PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Evaluating the impact of using mobile vaccination units to increase COVID-19 vaccination uptake in Cheshire and Merseyside, UK: A synthetic control analysis.
AUTHORS	Zhang, Xingna; Tulloch, John; Knott, Shane; Allison, Rachel; Parvulescu, Paula; Buchan, Iain; Garcia-Finana, Marta; Piroddi, Roberta; Green, Mark; Baird, Sophie; Barr, Ben

VERSION 1 – REVIEW

REVIEWER	Kayoko Shioda Yale University School of Public Health
REVIEW RETURNED	07-Mar-2023

GENERAL COMMENTS	This study evaluated the impact of mobile vaccination units (e.g., vaccine buses or pop-up clinics) on the weekly number of first-dose vaccines administered among people aged 18+ years in the Northwest of England in 2021. The authors conducted two main analyses (the synthetic control analysis and the weighted Poisson regression models), generating robust estimates of the intervention impact controlled for both community-level and individual-level characteristics. The impact of mobile vaccination units had not been thoroughly evaluated, and this study provides important evidence on this topic. The authors found that, while mobile vaccination units successfully increased vaccine uptake, the effect was smaller among older adults, Asian and Black ethnic groups, and most socioeconomically deprived populations, providing important policy implications. I only had a few comments. First, I was wondering how the authors determined the follow-up period (three weeks). Was this because vaccine buses were usually implemented for three weeks on average? Sorry if I missed it but I could not find the information on how long the buses and pop-up clinics ran. Second, for reproducibility, it would be great if the authors could post their codes online. Third, for the synthetic control analysis, I usually recommend a sensitivity analysis where the authors remove the most heavily-weighted control from the model and evaluate the intervention impact again, which will allow us to see how robust these estimates are. However, as the authors have already done other alternative methods to support their findings (e.g., individual-level Poisson model), I do not think that is necessary.

REVIEWER	Gary W Reinbold University of Illinois at Springfield
REVIEW RETURNED	17-May-2023

GENERAL COMMENTS	The authors have an interesting research question and a very good source of individual-level data and seem to be using appropriate methods and reaching appropriate conclusions. However, they need to explain their data and research methods a little more carefully and interpret their results a little more explicitly for their readers. In particular, it seems that only the aggregate effect is presented in absolute terms and all of the subgroup effects are presented only in relative terms compared to reference subgroups, which makes it difficult to understand the size of the absolute effect for each subgroup. Page 9, line 9: Do the authors mean “causal” instead of “casual”? If they do, that seems a little strong for quasi-experimental analysis. Page 9, lines 17-21: Couldn’t exclusion of these individuals also lead to overestimation of the effect? Page 9, lines 22-26: Isn’t that limitation true regardless of the length of the follow-up period? One can never know the unobserved counterfactual. Page 10, line 16: Omicron isn’t really an emerging strain anymore. Page 10, line 21: Again, wasn’t this strategy implemented at the end of 2021. That hardly seems “current.” Page 10, lines 52-57: I would think there would be a very large body of literature analyzing interventions to increase various types of vaccinations. If I were the authors, I would limit my literature review to (1) interventions to increase COVID vaccination rates, and (2) the use of mobile vaccination units to increase vaccination rates for any vaccines (although there apparently are no such studies in this second category). Page 11, lines 7-13: The first two sentences of this paragraph seem unnecessary. The authors had just stated that there are no prior studies of the impact of mobile vaccination units, so the effects of those units are necessarily unknown. Page 11, lines 26-30: This issue first comes up here in the body of the paper, so I’ll comment on it here. The fact that the study population is limited to people who are registered with a GP seems potentially significant. What percentage of the population in this area is registered with a GP? Do they differ in any important way from the people who are not registered? Page 11, lines 26-46: How often are these data updated? The discussion later implies that the data are updated weekly, but if that is true, on which day of the week are they updated? Did the authors match up the dates of data updating with the exact intervention dates to clearly identify pre- and post-intervention data? Page 12, lines 10-25: Can the authors please explain more about the timing and duration of these interventions? How long were the units at each site? How did you handle sites that were visited twice? Page 12, lines 27-31: Was this method also used to identify treatment and control groups for the individual analysis? If so, couldn’t that result in some individuals in the treatment group living further from a mobile unit than some individuals in the control group? Page 12, lines 43-45: Why did the authors decide to use seven weeks of preintervention data? Page 12, lines 45-47: Why were outcomes measured up to three weeks after the intervention? Wouldn’t any vaccinations administered by the units have been recorded in the first week? Page 12, lines 55-57: I understand that the authors matched on the weekly number of new first doses for the seven weeks prior to the intervention. Did they also match on the total number of pre-intervention first doses? If so, why? Page 13, lines 3-12: Robbins et al. developed their approach for a situation with multiple outcomes and cautioned that “failure to

