC's work fell outside the scope of SQW's assignment.

To build on this foundation, Oxford Insights was commissioned to evaluate the economic dimensions of the CDC's activities. Specifically, this work involved:

- **Assessing the economic benefits** generated by the CDC's data linkage initiatives;
- **examining the overall value for money** and return on investment (ROI) of these activities; and
- **exploring potential future economic benefits** and broader economic value through a series of illustrative case studies.

This report details the findings of this evaluation, which was conducted between September and December 2025.

Project methodology

A theory-based approach to economic evaluation

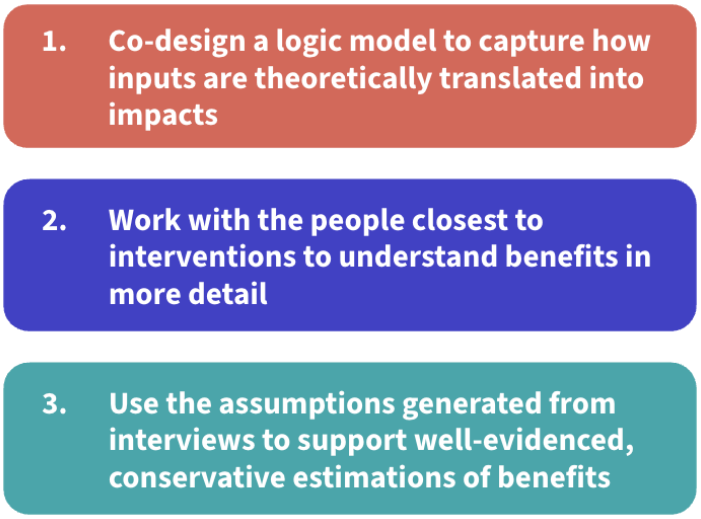


Figure 1. Approach to economic evaluation.

The evaluation adopted a theory-based approach to economic analysis, centred on a bespoke **Theory of Change (ToC)**, or logic model, which is included in the Annex of this report. This model mapped the specific pathways through which CDC interventions generate economic benefits, directly informing the selection of metrics and the foundational assumptions used throughout the study. This evaluation focused specifically on the following core outcomes:

- Researchers, SMEs, and public sector bodies have access to **more and better data** for research and decision making;

- **innovative projects** that support better public health interventions (which might take the form of new products, policies, or programmes) are enabled; and
- the Liverpool City Region receives **more funding for research and data-driven projects.**

Evidence gathering: interviews and desk research

Once the logic model was established and tested with stakeholders in a workshop, it served as the framework for identifying measurable benefits. The evidence base was built through engagement with stakeholders to capture the nature, scale, and attribution of impact. Specifically, we conducted **13 in-depth interviews** with a cross-section of experts and partners, including:

- CDC project and technical teams;
- NHS senior leadership; and
- Academic partners.

Insights from these consultations were used to calibrate the assumptions underpinning the economic modeling and to evaluate the extent to which benefits could be specifically attributed to the CDC. This attribution analysis was essential given that the CDC operates within a complex ecosystem of partners; the interviews allowed us to isolate the CDC's unique contribution while acknowledging the collaborative nature of the broader landscape.

To ground our quantitative analysis in economic best practice, we also conducted desk research on topics including:

- methods of estimating the value of data assets, particularly personal health data;
- the current market for health data assets and health data companies; and
- case study-specific topics, such as academic evaluations of home fire safety visits.

Synthesis and economic modelling

Findings from desk research and stakeholder interviews formed the empirical basis for estimating the value of economic benefits, underpinning Benefit–Cost Ratio (BCR) calculations. Throughout the analysis, we prioritised conservative estimates and prudent assumptions. By adopting this cautious baseline, we ensure that the findings are both robust and defensible, providing a high degree of confidence in the reported value.

This approach allows the report to not only quantify the CDC's economic contribution but also to articulate the specific mechanisms through which that value is realised.

Further details regarding the methodology, including an assessment of limitations, are provided in the Annex to this report.

Findings

The value of linked data

A major part of the CDC's work is supporting the development of shared and linked data assets for the public good. Because many of the CDC's activities feed into this work, estimating the value of the datasets the CDC has supported is one way of measuring the CDC's value. We took this approach to three datasets the CDC has helped build: CIPHA, M-RIC, and CGULL.

We focussed on three CDC-enabled datasets

1. ***Combined Intelligence for Population Health Action (CIPHA)***

CIPHA brings together health and social care data from different sources in the Cheshire & Merseyside Integrated Care Board. It currently includes data from GPs, community services, mental health services, hospitals, NHS Digital, social care, and national datasets. According to NHS England, there are 2,810,308 patients in C&M ICB registered with a GP, indicating a core linked dataset of 2.8 million person records within CIPHA.

2. ***Mental Health Research for Innovation Centre (M-RIC)***

Data linkage is also a major workstream of the Mental Health Research for Innovation Centre (M-RIC). M-RIC is working to develop a dataset linking data from GPs and mental

health services to improve care and develop new treatments. It currently holds records on about 921,000 patients.

3. ***Children Growing Up in Liverpool (CGULL)***

Children Growing Up in Liverpool (CGULL) is a longitudinal cohort study of children born in Liverpool since September 2025. The CDC supports the data systems that will underpin CGULL's data linkage programme to enrich its dataset.

Broadly, there are three ways of estimating the value of datasets like these:

1. **Market-based methods**, which base estimates on the market price of comparable data assets or the valuations of companies that derive their value from data;
2. **Cost-based methods**, which base estimates on the monetary costs incurred to collect and process the data; and
3. **Use-based methods**, which base estimates on the value of the potential benefits brought about by different uses of the data to different groups.[1]

For this project, we chose to use **market-based** and **cost-based methods** to estimate the value of the CDC's contributions in the form of new datasets. We made a methodological decision not to calculate use-based economic impacts of CDC-enabled datasets because data, by nature, can have almost countless potential use cases—some of which are still yet to be discovered. Many potential benefits also require significant additional investment to be fully realised, making estimates of the cost side of the equation difficult as well. That said, we have also included a few brief case studies of use-based economic impacts of the CDC's work to demonstrate its value and provide tangible examples of impact. These are not included in the CDC's overall BCR to avoid double counting or overestimating the economic benefits of the CDC based on more speculative assumptions.

[1] Frontier Economics (2021) *The Value of Data Assets: A report for the Department for Digital, Culture, Media and Sport*. https://assets.publishing.service.gov.uk/media/6399f93d8fa8f50de138f220/Frontier_Economics_-_value_of_data_assets_-_Dec_2021.pdf

Applying market-based approaches to valuing CDC-enabled datasets

Rationale

There are two potential ways of using market prices to estimate the value of a data asset:

1. Considering the market prices of **comparable data assets**; or
2. Considering the market valuations of **companies that rely heavily on data**.

The first method is the most direct because it reflects the value of data assets, rather than whole companies. However, determining the market price of a data asset requires multiple publicly available transactions of similar data, which are often unavailable due to business confidentiality. At best, market intelligence platforms may offer broad estimates of transaction details, but these are difficult to verify.

