UK Health Security Agency

Greater Manchester self-isolation pathfinder pilot

Impact evaluation: final report

Contents

Executive summary

The Greater Manchester (GM) Pathfinder aimed to reduce the spread of coronavirus (COVID-19) in areas that had seen enduring transmission throughout the pandemic by providing enhanced support to encourage cases and household contacts to self-isolate. Pilots in 10 local authorities ran between May and September 2021. This evaluation report presents the causal impact analysis of the Pathfinder across participating local authorities on compliance with self-isolation, COVID-19 testing, and engagement with national Test and Trace (T&T).

Research questions and outcome measures

Research question	Outcome measure
Has the availability of support led to an increase in compliance with self-isolation?	Primary: weekly percentage of cases per local authority with successful T&T isolation follow-up call outcome (day 4, 7, 10 calls all successful)
	Exploratory: weekly percentage of newly isolating cases per local authority without any non-household contacts
Has the availability of support led to an increase in testing?	Secondary: weekly PCR testing rate per local authority (number of PCR tests per week / local authority population)
	Exploratory: weekly LFD testing rate per local authority (number of LFD tests per week / local authority population)
Has the availability of support led to an increase in engagement with T&T?	Exploratory: weekly average number of contacts shared with T&T per local authority
	Exploratory: weekly percentage of cases sharing at least one contact with T&T per local authority

Evaluation design

Synthetic control method, in which programme impact is estimated by comparing outcomes in GM with a synthetic comparison group, constructed using data from local authorities that share similar characteristics and pre-intervention outcomes.

Results

There was no evidence that the Pathfinder programme impacted self-isolation compliance

The proportion of successful follow-up calls was significantly higher in the treated local authorities than in the comparison group. However, this difference was due to a dip in observations in some untreated local authorities which does not likely represent a true drop in compliance. When those weeks are excluded from analysis, there is not a significant difference between the 2 groups ($p = 0.37$). It is unclear whether the Pathfinder programme protected GM

from this dip or not, but even if it did, the interpretation would be that the programme improved data quality rather than improved self-isolation compliance.

Initial investigation by UKHSA indicates that the dip was related to logistical issues rather than indicative of a change in isolation behaviour. This interpretation is reinforced by other evidence relating to compliance with self-isolation, including random probability surveys on compliance behaviour undertaken by ONS.

Additionally, there was also no statistically significant difference in the percentage of new isolating cases with no non-household contacts. This result holds regardless of the analysis approach.

There was no evidence of impact on testing for COVID-19

There was no statistically significant difference between GM and the comparison group in the secondary outcome of PCR testing rate. This result holds whether the pilot period is considered to start when the first local authorities introduce the programme, or when all treated local authorities had introduced it. It also holds if any single local authority is removed from the set of possible comparison local authorities. For the exploratory outcome of LFD testing rate, the rate was lower in GM than in the synthetic comparison group. However, there is no comparison local authority or combination of local authorities which match GM well on this outcome, and the difference is evident before the Pathfinder started. Therefore, there is low confidence that the Pathfinder caused this difference.

There was no evidence of impact on engagement with T&T

There was no statistically significant difference in the average number of contacts shared, and the percentage of cases sharing at least one contact (exploratory outcomes). As above, these results hold regardless of the analysis approach.

Interim implementation and process evaluation

Prior to the causal impact analysis, an interim quantitative implementation and process evaluation (IPE) of the Pathfinder and rapid qualitative research in 2 local authorities was conducted by BIT (for details of the implementation evaluation, see 1). The IPE highlighted the pilot's achievements in surpassing targets for both reach and engagement. By 20 August 2021, the Pathfinder had reached more than 45,000 people (351% of the original target) and provided enhanced support to over 5,000 people (156% of the original target).

The findings from the interim IPE provide tentative explanations for the findings of the impact evaluation. The lack of observed impact in these latter outcomes may be explained by: (i) variation across local authorities in awareness of enhanced support and in the number of people supported across local authorities; as well as (ii) potentially insufficient changes in the financial support offered to residents, which is reported as the primary barrier to compliance (2). Further exploration of how insights from the IPE may explain findings from the causal analysis are

considered in Discussion. A qualitative IPE conducted across all 10 local authorities is currently ongoing and may shed further light on these aspects.

More broadly, the observed absence of impact in the outcome measures investigated here is not evidence of no impact on other outcomes that may have been affected by the pilot intervention. The current evaluation results should also be interpreted in light of several limitations pertaining to the design, data, and outcome measures used.

Policy recommendations for future pilots

1.1 Empower future pilots to tackle all major barriers to self-isolation compliance: For example, whilst local authorities in GM were able to tackle 4 of the top 5 barriers to self-isolation in GM (2), the pilots did not have the scope to address the primary barrier to non-adherence (that is, financial support).

1.2 Focus on – and budget for – increasing awareness of the pilot as a key mechanism in the theory of change: For some pilots, as with the Pathfinder, awareness of the intervention may be a key mechanism in the theory of change. In this case, support should be provided – both financially and through accountability metrics – to promote the pilot. For example, some local authorities in the Pathfinder successfully used SMS, emails or mail-outs to expand their reach.

1.3 Ensure pilots engage in consistent and regular reporting of progress during implementation: For example, Greater Manchester's weekly Summary Tracker provides a model for how pilots can monitor progress towards their goals, identify outliers for further investigation and encourage sharing of lessons during implementation.

1.4 Support pilots to be responsive and adaptive: For example, local adaptations in the pilots in Wigan and Oldham offer examples for how pilots can ensure that the support offered best meets the needs of their local community.

Recommendations for future pilot evaluations

1.1 Commission dedicated 'data exploration' stages before committing to a design or to an evaluation altogether: This will help ensure data is of sufficient quality before proceeding with evaluations.

1.2 Test and refine data sharing and access procedures during a dedicated 'data exploration' stage: This includes considerations around the way data is stored and shared and, as prepared by DHSC for this evaluation, accompanying data dictionaries which can facilitate evaluation design.

1.3 Commission mixed-methods evaluations which focus on evaluating impact as well as implementation and process: These rely on quantitative and qualitative methods to evaluate whether a programme worked, why it did or did not work, and the experiences of those involved. For example, the evaluation of the Pathfinder programme includes this quantitative impact evaluation, as well as a quantitative IPE and a separate qualitative IPE focusing on all ten pilots.

1.4 Build in evaluation as an integral part of the pilot design and planning: This includes (i) developing a logic model that can drive the development of both intervention and evaluation, (ii) ensuring that the pilot intervention and delivery are suitable for evaluation using available data, (iii) potentially implementing the pilot as a randomised controlled trial to ensure a robust evaluation.

Options for future research for this programme

1.1 Conduct separate evaluations of each pilot in GM to address variation across local authorities: This will test whether specific local authorities within GM have experienced statistically significant impact from the programme.

1.2 Conduct complementary difference-in-differences analyses (at the ward level for PCR and LFD testing rate, and person level for all other outcomes): This may increase statistical power due to the use of disaggregated observations, as well as potentially overcome limitations of the current matching approach.

1.3 Interpret impact evaluation findings together with the qualitative implementation and process evaluation, once completed: This will help contextualise the observed pattern of results in the impact evaluation and quantitative IPE. Conclusions from this IPE are expected in February 2022.

Intervention: the self-isolation pilot

Introduction

The Greater Manchester (GM) Pathfinder aimed to reduce the spread of COVID-19 in areas which had seen enduring transmission throughout the pandemic, by providing enhanced support to encourage cases and household contacts to self-isolate (for a simplified logic model, see Figure 2).

Greater Manchester's 10 local authorities participated in the programme: Bolton, Bury, Manchester, Oldham, Rochdale, Salford, Stockport, Tameside, Trafford, and Wigan. In some local authorities, the entire local authority participated; in others, participation was only in a number of target wards and the remainder of the local authority did not participate.[1] Target wards tended to be those with high COVID-19 case rates, low vaccination rates, and those with high deprivation rates (see Table 1 for more details on participation and a summary of the selection criteria).

The programme launched in May 2021 for an initial period of 12 weeks, with individual pilots launching between May and June (the first pilot launched in the week commencing 17 May 2021 in Bury and the last pilots launched in the week commencing 14 June 2021 in Salford, Tameside, and Trafford; see Table 1).[2] The interim evaluation of the programme implementation by BIT found that by 20 August 2021, the pilots had reached more than 45,000 people, 351% of the original target (for details of the implementation evaluation, see reference 1).

[1] Some local authorities changed their target area after launch, with several expanding to include a wider area (for example, the entire local authority, in Wigan) and one reducing the target area (from the entire local authority to several target wards, in Bury).
[2] In practice, some local authorities ran the programme for 11 weeks (Tameside, Salford), whereas others ran it for 13 weeks (Trafford).

Table 1. Geographical targets and pilot dates for each local authority. Decision criteria around geographical targets are as reported by local authority representatives

Local authority	Launch week	End week	Planned target area	Decision criteria to pick target wards (or entire local authority)
Bolton	07/06/2021	23/08/2021	Target wards (wards within BL3 postcode area)[3]	Not available[4]
Bury	17/05/2021	02/08/2021	Entire local authority[5]	• highest COVID-19 rates • lowest vaccination rates • highest deprivation rates
Manchester	24/05/2021	16/08/2021	Target wards (Longsight, Levenshulme, Cheetham, Crumpsall, Rusholme, Ardwick, Moss Side, Whalley Range)	• high case rates and evidence of rapid transmission, especially among younger age groups and South Asian communities • sustained and enduring high levels of community transmission, clusters and outbreaks • significantly lower than average uptake with vaccinations
Oldham	24/05/2021	09/08/2021	Entire local authority	High overall COVID-19 case numbers
Rochdale	07/06/2021	23/08/2021	Target wards (Milkstone and Deeplish, Central Rochdale, Kingsway, West Middleton)	Not available

[3] As of 29 June 2021 (3 weeks after launch), this was expanded to the entire local authority.
[4] Not available indicates that this level of detail was not made available.
[5] As of 21 June 2021 (5 weeks after launch), this was reduced to 5 target wards (East, Moorside, Radcliffe West, Besses, Sedgley). [5] If too resource-intensive, they would consider targeting areas with consistently high rates. Widespread transmission across local authority (no particular geographic hotspots). At individual-level, using a stepped approach (texts for all cases, phone calls for 25% hard to reach, door-knocking for 40% to 50% hardest to reach). As of 4 August 2021 (7.5 weeks after launch), this was expanded to all wards within M32 and M33 postcodes.

Local authority	Launch week	End week	Planned target area	Decision criteria to pick target wards (or entire local authority)
Salford	14/06/2021	30/08/2021	Target ward (Little Hulton)	Not available
Stockport	07/06/2021	30/08/2021	Entire local authority	Not available
Tameside	14/06/2021	30/08/2021	Entire local authority	High overall COVID-19 case numbers
Trafford	14/06/2021	13/09/2021	Target wards (Bucklow St Martins, Clifford)	Not available
Wigan	24/05/2021	16/08/2021	Target wards (Aspull, New Springs and Whelley; Atherleigh; Atherton; Hindley; Hindley Green; Standish with Langster; Wigan Central; Mossley Common and Tyldesley)[6]	Since proactive contact was not converted into enhanced support in original target wards, Wigan shifted on week closing 21 June 2022 their focus to local authority-wide CTAS referrals or contacts, local self-isolation referrals for support and proactive door-knocking for hardest to reach (for example, via phone) and high potential to not self-isolate.

[6] As of 21 June 2021 (4 weeks after launch), this was expanded to the entire local authority.

The national context for the Pathfinder

The Pathfinder was delivered in a period of significant change in the UK's response to COVID-19, as the UK Government implemented its 'Roadmap out of lockdown' for England, gradually easing restrictions that had directed the public's behaviours in varying forms since March 2020 (3).

However, most important for the Pathfinder, was a significant change to self-isolation requirements which took place while many of the pilots were still running. From 17 May, when the first pilot in Bury was launched, until 16 August, people in England were required to self-isolate if they were identified as a close contact of a positive coronavirus case. However, from 16 August, people who were double vaccinated or under 18 were no longer legally required to self-isolate if identified as a contact of a positive case. Rather, they were advised to take a PCR test to confirm their coronavirus status. Only double-vaccinated individuals who subsequently tested positive were required to self-isolate (4). Whilst 38% of adults in GMCA were double vaccinated by 23 May 2021 when many pilots started (5), this had risen to 67% by 16 August (6). This change – and variation in vaccination rates across the local authorities (7) – will have affected the pool of individuals for whom self-isolation was mandatory.

Through the pilot's delivery, 2 forms of support with self-isolation were available at the national level. First, NHS volunteer responders were available through a national call center to provide assistance with collecting shopping, medicines and loneliness (8). Second, a £500 Test and Trace Support Payment was potentially available for individuals meeting eligibility requirements (for example, low income, inability to work from home, or facing financial hardship) (9).

Overview of delivery model

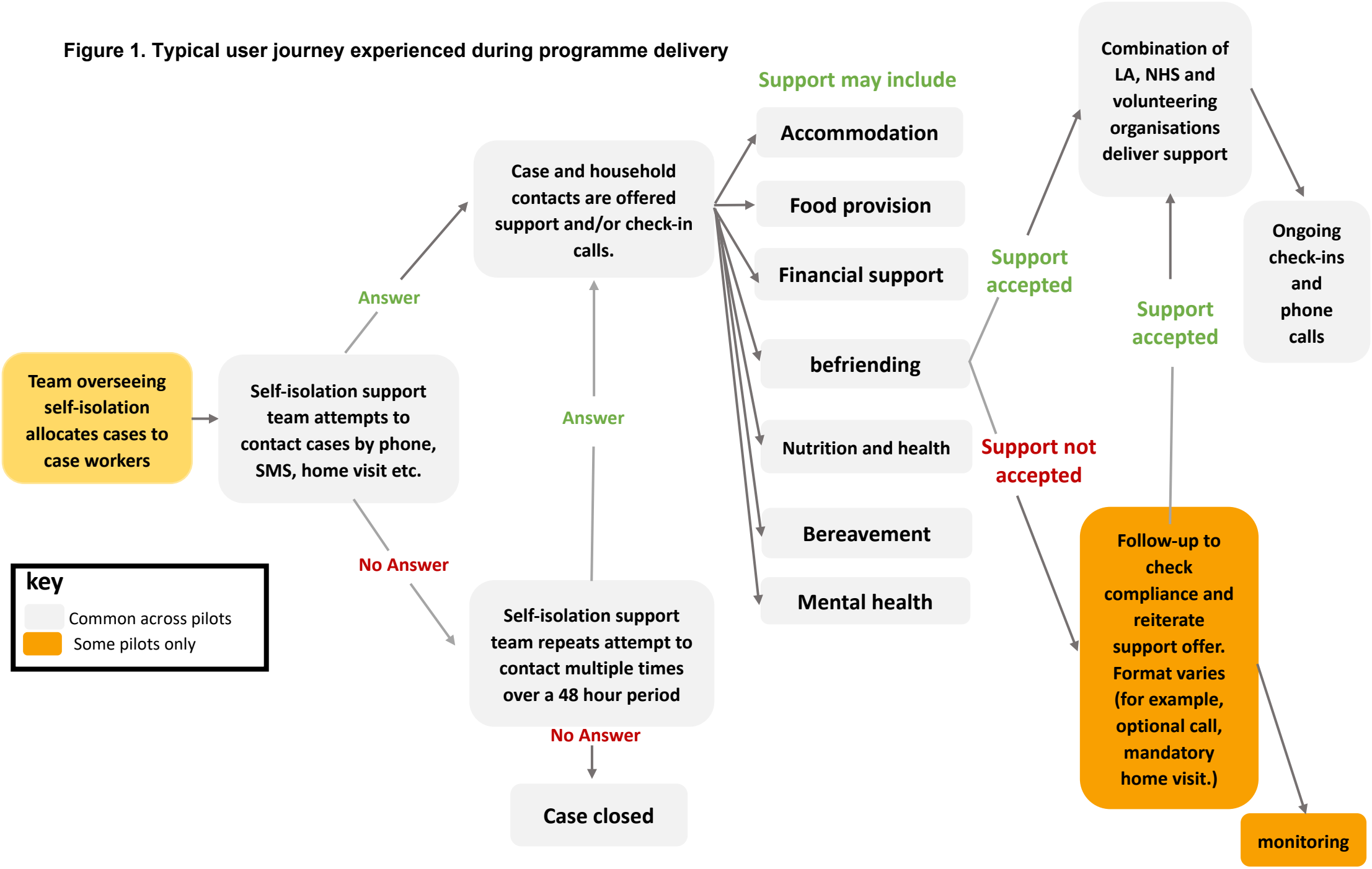
Introduction

The underlying hypothesis for the delivery model was that offering enhanced support (for example, providing assistance with food delivery or befriending) would tackle key barriers preventing people from self-isolating under the standard Test, Trace and Isolate model (TTI).

Responsibility for developing these local interventions and delivering the pilots resided with each local authority, usually undertaken by the teams responsible for delivering the standard pre-pilot local TTI programme.

Whilst there were local variations in the pilot design, certain features of the user journey were common across the local authorities (see Figure 1, below). In the next section, the overarching delivery model for local authorities is outlined, with examples of the ways the pilots varied among the local authorities. A more complete summary of key aspects of – and minor variations in – the delivery model for each local authority is provided in Annexe A.

Figure 1. Typical user journey experienced during programme delivery



Narrative of the overarching delivery model across the 10 local authorities

The team overseeing self-isolation support assigns cases and/or contacts in the same household to case workers. The composition of these self-isolation support teams varies across local authorities: in Bolton, case workers are drawn from community engagement teams, whilst in Manchester a specific 'Support to Self-Isolate Team' works closely with partners delivering humanitarian assistance in Manchester and local nurses. Caseworkers then make contact with cases, sometimes after a process of triaging to identify those most in need of support, or prioritisation based on a request for support at the national level. Contact was typically made through SMS (for example, Tameside), telephone (for example, Oldham), email, home visits (for example, Wigan) or a combination of these means.

To enable cases and contacts to self-isolate, they are then offered a range of support (see Table 2), which differs across the local authorities, allowing the pilots to work flexibly to meet local needs. For those cases or household contacts requesting support, the Self-Isolation Support Team either directly delivers the specific assistance (for example, Wigan) or works with local delivery partners (for example, Oldham).

Table 2. Services offered by each local authority as part of the enhanced self-isolation support pilot

Local authority	Alternative accommodation	Food provision	Medicine delivery	Signposting to financial support	Provision of financial support	Signposting to support services	Clinical support	Befriending	Nutrition or physical wellbeing	Additional support*	Welfare calls	Home visits
Bolton	R		R	R	R							R
Bury		R		R			R	R				R
Manchester		R	R	R			R					
Oldham		R	R	R	R					R	R	
Rochdale		R		R		R				R		
Salford		R	R	R	R					R		
Stockport		R	R				R			R		
Tameside		R	R	R	R	R						
Trafford		R		R		R						
Wigan		R				R						R

* Additional support included dog walking, provision for free school meals, rent subsidies, heating and library activities.

Evaluation aims and research questions

Based on the simplified logic model (see Figure 2 below), the GM Self-Isolation Pathfinder was primarily intended to increase compliance with self-isolation.

A direct pathway to impact is expected for resident cases and contacts within the GM target areas that are reached by local support teams and offered support plans. This is based on the hypothesis that this enhanced support might directly tackle barriers to self-isolation (which standard self-isolation support models fail to address) and thereby increase adherence to self-isolation. For example, the Pathfinder pilots offer support to address 4 out of 5 of the primary reasons for non-compliance with self-isolation identified in the 'Safely Managing COVID-19: Greater Manchester Population Survey': the need for assistance with food shopping or essential goods, mental health, loneliness and non-work responsibilities (2). However, with financial support to self-isolate set at a national level, it is important to note that the Pathfinder has less potential to address the top reason for non-compliance with self-isolation in Greater Manchester: work or financial reasons.

An indirect pathway to impact is expected for anyone that is not directly reached by the local support team but who hears of the Pathfinder activities through family, friends, and the wider community. This indirect pathway to increased engagement with T&T is based on the hypothesis that awareness of the availability of enhanced support to self-isolate might mitigate some of the concerns and challenges individuals have about self-isolating. For example, support for caring for children or elderly relatives, accommodation and food-shopping (10). In this way, the Pathfinder may also increase residents' willingness to get tested for COVID-19, as well as their engagement with Test and Trace (T&T), should they become cases or contacts.