	include a robust set of outcomes may result in omitted variable biases." Is this issue a potential concern here? Did the authors consider alternate approaches designed for single outcomes that construct a separate synthetic control for each treated unit? Page 13, lines 14-17: Do the authors believe that they were able to match on all important predictors of the outcome? Some studies indicate that other variables such as education, income, employment status, health status, and the presence of children in the household were important predictors of COVID vaccination uptake at the individual level. Did the authors assess the predictive power of their predictors on the preintervention outcomes? Page 13, lines 43-48: Was the Poisson regression analysis also limited to people registered with a GP? Page 13, lines 43-60: Did the authors feel that they were able to include all important potential confounders in the Poisson regression analysis? (I assumed that the reason they started with a synthetic control analysis instead of a multivariate regression analysis is because individual-level data weren't available for all necessary control variables.) Did they consider using a difference-in-differences analysis as a robustness check in addition to or instead of the Poisson regression analysis? Page 16, lines 30-39: The authors should better explain Table A2. Why is the point estimate for the mobile vaccination units 1.68 as compared with 1.23 in Table 2? Which groups actually have increased vaccination rates with the mobile vaccination units? The authors say that they found lower impacts for some groups, but were the impacts still positive? The authors might specifically interpret the 0.64 number for the interaction term for 30-65 year olds with mobile vaccination units and explain how that number can be used to derive the absolute effect of the intervention on that age group (which seems more important than the relative effect of the intervention on that age group as compared with the effect on the 18-30 year old age group). Page 16, lines 43-59: Similarly, the authors should better explain Table 3. They might specifically interpret the first number in the table as an example. The authors state that the table shows increased uptake for all groups in areas with the mobile vaccination units, but that isn't obvious to me from the table.

REVIEWER	Philip Britteon The University of Manchester
REVIEW RETURNED	25-May-2023

GENERAL COMMENTS	Title: Evaluating the impact of using mobile vaccination units to increase COVID-19 vaccination uptake: A synthetic control analysis Journal: BMJ Open Manuscript ID: bmjopen-2023-071852 Summary The authors use the synthetic control method to assess the impact of mobile vaccination units on vaccination uptake during the COVID-19 pandemic, by exploiting data on the geographic deployment of the mobile units in the North-West of England. They estimate the impact of mobile vaccination units on the vaccination rates of the population in immediate area around each mobile unit and assess how this effect differed between different subgroups within the targeted population (by age, income deprivation, ethnicity). The results suggest mobile units increased vaccination rates in the targeted areas. However, this effect differed across subgroups and

	intersections of the population in each area. In particular, the impact of the intervention was smaller for Asian and Black ethnic groups, the older population, and in the most deprived targeted areas. The paper is well written and addresses a well-motivated research question using a policy evaluation method that is appropriate for the study. My comments highlight limitations that could be addressed to improve the policy recommendations that can be drawn from the study, to strengthen confidence in the underlying assumptions of the model, and to refine the motivation and discussion of the paper. Comments (1) Exploiting distance from mobile vaccination unit The authors use a 1km radius around each mobile unit to define the intervention areas. This allows the authors to estimate the targeted effect of the intervention on those living closest to each mobile vaccination unit. However, I think it is important to also investigate whether there was a "dose" effect of the policy as the distance from mobile unit increased in order to strengthen the findings of the study. As a simple test, this could be done reestimating separate synthetic control models on two or more treatment zones (e.g. less than 1km, 1km-3km) and using the other wider regions as controls. This test would serve three purposes. First, it would allow the authors to check whether they were underestimating the true impact of the reform by restricting the intervention group to a small surrounding area. Second, it would allow the author to test the hypothesis that distance matters (i.e. did the effect dimmish and over what distance). Third, it could also be interpreted as a robustness check when testing the assumptions of the model (see point on spillovers below). Depending on how the treatment areas are defined, the authors should also test for spillover effects by excluding surrounding/adjacent areas to the intervention areas from the pool of potential control areas in a sensitivity test. Spillover effects from intervention to non-intervention areas in the synthetic control group could bias the estimates of the model (most likely downwards, if the mobile units also attracted people outside the 1km boundary). The authors could also test whether their results were sensitive how they defined distance in the analysis. The authors identify distance from each mobile unit based on geographical distance (less than 1km), but travel time could be a better measure in this instance if data on this is available. As the authors have data on travel time to nearest static clinic they would then be able to define the full intervention population as being equal to the population whose travel time to the mobile clinic was less than their travel time to static clinic. This would allow the authors to capture the full impact of the intervention (dependent on data availability, of course!). (2) Extending the follow-up period I was also wondering why the authors chose to restrict the follow-up period to only 3 weeks. This should be justified in the methods section. If data is available for a longer follow-up period, then this could then be used to exploit the "switch off" of the treatment effect to begin to unpick whether the mobile units attracted people who