For this reason, we also used the second market-based approach of considering the market valuations of companies that rely heavily on data. Market valuations of publicly traded companies are widely available, and the details of their data assets can also be more transparent due to the reporting requirements imposed on publicly traded companies. A key limitation of this method is that company valuations reflect all company assets, not only data, so applying a discount rate can help isolate the value attributable to data alone.

Dataset values based on annual contract prices of real world data companies

In recent years, a nascent industry known as the real world data (RWD) or real world evidence (RWE) industry has developed around providing real patient health data to pharmaceutical and healthcare companies to support drug development and regulatory approval.[2] Detailed information about the datasets that RWD companies sell is not publicly available, but because these datasets contain real patient information, they are likely comparable to CDC datasets in key ways:

- **Patient records have varying amounts of data:** Patients with complex medical histories who frequently interact with healthcare services will naturally have more data in their records than patients without any conditions requiring regular

[2] Dang A. (2023). Real-World Evidence: A Primer. *Pharmaceutical medicine*, *37*(1), 25–36. https://doi.org/10.1007/s40290-022-00456-6

treatment, and we can assume that large datasets drawn from real populations will include a mix of both kinds of patients records.

- **Patient data recording may be inconsistent:** Many patient records are kept primarily for internal use within a GP surgery or hospital, and data recording practices often vary across medical institutions. This can lead to data that varies widely in structure, coding, and terminology depending on where a patient was treated.

There is currently no standard, publicly available pricing model for data assets in the RWD industry. We conducted desk research to identify the average annual contract values for top RWD companies and the number of patient records held in the data assets they sell access to. By dividing each company's average annual contract value by the number of patient records in their datasets, we can calculate a price per record for each company. We then took the median of these, estimating the market's annual price per record of real world health data at £0.08. By multiplying this price per record by the number of patient records held in the CIPHA, M-RIC, and CGULL datasets, we find the total value of the three datasets to be £299,000 per year.

To determine the value of the CDC's contribution to these datasets, we must adjust this figure in 3 ways:

1. Because we used annual contract data, we must multiply the value of each dataset by the number of years it has existed in order to find its total value.

2. In the market, companies sell access to the same RWD datasets to multiple customers at the same time, just as many organisations can access CDC datasets at the same time. To account for this, we must multiply the value of each CDC dataset by the number of organisations that access it.

3. While the CDC's work enabled each of these datasets, stakeholder research made clear that they are the result of collaborative efforts. To account for this, we created attribution rates reflecting how much of each dataset the CDC could be considered solely responsible for.[3]

After performing the calculations above, we estimated the total price of the datasets attributable to the CDC at £20.9 million. Given the CDC's total cost of £5.3 million since its

[3] We recognise that simple percentage attribution rates cannot fully capture the complexity of collaborative work. This issue is discussed further in the Limitations section.

creation, this produces a BCR of 3.96, meaning that **for every £1 invested into the CDC, £3.96 of value has been created from its data assets alone.**

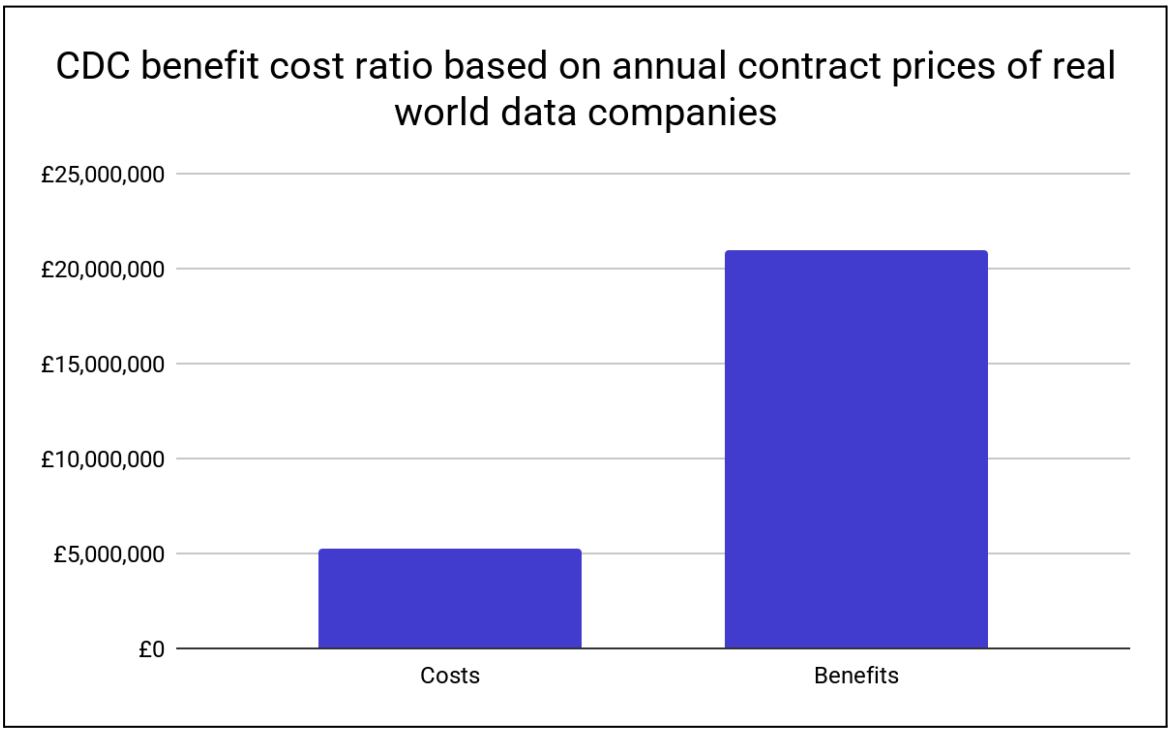


Figure 2. CDC benefit cost ratio based on market-based approach of annual contract prices of real world data companies.

Dataset values based on market valuations of companies with health data assets

Because of the limited information available on the contract prices and datasets RWD companies offer, we decided to complement this approach by exploring another way of estimating the market value of CDC-enabled datasets. This approach uses the market valuations and patient record numbers of publicly traded companies with significant health data assets to estimate a market price per record.

We found the market valuations and number of patient records held by 16 publicly traded companies focused on primary, secondary, and tertiary care records, based on previous research on the value of NHS patient records.[4] We updated these company valuations to 2025 and divided them by the number of patient records each company holds. This gave

[4] Wayman, Chris, and Hunerlach, Natasha. (2019) *Realising the value of health care data: a framework for the future.* https://www.ey.com/content/dam/ey-unified-site/ey-com/en-gl/insights/life-sciences/documents/ey-value-of-health-care-data-v20-final.pdf

us a median price per record of £54. This higher price per record compared to the annual contract price per record is likely due to multiple factors:

1. The annual contract price per record represents one year of access to a dataset, while the company valuation approach represents ownership of the dataset.
2. Company market valuations include all of a company's assets, not only data assets. This could include anything from proprietary software to office buildings.

The broadness of company market valuations is the main drawback of this approach, but we can adjust for this by creating a discounting rate that represents the percentage of a company's value that is derived from its data assets. It is extremely difficult to estimate a single percentage that would be consistent across every health data company. However, because all of the companies included are healthcare technology companies, we can assume that even many of their non-data assets, such as proprietary software, rely upon data. To account for this, we set the discounting rate at 50%, meaning these companies derive half of their value from their data assets.