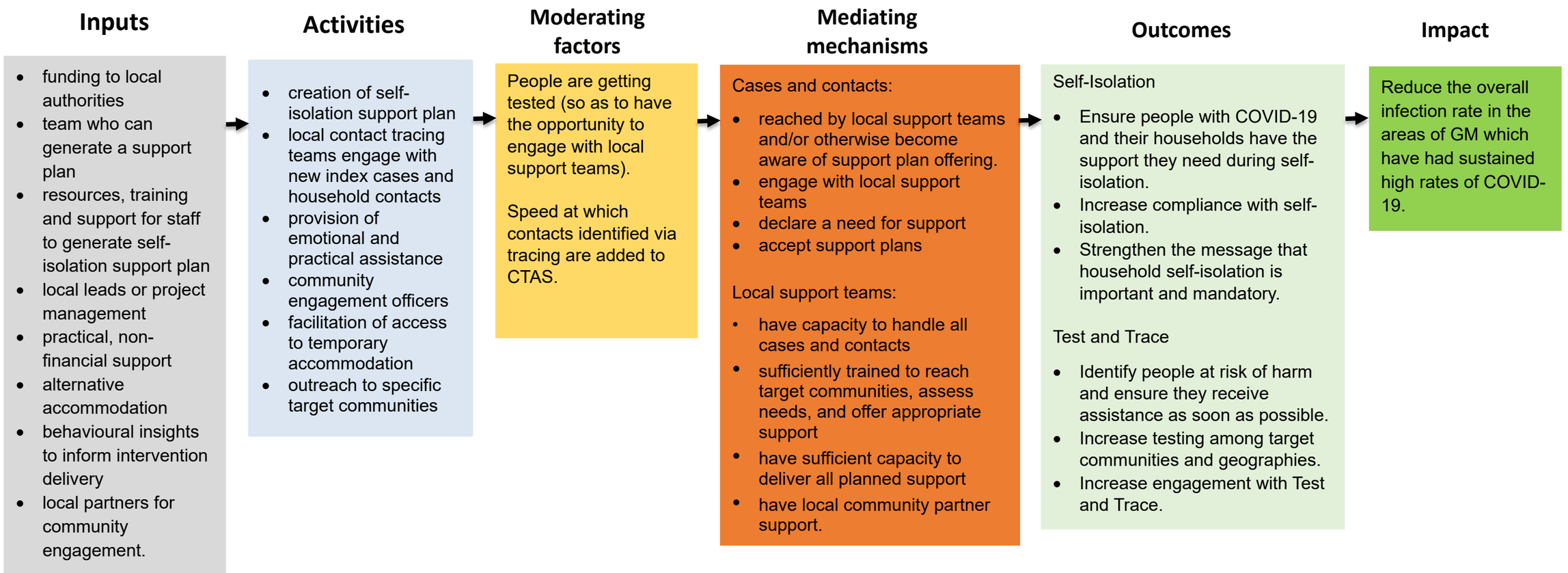
Based on this model, this evaluation seeks to understand the causal impact of the GM Self-Isolation Pathfinder across the 10 local authorities on compliance with self-isolation, as well as COVID-19 testing, and engagement with T&T. As it is infeasible to evaluate outcomes specifically for individuals that are reached by the Pathfinder, the evaluation is focused on evaluating the availability of support in GM, rather than the receipt of support (that is, an intention-to-treat approach).[7]

Consequently, there are 3 research questions for the evaluation:

1.4 Has the availability of support led to an increase in compliance with self-isolation?
1.5 Has the availability of support led to an increase in testing?
1.3 Has the availability of support led to an increase in engagement with Test and Trace (T&T)?

[7] The evaluation relies on data from the national T&T programme, whereas the support is offered via local authorities; therefore, the analysis is not able to identify those cases that have been offered and took up support.

Figure 2. Simplified logic model for the GM Pathfinder programme



Outcome measures

We measure the performance of the programme through 6 outcome measures, which are summarised below in Table 3. Outcomes broadly fall into 3 categories: (i) compliance with self-isolation, (ii) testing rates, and (iii) engagement with the Test and Trace process. These categories and outcomes were chosen in collaboration with DHSC, and aim to strike a balance between outcomes that we might expect to be directly affected by the intervention and outcomes which are able to be computed using the available administrative data.

For each category, we use 2 outcome measures to try and capture some of the multi-dimensional nature of compliance, testing and engagement. However, given that the use of multiple outcome measures raises the possibility of a false positive result, we select one outcome as the primary outcome, the main measure of the programme's effects. This is the proportion of cases with successful call outcomes on all of the day 4, 7 and 10 calls. This was chosen as the primary outcome measure because it is the main measure of compliance with self-isolation that the national T&T system relies on.

Data for outcome measures comes from the DHSC Test and Trace Dashboard, which includes individual-level data on individuals in touch with the Contact Tracing and Advice Service (CTAS), PCR and LFD tests and outcomes, and COVID-19 vaccinations.

Table 3. Summary of outcome measures for each research question.

Research question	Outcome measure	Definition	Limitations
Has the availability of support led to an increase in compliance with self-isolation?	Primary: Weekly % of cases per local authority with successful T&T isolation follow-up call outcome (day 4, 7, 10 calls all successful)	(Number of cases with day 4, 7, 10 isolation follow-up calls successful) / (number of cases reached and called)[8]	Only measures outcomes for a subset of cases that are reached and selected for follow-up calls by T&T
	Exploratory: Weekly % of newly isolating cases per local authority without any non-household contacts	(Number of newly isolating cases without any non-household contacts) / (number of newly isolating cases)[9]	Measures compliance during self-isolation period which is likely before support is offered in GM; only measures outcomes for a smaller subset of newly isolating cases
Has the availability of support led to an increase in testing?	Secondary: Weekly PCR testing rate per local authority	(Number of PCR tests per week) / (local authority population)	May be influenced by presence of surge testing in treatment and/or comparison areas
	Exploratory: Weekly LFD testing rate per local authority	(Number of LFD tests per week) / (local authority population)	As above

[8] Denominator includes only cases that have been called, as not all cases are selected to be called by T&T. Given that calls can occur beyond each week, the reference week is when the person enters CTAS (with call outcomes assigned retrospectively to this week).

[9] This measure is only available for a subset of cases that were previously identified as contacts and have since tested positive for COVID-19 to become cases ("newly isolating cases"). Once they become cases, their contacts are usually self-reported during initial T&T calls or online, and reflect the time period between day 2 and the end of their previous required isolation period, calculated from their exposure date as a contact.

Research question	Outcome measure	Definition	Limitations
Has the availability of support led to an increase in engagement with T&T?	Exploratory: Weekly average number of reported contacts shared per local authority	Includes household, non-household, complex contacts, and potentially others that are not in these 3 categories.	Cannot distinguish having had more contacts from willingness to engage with T&T and declare contacts; interpretation depends on social restrictions remaining stable across areas and within evaluation time period
	Exploratory: Weekly % of cases sharing at least one contact per local authority	(Number of cases reporting at least one contact) / (number of cases reached and called)[10]	As above

Note: All data on self-isolation compliance, testing, and contact tracing is sourced from the DHSC Dashboard; local authority population data is sourced from ONS mid-2019 population estimates (see Annexe B: Technical details).

Has the availability of support led to an increase in compliance with self-isolation?

Primary outcome: weekly percentage of cases in the local authority with successful T&T self-isolation follow-up call outcomes

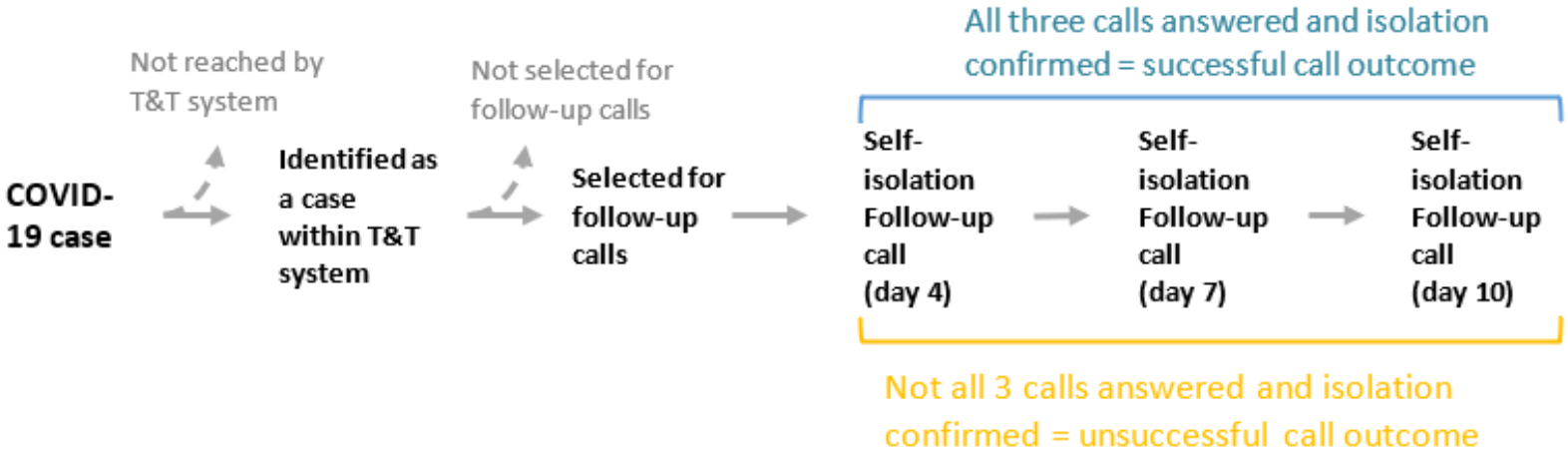
Follow-up calls are conducted by the national T&T system on days 4, 7, and 10 and require the case to answer and confirm isolation. A successful outcome here is defined by DHSC as having all 3 calls answered and confirmed (see Figure 3 below).[11]

Importantly, any cases that are identified by local T&T teams in GM as being in breach of self-isolation are not passed back to the national T&T system. This means that the programme in GM should not affect this metric by merely detecting more cases that are in breach. An

[10] Denominator includes only cases that have been called, as not all cases are selected to be called by T&T.
[11] Of the proportion of COVID-19 cases in a given area, not all cases are reached by the national T&T system and, of those reached, not all are selected for follow-up calls. Therefore, the denominator for this outcome consists of all cases that are reached and selected for follow-up calls.

advantage of this metric is that it focuses on behaviour on days 4, 7, and 10 of isolation, which are likely to be after the offer of programme support is made, and therefore has a meaningful opportunity to be directly impacted by the programme.

Figure 3. Diagram illustrating the definition of the primary outcome



Note: not all COVID-19 cases are reached by the national T&T system, and of those that are reached, not all are selected for follow-up calls. Therefore, the primary outcome only includes the subset of cases that are reached and selected for follow-up calls.

Text version of Figure 3

COVID-19 case was either not reached by T&T system or identified as a case within the T&T system.

If the latter, the case either was or was not selected for follow-up calls.

If selected, the case received a self-isolaton follow-up call on day 4, a self-isolation call on day 7, a self-isolation follow-up call on day 10.

Either all 3 calls were answered and isolation confirmed = a successful outcome; or not all 3 calls were answered and isolation was not confirmed = unsuccessful call outcome.

Exploratory outcome: weekly percentage of newly isolating cases without any non-household contacts in the local authority

This measure captures cases that were previously identified as contacts and have since tested positive for COVID-19 to become cases. Given that these individuals should have been self-isolating since being identified as contacts (or since they noticed symptoms, whichever is first), any non-household contacts that they report to the national T&T system (via the initial phone call or online) at the point of being identified as a case reflect non-compliance with self-isolation (see Figure 4 below).

Although this is a direct measure of compliance with self-isolation, a key disadvantage is that the self-isolation period in which compliance is measured is likely to start before support is offered in GM; this means that the programme is more likely to have only an indirect impact on this measure.[12] Additionally, it only captures compliance for a specific subset of cases. This outcome is therefore considered exploratory.

Figure 4. Diagram illustrating the definition of the exploratory outcome percentage of new isolating cases without any non-household contacts

Note: Non-household contacts are reported with respect to the initial self-isolation period, between day 2 of their becoming a contact and the end of this initial self-isolation period.

Text version of Figure 4

There are 3 stages:

1. Case identified as a contact within the T&T system (required to self-isolate).
2. Case identitified as a case within the T&T system (any contacts prior to self-isolation period reported via T&T call or online).
3. New self-isolation period begins.

Has the availability of support led to an increase in testing?

Secondary outcome: weekly PCR testing rate in the local authority

This measure reflects a direct aim of the Pathfinder programme's logic model (see Figure 2 above). One potential limitation of this measure is that it may be influenced by surge testing in the treated or control units that is unrelated to the Pathfinder programme. To mitigate this, the

[12] Non-household contacts are counted for these cases between day 2 and the end of their previous required isolation period, calculated from their exposure date as a contact. However, support in GM is offered once individuals are identified as cases.

evaluation has factored in the extent of surge testing in each local authority in the analysis procedure (see the Evaluation design section for details).

Exploratory outcome: weekly LFD testing rate in the local authority

This captures another form of testing that is governed by separate government policies and communications, and which may be impacted by the Pathfinder.

Has the availability of support led to an increase in engagement with Test and Trace?

Exploratory outcome: weekly average number of reported contacts shared per local authority

Given that engagement is generally challenging to quantify, this measure provides a potentially valuable insight into engagement. However, one drawback is that it cannot distinguish true behaviour (having had more contacts) from reported behaviour (willingness to engage with T&T and declare contacts). Its interpretation also depends on social restrictions remaining stable across the intervention and comparison areas, and throughout the time period of evaluation.

Exploratory outcome: weekly percentage of cases sharing at least one contact per local authority

This is a binary version of the above metric and reflects a measure that the national T&T system relies on to assess engagement.

Evaluation design

Synthetic control method

The evaluation design relies on the synthetic control method, which combines elements from difference-in-differences (DID) and matching designs. The DID design estimates intervention effects by comparing the change in outcomes in the treatment group before and after the intervention, and comparing this to the change in outcomes in a comparison group (11). The DID design rests on the counterfactual assumption that – in the absence of intervention – outcome trends would be parallel between treatment and comparison groups in the post-intervention period. To strengthen this assumption, the DID design requires manually selecting a comparison group that demonstrates parallel outcome trends with the treatment group in the pre-intervention period.

Instead of manual selection, the synthetic control method uses a formal, data-driven approach to construct a comparison. Specifically, similar to matching, it constructs a weighted combination of potential control areas (called the 'donor pool'), with the weights chosen to maximise similarity between the treated group and comparison on outcome values and covariates ('matching' variables) during the pre-intervention period (see the Annexe B: Technical details for details on how weights were estimated). This weighted comparison is the 'synthetic comparison'. The treatment effect is estimated by the difference in outcomes between the treatment and the synthetic comparison groups in the post-intervention period.

The evaluation uses a synthetic control approach recently developed for contexts with multiple treated units (in this case, local authorities within GM), and implemented in the R package microsynth (12).

The geographical unit of analysis is the local authority (excluding any wards in which the programme was not implemented; see Annexe B: Technical details)[13], and the time unit of analysis is the week. The estimated treatment effect therefore reflects the effect of treatment for an average treated local authority over the post-intervention weeks.

It may be the case that, in local authorities where not all wards were treated, the treated wards do not represent the local authority as a whole. This does in principle affect our assessment of how these results might extend outside Greater Manchester, however, in practice we do not believe that this is a major problem, for 2 reasons:

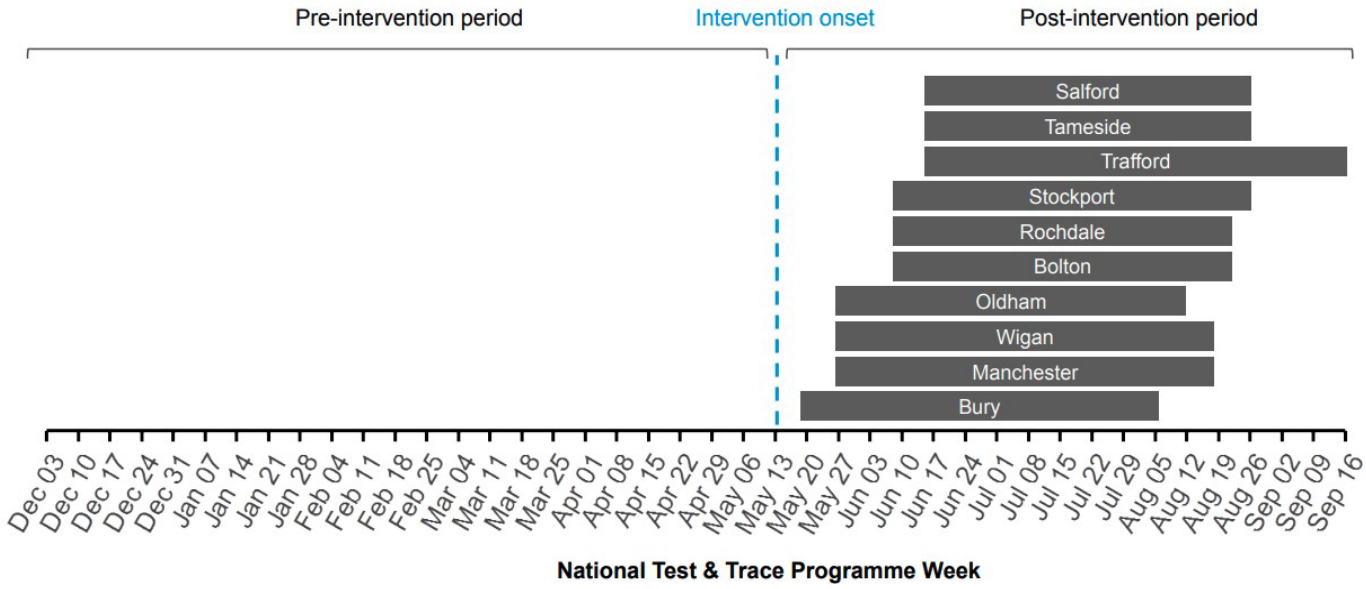
[13] This means that outcome data for local authorities in GM which selected target wards reflects only those target wards and not the entire local authority; where possible, the same approach was taken for other matching covariates (see Matching variables section for details). Analysing data at the more granular ward level was not feasible due to the sparseness of COVID-19 cases in some wards, which meant that there are time periods for which some outcomes are not defined (for example,the primary outcome for compliance with self-isolation). Given the importance of time series granularity for the synthetic control method (for example, see Abadie 2021) (11), we aggregated data across geography, rather than across time.

- half of treated local authorities treated all of their wards
- the synthetic comparison group is constructed to be comparable to the treated group as defined above; as such, untreated wards in otherwise treated local authorities do not affect the choice of comparison

Study period

The pre-intervention period consists of 24 weeks, defined from 3 December 2020 until 13 May 2021 inclusive (corresponding to T&T programme weeks 28 to 51)[14]. The post-intervention period consists of 18 weeks, defined from 20 May 2021 until 16 September 2021 inclusive (weeks 52 to 69), and covers all of the official pilot periods. Figure 5 illustrates the study period and the duration of each pilot.

Figure 5. Study period and pilot launch dates for each local authority



Treated areas

Table 4 summarises the final target areas for analysis in each local authority.[15]

[14] Data does exist prior to this period; however, in the earliest phase of the pandemic it becomes harder to construct a good quality match for the treated local authoritiess in GM because the outcomes have very high variance, depending on how quickly different local authorities established their Test and Trace capability. This is largely unrelated to the intervention under consideration so we do not include the early period in our analysis.

[15] Several local authorities in the Pathfinder programme re-defined their target areas for intervention after programme launch, either due to high case rates or insufficient uptake. See Annexe B: Technical details for details on how the final target areas for analysis were determined.

Table 4. Summary of target areas for analysis for each local authority in Greater Manchester

Local authority	Target areas for analysis
Bolton	Entire local authority
Bury	Target wards (East, Moorside, Radcliffe West, Besses, Sedgley)
Manchester	Target wards (Longsight, Levenshulme, Cheetham, Crumpsall, Rusholme, Ardwick, Moss Side, Whalley Range)
Oldham	Entire local authority
Rochdale	Target wards (Milkstone and Deeplish, Central Rochdale, Kingsway, West Middleton)
Salford	Target ward (Little Hulton)
Stockport	Entire local authority
Tameside	Entire local authority
Trafford	Target wards (Bucklow St Martins, Clifford)
Wigan	Entire local authority

Defining the donor pool

The donor pool is defined as the set of untreated comparison units that make up the synthetic comparison group via weighted averaging. To be included in the donor pool for the current evaluation, local authorities must have met the following inclusion criteria: inside England, outside of Greater Manchester, not implementing a similar pilot programme (as reported by DHSC)[16], and have a Rural-Urban classification of Major Urban Conurbation (that is, matching that of local authorities in GM). Applying these criteria left 47 out of 317 local authorities (15%) for the donor pool, using 2019 local authority boundaries. See Annexe B: Technical details for further details on the donor pool.

Matching variables

To enable consistent comparison and interpretability of intervention impact across all outcome measures, the same set of matching variables and estimated weights were used for all analyses.[17] Matching on multiple variables entails a level of compromise between them; in practice, a high-quality match was observed for all outcomes except for one exploratory outcome (LFD testing rate; see the Results section). We therefore judged this to be an overall superior approach for comparability purposes. Matching variables were selected on the basis of

[16] Based on data provided by DHSC, 52 local authorities were excluded for this reason.
[17] This is a deviation from the Evaluation Protocol which had indicated that a separate synthetic comparison group (that is, set of weights) would be constructed for each outcome, and that for each analysis, each relevant outcome would be substituted as a pre-intervention matching variable. However, a single consistent synthetic comparison group with all pre-intervention outcomes was constructed instead to enable comparison and interpretability of results across analyses.

being predictive of variation in the outcome; this includes the outcomes themselves during the pre-intervention period, as well as other relevant covariates. Table 5 lists the matching variables and the brief rationale for inclusion (see also Annexe B: Technical details).