After performing the calculations above and applying attribution rates to account for the CDC's partial contribution to these datasets, we estimated the total price of the datasets attributable to the CDC at £25 million. Given the CDC's total cost of £5.3 million since its creation, this produces a BCR of 4.74, meaning that **for every £1 invested into the CDC, £4.74 of value has been created from its data assets alone.**

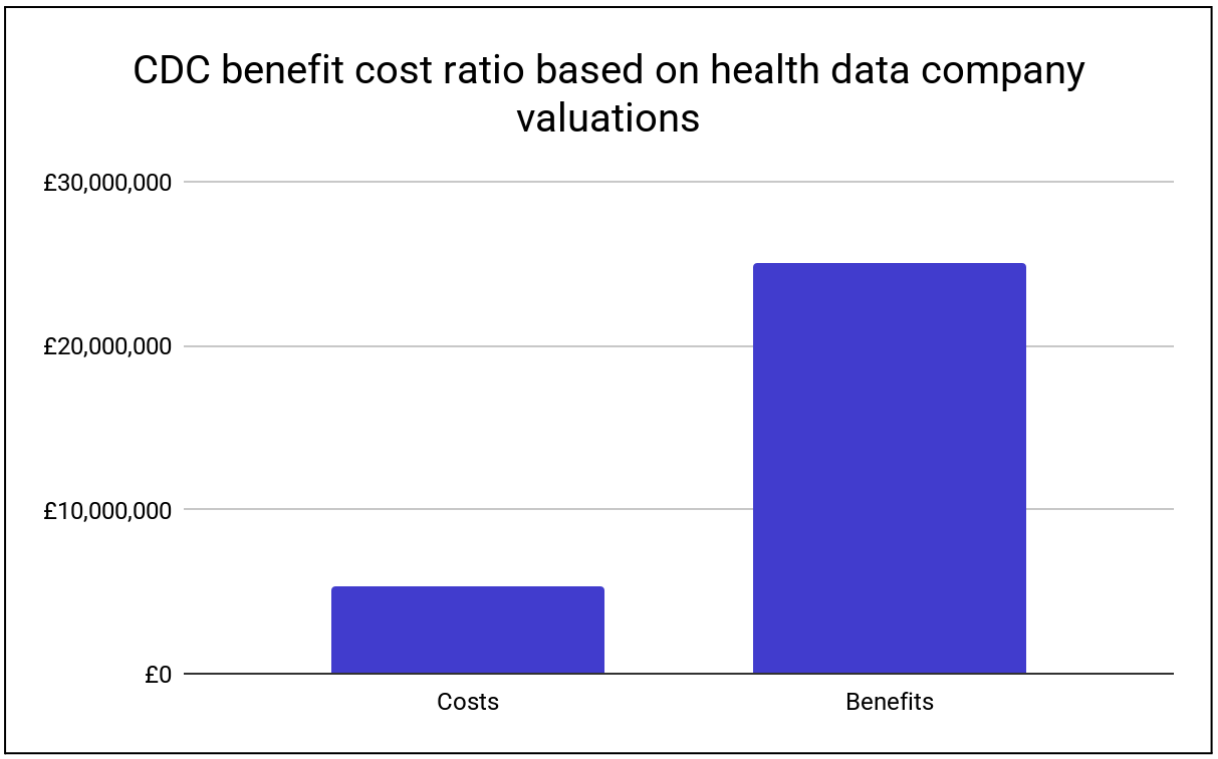


Figure 3. CDC benefit cost ratio based on market-based approach of valuations of companies with significant health data assets.

Applying a cost-based approach to valuing CDC-enabled datasets

Rationale

Cost-based methods represent an alternative approach to estimating the value of datasets. Cost-based methods attempt to estimate the value of datasets based on the monetary costs incurred to collect and process the data. There are two methods of measuring these costs:

1. **Historic cost**, or the actual cost incurred in creating the dataset; and
2. **Replacement cost**, or how much it would cost to create the dataset anew if it had not already been created.

Historic cost approaches are not entirely appropriate in this case because the health data in question is being collected as a byproduct of business processes.[5] GP data, for example,

[5] Frontier Economics (2021) *The Value of Data Assets: A report for the Department for Digital, Culture, Media and Sport*. p. 15. https://assets.publishing.service.gov.uk/media/6399f93d8fa8f50de138f220/Frontier_Economics_-_value_of_data_assets_-_Dec_2021.pdf

is being collected in the course of writing patient notes and prescribing medicine, meaning the dedicated data collection costs are minimal and nearly impossible to separate from the wider costs of providing care. Therefore, using a historic cost approach to value GP data would likely severely underestimate its value.

Instead, we can consider the **replacement cost of CDC-enabled datasets**. This method is more appropriate because a key benefit of the CDC's work is that, by linking existing data and allowing it to be used for new purposes, it eliminates the need to collect this data anew for each potential research project. By estimating the potential cost of creating new datasets of a similar type, size, and level of detail, we can thus estimate the potential value of CDC-enabled datasets.

Dataset values based on replacement cost

To estimate the potential cost of creating new health datasets, we sought to identify large health datasets comparable to CIPHA and M-RIC.[6] The closest equivalent to collecting new health data across many years, as these datasets do, is a longitudinal cohort study, in which data is collected from a group of subjects over a defined period of time.[7] Many longitudinal cohort studies exist, so we selected a sample from the UK Longitudinal Linkage Collaboration's list of partner studies.[8] We then identified the number of participants and estimated costs for each study and calculated an estimated study cost per participant.

While CDC-enabled datasets are comparable to those created by longitudinal cohort studies in terms of type of data, we know that they are less comparable in terms of level of detail or research readiness. As noted elsewhere, the data held about a patient in CIPHA or M-RIC is entirely reliant on how often and for which purposes that patient has engaged with health services. There can also be differences in data among patients who engage with the same frequency and for the same purpose, as different clinicians at different practices may have slightly different ways of collecting data. This contrasts with longitudinal cohort studies, in which the data to be collected from each participant is set out in the study design, and data dictionaries are strictly followed by each researcher. To account for this difference, we applied a discounting factor of 5% to our dataset

[6] We excluded CGULL from this valuation approach because it is itself a longitudinal cohort study, so does not rely solely on data linkage to create value in the same way as the other datasets.

[7] Taur S. R. (2022). Observational designs for real-world evidence studies. *Perspectives in clinical research, 13*(1), 12–16. https://doi.org/10.4103/picr.picr_217_21

[8] https://ukllc.ac.uk/partner-studies

valuations, based on the broad assumption that the data in CIPHA and M-RIC is only 5% as research ready as that collected in longitudinal cohort studies.[9]

We also recognise that participant recruitment and retention represents a large portion of the costs of longitudinal cohort studies, while CIPHA or M-RIC do not face these additional costs. To account for these costs, we applied a discounting factor of 50%, thus excluding half the costs of these studies from our valuations.

After performing the calculations above and applying attribution rates to account for the CDC's partial contribution to these datasets, we estimated the total price of the datasets attributable to the CDC at £46.8 million. Given the CDC's total cost of £5.3 million since its creation, this produces a BCR of 8.87, meaning that **for every £1 invested into the CDC, £8.87 of value has been created from its data assets alone.**

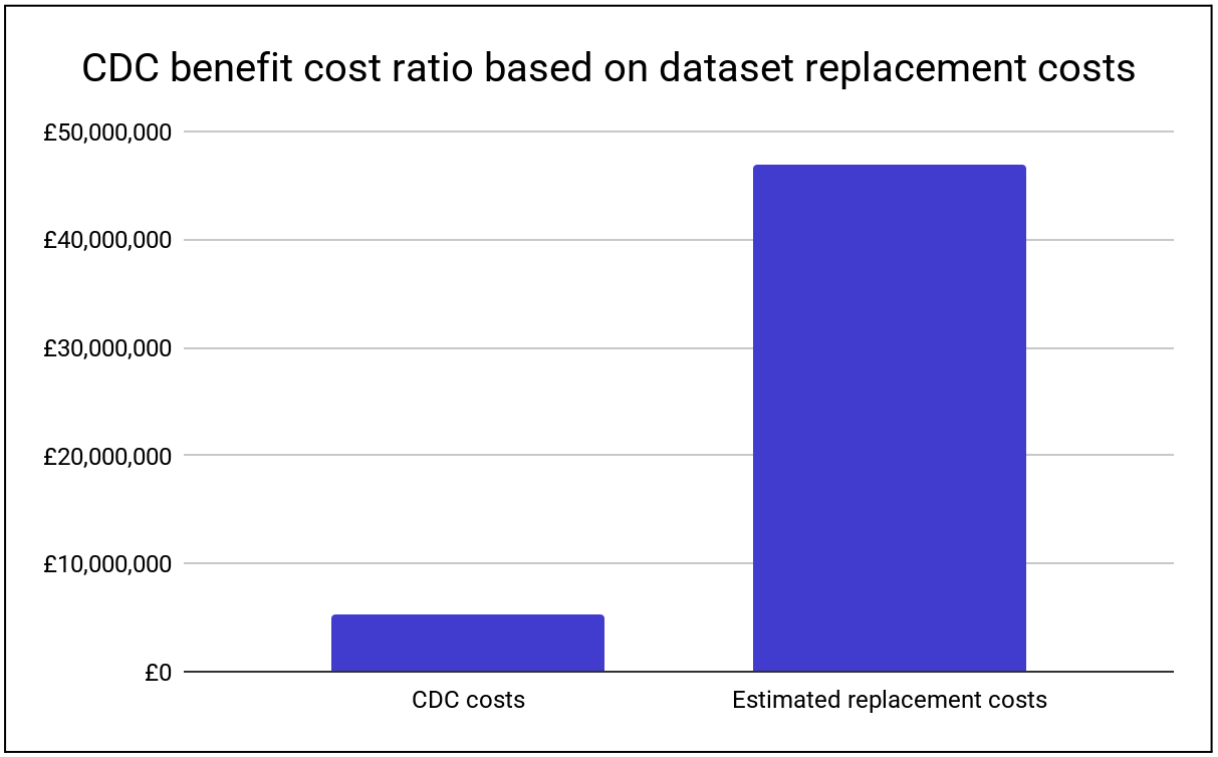


Figure 4. CDC benefit cost ratio based on cost-based approach estimating replacement costs of comparable datasets.

[9] To be clear, we are not suggesting that patient data collected by clinicians is poor quality. Rather, it is not collected for the purpose of research, and thus cannot be expected to be as rigorously standardised across patients.

Increased research funding

In addition to the above economic estimates, the CDC also has proven, direct economic impacts in the form of research funding that it has helped attract to the LCR. To avoid double counting or combining incompatible approaches, we have not included this additional research funding in the above BCRs for the CDC. However, it is still important to recognise the value of the CDC's work in this area, so we have sought to isolate the amount of new research funding for the region that can be solely attributed to the CDC.

The CDC contributes to winning research funding in multiple ways:

- By funding staff time to write grant applications;
- By leveraging the expertise of senior staff as principal or co-investigators; and
- By leveraging the expertise of the CDC more broadly, such as the organisation's technical and PPIE experience.

> *'To build that [data sharing] system you need a social licence. So that [data access] work package in M-RIC wouldn't have even been funded had they not leveraged the CDC to say we're already expert on that.'*
>
> **Interviewee working on M-RIC**

Multiple interviewees were explicit in their belief that they would not have won competitive funding bids if not for the CDC's backing (see quote).

In some cases, projects that won funding in part by leveraging the CDC have since been awarded additional funding. For example, M-RIC recently won additional work funded by the National Institute for Health and Care Research (NIHR).[10] Another project that benefited from the CDC's work as a convener suggested that the relationships the CDC had helped build for them could open new funding opportunities for them in future. While these less direct benefits are not included in our BCR, their intangible impacts should not be discounted.

To calculate how much research funding could be attributed to the CDC alone, we worked with CDC stakeholders to define projects in which the organisation had contributed to successful grant applications in the above ways. These projects totaled over £55 million in

[10] M-RIC (2025) 'Mersey Care awarded funding to transform mental health research through a Secure Data Environment'. 10 November 2025. https://mric.uk/mersey-care-awarded-funding-to-transform-mental-health-research-through-a-secure-data-environment/

funding. We sought to estimate an attribution rate for each project indicating how responsible the CDC was for each grant. However, we are sensitive to the fact that individual grants are often distributed across multiple partner organisations, who also contribute to winning funding in different ways. We recognise that simple percentage attribution rates cannot capture the full complexity of these kinds of collaborations.[11] As with our other calculations, we sought to reach agreement on assumptions that were conservative and defensible, even if they do not reflect as much complexity as we would like. We approached this by assigning standard values of 5% to two of the ways the CDC might contribute to a funding application: 1) by supplying a staff member as a principal investigator or co-investigator; and 2) by being otherwise named as an organisation in the funding application. This gave us an attribution rate for each grant, which we multiplied by the grant's value to get the funding amount that can be attributed solely to the CDC. The sum of these figures came to a total of **£6.4 million in funding leveraged for the LCR solely attributable to the CDC.**

Case studies of use-based economic impacts

As explained above, it is difficult to estimate the CDC's economic impact based on every potential use of the datasets it has enabled. This is partly because by its nature, data can be put to almost countless possible uses, including many that have yet to be created. However, we have calculated a few use-based estimates of the CDC's economic impact below to provide tangible examples of the diverse ways in which the CDC has effected change.

We purposely excluded these impacts from the CDC's overall BCR to avoid double counting or combining inconsistent approaches.

CIPHA enabled the expansion of telehealth remote monitoring in Liverpool, resulting in avoided costs for the NHS

Background

The telehealth team in Liverpool offers remote monitoring services to patients with chronic long term conditions such as COPD, diabetes, and heart failure. This service gives enrolled patients a tablet connected to vital signs measurement equipment as well as scheduled educational videos and physical and mental health questionnaires.

[11] This issue is discussed further in the Limitations section below.

Information from the tablets is transferred to a clinical hub, which raises alerts for staff if needed. The same platform allows staff to view clinical overviews and GP records for each patient. An academic evaluation of this programme found an average percentage decrease of 22.7% in emergency admissions for telehealth patients compared to an unenrolled control group with the same disease morbidity and number of emergency admissions in the last 12 months.[12]

How did the CDC add value?