Table 5. Matching variables in the analysis and the rationale for inclusion

Matching variable	Rationale for inclusion
Proportion of successful T&T isolation follow-up call outcomes at key pre-intervention weeks (primary outcome lag)	This captures variation in the outcome. Values at key pre-intervention weeks are used instead of the full pre-intervention time series to avoid precluding the influence of other covariates that may be relevant (see Technical Annexe B for details) (13).
Proportion of new self-isolating cases with no non-household contacts at key pre-intervention weeks (exploratory outcome lag)	As above
PCR testing rate at key pre-intervention weeks (secondary outcome lag)	As above
LFD testing rate at key pre-intervention weeks (exploratory outcome lag)	As above
Average number of reported contacts at key pre-intervention weeks (exploratory outcome lag)	As above
Proportion of cases sharing at least one contact at key pre-intervention weeks (exploratory outcome lag)	As above
COVID-19 case rate at key pre-intervention weeks	This captures any variation in the number of cases that comprise the outcomes and in the behaviours related to self-isolation, testing, and engagement with T&T.
Total number of weeks in pre-intervention period with any surge testing, with postcode surge testing, and with site surge testing	This captures any variation in the testing outcomes and in the number of cases comprising the other outcomes.

Matching variable	Rationale for inclusion
Cumulative full vaccination rate at the end of the pre-intervention period	This captures any variation in behaviours related to self-isolation, testing, and engagement with T&T, including those arising from self-isolation policy changes related to vaccination[18]
Proportion of each major ethnic subgroup in local authority (Asian, Black, Mixed, Other, White)	This captures any variation in population composition that may be linked to COVID-19 case rates and behaviours related to self-isolation, testing, and engagement with T&T
Proportion of each age group in local authority (age 0 to 9, 10 to 19, 20 to 29, 30 to 39, 40 to 49, 50 to 59, 60 and over)	As above
IMD 2019 average score for local authority	This captures any deprivation-related variation across local authorities that may be linked to COVID-19 case rates and behaviours related to self-isolation, testing, and engagement with T&T

Note: For local authorities in GM which selected specific target wards for the Pathfinder, all data on pre-intervention outcomes, COVID-19 case rate, and cumulative vaccination rate is specific to the set of target wards; data on the age and ethnic subgroup breakdowns and the IMD score for each local authority instead reflect the entire local authority, due to data availability limitations.

Treatment effect estimation and statistical inference

To obtain an estimate of the treatment effect for each analysis, the summed outcome value across all treated units (local authorities in GM) was calculated, subtracted from the corresponding value for the synthetic comparison group, and summed across post-intervention time periods. To conduct statistical inference on this treatment effect estimate, non-parametric permutation testing[19] was used. See Annexe B: Technical details for further details.

[18] For example, Self-isolation removed for double-jabbed close contacts from 16 August.

[19] This is a procedure to estimate the statistical significance of the results which does not rely on making assumptions about how the data might be distributed, and as such is considered more reliable in the context of a synthetic comparison group.

To facilitate interpretation, we also provide overall post-intervention outcomes in GM and the synthetic comparison group. For the outcome in GM, this is calculated by taking the weekly outcome, for each local authority, then averaging these weekly outcomes across local authorities in GM and across post-intervention weeks. For the outcome in the synthetic comparison group, microsynth's output of the relative 'percentage difference' in outcomes between GM and the synthetic comparison group is then applied to the former average to obtain the corresponding figure for the synthetic comparison group.

Data sources

Table 6 summarises the data sources used in the evaluation (see Annexe B: Technical details for details).

Table 6. Data sources used in the evaluation

Source	Purpose
DHSC Test and Trace Dashboard	To construct all the outcome measures and additional covariates for the analysis
Index of Multiple Deprivation 2019 data	To obtain an estimate of local authority-level deprivation as a covariate for matching
ONS ward-level population estimates (mid-2019 data)	To construct local authority -level population estimates (total count and proportions by age group), excluding non-treated wards, for covariates for matching
ONS regional ethnic diversity data (Census 2011 data using mid-2018 local authority boundaries)	To provide the ethnic population breakdown for each local authority, as a proportion of the total local authority population, as a covariate for matching
Rural Urban Classification of local authorities (Census 2011 data using mid-2019 local authority boundaries)	To restrict the donor pool to local authorities with a similar classification as those in GM

Results

We summarise the results in Table 7 below, before presenting more detail on each finding in subsequent sections.

Table 7. Summary of results

Research question	Outcome measure	Results
Has the availability of support led to an increase in compliance with self-isolation?	Primary: Weekly % of cases per local authority with successful T&T isolation follow-up call outcome	No statistically significant difference, aside from a data anomaly in the weeks following implementation in some of the untreated comparison local authorities.
	Exploratory: Weekly % of newly isolating cases per local authority without any non-household contacts	No statistically significant difference
Has the availability of support led to an increase in testing?	Secondary: Weekly PCR testing rate per local authority	No statistically significant difference
	Exploratory: Weekly LFD testing rate per local authority	Inconclusive due to a poor pre-intervention match between GM and the synthetic comparison group: statistically significant decrease in GM compared to synthetic comparison group, but this begins prior to intervention onset
Has the availability of support led to an increase in engagement with T&T?	Exploratory: Weekly average number of reported contacts shared per local authority	No statistically significant difference
	Exploratory: Weekly % of cases sharing at least one contact per local authority	No statistically significant difference

Has the availability of support led to an increase in compliance with self-isolation?

Primary outcome: Weekly percentage of cases per local authority with successful T&T self-isolation follow-up call outcomes

We do not find evidence that the intervention increased self-isolation compliance. In GM, the average successful T&T follow-up call outcome rate in the post-intervention period is 72%.[20] By comparison, the average outcome in the synthetic comparison group is 68%. In this case, GM has statistically significantly higher post-treatment outcomes than the comparison group ($p = 0.0036$). However, this is largely down to a sharp dip in recorded successful T&T call outcomes around July 1 to 29 (7 to 11 weeks), after intervention onset which is present in non-treated local authorities[21] but not in GM (Figure 6, top panel).[22]

If the non-treated local authorities which record a large post-intervention dip are removed from the donor pool, then the apparent effect of the intervention becomes negative[23], and is significant ($p < 0.0001$). If we instead remove weeks 52 to 55 from the data (those weeks where the intervention was implemented in some wards and local authorities but not others) then this changes to a null result ($p = 0.14$). The result is not materially changed by leaving any single local authority out of the donor pool.

Removing the weeks in which the dip happens (covering the period July 8 t o28) rather than the local authorities also produces a null result ($p = 0.37$).

We confirmed that GM and the synthetic comparison group were balanced on matching covariates (see Annexe B: Technical details for details, and for the allocation of weights across the local authorities in the donor pool).

The correct interpretation of these results depends on the reason for the dip. Unfortunately, the cause of the dip is not obvious from the data alone. This dip is not observed in the other outcomes which may suggest that there is some inaccuracy in the data. Secondly, the dip is 'all-or-nothing' – some local authorities record dips of 40 to 50 percentage points and some local

[20] This is calculated by taking the weekly percentage, for each local authority, of T&T call outcomes which are successful and then averaging these weekly percentages across local authorities in GM and across post-intervention weeks. The relative '% difference' in outcomes between GM and the synthetic comparison group from the microsynth output is then applied to this average to give the corresponding figure for the synthetic comparison group.

[21] Donor pool local authorities with the dip are: Hertsmere, Watford, Broxbourne, Gravesham, Dartford, Sunderland, Elmbridge, Solihull, Wolverhampton, Spelthorne, Sandwell, Newcastle Upon Tyne, Woking, Walsall, Birmingham, Dudley, Runnymede, Epsom and Ewell, Gateshead, Three Rivers and South Tyneside.

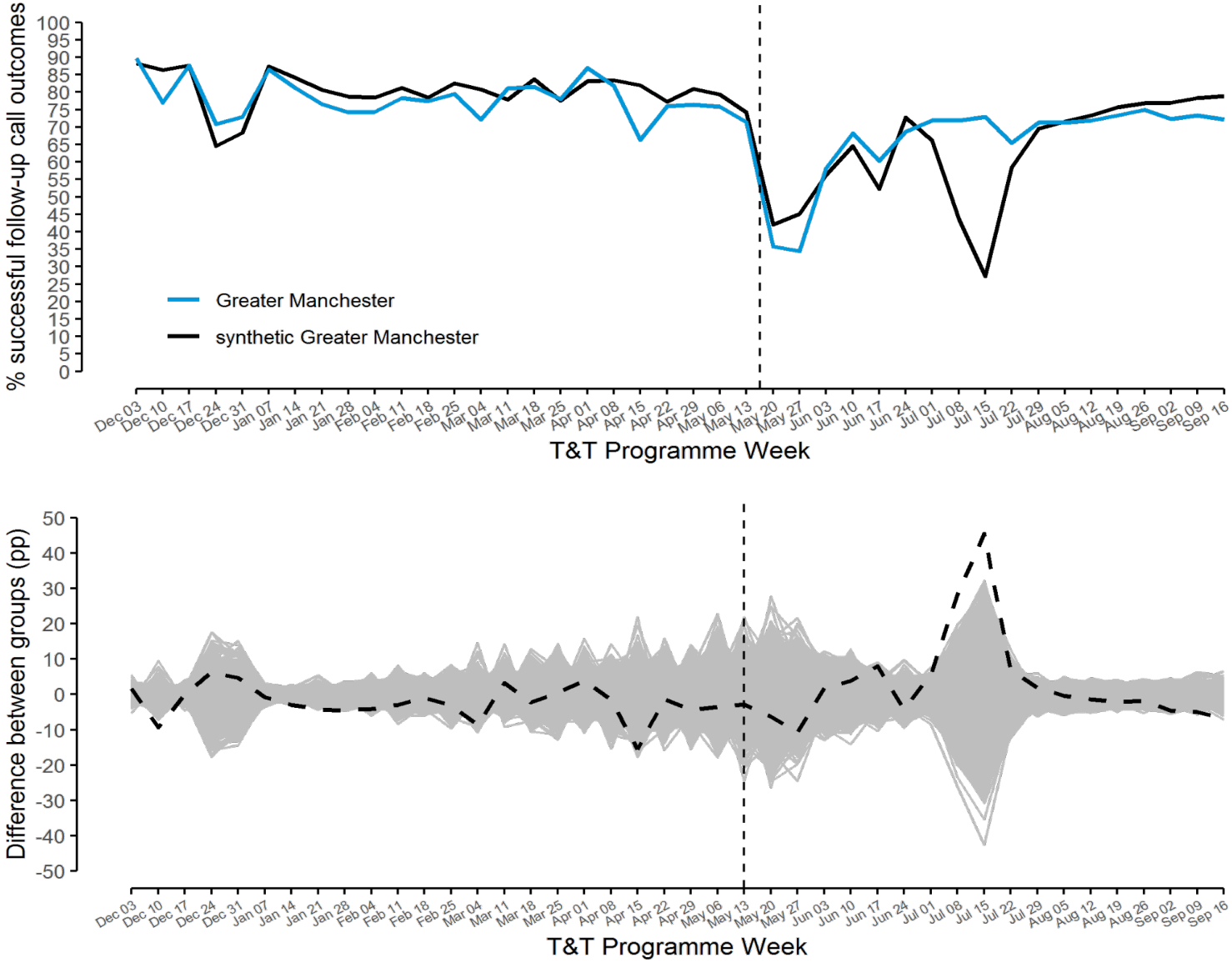
[22] Note that the estimated treatment effect is the average difference between GM and the synthetic comparison group across the post-intervention weeks, whereas the top panel in Figure 6 shows the weekly outcome to visualise the time trend for GM and synthetic comparison group; the bottom panel shows the weekly difference between groups, as well as the analogous difference for each placebo permutation, to visualise the associated uncertainty.

[23] This analysis point is only made to emphasise that the correct interpretation of the result depends on the reason for the dip; it should not be used as an estimate of the (in)effectiveness of the programme.

authorities experience no dip at all, but no local authority in the donor pool records anything between these 2 extremes.

Initial investigation by UKHSA indicates that the dip was related to logistical issues rather than indicative of a change in isolation behaviour. This interpretation is reinforced by other evidence relating to compliance with self-isolation, including random probability surveys on compliance behaviour undertaken by ONS. As such, considering the set of analysis results, we conclude that we do not have any evidence that the Pathfinder programme affected self-isolation compliance, but it is possible that it protected the treated local authorities from these logistical issues which affected data quality.

Figure 6. Primary outcome analysis results: percentage of cases with successful T&T self-isolation follow-up call outcomes



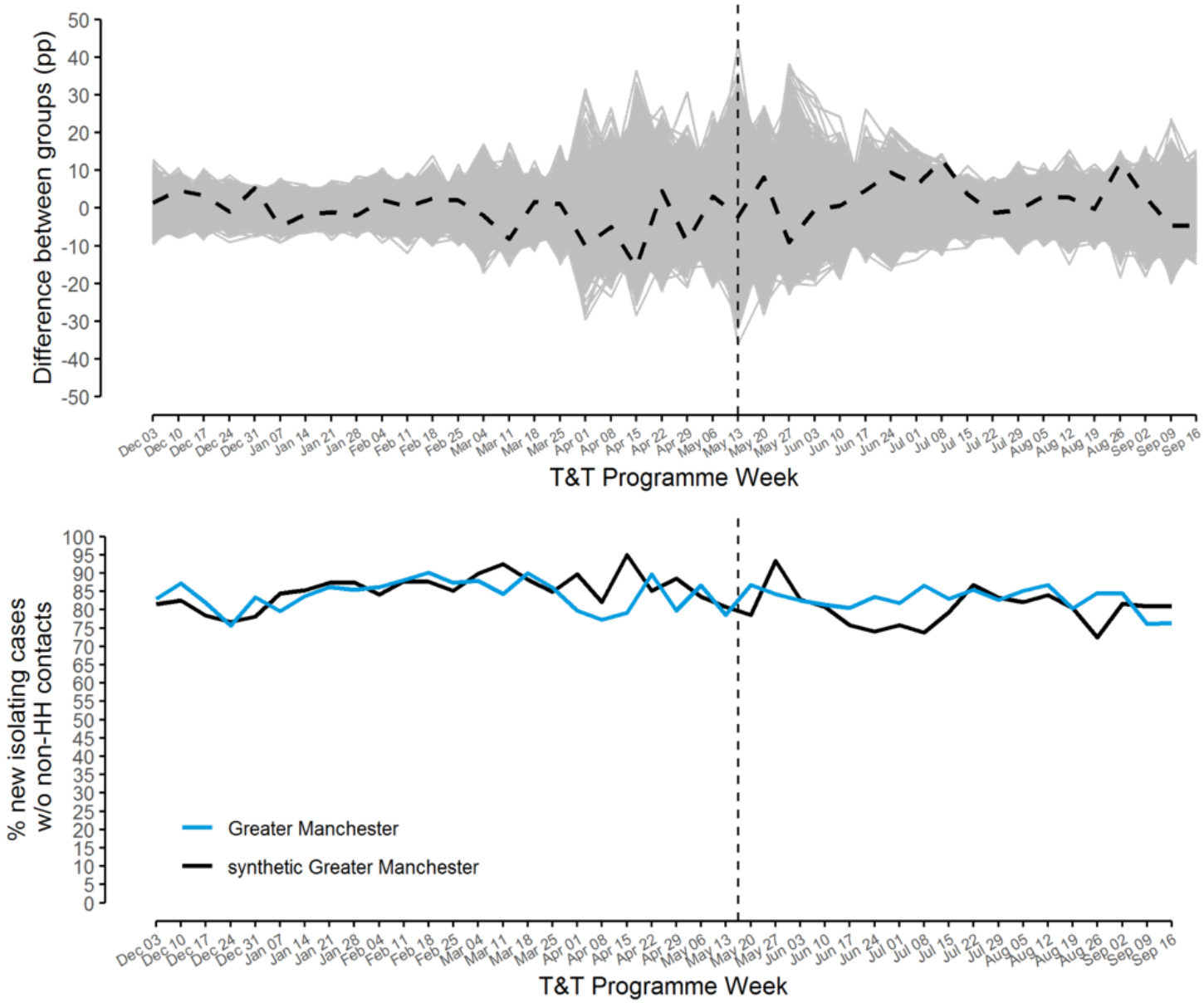
Note: Top panel: Outcome for GM (across all treated local authorities; blue line), and the synthetic comparison group (black line). The outcome is rescaled by the number of treated local authorities. Bottom panel: the difference between GM and synthetic comparison lines (dashed black line), and the analogous difference for each placebo permutation (grey lines). The vertical dashed line indicates the onset of the post-intervention period. The estimated treatment effect is the average difference between groups across the post-intervention weeks, but the weekly outcome is shown here to visualise the time trend for GM and the synthetic comparison group.

Exploratory outcome: Weekly percentage of newly isolating cases per local authority without any non-household contacts

This outcome measure only pertains to cases that were previously identified as contacts and have since tested positive for COVID-19 to become cases (hence 'new' isolating cases).[24] In GM, the average percentage of new isolating cases without any non-household contacts was 84% in the post-intervention period; by comparison, this figure was 81% in the synthetic comparison group. We do not find a significant difference in this outcome between GM and the synthetic comparison group ($p = 0.20$). This result does not change if we assume the intervention began at week 56 when all wards and local authorities had launched their pilots, and remove weeks 52 to 55 from the data ($p = 0.56$). Figure 7 presents the time series for each group and the difference between groups.

[24] Given that this outcome is defined for the subset of cases which are newly self-isolating after previously being contacts, there are some local authorities which have no such cases in certain weeks in the analysis; this 'missing' data was imputed for analysis (see Annexe B: Technical details for details).

Figure 7. Exploratory outcome analysis results: percentage of new isolating cases without any non-household contacts



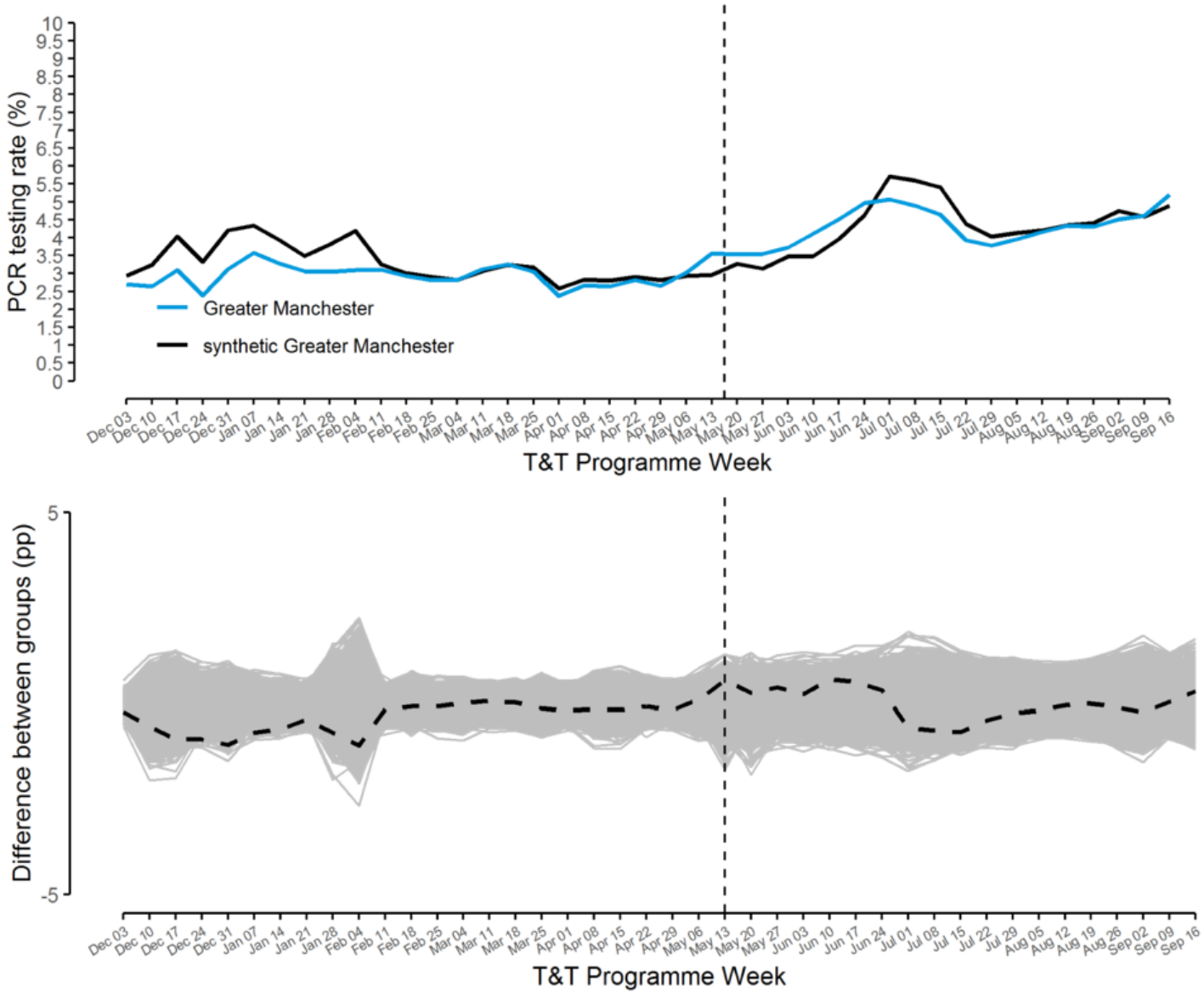
Note: Top panel: Outcome for GM (across all treated local authorities; blue line), and the synthetic comparison group (black line). The outcome is rescaled by the number of treated local authorities. Bottom panel: the difference between GM and synthetic comparison lines (dashed black line), and the analogous difference for each placebo permutation (grey lines). The vertical dashed line indicates the onset of the post-intervention period. The estimated treatment effect is the average difference between groups across the post-intervention weeks, but the weekly outcome is shown here to visualise the time trend for GM and the synthetic comparison group.