The creation of CIPHA, which the CDC initiated and supported, allowed development of an enhanced case finding tool to target telehealth interventions. Whereas previously eligible patients were identified individually by GPs or other service providers, the enhanced case finding tool allowed the telehealth team to proactively identify patients who might benefit from telehealth remote monitoring. This allowed the telehealth team to expand their caseload from 250 patients in early 2024 to 2,000 patients by 2025.

How might we estimate the economic benefit?

If we assume that the 22.7% reduction in emergency admissions remains consistent for the expanded caseload, we calculate that expanding the telehealth programme led to 397 additional emergency admissions avoided. At a cost of £367.50 per A&E visit,[13] this is an **avoided cost of £145,897.50.**

CIPHA enabled more accurate targeting of Safe & Well home visits, potentially preventing fires and providing a BCR of 1.95 if scaled to the whole region

Background

Cheshire Fire and Rescue Service (CFRS) regularly conduct Safe and Well visits that involve visiting a household identified as at higher risk of an accidental dwelling fire (ADF) and sharing fire prevention information alongside signposting and referrals to other health and social care services. Safe and Well visits from Fire and Rescue Services

[12] van Berkel C, Almond P, Hughes C, et al. (2019). Retrospective observational study of the impact on emergency admission of telehealth at scale delivered in community care in Liverpool. *UKBMJ Open* 9:e028981. doi: 10.1136/bmjopen-2019-028981

[13] The King's Fund. (2025). 'Key facts and figures about the NHS.' 2 July 2025. https://www.kingsfund.org.uk/insight-and-analysis/data-and-charts/key-facts-figures-nhs

already have proven benefits, with ADFs significantly less likely to occur in homes that have had these visits.[14]

How did the CDC add value?

The creation of CIPHA and project management support from the CDC allowed development of the Safe and Well pilot programme, which helps FRSs prioritise the highest risk households for Safe and Well visits. Instead of targeting households based on whether a resident was over 65 years old and using out-of-date Experian Mosaic predictive data, the Safe and Well programme allowed FRSs to target households using near real-time NHS data on a number of health factors.

Interviewees from CFRS explained that more accurate and granular data on which to base risk scoring means that households that wouldn't be prioritised under the broad approach are more quickly prioritised under the new approach. This is reflected in early evaluation figures. CFRS found that under the previous approach, 80% of households visited had at least one vulnerability identified. Under the new approach, that increased by 15% to 95%. Given that the pilot programme visited 484 households, this means the FRS could have visited an additional 72.6 vulnerable households instead of the same amount of households with no vulnerabilities.[15]

How might we estimate the economic benefit?

We know from academic research that in 2022/23, there were 709 ADFs in vulnerable households in Merseyside that had not recently received a Safe and Well visit. Out of an area population of 1,442,081, this represents an ADF prevalence of 0.049% for vulnerable households that did not receive a visit. In the same year, there were 51 ADFs in vulnerable households that had recently received a Safe and Well visit, representing a reduced ADF prevalence of 0.0035% for this group.[16]

If we assume these rates are relevant for Cheshire in 2025 when the new targeting system was piloted, we can multiply the likelihood of fire in vulnerable homes with and without a Safe and Well visit by the number of vulnerable homes that received a visit under the pilot and might not have without the pilot. Because the pilot was small and

[14] Waring, S., Fielding, J., & Thomas, M. (2025). Examining the effectiveness and economic benefits of home fire safety visits. *Journal of Risk Research*, 1-17. https://doi.org/10.1080/13669877.2024.2447261

[15] Data collected from interview with Cheshire Fire and Rescue Service, 13 November 2025.

[16] Waring, S., Fielding, J., & Thomas, M. (2025). Examining the effectiveness and economic benefits of home fire safety visits. *Journal of Risk Research*, 1-17. https://doi.org/10.1080/13669877.2024.2447261

ADFs are still very rare, this results in 0.036 likely ADFs in these vulnerable homes without the pilot and 0.003 likely ADFs with it.

These figures are small, but their impact is clearer if we estimate the impact of the programme if the pilot were scaled up to the whole Cheshire region. If we assume that the new targeting approach were applied to all Safe and Well visits by CFRS and that CFRS continues to perform about 30,000 visits per year, a similar 15% increase in visiting vulnerable households instead of less vulnerable ones would lead to 4,500 visits to vulnerable households who might have been missed otherwise. Based on the ADF prevalence numbers given above, this could lead to 2.05 fewer ADFs that year. The Home Office suggests an economic cost of £44,931 per ADF, meaning **2.05 fewer ADFs represents an avoided cost of £92,256. Given CDC project costs of £40,211.39 and an attribution rate of 85%, the BCR can be estimated at 1.95.**

This is a conservative estimate because the CDC's costs for these projects are related to initial setup and testing, meaning costs would likely decrease over time.

CIPHA supported a successful mass event pilot during the COVID-19 pandemic

Background

The CDC originally set up the CIPHA linked data platform in early 2020 as a response to the COVID-19 pandemic. The ability to link health data from multiple sources in the Liverpool City Region supported the introduction of asymptomatic lateral flow testing in the region. This in turn led to Liverpool being included in the Events Research Programme (ERP), a UK central government programme with the goal of understanding how mass events could return while limiting COVID-19 transmission.[17] As part of this programme, Liverpool hosted three large events in April 2021: the Good Business Festival Liverpool, First Dance at the Circus Nightclub, and the Sefton Park Pilot music festival. Attendees were required to have a recent negative lateral flow test result to

[17] Department for Culture, Media and Sport. (2021). Events Research Programme: Phase I findings. 1 July 2021. https://www.gov.uk/government/publications/events-research-programme-phase-i-findings/events-research-programme-phase-i-findings

enter each event. CIPHA enabled these events to use a 'test to ticket' system that matched tickets to attendees' test results and invalidated tickets of attendees who tested positive. These events were attended by over 13,000 people and linked to only 15 cases of COVID-19. Public sentiment about the events and the approach to testing was largely positive.

How might we estimate the economic impact?

Under normal circumstances, calculating the economic impact of live events in a region would focus only on visitors from outside the region that travelled specifically for the event, as this would be considered truly 'additional' spending that would not have occurred otherwise. However, in spring 2021, mass events were banned. Any money spent on the Liverpool ERP events was unlikely to be spent in the Liverpool live events sector because it was essentially shut down. Because of this, we can consider attendee spending on these events as an economic injection into a sector that had otherwise been shut down for over a year.

First, we found attendance numbers from the ERP's Phase I report: there were 5,900 attendees at the Sefton Park Pilot outdoor music festival, 240 attendees at the ACC Liverpool conference, and 7,100 attendees at the Circus Nightclub events.[18] We found ticket prices for each event in news and marketing materials: Sefton Park tickets were £29.50,[19] Circus Nightclub tickets were £32.50,[20] and conference tickets were free. Finally, we found estimates of spending not related to tickets or accommodation for each type of event. A 2019 report estimated that the average UK festival attendee spends £67 per day on food and merchandise.[21] A 2025 report from Visit Britain estimated that UK-based business conference attendees spend an average of £45 on food, drink, and local

[18] Department for Digital, Culture, Media & Sport. *Events Research Programme: Phase I findings*. 1 July 2021. https://www.gov.uk/government/publications/events-research-programme-phase-i-findings/events-research-programme-phase-i-findings#

[19] Barrett, C. (2021). 'Festival Republic to stage 5,000-capacity Sefton Park Pilot music festival.' *Access All Areas*. 19 April 2021. https://accessaa.co.uk/festival-republic-to-stage-5000-capacity-sefton-park-pilot-music-festival/

[20] 'Liverpool gets ready for The First Dance.' Liverpool Express. 21 April 2021. https://liverpoolexpress.co.uk/liverpool-gets-ready-for-the-first-dance/

[21] Elsworthy, E. (2019). 'UK festival-goers to spend £1.2bn this summer, study finds.' The Independent. 3 July 2019. https://www.independent.co.uk/money/music-festival-uk-money-spending-habits-summer-shopping-study-a8984786.html

travel.[22] A 2019 survey found UK consumers spend an average of £69.64 on nights out.[23] This included entry fees, so we subtracted Circus Nightclub ticket fees for an estimated non-ticket spend of £37.44 per attendee.

If we multiply attendee spend by the number of attendees for each event, we estimate a **total attendee spend of £1,076,724 across the three Liverpool ERP events.** While total cost data for the programme is unavailable, the impact of a cash injection of this size should not be underestimated, given the dire straits of the Liverpool event sector in spring 2021. The work of the CDC in enabling the test to ticket system was critical in allowing these events to happen safely, as the government's final report noted: 'This test to ticket matching requirement underpins the testing and tracing infrastructure. Without it, it is impossible to reliably associate attendees' test results to events and therefore to reliably operate outbreak prevention and control for events.'[24]

Programmes like this highlight the ways in which the CDC's work can help improve decision-making in both the public and private sectors, leading to substantial socioeconomic benefits. Though the particular circumstances of COVID-19 regulations were unique, one might consider how data-driven decision-making facilitated by similar work from organisations like ADR UK might shed light on the possible benefits of data.

Limitations

It is important to acknowledge several limitations of this work.

First, a solely economic evaluation cannot capture the full scope of value that the CDC provides. While this project aimed to quantify as much of the CDC's impact as possible, some contributions are difficult to express in monetary terms. Intangible benefits—such as public trust, credibility, and strategic leadership within the community—are critical to the CDC's mission but resist precise valuation. In particular, many stakeholders noted the value of the CDC's work to create a learning health system, or a system that brings

[22] Visit Britain. (2025). *Business Events Delegate Survey*. May 2025. https://www.visitbritain.org/research-insights/business-events-research

[23] Shead, S. (2019). 'How much do you spend on a night out?' *BBC*. 18 December 2019. https://www.bbc.co.uk/news/business-50821110

[24] Department for Digital, Culture, Media & Sport. *Events Research Programme: Phase I findings*. 1 July 2021. https://www.gov.uk/government/publications/events-research-programme-phase-i-findings/events-research-programme-phase-i-findings#

researchers, policy teams, and data together to create positive feedback loops to improve healthcare.

This kind of systems thinking is clearly valuable to stakeholders, but most forms of economic valuation assume a more linear approach. This challenge was most apparent when trying to come to agreement on attribution rates for CDC datasets, funding, and projects. Stakeholders made clear that the concept of assigning percentage values of responsibility to each organisation involved in a project ran contrary to the idea that collaboration allows these organisations to create new, innovative datasets and projects that could not be created otherwise (see quote).

> *'It's a bit like asking, 'What's the economic evaluation of a spark plug?' Probably not very much, but you won't get very far in your car without it.'*
>
> ***NHS interviewee***

A second important limitation of our work is that, because of the difficulties of use-based methods of valuing data, we are primarily estimating the value of CDC-enabled datasets as private goods, meaning their benefit comes from the profit they would make if sold. However, there are clearly many potential public benefits that come from the CDC's work, such as improved health outcomes due to discoveries of new treatments. By not calculating these public benefits, we have taken an inherently conservative approach, meaning these are almost certainly underestimates. It is worth emphasising the fact that the CDC has achieved large positive BCRs despite the exclusion of public benefits.

Finally, we were unable to estimate the impacts of some of the CDC's projects because they are still underway and benefits take time to be realised. For example, the CDC's Digital Commons could yield enormous public benefit, but we excluded it from our calculations because it is not yet fully live.

Conclusion

The findings of this evaluation confirm that the CDC has substantial economic value, even when using conservative estimates of its overall benefits and attribution rates. The quantitative analysis demonstrates a strong return on investment, with BCRs ranging from 3.96 to 8.87. While the upper end of this range may seem high, it is consistent with valuations of similar international data linkage initiatives. For example, the UK's Administrative Data Research network has BCRs between 3.26 and 5.78, and Australia's

Population Health Research Network reports BCRs between 12.7 and 16.5 across various scenarios.[25] These examples show that integrated data infrastructures consistently deliver substantial policy and economic benefits.

The CDC's value goes beyond its measurable economic impact. It functions as a strategic "spark plug," the enabling infrastructure without which the region's complex data ecosystem cannot operate.

In particular, one of the CDC's most foundational contributions is the creation of a durable social licence to operate. By embedding public and patient involvement (PPIE) at its core and cultivating a trusted regional identity, the CDC has established the legitimacy and confidence required for data-driven innovation to occur at scale. This enabling role underpins all other impacts, from fostering a regional "learning health system" to positioning Liverpool alongside international leaders in responsible data stewardship.

Ultimately, the economic returns identified in this report reflect early-stage outcomes. As the CDC's current project portfolio moves from pilot to full-scale implementation—specifically as the Digital Commons infrastructure is deployed—the volume of realised benefits is expected to grow. To capture this progress, we recommend that the CDC continues to track its impact using the market-based, cost-driven, and use-case methodologies established in this report, ensuring a consistent record of the value delivered as the programme matures.

[25]Kendall, J., Martinescu, L., Rahim, S., and Tunny, G. (2024) Interim evaluation of Administrative Data Research UK: Summary report. https://oxfordinsights.com/wp-content/uploads/2025/03/Interim_evaluation_of_ADR_UK_-_Summary_report__Nov_2024.pdf
Lateral Economics. (2017).Population Health Research Network (PHRN) Impact and Return on Investment. https://lateraleconomics.com.au/wp-content/uploads/final-report-phrn-lateral-economics-oct-2017_exec-summary-final-clean.pdf

Annex I: Explanation of variables

Dataset value: Company valuation approach

Number of patient records: CIPHA dataset

According to NHS England[26], there were 2,810,308 patients in Cheshire and Merseyside Integrated Care Board registered with a GP as of October 2025. Because CIPHA updates when GPs update their data, we can assume that all of these patients will have at least a basic patient record ready for linkage in CIPHA.

Number of patient records: M-RIC dataset

Figure provided by CDC stakeholders.

Number of patient records: CGULL dataset

Figure provided by CDC stakeholders.

Price per record based on company valuation approach

For this figure, we took inspiration from a 2019 EY report[27] on the value of all NHS records. This report identified a number of companies with significant patient health data assets, the companies' valuations, and the estimated number of patient records held by each company to reach an estimated value per record. Due to the limited scope of this project, we used the same list of companies for the basis of our work, but limited the companies analysed to those in the electronic health record (EHR) and episodic or electronic medical records (EMR) domains (as opposed to the genomic, oncology, and drug development domains). We did this because we could not assume that the CIPHA and M-RIC datasets would contain linked patient genomic data, which is crucial to genomic, oncology, and pharmaceutical research.