Has the availability of support led to an increase in testing?

Secondary outcome: Weekly PCR testing rate per local authority

In GM and the synthetic comparison, the average PCR testing rate in the post-intervention period was the same (3.6%). There is no discernible change in the PCR testing rate between GM and the synthetic comparison group ($p = 0.82$). Whilst local authorities in GM generally had lower PCR testing rates than other local authorities, this was explained both by lower pre-intervention levels of PCR testing and other covariates, rather than by the Pathfinder programme. This result does not change if we assume the intervention began at week 56 when all wards and local authorities had launched their pilots, and remove weeks 52 to 55 from the data ($p = 0.69$). Figure 8 presents the time series for each group and the difference between groups.

Figure 8. Secondary outcome analysis results: PCR testing rate



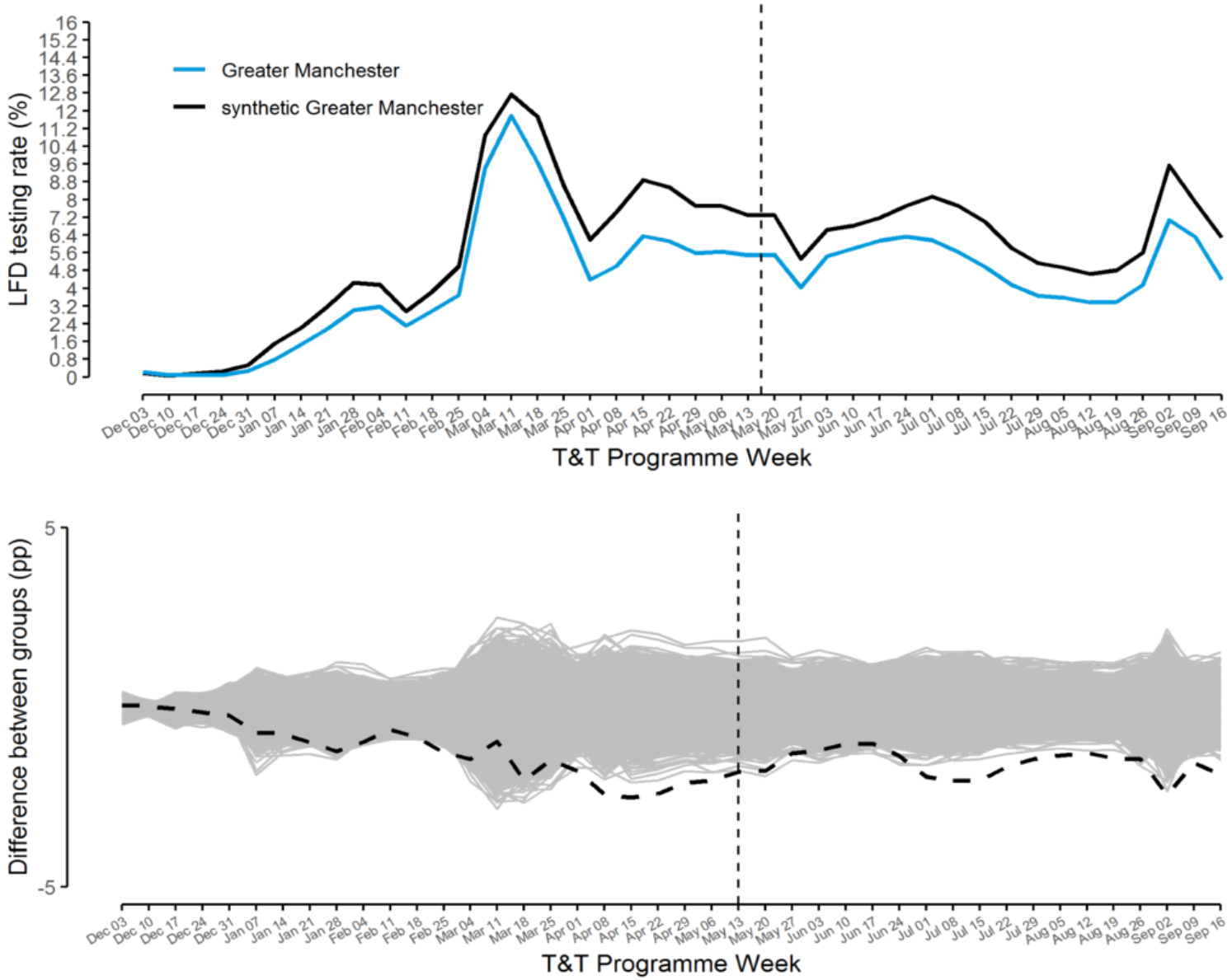
Note: Top panel: Outcome for GM (across all treated local authorities; blue line), and the synthetic comparison group (black line). The outcome is rescaled by the number of treated local authorities. Bottom panel: the difference between GM and synthetic comparison lines (dashed black line), and the analogous difference for each placebo permutation (grey lines). The vertical dashed line indicates the onset of the post-intervention period. The estimated treatment effect is the average difference between groups across the post-intervention weeks, but the weekly outcome is shown here to visualise the time trend for GM and the synthetic comparison group.

Exploratory outcome: Weekly LFD testing rate per local authority

In GM, the average LFD testing rate across the post-intervention period was 4.6%, compared to 5.6% in the synthetic comparison group. The analysis indicates a significant reduction in LFD testing rates; however, the matching procedure did not produce a good-quality match for the local authorities in GM for this outcome. In particular, the matched synthetic comparison group showed a difference in LFD testing rates before the intervention started, so we have only limited confidence in the apparent reduction in LFD testing rates, even though the p-value is less than 0.0001 for both the main analysis and the robustness check in which we assume the intervention began at week 56 rather than week 52.[25] Figure 9 presents the time series for each group and the difference between groups.

[25] This analysis is nevertheless included here as it forms part of the pre-planned set of outcomes, agreed together with DHSC.

Figure 9. Exploratory outcome analysis results: LFD testing rate



Note: Top panel: Outcome for GM (across all treated local authorities; blue line), and the synthetic comparison group (black line). The outcome is rescaled by the number of treated local authorities. Bottom panel: the difference between GM and synthetic comparison lines (dashed black line), and the analogous difference for each placebo permutation (grey lines). The vertical dashed line indicates the onset of the post-intervention period. The estimated treatment effect is the average difference between groups across the post-intervention weeks, but the weekly outcome is shown here to visualise the time trend for GM and the synthetic comparison group.

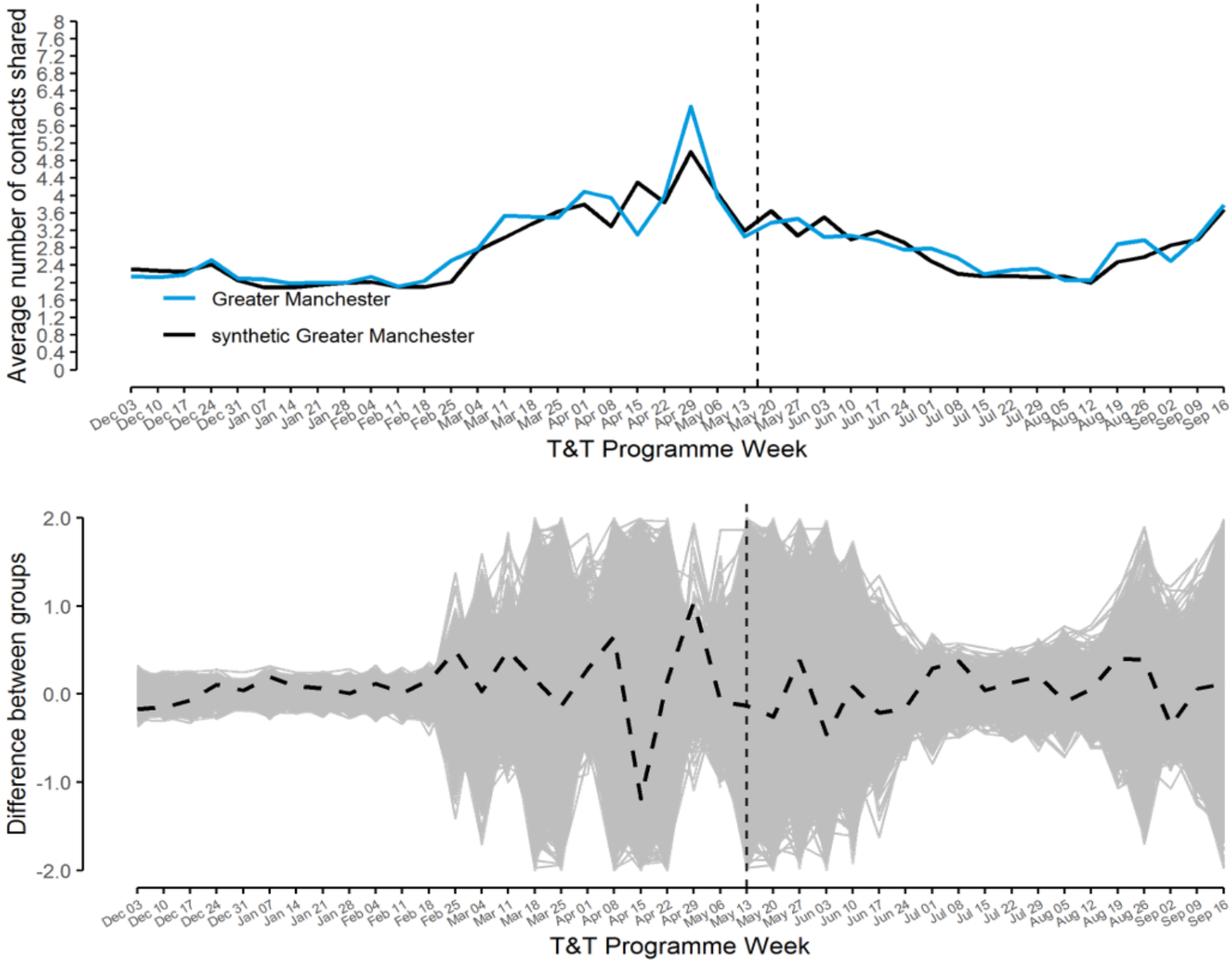
Has the availability of support led to an increase in engagement with Test and Trace?

Exploratory outcome: Weekly average number of contacts shared per local authority

In GM, the average number of contacts shared across the post-intervention period was 2.9, compared to 2.8 in the synthetic comparison group. We do not observe a significant change in this outcome, either on the main analysis ($p = 0.78$) or the robustness check in which we assume the intervention began at week 56 rather than week 52 ($p = 0.54$). Figure 10 presents the time series for each group and the difference between groups.

This outcome is quite variable in certain periods, where there is great variation across local authorities and individuals in the number of contacts reported, sometimes with small numbers of cases. This generally coincides with times with relatively fewer restrictions when the average number of contacts was higher.

Figure 10. Exploratory outcome analysis results: average number of contacts

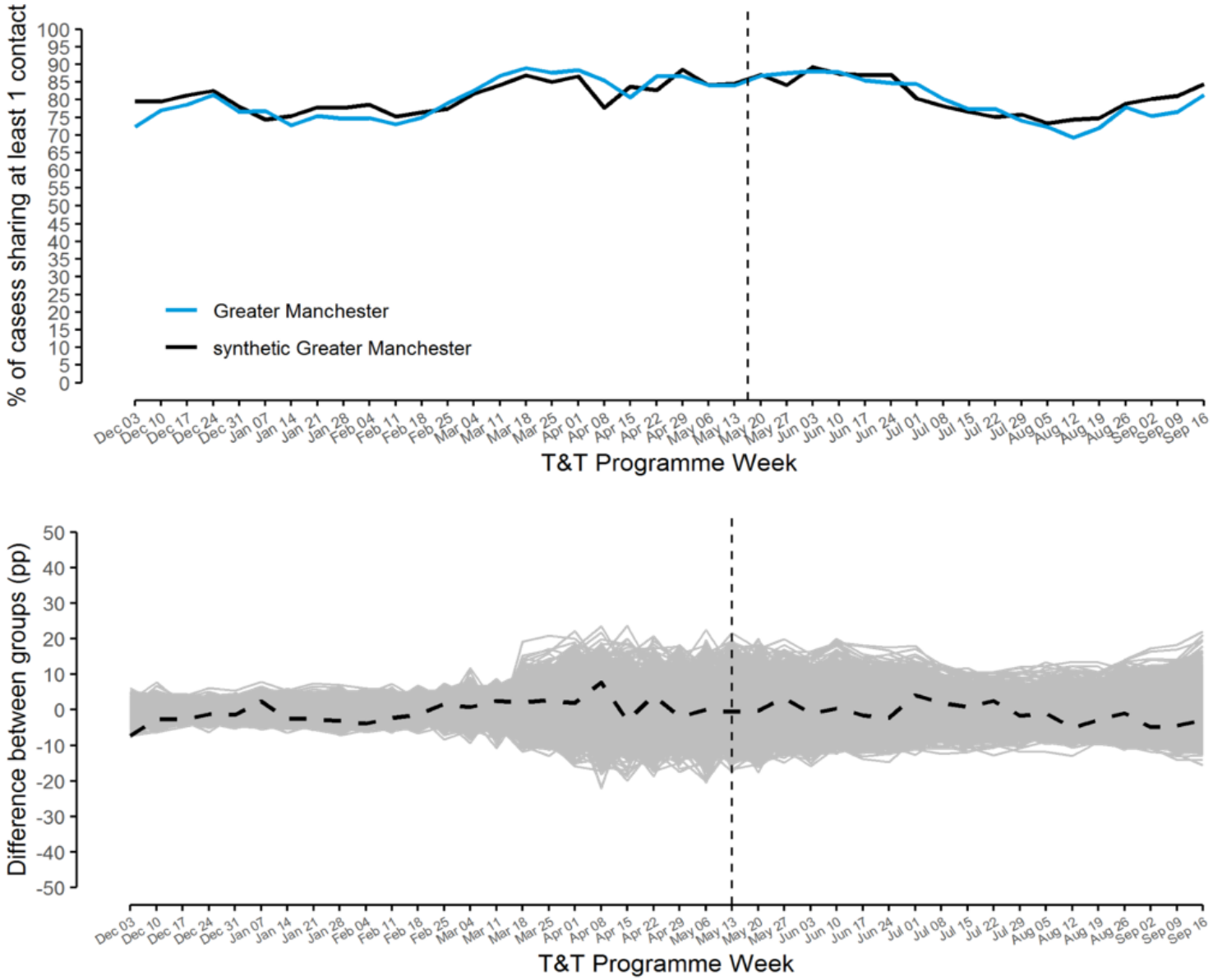


Note: Top panel: Outcome for GM (across all treated local authorities; blue line), and the synthetic comparison group (black line). The outcome is rescaled by the number of treated local authorities. Bottom panel: the difference between GM and synthetic comparison lines (dashed black line), and the analogous difference for each placebo permutation (grey lines). The vertical dashed line indicates the onset of the post-intervention period. The estimated treatment effect is the average difference between groups across the post-intervention weeks, but the weekly outcome is shown here to visualise the time trend for GM and the synthetic comparison group.

Exploratory outcome: Weekly percentage of cases sharing at least one contact per local authority

In GM, the average percentage of cases sharing at least one contact was 80% across the post-intervention weeks; by comparison, this figure was 81% in the synthetic comparison group. As with the other engagement outcome, there is no significant difference between groups ($p = 0.67$). Likewise, there is no significant difference for the robustness check in which we assume the intervention began at week 56 rather than week 52 ($p = 0.37$). Figure 11 presents the time series for each group and the difference between groups.

Figure 11. Exploratory outcome analysis results: percentage of cases sharing at least one contact in the local authority



Note: Top panel: Outcome for GM (across all treated local authorities; blue line), and the synthetic comparison group (black line). The outcome is rescaled by the number of treated local authorities. Bottom panel: the difference between GM and synthetic comparison lines (dashed black line), and the analogous difference for each placebo permutation (grey lines). The vertical dashed line indicates the onset of the post-intervention period. The estimated treatment effect is the average difference between groups across the post-intervention weeks, but the weekly outcome is shown here to visualise the time trend for GM and the synthetic comparison group.

Robustness checks

For each outcome, we re-analysed the data as if the intervention started at week 56 rather than week 52, which corresponds to when all treated local authorities had launched their pilots (that is, we removed post-intervention weeks 52 to 55 from the analysis). For the primary outcome analysis, this changed the effect from statistically significant and positive to non-significant. In all other cases, there was no material difference in the results.

Additionally, for each outcome, we re-analysed the data after iteratively excluding each untreated local authority from the donor pool (a ‘leave-one-out’ analysis) (11). In all cases, there was no material difference in the results.

As detailed above, we also repeated the primary outcome analysis but excluding the local authorities which show a sharp post-treatment dip in the primary outcome. Given that this produces a result with an opposite sign, we cannot make further conclusions about the nature of this result without further knowledge of what caused this dip, and in particular whether the recorded data reflects what actually happened in those local authorities. Lastly, we repeated the primary outcome analysis after removing the weeks in which the dip happens (covering the period July 8 to 28) rather than the local authorities. This produces a null result.

Discussion

Summary of findings

This evaluation aimed to understand the causal impact of the GM Self-Isolation Pathfinder on compliance with self-isolation, COVID-19 testing, and engagement with T&T.

We do not find evidence of impact on compliance with self-isolation

Aside from a data anomaly in the weeks following implementation in some of the untreated comparison local authorities, there is no significant difference in successful Test and Trace call outcomes between GM and the comparison group.

We also do not observe a statistically significant difference in the exploratory outcome of percentage of new isolating cases without any non-household contacts. This result holds whether the pilot period is considered to start when the first local authorities introduce the programme, or when all treated local authorities had introduced it. This also holds if any single local authority is removed from the set of possible comparison local authorities.

We do not find evidence of impact on COVID-19 testing

There was no statistically significant difference between GM and synthetic comparison group on the secondary outcome of PCR testing rate. These results hold regardless of the analysis approach.

For the exploratory outcome of LFD testing rate, the rate was lower in GM than in the synthetic comparison group; however, the matching procedure did not produce a suitable set of matching local authorities for LFD testing rates. As such, with the best available synthetic comparison group, the difference in LFD testing rates is apparent before the treatment started. Consequently, the treatment cannot reasonably be assumed to have caused this difference.

We do not find evidence of impact on engagement with Test and Trace

There was no statistically significant difference between GM and synthetic comparison group in the average number of contacts shared and the percentage of cases sharing at least one contact (exploratory outcomes). These results hold regardless of the analysis approach.

Potential implementation factors underlying the results

To aid interpretation of these results, it is helpful to refer to the separate quantitative implementation and process evaluation (IPE) of the Pathfinder and rapid qualitative research in 2 local authorities conducted by BIT (for details of the implementation evaluation, see 1). They

highlight both the pilot's significant achievements in implementation and offer insights into the results of the impact evaluation. By 20 August 2021, the pilot had:

- reached more than 45,000 people, 351% of the original target
- provided enhanced support to more than 5,000 people, 156% of the original target

Two rapid qualitative research projects in Oldham (14) and Wigan (15) also indicated that local authorities had reported that the support available had been invaluable in identifying and reaching people in vulnerable situations.

Whilst the IPE highlights the pilot's significant achievements, it may also offer explanations for why the Pathfinder may not have affected the impact evaluation outcomes. An underlying hypothesis for the pilot's potential to positively impact these outcomes is that the enhanced offer of support would address reasons identified for non-compliance with self-isolation requests.[26]

That is, if people knew that their needs during self-isolation would be met, then people might be more inclined to self-isolate, test, and share their contacts' details. The 3 tentative suggestions for why the pilot might not have addressed the mechanisms underlying this hypothesis are listed below.

First, awareness of the enhanced support varied significantly across local authorities on the pilot

There is noticeable variation in the 'reach' (that is, the number of people contacted by the pilots) across the local authorities: whilst Tameside achieved 2,500% reach of their initial target (949 individuals), Stockport achieved 57% of its target (1085 individuals) (1). This variation is also seen in the data provided by 6 local authorities on the number of eligible people reached. Whilst on average these 6 local authorities reached 80%, this ranged from 22% in Oldham to 100% in Manchester, Tameside and Rochdale. The different approaches to outreach account for this: Tameside used mass texting, contacting every individual eligible for the pilot; Stockport relied purely on telephone and leaflets to engage cases and contacts (1). For the hypothesis behind the impact evaluation outcomes to be true, it requires residents in pilot areas to know that their needs in self-isolation will be met. The variation in the pilot's reach across local authorities may indicate that such knowledge was not universal, and consequently may partly account for the null result in the impact evaluation.

Second, there is notable variation in the number of people supported across the local authorities

Wigan supported 356% of its target (381 people), Oldham supported almost 300% of its initial target (297 individuals), and Salford supported under 20% (301) (1). Qualitative research in Wigan and Oldham suggest that, in addition to reach, differences in process and implementation might account for some of the variation in the uptake of support. Whilst most local authorities employed telephone scripts for contact tracing and self-isolation support, Oldham employed strength-based and conversational-style calls which focused on welfare and

[26] GMCA (2021a) (3) identified the 5 top reasons for non-compliance as: financial support, the need for assistance with food shopping or essential goods, mental health, loneliness and non-work responsibilities.

minimised the administrative burden on the contact or case during the call (15). Similar efforts to reduce administrative tasks and use motivational interviewing were found in Wigan's model (14). As with reach, variation in support offered across the local authorities might affect awareness of the extent to which the pilot addressed key barriers to isolation, with the consequent impact on the impact evaluation outcomes.

Finally, whilst the pilots address many of the reasons for non-compliance, they may not adequately address the primary barrier: financial support (2)

The overwhelming demand for support from the pilots related to financial support, with sign-posting and direct provision accounting for 82% of all support offered. (This compares with 12% for food provision, 2% for medical provision and 3% for all other forms of support.) However, with policy on self-isolation support payments set nationally, the pilots had reduced scope to actively address this barrier. Failure to reassure people that their financial - and employment needs – would be met might have limited the extent to which the pilots could encourage people to come forward to get tested or share contact details.

However, these tentative explanations require a significant caveat. Until the completion of the comprehensive qualitative IPE commissioned by Greater Manchester – which will engage with all 10 local authorities participating in the pilot – it will be hard to assess the extent to which differences in the implementation of pilots across Greater Manchester contributed to the absence of observed impact in the current evaluation.

Analysis limitations

Aside from the potential implementation factors mentioned above, there are some limitations with the study design, data and the outcome measures that may underlie the observed pattern of results.

Design limitations

There are several limitations which are a function of the evaluation's design:

1.1 We have chosen a synthetic comparison group which is designed to match GM on as many variables as possible, including pre-intervention values of the 6 outcome measures, rural or urban classification and other variables (for a full list see the Matching variables section earlier in this report and Annexe B: Technical details). However, as with all such analyses, it is possible that the synthetic comparison group does not match GM on unobserved variables.

1.2 Matching on more than one variable does entail a level of compromise between them. In this case, while the match quality on almost all variables was good, there was a systematic difference in the LFD testing rate between GM and the synthetic comparison group in the pre-intervention period.

1.3 Given the comparatively small number of treated local authorities and the matching constraints (which produce a similarly small matched comparison group), we are only able to detect relatively large effects (of the order of a relative change of 10% in the primary outcome). If the programme had a smaller effect than this, then the analysis may not have been able to detect it as statistically significant.

1.4 We do not include untreated wards in GM in either the treated or comparison group because it is unclear how much the treatment effect might have "spilled over" and affected these wards. However, the treated wards in local authorities which did not target all wards may be systematically different to untreated wards in those local authorities.[27] It is also possible that there are spillovers from treated to untreated local authorities, which might reduce the apparent difference between them.

1.5 We are not able to identify in the data set the specific cases in each local authority that have been offered support and have taken it up. Although this does not bias the analyses, it does add noise to the estimates, thereby reducing our ability to detect a statistically significant effect.

1.6 The analysis may not adequately account for intervention variation across local authorities in GM in terms of the timing of pilot launches, extent of target wards, and delivery approaches. Although we address this variation by testing the results' robustness to changes in the intervention onset time, and define outcome data that is specific to the target wards, there may be other variation that is unaccounted for. We note that flexibility in local delivery is an intentional feature of the Pathfinder programme; however, as this is an evaluation of the causal impact of the Pathfinder programme as a whole, the design has accordingly focused on estimating impact across all pilots in GM.

1.7 Lastly, the post-intervention evaluation period is relatively short, covering at most 6 weeks after the end of the earliest pilot in Bury. Potential impact may take longer to emerge in the behavioural outcome measures that we focused on. By contrast, potential impact on attitudes and perceptions towards compliance with self-isolation, testing, and engagement with T&T may occur earlier. The qualitative Implementation and Process Evaluation (IPE) may uncover such early changes among community members in GM.

Data limitations

DHSC informed us that there are some data quality issues affecting the T&T call outcome data from late May until the end of July, which are not expected to differ across geographies. This is likely to add noise to the estimates.

[27] Since we are matching local authorities on pre-intervention data, this will not bias our analysis. However, it may have implications for how representative the effects we observe are, and they may not extend to wards which are unlike those in the treated group in GM.

Additionally, initial investigation by UKHSA indicates that the post-intervention dip in successful call outcome rates in some untreated local authorities was related to logistical issues rather than indicative of a change in isolation behaviour. This interpretation is reinforced by other evidence relating to compliance with self-isolation, including the random probability surveys on self-isolation compliance undertaken by ONS. This adds some uncertainty to the interpretation of the primary outcome analysis; however, as concluded above, the set of all analyses conducted suggests that there is no evidence that the Pathfinder programme impacted self-isolation compliance.

Limitations of each outcome measure

The national T&T system only selects a subset of cases for follow-up calls to check compliance with self-isolation (the primary outcome). We assume that this selection cannot be linked to the presence or absence of the Pathfinder programme, and therefore does not bias the analysis. However, it does add noise to the estimates, thereby reducing our ability to detect a statistically significant effect.

For the exploratory outcome focusing on the percentage of new isolating cases without any non-household contacts, the self-isolation period in which compliance is measured is likely to start before support is offered in GM. This means that the programme is more likely to have only an indirect impact on this measure.

The PCR testing rate (and to an extent, the LFD testing rate) may be influenced by the presence of surge testing in an area. We are unable to match local authorities on when surge testing was in force as it was generally only in a small number of locations at any one time. However, we do match local authorities on the basis of the total amount of each kind of surge testing in the pre-intervention period.

The outcomes that focus on the number of reported contacts are only proxies for true behaviour; whilst they capture both engagement with the T&T process and some information about the number of contacts individuals have actually had, they are only imperfect information about the latter.

More broadly, the observed absence of impact in the outcome measures investigated here is not evidence of no impact on other outcomes that may have been affected by the pilot intervention (for example, improved reach of people in vulnerable situations, as reported in a case study from the quantitative IPE).

Recommendations

Policy recommendations for future pilots

1.1 Empower future pilots to tackle all major barriers to self-isolation compliance: The pilots were able to address many of the locally (2) – and nationally (16) – identified reasons for non-adherence to self-isolation. However, the pilot was unable to tackle a key factor: financial barriers or loss of income. With the £500 self-isolation support payment only covering around 1 in 8 workers (17) and statutory sick pay covering only 25% of the average worker's wages, over 2 million low-paid workers are excluded (17). This may be particularly important for Greater Manchester where 19.2% of jobs are low paid (18).

1.2 Focus on – and budget for – increasing awareness of the pilot as a key mechanism in the theory of change: Central to the theory of change for both the primary and secondary outcomes is the notion that availability of enhanced support will address key barriers preventing people from isolating. Even with an intervention tackling all these barriers, its impact may be reduced if the population is unaware that such support exists. As noted above, the quantitative IPE indicates significant variation in the pilot reach between local authorities (1). For pilots where awareness of the operating model will play a central role in the engagement with a target population, increased emphasis and budget on advertising and promotion should be prioritised. The qualitative IPE will seek to further understand what lessons can be learned from local authorities which used mass SMS, email or mail-outs to widen their reach.

1.3 Ensure pilots engage in consistent and regular reporting of progress during implementation: Greater Manchester's weekly Summary Tracker allowed all local authorities to monitor progress towards their goals, identify outliers for further investigation and encourage sharing of lessons during implementation at weekly review meetings. Using Greater Manchester's approach as a model, future pilot programmes should make such a process a condition of the grant, and include additional requirements for grantees to identify and define key metrics to ensure consistency in the data collected across pilots.

1.4 Support pilots to be responsive and adaptive: Preliminary qualitative research in Wigan and Oldham indicate the benefits of enabling local areas to adapt the delivery model to their local needs. For example, by enabling Oldham flexibility in how to manage its contact tracing calls it was able to pilot a non-script based approach that drew on local knowledge and minimised bureaucracy (14). In the case of Wigan, the autonomy to adapt their delivery model allowed them to move away from indiscriminately offering support to all cases or contacts to re-focusing resources on priority groups (for example, those who had requested support) (15). Future pilots would benefit if further support were offered to enable local authorities to understand the local needs and adapt delivery models to meet them.

Recommendations for future pilot evaluations

1.1 Commission dedicated ‘data exploration’ stages before committing to a design or to an evaluation altogether: Data quality issues are not always possible to foresee, so this will help ensure data is of sufficient quality before proceeding with evaluations.

1.2 Test and refine data sharing and access procedures during a dedicated ‘data exploration’ stage: This will help ensure that data is easily accessible before proceeding with evaluations. This includes considerations around the way data is stored and shared (for example, such that it can be easily exported as a comma-separated values (CSV) file) and, as prepared by DHSC for this evaluation, accompanying data dictionaries which can facilitate evaluation design.

1.3 Commission mixed-methods evaluations which focus on evaluating impact as well as implementation and process: Such evaluations rely on quantitative and qualitative methods to evaluate whether a programme worked, why it did or did not work, and the experiences of delivery stakeholders and programme participants. For example, the evaluation of the Pathfinder programme includes this quantitative impact evaluation, as well as a quantitative IPE and a separate qualitative IPE focusing on all ten pilots. These aspects can inform each other and therefore substantially improve interpretation of findings. This enables a rich, comprehensive, and coherent understanding of programme delivery, potential impact, and possible barriers and facilitators to impact. However, we recognise that such evaluations carry higher costs necessary to accommodate the additional research and interpretation.

1.4 Build in evaluation as an integral part of the pilot design and planning: Evaluation planning should be embedded at the start of pilot design and planning. This includes but it is not limited to (i) developing a logic model that can drive the development of both intervention and evaluation, (ii) ensuring that the pilot intervention and delivery (reach, dose, and target population) are suitable for being evaluated using available data, (iii) potentially implementing the pilot as a randomised controlled trial to ensure a robust evaluation (for example, delivering the intervention only to a randomly selected sets of postcodes). Additionally, this includes accommodating evaluation support within pilot budgets and timelines for participating local authorities. This helps reduce the burden on local authorities as they help with data collection and evaluation planning.

Options for future research for this programme

1.1 Conduct separate evaluations of each pilot in GM: This will test whether specific local authorities within GM have experienced statistically significant impact from the programme. These results should then be linked to and interpreted with the quantitative IPE findings which show substantial variation in reach and offered support across pilots. However, challenges with this approach include the statistical power implications of reducing the size of the treatment group (from 10 local authorities to one local authority)

and adjusting for multiple analyses in each local authority, as well as coherently interpreting the full set of results across local authorities.

1.2 Conduct complementary difference-in-differences analyses (at the ward level for PCR and LFD testing rate, and person level for all other outcomes): This will complement the synthetic control method. It may increase statistical power due to the use of disaggregated observations, as well as potentially overcome limitations of the current matching approach.

1.3 Interpret impact evaluation findings together with the qualitative Implementation and Process Evaluation once completed: Conclusions from this IPE are expected in February 2022. This will help contextualise the observed pattern of results. For example, it will provide evidence as to whether attitudes and perceptions towards self-isolation compliance, testing, and engagement have changed among community members during the programme. It will also help identify potential barriers to behavioural change that may explain the findings from both the impact evaluation and quantitative IPE.

References

1. BIT (2021a). 'Interim report: implementation and process evaluation'
2. GMCA (2021a). 'Safely managing COVID-19: Greater Manchester population survey - June 2021' (accessed on 17 November 2021)
3. (2021a).'Guidance: COVID-19 response, spring 2021 (summary)' (accessed on 17 November 2021)
4. GOV.UK (2021b). 'Self-isolation removed for double-jabbed close contacts from 16 August' (accessed on 17 November 2021)
5. GMCA. (2020b). Press conference data 25 May 2021 (accessed on 17 November 2021)
6. GMCA. (2021c). COVID-19 data for Greater Manchester (accessed on 17 November 2021)
7. Tinman J. (2021). 'Why Salford has one of the lowest vaccination rates around and how the council is combating' Manchester Evening News (accessed on 17 November 2021)
8. NHS Volunteer Responders (2021). Homepage (accessed on 17 November 2021)
9. GOV.UK (2021c). 'Guidance: Test and Trace support payment scheme: claiming financial support' (accessed on 17 November 2021)
10. Reicher S. (2021). 'Contrasting figures on adherence to self-isolation show that support is even more important than ever' BMJ (accessed on 17 November 2021)
11. Abadie A. (2021). 'Using synthetic controls: Feasibility, data requirements, and methodological aspects'. Journal of Economic Literature 2021: volume 59, issue 2, pages 391 to 425
12. Robbins MW, Davenport S (2021). 'Microsynth: synthetic control methods for disaggregated and micro-level data in R'. Journal of Statistical Software 2021: volume 97, issue 1, pages 1 to 31
13. Kaul A, Klößner S, Pfeifer G and Schieler M (2021). 'Standard synthetic control methods: the case of using all pre-intervention outcomes together with covariates' Journal of Business and Economic Statistics, pages 1 to 34
14. BIT (2021b). 'Memo: Lessons from the Oldham Pilot'
15. BIT (2021c). 'Note: Lessons from the Wigan Pilot'
16. Smith SE, Potts HWW, Amlot R, Fear NT, Michie S and Rubin GJ (2020). 'Adherence to the test, trace and isolate system: results from a time series of 21 nationally representative surveys in the UK (the COVID-19 Rapid Survey of Adherence to Interventions and Responses [CORSAIR] study)'
17. Brewer M, Gustafsson M (2021). 'Time out: reforming statutory sick pay to support the COVID-19 recovery phase' Resolution Foundation. Retrieved on 17 November 2021
18. D'Arcy C, Gardiner L and Rahman F. 'Low pay in Greater Manchester: a report for the Greater Manchester Independent Prosperity Review' Resolution Foundation, March 2019
19. Abadie A, Diamond A, Hainmueller J. 'Comparative politics and the synthetic control method' American Journal of Political Science 2015: volume 59, issue 2, pages 495 to 510
20. Robbins MW, Saunders J, Kilmer B. 'A framework for synthetic control methods with high-dimensional, micro-level data: evaluating a neighborhood-specific crime intervention'

Journal of the American Statistical Association 2017: volume 112, issue 517, pages 109 to 126

21. McClelland R, Gault S. (2017). 'The synthetic control method as a tool to understand state policy' Washington, DC: The Urban Institute
22. Abadie A, Diamond A, Hainmueller J (2010). 'Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program' Journal of the American Statistical Association 2010: volume 105, issue 490, pages 493 to 505

Authors

Clare Delargy — Principal Advisor

Clare is a Principal Advisor who has a strong track-record of leading and managing public policy projects across a range of areas including COVID-19, volunteering, loneliness, and charity management. At present, Clare leads BIT's work on COVID-19 and deliberative democracy. She has a Master's in Public Policy from the University of Chicago and a Bachelor's in Psychology from Trinity College Dublin.

James Lawrence – Head of Quantitative Research

James is the Head of Quantitative Research within BIT's Research and Evaluation team and has managed and delivered data science and evaluation projects in policing, road safety and targeting government inspections. He has also conducted and published statistical ecology research into the population dynamics of North American wildlife, using statistical simulation techniques. James holds a MMath and BA from the University of Cambridge. In his master's thesis, he examined the effects of measurement error on classification problems.

Tim Ramsey – Senior Advisor

Tim is a Senior Advisor in the Health team, leading COVID-19 projects in the UK and USA. He has 8 years' experience of running complex projects in the UK and overseas, including in behavioural science and qualitative research. Tim has a Double First in Classics and Arabic from the University of Oxford, an MBA from the Quantic School of Business and Technology and a Master's degree in Psychology from Harvard University.

Dr Giulia Tagliaferri – Senior Research Advisor

Giulia is a Senior Research Advisor in the Communities, Education and Health team. She leads BIT's analytical work in complex quantitative evaluations (observational data, quasi-experimental methods, thorny RCTs) in education, health and social capital. Prior to joining BIT, Giulia worked for the Centre for Vocational Education Research (CVER) at the London School of Economics. She completed her PhD in Economics at Queen Mary University of London in 2018, where her research focused on the economics of identity. She holds an MSc and a BA in Economics and Social Sciences from Universita' Bocconi in Milan.

Dr Lev Tankelevitch – Research Advisor

Lev is a Research Advisor specialising in trial design, implementation, and quantitative data analysis in policy areas including public health and education. Prior to joining BIT, Lev was a research fellow at the Parliamentary Office of Science and Technology, where he worked on policy briefings about sleep health. Lev holds a DPhil in Neuroscience from the University of Oxford, where he studied the interaction between learning and attention. He also holds an MSc in Neuroscience from the University of Oxford, and a BSc in Psychology Research from the University of Toronto.

Annexe A. Summary of each local authority's delivery model

Bolton

Bolton's Response Hub led the primary delivery of their pilot. In addition to proactive calling of contacts and cases, Bolton developed a 'Self-isolation Support Pack', available in both hard and soft copies, with guidance and support on how to self-isolate. Based on the results of a doorstep survey in one postcode (BL3), Bolton introduced further innovations into the design including: email, letter and text services to contact potential recipients, additional language support to engage residents for whom English was not their first language and a dedicated community engagement team to make door-to-door visits.

Bury

Bury proactively called index cases and contacts, offering translation services in multiple languages, including Urdu, Punjabi, and Spanish. They produced a brochure on self-isolation rules and support available which was sent to every household. If cases or contacts did not reply after multiple attempts, a Housing Officer or Covid Marshall would conduct a home visit to offer support in a strengths-based manner directly and ensure compliance with self-isolation. Specific support outlined in relation to the Bury Community Support Network in relation to humanitarian assistance and Getting Help Helpline in relation to mental wellbeing. Integrated with local Community Hubs to provide as much local support as possible to individuals and used the engagement to proactively identify local vaccination pop-up sites for each individual to encourage uptake. Specific engagement through Education Service to support individuals from Roma-Gypsy Traveller community on both practical and social support if required to isolate through trusted sources and building on existing relationships.

Manchester

Calls to cases and contacts were led by the local Trace Support to Self-isolate Team (TSST) which works closely with other teams delivering humanitarian support, including both the City Council Community Response Helpline and the Food Partnership. The TSST is supported by local contact tracing nurses to whom cases were escalated if there were medical or safeguarding concerns. To handle the high volume of demand in the middle of June, the team introduced a prioritisation matrix which focused support on:

1.1 Those who requested support and were resident in wards participating in the GM pilot
1.2 Those who requested support living outside of the pilot area
1.3 Older people and those with young families
1.4 Those who do not fulfil criteria 1 to 3

Manchester also sent text messages to cases and contacts which provided information on the support available and how they could access it.

Oldham

Oldham adopted a borough-wide approach to enable greater flexibility in the model to respond to demand, whilst still supporting those areas with high levels of enduring transmission. The team was composed of members of “Access Oldham” which enabled them to address additional concerns on the call relating to council tax, benefits and other services. Oldham focused on “localising” the approach - using local numbers, recruiting people with local knowledge and emphasising their connection with the council to build trust. Rather than using scripts, they focused on keeping the conversation free-flowing and natural, undertaking any administration associated with self-isolation before or after the call. The self-isolation support team helped residents to develop personalised plans and connected individuals to support provided by local voluntary sector organisations, if required.

Rochdale

The Isolation Team aimed to contact all cases or contacts in the target wards within 24 hours in order to undertake an assessment of their needs and additional support. The support plan developed would also include any assistance required by household members (for example, other occupants or children). For cases or contacts who did not respond to the phone calls, home visits took place. In order to maximise engagement with potential cases or contacts, Rochdale sought to build engagement and comms though community settings (that is, schools, religious settings), local businesses and trusted messengers (for example, GPs, religious leaders).

Salford

The Tracing Team made contact with cases or contacts by phone, email or text message using scripts to assess their support needs. Residents who identified they needed support were referred directly into the Spirit of Salford Helpline, which has been in place throughout the pandemic offering humanitarian assistance including financial support, Welfare advice, food, mental wellbeing support and access to volunteers. During the pilot stage outreach staff were out every day speaking to residents about the key messages and national guidance including isolation, testing, vaccinations and accessing support.

Stockport

Stockport’s Self-isolation Support Team contacted cases/contacts to identify support needs and connected them with the council or voluntary organisation providing support. In June, Stockport engaged with colleagues in the local library service to secure online resources for children and

adults in addition to local book delivery. In July, a leaflet with information about self-isolation and testing was produced to try and increase engagement. Further, to address the high demand for financial support, Stockport increased the size of the team administering self-isolation payments.

Tameside

Mass text messaging of all eligible cases and contacts was used to raise awareness of the pilot programme, with the local self-isolation support team focusing telephone calls on harder to reach cases. These calls used scripts to identify needs; anyone who requested support was passed onto a call centre for follow-up conversations. Tameside worked with local third sector partners to provide additional support and outreach in the local communities. From the start of July, those cases which the team are unable to reach receive a letter about self-isolation through the post.