After narrowing the company list to those in the EHR and EMR domains, we then conducted desk research to update the estimated valuation of each company for 2025. For

[26] NHS England. (2025). *Patients Registered at a GP Practice, October 2025*. 16 October 2025. https://digital.nhs.uk/data-and-information/publications/statistical/patients-registered-at-a-gp-practice/october-2025

[27] Wayman, Chris, and Hunerlach, Natasha. (2019) *Realising the value of health care data: a framework for the future.* https://www.ey.com/content/dam/ey-unified-site/ey-com/en-gl/insights/life-sciences/documents/ey-value-of-health-care-data-v20-final.pdf

companies that have been acquired or delisted since 2019, we took the most recent valuation and adjusted it for inflation so that all currencies were in 2025 GBP. We then divided the estimated number of records by the updated valuations to get an updated value per record estimate for each company. We used the median valuation as the price per record for further calculations.

Health data company valuations considered

Company	Estimated value (2025 £GBP)	Number of records	Estimated value per record (2025 £GBP)
Computer Programs and Systems, Inc	£259,201,010	18,000,000	£14
Cerner Corporation	£24,527,692,680	100,000,000	£245
AllScripts Healthcare Solutions, Inc.	£612,029,476	16,000,000	£38
athenahealth, Inc.	£5,517,868,100	106,000,000	£52
NextGen Healthcare, Inc	£1,465,929,880	240,000,000	£6
EMIS Group plc	£1,271,000,000	40,000,000	£32
Pharmagest Interactive SA	£564,897,928	135,000	£4,184
Alibaba Health Information Technology Limited	£8,475,045,120	28,000,000	£303
IQVIA Holdings Inc.	£28,849,615,200	530,000,000	£54
Inovalon Holdings Inc.	£6,799,087,680	240,000,000	£28
Medidata Solutions Inc.	£5,613,224,520	3,800,000	£1,477
Guardant Health Inc.	£9,855,394,200	70,000	£140,791
Syneos Health Inc.	£7,825,000,000	100,000,000	£78
Evolent Health Inc.	£329,812,375	2,700,000	£122
Inovalon Holdings Inc.	£6,798,770,440	240,000,000	£28

Discounting rate for value of company data assets relative to company value

As noted elsewhere in this report, a limitation of the company valuation approach to valuing data is that many things besides data assets can contribute to a company's valuation. These health data companies in particular are likely to have proprietary software or analysis services that they derive value from, in addition to their data assets. For this reason, we applied a discounting factor to better estimate the value of a company's data assets alone. We applied a factor of 50%, meaning these companies derive half of their value from their data assets. There is limited research available valuing data in this way, but we think this is a reasonable assumption given that these are health technology companies, meaning data is a critical asset that underlies many of their other services.

CDC attribution rate: CIPHA

We set the attribution rate for the CIPHA dataset at 5%, meaning the CDC is solely responsible for 5% of the value of this dataset. The CDC was very influential in the creation of CIPHA, with several of its volunteers writing the successful business case for CIPHA and designing its architecture in 2020. However, CIPHA is now hosted and maintained by the Cheshire and Merseyside Integrated Care Board as part of the NHS. The low attribution rate seeks to recognise these changes while remaining a conservative estimate.

CDC attribution rate: M-RIC

We set the attribution rate for the M-RIC dataset at 85%, meaning the CDC is solely responsible for 85% of the value of this dataset. The CDC was critical in gaining the funding for M-RIC's creation, and the CDC funds many of the project management and technical staff who are building the M-RIC dataset.

CDC attribution rate: CGULL

We set the attribution rate for the CGULL dataset at 50%, meaning the CDC is solely responsible for 50% of the value of this dataset. The CDC is responsible for implementing and maintaining all the data systems that support CGULL, and will be responsible for linking this data and making it available as the study progresses.

CDC costs

The datasets discussed in this report are the result of years of work from different aspects of the CDC. Taken together, these datasets represent nearly every activity the CDC does,

from applying for funding, to designing data sharing agreements aligned with the public's priorities, to creating the technical architecture needed for data sharing and linkage. Because of this, we have used the CDC's total grant cost of £5,278,535.

Dataset value: Annual contract values approach

Price per record based on annual contract values of real world data companies

We used the CBInsights market intelligence platform, which includes RWD as a company category, to find the average annual contract values for top RWD companies. We got the most specific ranges from a 2023 CBInsights article on the topic[28], and we supplemented this with additional companies listed in the CBInsights RWD industry scorecard. For each company, we took the median of the average annual contract value range provided. We then conducted research on each company's website to estimate the number of records each company offers access to as part of its contracts. We divided the annual contract values by the estimated number of records to reach an annual contract price per record for each company. For the companies from the 2023 article, we adjusted for inflation so that all values were in 2025 GBP. We used the median of these prices in our further calculations.

RWD company contract values considered

Company	Median annual contract value (2023 $USD)	Number of records available	Cost per record (2025 $USD)
Tempus	$1,250,000	8,500,000	$0.162
Flatiron	$450,000	4,000,000	$0.124
PurpleLab	$600,500	330,000,000	$0.002
Health Verity	$1,350,000	245,000,000	$0.006
Clarify	$675,000	300,000,000	$0.002
Cota	$1,150,000	2,000,000	$0.634
Holmusk	$300,500	32,000,000	$0.010
Atropos Health	$800,000	300,000,000	$0.003
Dandelion	$600,500	10,000,000	$0.060

[28] CBInsights. (2023). 'Here's how much pharma executives are paying for real-world data — and who they're buying data from.' 10 July 2023. cbinsights.com/research/pharma-real-world-data-vendors-cost/

ipm.ai	$300,500	300,000,000	$0.001

Years of access

Because we calculated the value of an annual contract to access data comparable to CIPHA, M-RIC, and CGULL, we had to multiply the total price of each dataset by the number of years access to it has been available. For CIPHA, this was 5 years. For M-RIC and CGULL, we used 1 year.

Number of organisations using each dataset

We understand from CDC stakeholders that almost all GP practices in the Cheshire and Merseyside ICB have agreed to data sharing with CIPHA, meaning approximately 337 practices can access the platform.[29] Besides these organisations, there are 24 other organisations listed in the data sharing register for CIPHA.[30]

Dataset value: Cost-based approach

Study cost per participant

To create an estimate of the costs to create comparable datasets as CIPHA and M-RIC from scratch, we looked at a sample of longitudinal cohort studies in the UK. We based our sample on the list of UK Longitudinal Linkage Collaboration's partner studies.[31] We excluded studies focused on genomic data because most CIPHA and M-RIC records will not include genomic data. We also excluded registries because these have a different structure to traditional longitudinal cohort studies. This gave us a list of 12 studies. We then researched each study to understand how many participants were included and how much each study cost. Where we could not find annual study costs, we extrapolated the funding data available across the length of the study. However, because these studies can run for years or decades, we only calculated the costs of the first 15 years of the study. This is to align with the length of patient medical records that are likely to be digitised in the CDC-enabled datasets. We then divided study cost by the number of participants to

[29] NHS Cheshire and Merseyside. (n.d.). 'We are NHS Cheshire and Merseyside'. Accessed 15 December 2025. https://www.cheshireandmerseyside.nhs.uk/

[30] Data into Action. (2025). 'Data use'. Accessed 15 December 2025. https://dataintoaction.cheshireandmerseyside.nhs.uk/data-use/

[31] https://ukllc.ac.uk/partner-studies

calculate an estimated study cost per participant. We used the median study cost per participant in our further calculations.