Trafford

Trafford's original scope was to contact cases in the geographical location of M15 and M16 postcodes to explicitly ask them if they have any support requirements that would enable them to fully carry out their isolation period. The scope of this progressed into additional geographical locations. The primary focus was on early engagement with index cases and their households to develop a person-centred self-isolation support plan, practical and emotional support and alternative accommodation. The Trafford model was based on a person-centred and focussed on supporting people and having asset based conversations to develop a plan which would be delivered internally and alongside partners. Such enhanced support might include frequent calls, practical support, support with financial elements and other clinical or health needs. Cases and contacts were contacted following scripted questions to allow individuals to establish what types of support communities can provide. This replaced the national standard question of "Do you require any support whilst self-isolating?"

Wigan

Wigan initially implemented a 'broad-brush' approach to self-isolation support across a limited number of postcodes, following-up all contract tracing calls with an in-person visit. However, the team revised this approach to target resources to those in need across the whole borough, prioritising those who had requested – or were suspected of being in need of – support and those whom they were unable to contact. Calls and home visits focused on people's welfare, building trust and using motivational interviewing techniques to elicit information. Wigan aimed to reach cases as early as possible in their self-isolation period, with a home visit undertaken with 24 hours of referral and support available at the weekend. For every case or contact in the borough, Wigan delivered information packages about the support available.

Annexe B. Technical details

Synthetic control method

The synthetic control method affords several advantages (1, 12, 19):

- the data-driven approach obviates the need for a manual and potentially subjective selection of comparison units
- explicitly matching on key relevant variables can potentially increase the similarity of treatment and comparison groups, strengthening the parallel trends assumption
- the weighting approach can provide a suitable comparison group even when every individual untreated unit fails to demonstrate pre-intervention parallel trends with the treatment group

The synthetic control approach has usually been applied in contexts with one treated unit, which is compared to multiple untreated units. However, in the current evaluation there were multiple treated units (local authorities within GM), which can improve the construction of the synthetic comparison group due to the increased granularity of the data (20). Therefore, the evaluation used a synthetic control approach recently developed for such meso-level data, and implemented in the R package microsynth (12).

The geographical unit of analysis is the local authority (excluding any wards in which the programme was not implemented; see the Treated areas' section, below, for the approach to local authorities which re-defined their target area after launch), and the time unit of analysis is the week.[28] The estimated treatment effect therefore reflects the effect of treatment for an average treated local authority over the post-intervention weeks.

Treated areas

For most local authorities, the planned target areas remained consistent across the programme period. However, several local authorities in the Pathfinder programme re-defined their target areas for intervention after programme launch, either due to high case rates or insufficient uptake. For 3 of these local authorities, which re-defined their target areas within approximately one month of launch, the updated target area was used for the entire post-intervention period (that is, it was retrospectively updated and we adopt an intention-to-treat approach); for one local authority, which re-defined its area within approximately 2 months of launch, the original target area was retained for the entire post-intervention period (see Table 8 below for a

[28] Analysing data at the more granular ward level was an alternative approach that was considered. This aligns with the fact that some local authorities chose target wards for the programme, whereas others ran the programme in all wards covering the entire local authority. However, the sparseness of COVID-19 cases in some wards meant that there are time periods for which some outcomes are not defined (for example, the primary outcome for compliance with self-isolation). Given the importance of time series granularity for SCM, we therefore aggregated data across geography, rather than across time (11).

summary of the final target areas for analysis for each local authority). This approach was adopted to enable a consistent intervention area for the entirety of the post-intervention period (necessary for the synthetic control evaluation approach here), while maximising the number of people that could be considered exposed to the intervention.

Table 8. Summary of target areas for analysis for each local authority in Greater Manchester

Local authority	Target areas for analysis
Bolton	Entire local authority[29]
Bury	Target wards (East, Moorside, Radcliffe West, Besses, Sedgley)[30]
Manchester	Target wards (Longsight, Levenshulme, Cheetham, Crumpsall, Rusholme, Ardwick, Moss Side, Whalley Range)
Oldham	Entire local authority
Rochdale	Target wards (Milkstone and Deeplish, Central Rochdale, Kingsway, West Middleton)
Salford	Target ward (Little Hulton)
Stockport	Entire local authority
Tameside	Entire local authority
Trafford	Target wards (Bucklow St Martins, Clifford)[31]
Wigan	Entire local authority[32]

Study period

The pre-intervention period consisted of 24 weeks, defined from 3 December 2020 until 13 May 2021 inclusive (corresponding to T&T programme weeks 28 to 51).[33] The post-intervention period consisted of 18 weeks, defined from 20 May 2021 until 16 September 2021 inclusive (weeks 52 to 69). The start of the post-intervention period (week commencing 20 May 2021) is the week closest to the earliest pilot launch date of 17 May 2021 in Bury. The end of the post-intervention period (week commencing 16 September 2021) covers the last pilot end date of 13 September 2021 in Trafford.

[29] This was originally 12 target wards, but as of 29 June 2021 (3 weeks after launch), it was expanded to the entire local authority; the target area was retrospectively updated for the evaluation.
[30] This was originally the entire local authority, but as of 21 June 2021 (5 weeks after launch), it was changed to 5 target wards; the target area was retrospectively updated for the evaluation.
[31] This is the original target area, but as of 4 August 2021 (7.5 weeks after launch), this was expanded to all wards within M32 and M33 postcodes; the original target area was retained for evaluation.
[32] This was originally 10 target wards, but as of 21 June 2021 (4 weeks after launch), this was expanded to the entire local authority; the target area was retrospectively updated for the evaluation.
[33] Data does exist prior to this period; however, in the earliest phase of the pandemic it becomes harder to construct a good quality match for the treated local authorities in GM because the outcomes have very high variance, depending on how quickly different local authorities established their Test and Trace capability. This is largely unrelated to the intervention under consideration so we do not include the early period in our analysis.

Although the programme was intended to run for 12 weeks, the evaluation period was extended to include the entirety of the programme for all local authorites, as well as to enable a period of several weeks for most local authorities to be able to observe any lingering impact after the official pilot end dates.

Defining absolute (rather than relative) pre-intervention and post-intervention periods for all local authorites was necessary to implement the synthetic control matching procedure, which attempts to match consistent time series in the donor pool to those in the treatment group. Although this means that there are weeks in the post-intervention period in which certain pilots were not officially active, the evaluation estimated average treatment effects across the post-intervention period. Additionally, as a robustness check for the analysis of primary and secondary outcomes, the intervention onset was set to the week closest to the launch date of the last local authorites (14 June 2021, in Salford, Tameside, and Trafford), instead of the first (17 May 2021, in Bury), to test how sensitive results were to this detail.

Defining the donor pool

It is necessary to exclude from the donor pool any units which may be affected by the self-isolation pilot programme, by similar interventions, or by large idiosyncratic shocks to the outcomes (11). It is also important to restrict the donor pool to units with characteristics 'similar' to the treated unit. This avoids the risk of interpolation biases (which may arise during the weighted averaging procedure), and overfitting, which "arises when the characteristics of the unit affected by the intervention […] are artificially matched by combining idiosyncratic variations in a large sample of unaffected units" (19). Therefore, to be included in the donor pool local authorities had to meet the following defined inclusion criteria:

- inside England, due to other UK countries maintaining different COVID-19 policies
- outside of Greater Manchester
- not implementing a similar pilot programme (as reported by DHSC)[34]
- with the same Rural-Urban classification as local authorities within Greater Manchester (that is, Major Urban Conurbation), as this broadly captures relevant characteristics that affect COVID-19 prevalence and self-isolation behaviour (for example, demographics, population density, employment trends)

Applying these criteria left 47 out of 317 local authorites (15%), using 2019 local authority boundaries.[35] Given the relevance of these criteria for all evaluation outcomes, the same donor pool was used for all analyses.[36]

[34] Based on data provided by DHSC, 52 local authoritie are excluded for this reason.
[35] This figure also excludes 2 additional local authorities (City of London and Thurrock) which were excluded due to having missing primary outcome data for some weeks in the evaluation period.
[36] An alternative synthetic comparison group was also constructed for the primary outcome analysis as a robustness check; see Robustness checks section for details.

The broad comparability of treatment and donor pool units was confirmed by examining time trends in the outcomes (with a particular focus on the primary outcome), COVID-19 case rate, and PCR testing rate, as well as the cumulative vaccination rate (at the end of the pre-intervention period), deprivation scores, and age and ethnicity population breakdowns (see Annexe C). Specifically, values for the donor pool units span the range of values observed for treatment units in GM. The only variable where the 2 were not deemed comparable was the pre- (and post-) intervention LFD testing rate. It does not appear to be possible to obtain a good match on LFD testing rate without compromising either the effective sample size or the quality of the match on other variables.

Estimating weights for the synthetic comparison group

The microsynth implementation constructs a synthetic comparison group by finding weights which satisfy 3 types of constraints (12, 20):

1.1 The sum of the weights should equal the number of units in the treatment group.
1.2 The weighted synthetic comparison group, aggregated across donor units, should match the aggregated treatment group across any time-invariant covariates.
1.3 The weighted synthetic comparison and treatment groups should match across all pre-intervention time points for the outcome (and any other time-varying covariates).

The implementation first attempts to exactly satisfy all of the above constraints. If this isn't feasible, microsynth attempts to find weights which exactly satisfy constraints (1) and (2), while minimising the degree to which constraint (3) is not satisfied. It also adds another constraint, which it tries to exactly satisfy: the outcome (and any other time-varying covariates), aggregated across pre-intervention timepoints, should match between weighted synthetic comparison and treatment groups. If this also isn't feasible, it exactly satisfies only constraint (1) above (always feasible) and minimises the degree to which the other constraints are not satisfied.

In the current evaluation of the primary outcome, this latter approach was necessary to construct a suitable synthetic comparison group.

To find weights for this latter approach, quadratic programming is used to find a solution to the following:

$$\text{Minimise } W \text{ in: } (X^{M}\top W - T^{M})\top(X^{M}\top W - T^{M})$$

$$\text{Subject to: } (X^{E})\top W = T^{E},$$

$$\text{and } W \geq 0$$

Where:

- W is the transposed matrix of weights $(w, . , w_{J_0})$ ⊤, for J_0 untreated units; the latter is a subset of J which is the set of all units (treated and untreated), indexed with j such that the first J_0 are untreated and the final $J - J_0$ are treated

- X is the transposed matrix of variables $(x_1, . , x_{J_0})$ ⊤ defining the 3 constraints above for J_0 untreated units (that is, the first weight sum constraint and any matching variables)

- T (not to be confused with ⊤, the transpose operator) is a vector which contains the target totals of the 3 constraints for the treatment group (specifically, it contains the elements of t_x, such that:

$$t_x = \sum_{j=J_0+1}^{J} x_j)$$

- X^M is the matrix which contains the matching variables from X that are subject to "minimisation", rather than exact matching (that is, all the matching variables in the current design)

- T^M is a vector which contains the target totals of constraints for the treatment group from T that are subject to "minimisation", rather than exact matching (that is, all the matching variables in the current design)

- X^E and T^E are analogues of X^M and T^M that correspond to constraints that are to be exactly satisfied (that is, only the first weight sum constraint in the current design)

Matching variables

Table 9 lists the matching variables used to construct the synthetic comparison group. For the purpose of matching, including all pre-intervention timepoints for outcome variables has been shown to preclude the influence of other covariates that may be relevant, and therefore this is not recommended (13). In line with recommendations, we therefore selected several key pre-intervention timepoints for each outcome to use as covariates; these are selected to capture the outcome trend in GM local authorities across the pre-intervention period (see vertical dashed lines in Annexe C, Figures 12 to 17) (21). The same is done for another other time-varying covariate: the COVID-19 case rate (see vertical dashed lines Annexe C, Figure 18).

Other matching variables included the cumulative full vaccination rate at the end of the pre-intervention period (T&T programme week 51); the total number of weeks in the pre-intervention period with any surge testing, with postcode surge testing, and with site surge testing (22); the ethnic and age proportion breakdowns of each local authority population; and the IMD 2019 average score for each local authority as a measure of local deprivation.

Table 9. List of variables for matching, their definition, and source

Matching variable	Definition	Source
% successful T&T isolation follow-up call outcomes at key pre-intervention weeks	(number of cases with day 4, 7, 10 isolation follow-up calls successful in week) / (number of cases reached and called in week)	DHSC Dashboard
% of new self-isolating cases with no non-household contacts at key pre-intervention weeks	(number of new self-isolating cases with no non-household contacts in week) / (number of new self-isolating cases in week)	DHSC Dashboard
PCR testing rate at key pre-intervention weeks	(number of PCR tests in week) / (local authority population)	DHSC Dashboard; ONS mid-2019 population estimates
LFD testing rate at key pre-intervention weeks	(number of LFD tests in week) / (local authority population)	DHSC Dashboard; ONS mid-2019 population estimates
Average number of reported contacts at key pre-intervention weeks	(sum of reported contacts) / (number of cases)	DHSC Dashboard
% of cases sharing at least one contact at key pre-intervention weeks	(number of cases sharing at least one contact) / (number of cases)	DHSC Dashboard
COVID-19 case rate at key pre-intervention weeks	(number of positive PCR tests in week) / (local authority population)[37]	DHSC Dashboard; ONS mid-2019 population estimates
Total number of weeks in pre-intervention period with site surge testing	Number of weeks in pre-intervention period with site surge testing[38]	DHSC data

[37] Population estimates at the local authority-level were constructed from ONS ward-level estimates to be able to exclude wards in GM that were not included in the pilot programme.
[38] This includes local and regional testing sites, and mobile testing units. Most local and regional testing sites are active for a full period of several weeks, whereas mobile testing units are deployed per day.

Matching variable	Definition	Source
Total number of weeks in pre-intervention period with postcode surge testing	Number of weeks in pre-intervention period with postcode surge testing[39]	DHSC data
Total number of weeks in pre-intervention period with area surge testing	Number of weeks in pre-intervention period with area surge testing[40]	DHSC data
Cumulative full vaccination rate at the end of the pre-intervention period	(number of people with second-dose vaccinations in local authority by week 51) / (local authority population)	DHSC Dashboard; ONS mid-2019 population estimates
% of each major ethnic subgroup in local authority (Asian, Black, Mixed, Other, White)	(number of people in ethnic subgroup) / (local authority population)	Census 2011 (using mid-2018 local authority boundaries)
% of each age group in local authority (age 0 to 9, 10 to 19, 20 to 29, 30 to 39, 40 to 49, 50 to 59, 60 and over)	(number of people in age group) / (local authority population)	ONS mid-2019 population estimates
IMD 2019 average score for local authority	Population-weighted average of the IMD 2019 scores across LSOAs	English indices of deprivation 2019

Note: For local authorities in GM which selected specific target wards for the Pathfinder, all data on pre-intervention outcomes, COVID-19 case rate, and cumulative vaccination rate is specific to the set of target wards; data on the age and ethnic subgroup breakdowns and the IMD score for each local authority instead reflect the entire local authority, due to data availability limitations.

[39] This reflects local testing in which all residents in an area will be approached and encouraged to take a test via door-to-door communication.
[40] Areas can have site testing only, postcode testing only, both or, in rare cases, neither and depend on another type of analysis (for example, wastewater analysis).

The constructed synthetic comparison group

The make-up of the synthetic comparison group is shown in Table 10. This group is used for all outcome measures. Non-zero weights are assigned to Watford, Walsall, Sunderland, South Tyneside, Runnymede, Ealing, Dartford, and Broxbourne. Around 31% of the weight is concentrated in the greatest-weighted local authority (Sunderland). The effective sample size of the comparison group is 5.18 local authorities. Note that the sparsity of the weights across local authorities is commonly observed and enables the constructed synthetic comparison group to be more interpretable (11).

Table 10. Synthetic comparison weights assigned to each local authority in the donor pool

Donor pool local authority	Weight assigned (%)
Watford	10.4
Walsall	20.3
Sunderland	31.2
South Tyneside	16.2
Runnymede	2.5
Ealing	8.7
Dartford	9.5
Broxbourne	1.1
Wolverhampton, Woking, Westminster, Waltham Forest, Tower Hamlets, Three Rivers, Sutton, Spelthorne, Southwark, Solihull, Sandwell, Richmond Upon Thames, Redbridge, Newcastle Upon Tyne, Merton, Lewisham, Kensington and Chelsea, Islington, Hillingdon, Herstmere, Havering, Harrow, Haringey, Hammersmith and Fulham, Greenwich, Gravesham, Gateshead, Epsom and Ewell, Enfield, Elmbridge, Dudley, Croydon, Camden, Bromley, Brent, Birmingham, Bexley, Barnet, Barking and Dagenham	< 0.01

Table 11 summarises the balance between the treatment group and the synthetic comparison group on key (weighted) covariates. The biggest differences are:

- the synthetic comparison group has a greater proportion of people of Black, Mixed and ‘Other’ ethnicities (4.9%, 2.7% and 1.2% versus 2.2%, 2.0% and 0.8% in the treated group)
- the treated group has an IMD average of 29 compared to 16 in the comparison group

- the treated group experienced more site-based surge testing (2.4 weeks) than the comparison group (0.9 weeks)
- the case rates were different at various points in the pre-intervention time period

Whilst these differences are not completely trivial, they would be unlikely to explain large differences in any of the outcomes we consider in this evaluation.

Table 11. Balance between the treatment group and the synthetic comparison group on key (weighted) covariates

Variable	Treated local authorities average	Comparison group weighted average
Age 0 to 9	14.5%	13.5%
Age 10 to 19	13.1%	11.5%
Age 20 to 29	13.8%	12.1%
Age 30 to 39	14.4%	14.6%
Age 40 to 49	12.4%	13.4%
Age 50 to 59	12.5%	13.0%
Age 60 and over	19.3%	21.9%
Ethnicity: Asian	9.7%	9.4%
Ethnicity: Black	2.1%	4.9%
Ethnicity: Mixed	2.0%	2.7%
Ethnicity: Other	8.1%	12.4%
Ethnicity: White	85.3%	81.8%
IMD	29	16
Surge testing: total weeks	2.4	0.9

Treatment effect estimation

treated units was calculated, subtracted from the corresponding value for the synthetic to obtain an estimate of the treatment effect ( ), the summed outcome value across all comparison group, and summed across post-intervention time periods:

$$\theta = \sum_{t=T_0+1}^{T} \left(\sum_{j=J_0+1}^{J} Y_{tj} - \sum_{j=1}^{J_0} w_j Y_{tj} \right)$$

Where:

- T is the set of all time periods, indexed with t such that the first T_0 are pre-intervention and the final $T - T_0$ are post-intervention

- J is the set of all units (treated and untreated), indexed with j such that the first J_0 are untreated and the final J- J_0 are treated
- Y_{tj} is the outcome at time t for unit j
- w_j is the estimated weight for unit j

Statistical inference

Non-parametric permutation testing was used to conduct statistical inference on the statistic (placebo treatment effects was created by repeatedly:θ) obtained in the 'Treatment effect estimation' section, above. First, a null distribution of

- selecting randomly (without replacement) and assigning a number of non-treated local authorities from the total pool of local authorities to a placebo treatment group (the same size as the actual treated group)
- constructing a synthetic comparison group using the weight estimation procedure described previously
- estimating a placebo treatment effect for each such 'permuted' group

This was repeated for 5,000 permutations.

A p-value for the observed treatment effect is obtained by calculating the proportion of placebo treatment effects that exceed the magnitude of the observed treatment effect. For 2-sided tests:

$$p = 2min\left\{\frac{\#k:\theta^k<\theta}{K}, \frac{\#k:\theta^k>\theta}{K}\right\}$$

Where θ is the observed treatment effect, and θ^k is the placebo treatment effect for permutation k.

Robustness checks

Several robustness checks were conducted to test how sensitive the primary and secondary outcome analysis results are to the analytical strategy:

1.1 Given that GM local authorities launched pilots on different dates, to test the importance of the intervention onset definition for the analysis of primary and secondary outcomes, the intervention onset date was changed from week 51 to week 56 (commencing 17 June 2021), which is the week closest to the last launch date (14 June 2021, in:

1.2 Salford, Tameside, and Trafford), instead of the first (17 May 2021, in Bury). Weeks 52 to 55 (inclusive) were excluded from this analysis to maintain a consistent pre-intervention period across analyses.

1.3 To test the importance of each individual donor pool unit for the results, a leave-one-out procedure was conducted for the primary outcome, where each donor pool unit was iteratively left out and the analysis re-run (as suggested in Abadie, 2021) (11).

1.4 Given the uncertainty around the large post-intervention dip in the synthetic comparison group for the primary outcome analysis, an alternative synthetic comparison group was constructed in which local authorities with such post-treatment dips for the primary outcome were removed from the donor pool (this reduced donor pool had 26 local authorities). The excluded local authorities were: Hertsmere, Watford, Broxbourne, Gravesham, Dartford, Sunderland, Elmbridge, Solihull, Wolverhampton, Spelthorne, Sandwell, Newcastle Upon Tyne, Woking, Walsall, Birmingham, Dudley, Runnymede, Epsom and Ewell, Gateshead, Three Rivers and South Tyneside.

Data sources

DHSC Test and Trace Dashboard

The Test and Trace Dashboard includes individual-level data on individuals in touch with the Contact Tracing and Advice Service (CTAS; sourced from Public Health England), PCR and LFD tests and outcomes (sourced from the National Pathology Exchange System), and COVID-19 vaccinations (sourced from Public Health England's vaccination data set, which is itself sourced from the National Immunisation Management System). The dashboard has national coverage, and data is available for periods before, during, and after the Pathfinder period. It was used to construct the outcome measures and additional covariates for the analysis.

Index of Multiple Deprivation 2019 data

Index of Multiple Deprivation 2019 data was used to obtain an estimate of local authority-level deprivation as a covariate for matching.[41]

ONS ward-level population estimates

ONS data on ward-level population estimates from mid-2019 were used to construct local authority-level population estimates (total count and proportions by age group) to define covariates for matching.[42] This aggregation was done manually so as to be able to exclude non-treated wards in GM from population estimates. Mid-2019 data was selected to match ward boundaries used in the DHSC Dashboard data.

ONS regional ethnic diversity data

This data was used to provide the ethnic population breakdown for each local authority, as a proportion of the total local authority population, as a covariate for matching. The original data is sourced from the 2011 Census and prepared by ONS.[43] The latest version of this data uses mid-2018 local authority boundaries, meaning that 5 newer local authorities in the DHSC

[41] English indices of deprivation 2019
[42] Data set: ward-level population estimates
[43] Regional ethnic diversity

Dashboard data set (with 2019 boundaries) do not have ethnic population breakdowns available. However, these local authorities are all classified as rural and are therefore not included in the analysis for this reason.[44]

Rural urban classification of local authorities

This data is sourced from the 2011 census but uses mid-2019 local authority boundaries.[45] It was used to restrict the donor pool to local authorities with a similar classification as those in GM.

Data cleaning

CTAS data

This data covered the individual-level call outcomes and contact tracing data. The raw data had n = 19,602,385 observations (no duplicates):

- 172 observations (< 0.01%) were excluded due to missing valid ward information
- 17,729 observations (0.09%) were as they concerned Scotland, Wales, and Northern Ireland (of the wards in England, the data set was missing data for Isles of Scilly and the City of London due to small-sample disclosure controls)

The cleaned data set had n = 19,584,484 individual-level observations which were aggregated into local authority-week observations for all of the outcome measures (except PCR and LFD testing rates; see below), excluding non-treated wards in GM.

For the primary outcome (weekly percentage of cases with successful T&T self-isolation follow-up call outcomes), within the analysis weeks (week 28 to 69), 2 local authorities within the donor pool (City of London and Thurrock) were missing outcome data for some weeks and were excluded from the analysis.

The exploratory outcome, weekly percentage of new isolating cases without any non-household contacts, is only defined for the subset of cases that are flagged as previously self-isolating contacts that subsequently became cases (that is, new isolating cases). Therefore, it limits the number of cases available for analysis. Within the analysis weeks (week 28 to 69), 2 local authorities in the treated group (Salford and Trafford)[46] and 16 local authorities within the donor pool were missing outcome data for some weeks. As microsynth is unable to directly handle missing data, missing outcome data was imputed using linear interpolation. If observations were missing at the beginning or end of the data, the last observation was carried forward or the first carried backwards (this never resulted in pre-treatment data carrying forwards to post-treatment or vice-versa, though sometimes the interpolation did cross the pre/post boundary). For

[44] These local authorities are Bournemouth, Christchurch and Poole; Dorset; East Suffolk; West Suffolk; Somerset West and Taunton.
[45] Rural Urban Classification lookup tables for all geographies
[46] Salford was missing data for 5 weeks during the pre-intervention period and one week during the post-intervention period, and Trafford was missing data for 3 weeks during the pre-intervention period.

example, a series of [(missing), 75%, (missing), (missing), 45%,60%, (missing)] would become [75%, 75%, 65%, 55%, 45%, 60%, 60%] under this scheme.

For the engagement with T&T outcomes depending on the number of contacts shared, a number of cases are missing this information. In the treated group, missingness ranged from 6.4% to 15.7% across weeks; in the donor pool, it ranged from 5.5% to 13.7%. We confirmed that time trends in missingness were comparable across treated and donor pool local authorities. It is unclear why this data is missing, and as we are not confident that this missingness necessarily reflects unwillingness to engage with Test & Trace (rather than an unrelated aspect of the contact tracing system), we excluded these cases from the analysis.

Testing data

This data covered the number of PCR and LFD tests performed and the number of positive tests per ward and week. The raw data had n = 1,589,206 observations (no duplicates):

- 704 observations (0.04%) were excluded due to missing valid ward information
- 12,228 observations (0.8%) were excluded as they concerned Scotland, Wales, and
- Northern Ireland (of the wards in England, the data set was missing data for Isles of Scilly and the City of London due to small-sample disclosure controls)
- 13,872 observations (0.9%) were excluded due missing programme week information
- (distributed evenly across wards)
- 135,288 observations (8.7%) were excluded due to invalid programme week information (for example, negative programme week numbers)

The cleaned data set had n = 1,427,114 ward-week observations, which were aggregated into local authority-week observations for the PCR and LFD testing rate outcomes, excluding non-treated wards in GM.

Vaccination data

This data covered the number of first-dose and second-dose vaccinations completed per ward and day. The raw data had n = 3,557,177 observations (no duplicates):

- 590 observations (0.02%) were excluded due to missing valid ward information
- 54,219 observations (1.5%) were excluded as they concerned Scotland, Wales, and Northern Ireland

The cleaned data set had n = 3,502,368 ward-day observations, which were aggregated into ward-week observations (using the T&T programme week), and then into local authority-week observations for the vaccination rate covariate (the second-dose cumulative vaccination rate at the end of pre-intervention period), excluding non-treated wards in GM.

Annexe C. Additional figures

Descriptive figures comparing GM and donor pool local authorities

The figures below serve 2 purposes. First, they provide an overview of the pre-intervention outcome trends and time-invariant covariates across GM and donor pool local authorities, demonstrating the overlap between the 2 groups. For all variables, local authorities in the donor pool span the range of values observed across local authorities in GM, indicating that the donor pool is broadly comparable to GM for the purpose of matching. Second, for the time-varying outcomes and matching covariates, they show which pre-intervention timepoints were selected for matching. These timepoints were selected to capture the pre-intervention time trends in the local authorities in GM to facilitate matching.

Figure 12. Percentage of successful T&T follow-up call outcomes during the pre- and post-intervention period for treated local authorities (blue) and donor unit local authorities (grey)

Vertical dashed lines indicate key timepoints selected for inclusion as matching covariates (see the Matching variables section). We include the post-intervention period here in contrast with later plots to illustrate the 'dip' which affects some donor local authorities but not others (see the Results section for details).

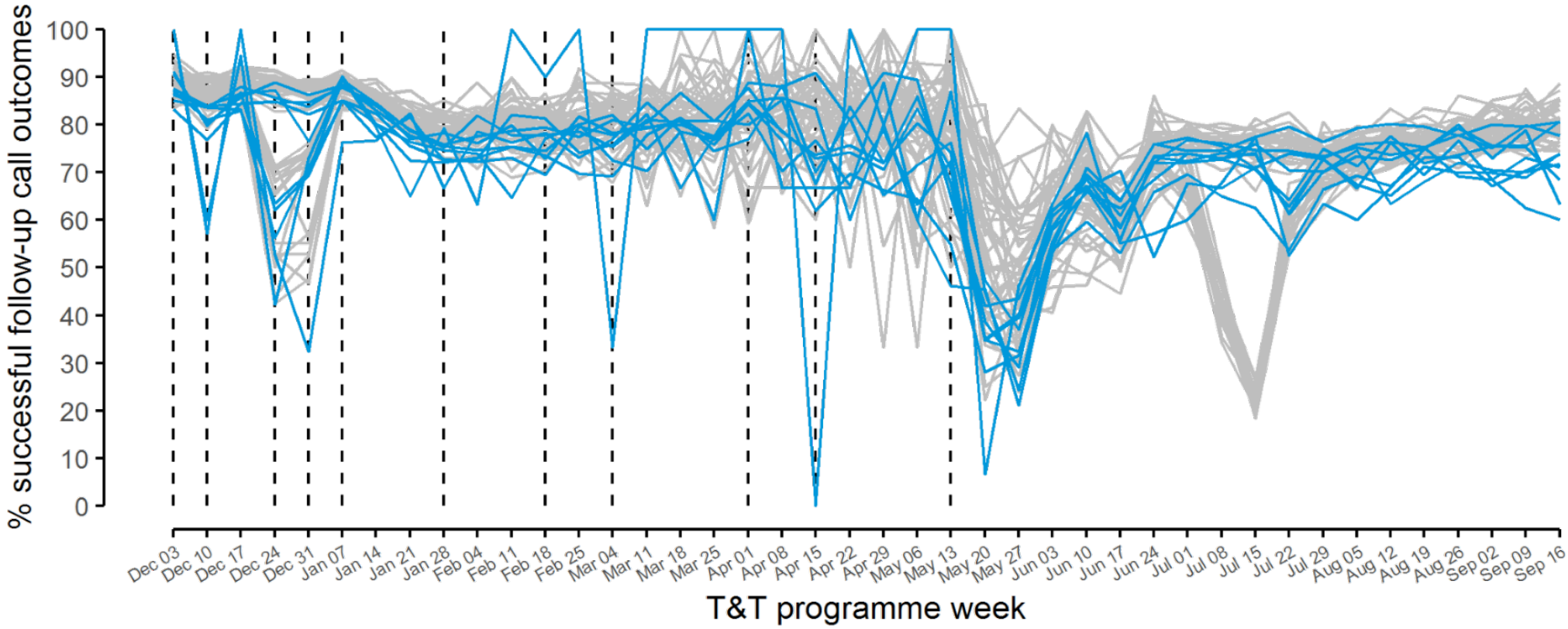


Figure 13. Percentage of new isolating cases without any non-household contacts during the pre-intervention period for treated local authorities (blue) and donor unit local authorities (grey)

Vertical dashed lines indicate key timepoints selected for inclusion as matching covariates (see the Matching variables section).

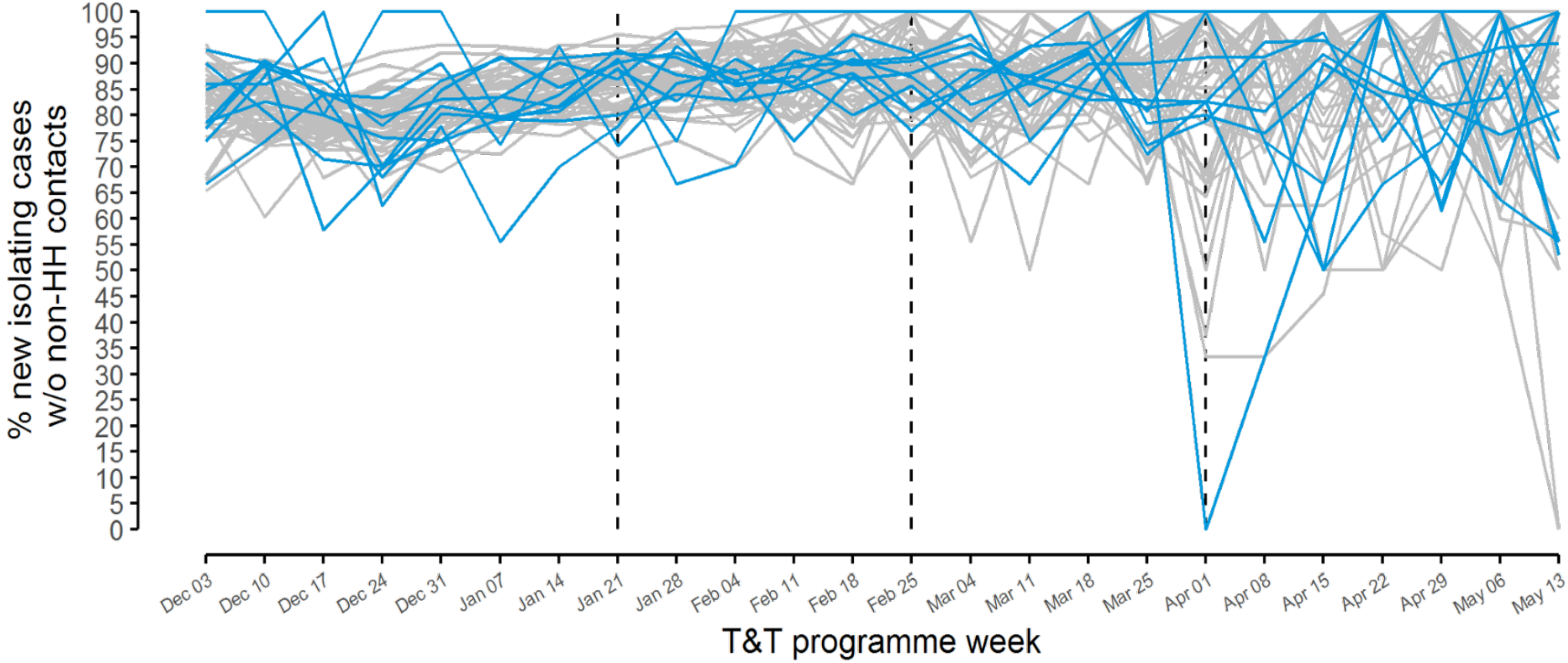


Figure 14. PCR testing rate during the pre-intervention period for treated local authorities (blue) and donor unit local authorities (grey)

Vertical dashed lines indicate key timepoints selected for inclusion as matching covariates (see the Matching variables section).

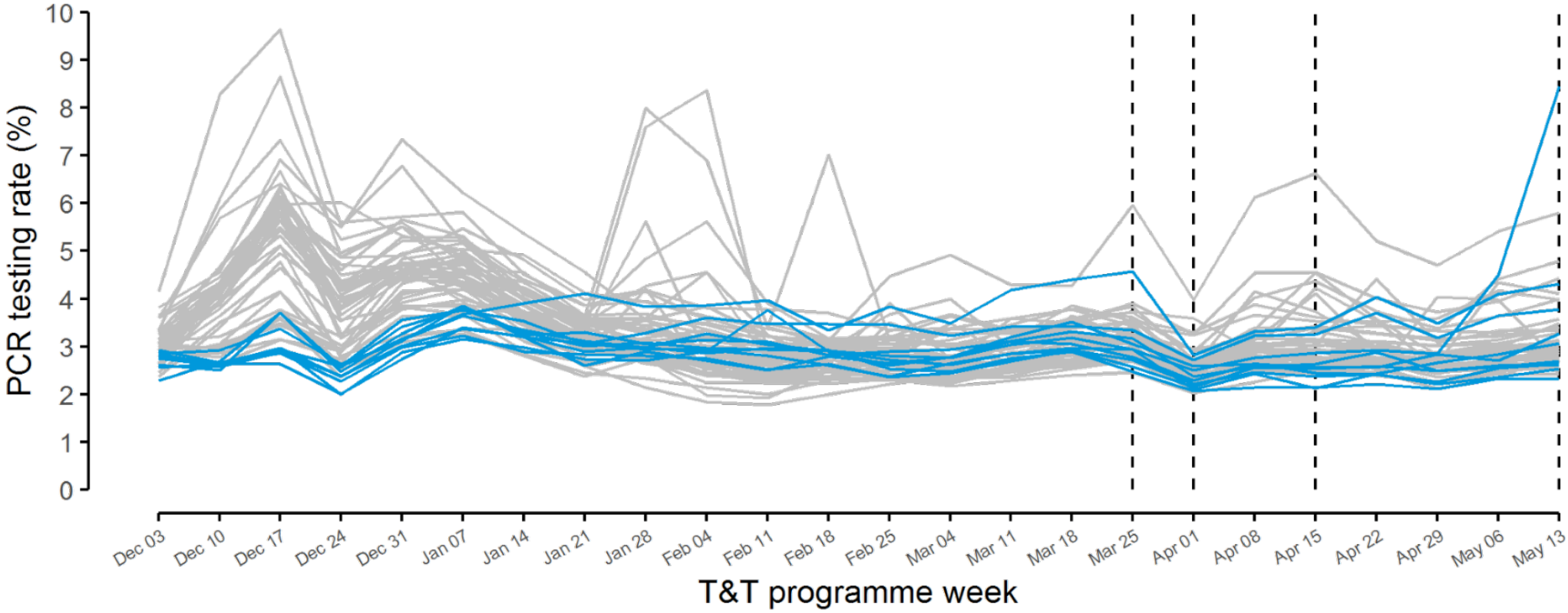


Figure 15. LFD testing rate during the pre-intervention period for treated local authorities (blue) and donor unit local authorities (grey)

Vertical dashed lines indicate key timepoints selected for inclusion as matching covariates (see the Matching variables section).

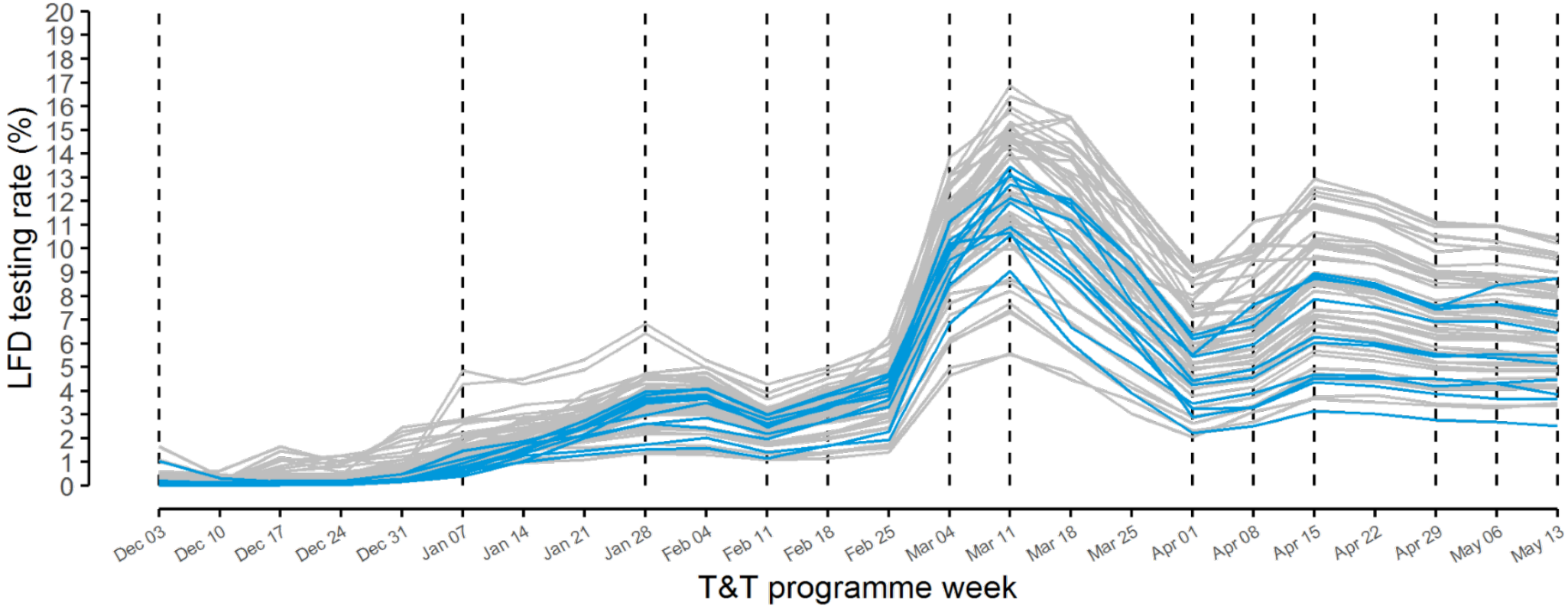


Figure 16. Average number of contacts shared during the pre-intervention period for treated local authorities (blue) and donor unit local authorities (grey)

Vertical dashed lines indicate key timepoints selected for inclusion as matching covariates (see the Matching variables section).