Cost discounting to exclude recruitment and retention costs

The costs of longitudinal studies are so large in part due to recruitment and retention costs, which the CIPHA and M-RIC datasets do not face. To account for this, we applied a 50% discounting rate to study costs.[32]

Discount rate to adjust for less research ready data

As noted elsewhere, the data held about a patient in CIPHA or M-RIC is entirely reliant on how often and for which purposes that patient has engaged with health services. There can also be differences in data among patients who engage with the same frequency and for the same purpose, as different clinicians at different practices may have slightly different ways of collecting data. This contrasts with longitudinal cohort studies, in which the data to be collected from each participant is set out in the study design, and data dictionaries are strictly followed by each researcher. To account for this difference, we applied a discounting factor of 5% to our dataset valuations, based on the broad assumption that the data in CIPHA and M-RIC is only 5% as research ready as that collected in longitudinal cohort studies.

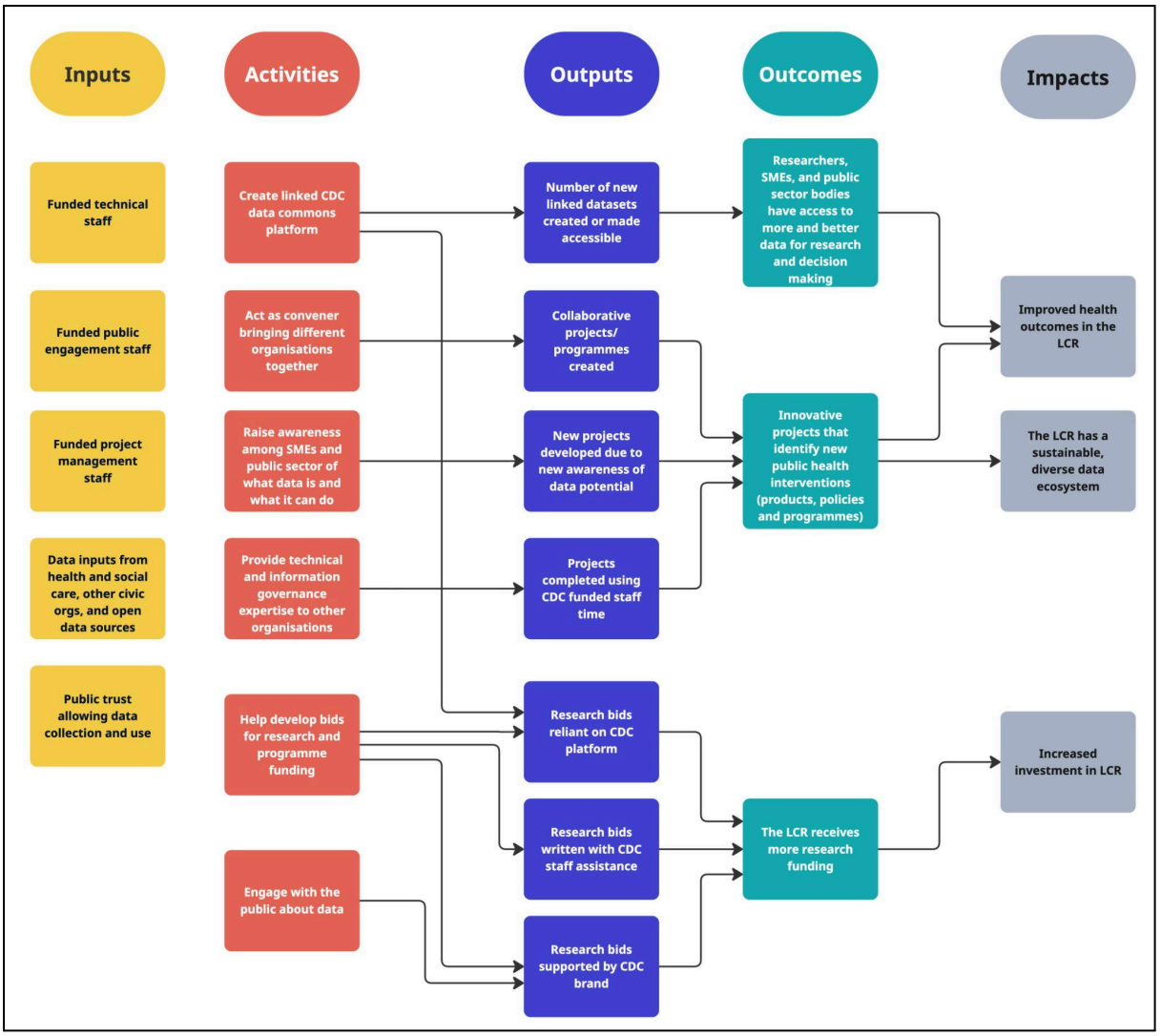
Annex II: Economic evaluation logic model for the CDC

To ground our work, we have created a logic model, clearly mapping inputs, activities, outputs, and outcomes. We began with SQW's logic model from their interim review,[33] but narrowed its scope to focus on the outputs and outcomes most relevant to an economic evaluation. We held workshops and discussions with key CDC staff and stakeholders to gain their insights, and created the working logic model shown below.

Figure 5. Working logic model for economic evaluation of the CDC.

[32] Nicholson, L. M., Schwirian, P. M., Klein, E. G., Skybo, T., Murray-Johnson, L., Eneli, I., Boettner, B., French, G. M., & Groner, J. A. (2011). Recruitment and retention strategies in longitudinal clinical studies with low-income populations. *Contemporary clinical trials*, 32(3), 353–362. https://doi.org/10.1016/j.cct.2011.01.007

[33] SQW (2025) *Review of the Liverpool City Region Civic Data Cooperative: Interim Report*, pp. 10-12.


Inputs
Activities
Outputs
Outcomes
Impacts
Funded technical staff
Funded public engagement staff
Funded project management staff
Data inputs from health and social care, other civic orgs, and open data sources
Public trust allowing data collection and use
Create linked CDC data commons platform
Act as convener bringing different organisations together
Raise awareness among SMEs and public sector of what data is and what it can do
Provide technical and information governance expertise to other organisations
Help develop bids for research and programme funding
Engage with the public about data
Number of new linked datasets created or made accessible
Collaborative projects/ programmes created
New projects developed due to new awareness of data potential
Projects completed using CDC funded staff time
Research bids reliant on CDC platform
Research bids written with CDC staff assistance
Research bids supported by CDC brand
Researchers, SMEs, and public sector bodies have access to more and better data for research and decision making
Innovative projects that identify new public health interventions (products, policies and programmes)
The LCR receives more research funding
Improved health outcomes in the LCR
The LCR has a sustainable, diverse data ecosystem
Increased investment in LCR