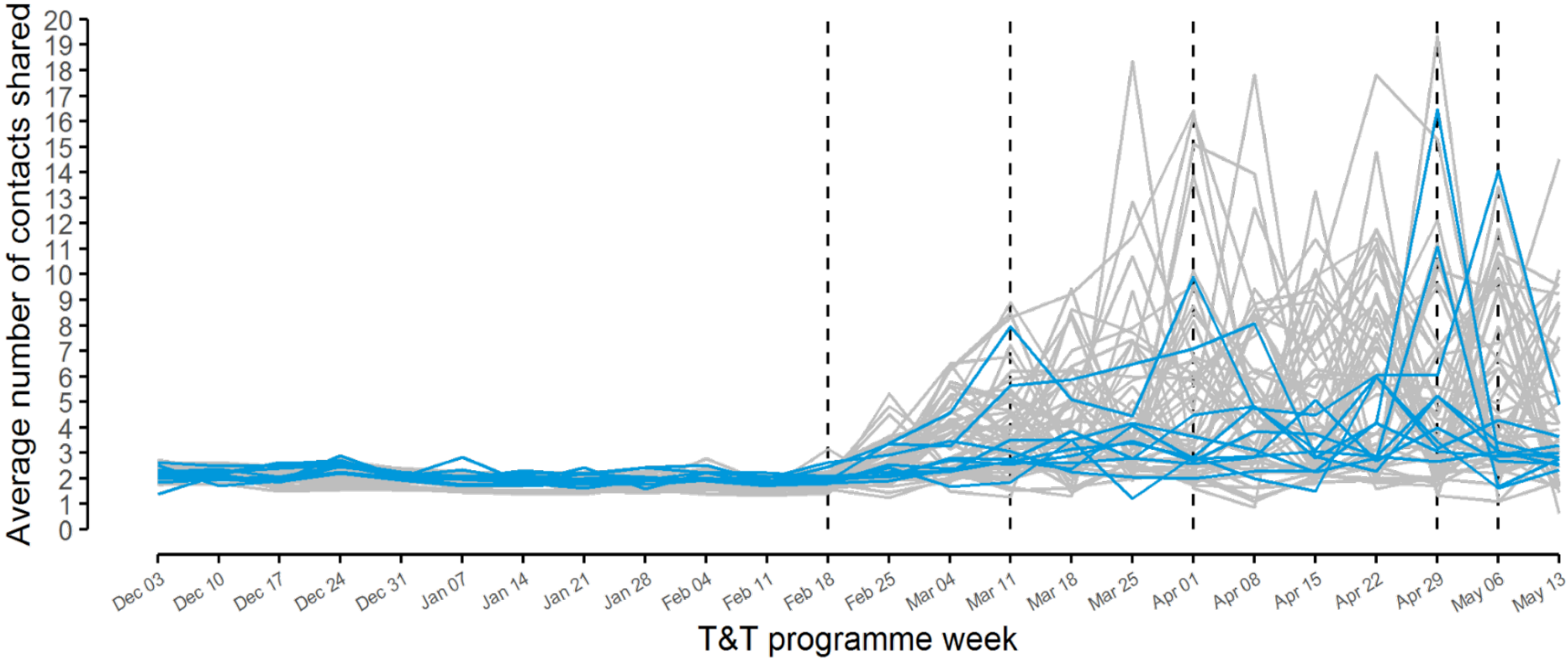


Figure 17. Percentage of cases sharing at least one contact during the pre-intervention period for treated local authorities (blue) and donor unit local authorities (grey)

Vertical dashed lines indicate key timepoints selected for inclusion as matching covariates (see the Matching variables section).

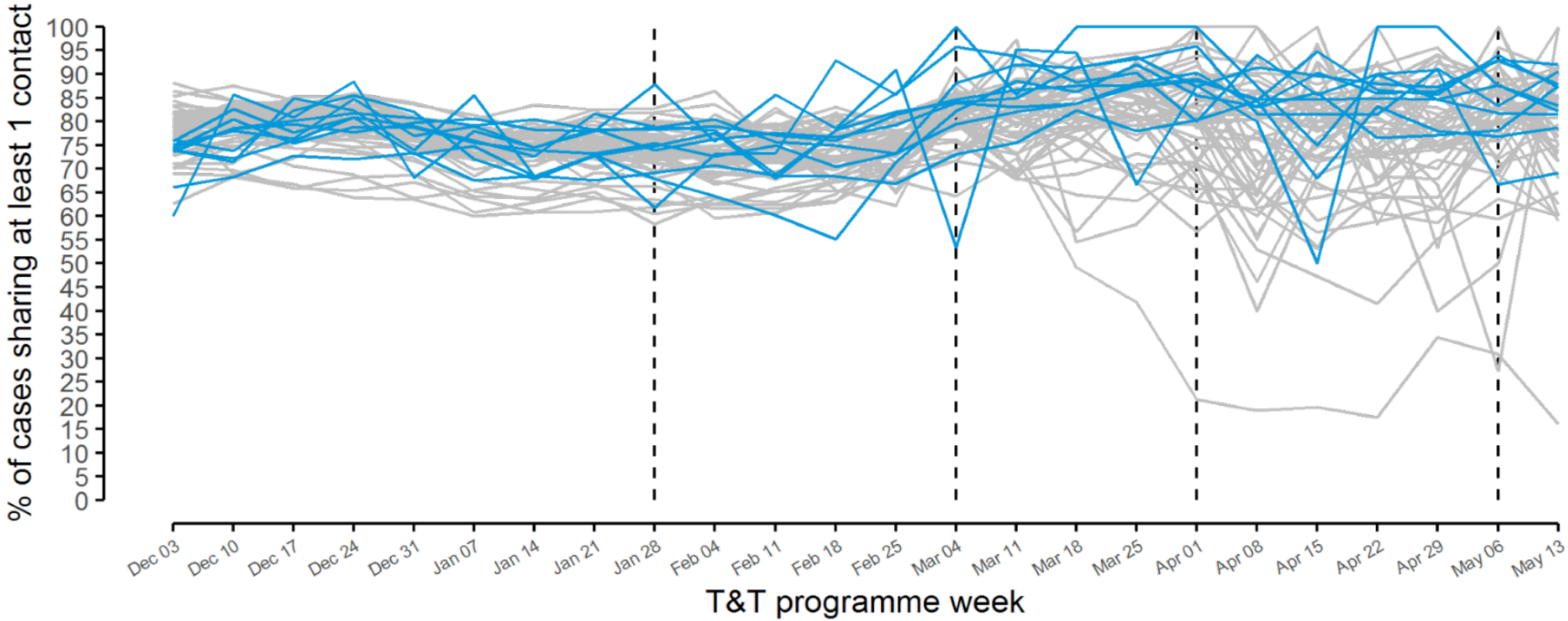


Figure 18. COVID-19 case rate (% positive PCR tests) during the pre-intervention period for treated local authorities (blue) and donor unit local authorities (grey)

Vertical dashed lines indicate key timepoints selected for inclusion as matching covariates (see the Matching variables section).

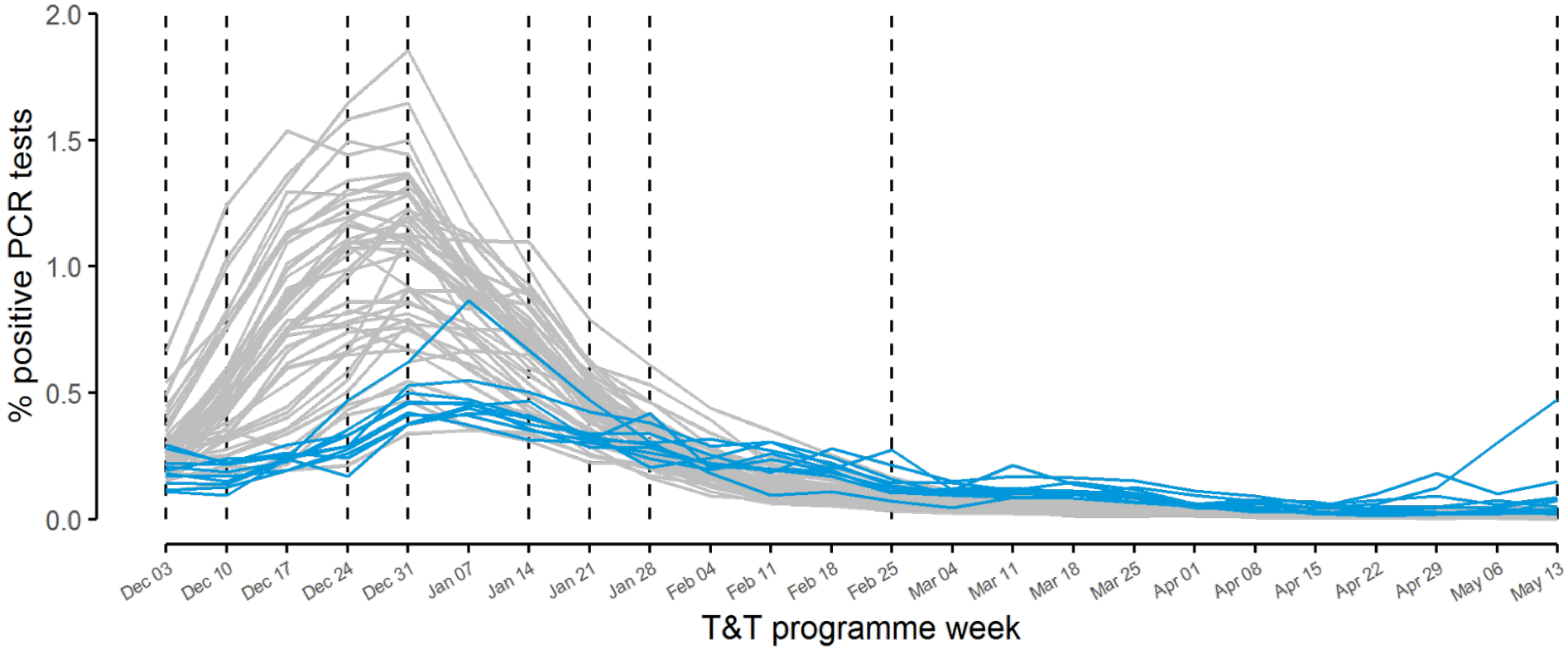


Figure 19. Left panel: Cumulative full (2 doses) vaccination rate at end of pre-intervention period (T&T week 13 May 2021) for treated local authorities (blue) and donor unit local authorities (grey)

Right panel: IMD average score (2019) for treated local authorities (blue) and donor unit local authorities (grey).

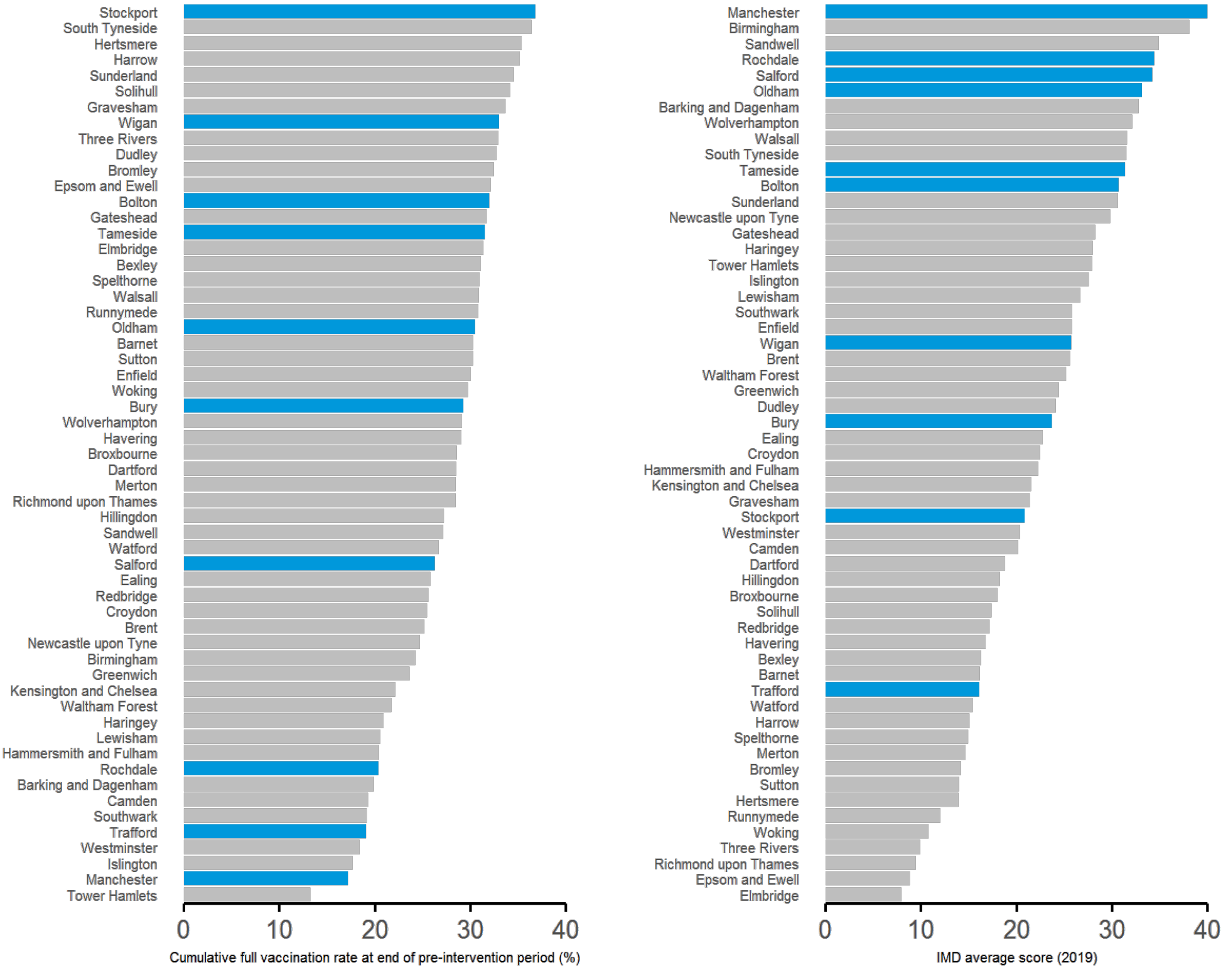


Figure 20. Age group breakdown as a proportion of total population for treated local authorities (highlighted in black) and donor unit local authorities

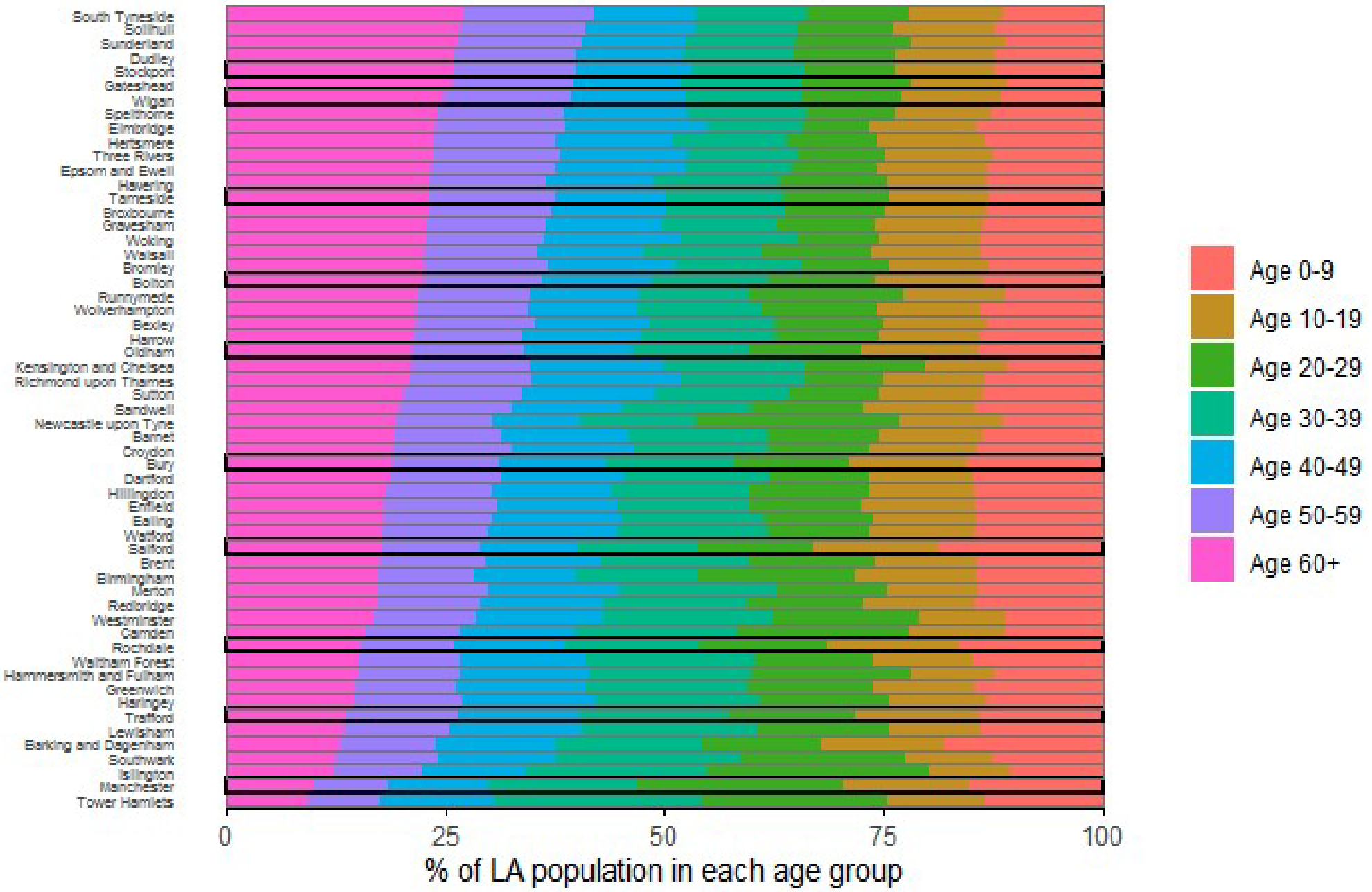
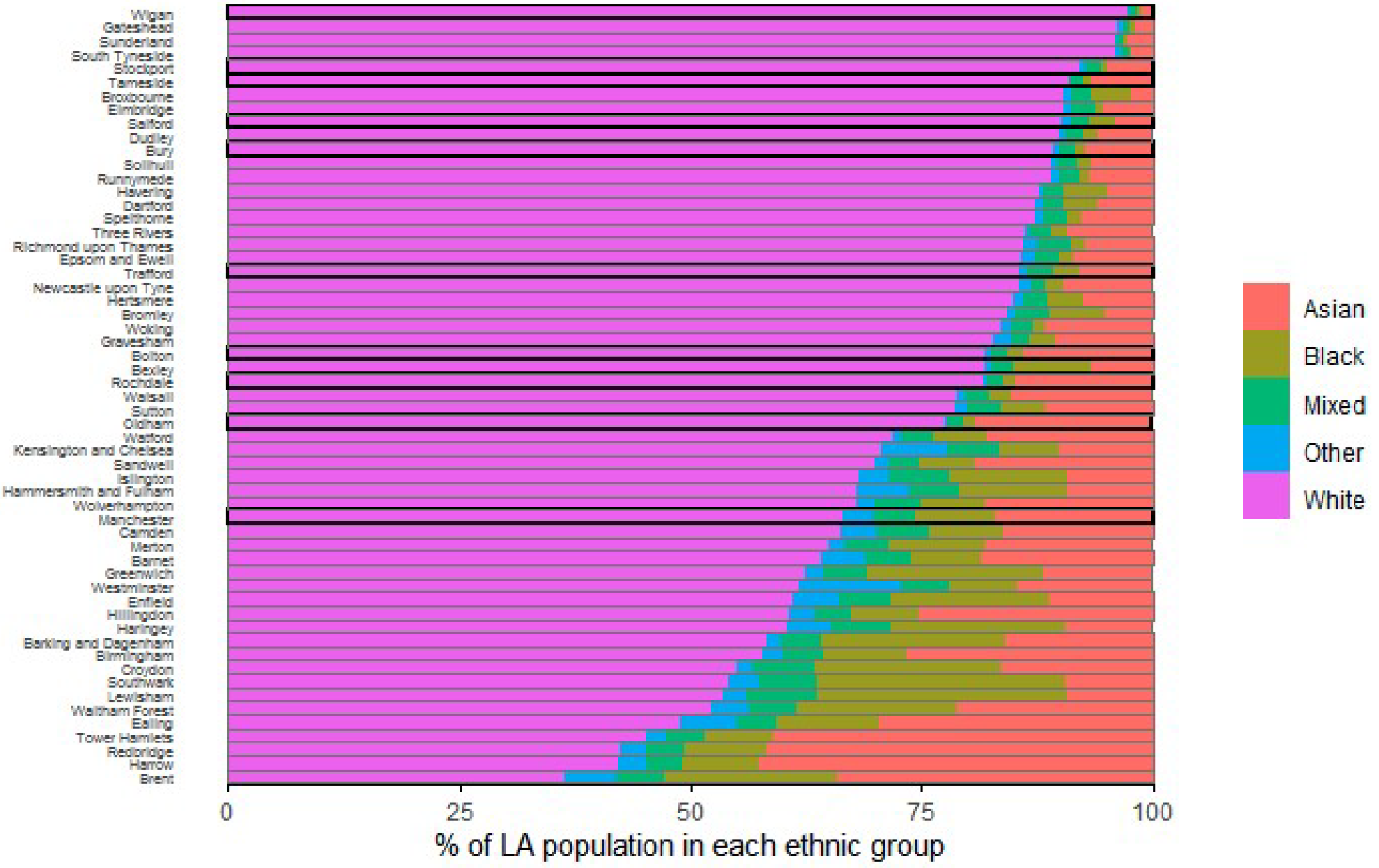


Figure 21. Ethnic group breakdown as a proportion of total population for treated local authorities (highlighted in black) and donor unit local authorities



About the UK Health Security Agency

UKHSA is responsible for protecting every member of every community from the impact of infectious diseases, chemical, biological, radiological and nuclear incidents and other health threats. We provide intellectual, scientific and operational leadership at national and local level, as well as on the global stage, to make the nation health secure.

UKHSA is an executive agency, sponsored by the Department of Health and Social Care.

Published: October 2023
Publishing reference: GOV-15038

UKHSA supports the UN Sustainable Development Goals