	would otherwise not have been vaccinated - a current limitation of the study and an important policy question. The question would then be: did the vaccination rate return to the same or a lower level than the control group after the intervention? However, this would require information on how long each unit was deployed in each area which may not be available to the authors. Extending the follow-up period until the time that the mobile unit was removed would also allow the authors to test whether vaccinations diminished during their tenancy, allowing the authors to infer policy recommendations on the length of deployment. (3) Exploiting the timing of deployment of each mobile unit The authors could also consider whether the timing of the deployment of each mobile unit affected the treatment effect by stratifying the analysis based on whether the unit was deployed in the first or second half of their sample. This would help to understand whether differences in the severity of the pandemic and the stage of the rollout of vaccination program affected the results to help draw conclusions on the applicability of their findings in different settings. It was also unclear how the authors dealt with the fact that mobile units appeared to have been deployed to the same areas more than once – “visited 37 sites on 54 occasions”. Does the estimated effect capture the impact on first time deployments only or does it include the impact of second, third time visits as well? If the latter, then this could be explored further in the paper by looking at whether the impact of the intervention reduced in subsequent visits. (4) Defining the synthetic control group in the inequalities analysis The individual level analysis allows the authors to easily stratify their analysis by different population characteristics. This inequalities analysis is a key part of the paper. However, I have questions about how the authors define the control group in these subgroup analyses. Do the authors use the same synthetic control group as in the main analysis? I think it would be more informative to identify new synthetic control groups for each subgroup so that you are comparing like-for-like across the intervention and non-intervention areas. This dawned on me when looking at the results in Figure 3 where the authors find an adverse impact of mobile vaccination units on certain subgroups despite there being no theoretical reason why mobile units should decrease vaccination uptake. These negative findings suggest that other unobserved differences over time might be influencing the results (e.g. 30-65 year old, black men and women, in the most deprived intervention areas had different underlying trends to the same demographic in the synthetic control group from the main analysis). Estimating a new synthetic control group for each subgroup would provide an arguably more meaningful comparison (providing that there is sufficient overlap to estimate the synthetic control method for each subgroup). If not, then the authors should at least present the trends for each sub-group in the intervention and non-intervention areas to allow the reader to assess whether they believe that the control group is a good comparison. (5) Investigating differences in deployment strategies

	The authors could also investigate whether there were differences in treatment effects between local authorities given that the decision to deploy units was made at this level. However, I would only suggest testing this if it is possible for the authors to meaningfully interpret the output from these results (i.e. if they are able to distinguish between the different approaches taken by each locality when deciding where to deploy each mobile facility). Alternatively, it might be possible to infer information on the deployment strategies of each mobile unit from the data (i.e. look at differences in the area level characteristics of the areas where each unit). The authors could then stratify the analysis accordingly to draw further policy recommendations on where decision makers might want to choose to deploy vaccination units in the future to achieve different aims (e.g. increase total vaccination rate, reduce vaccination inequalities). This is subtly different from stratifying the analysis based on individual level characteristics. However, I recognise that this could form a new paper in itself. (6) Defining the outcome variable It would be helpful if the authors clearly stated the outcome variable used in the study in the data/methods section and in the tables/figures (i.e. how they define and assess “vaccination uptake”). This was not initially clear when reading through the paper. (7) Adjusting for ceiling effects The authors should also justify why they chose to estimate the impact of the intervention on weekly first dose vaccinations rather than the total first dose vaccinations. There are advantages to using the cumulative total vaccination rate when constructing the synthetic control group. Namely, using total vaccinations allows one to construct a control group that is more likely to follow the same path in the absence of the intervention when ceiling effects come into play (e.g. if people are less and less forthcoming to be vaccinated as the vaccination rate increases). This potential issue is highlighted in Table 1, which highlights that the first dose vaccine uptake rate was higher in the non-intervention areas (69%) than in the intervention areas (53%). It could therefore be the case that the slight decline in the weekly vaccination rate in the control group (shown in Figure 2) may have been driven by a ceiling effect that would not have occurred in the intervention areas, raising concerns that the observed treatment effect is in fact an overestimate of the true effect of the intervention. Using total first rate vaccinations as the outcome to construct the synthetic control group would overcome this limitation. There is also the issue that when comparing effects across subgroups in Figure 3 that differences in the impact of the intervention could also be driven differences in the underlying vaccination rate and ceiling effects. It would be helpful to include descriptive statistics (either trends over time in a figure or averages in a table) of the total vaccination rate by subgroup. (8) Including additional descriptive statistics I would have also liked to see descriptive statistics on the full non-intervention population, not just those included in the synthetic control group in Table 1. This would help the reader to infer

	information on how the intervention areas differed from the full population (i.e. how local authorities may have made decisions on where to deploy the mobile units). (9) Defining the start date The authors should also justify their choice of start date (7 weeks prior to the deployment of the mobile unit). The synthetic control method performs better with a longer pre-policy period so unless there is a reason for using this cut-off point (which I suspect there might be) then a longer pre-policy period should be used. (10) Weighting by population size The authors do not weight by population size in model 1. I was confused by their explanation as to why this was the case: “a weighted regression model to estimate this effect would potentially underestimate standard errors and p-values as this would not account for complex aspects of the process used to generate the synthetic control”. This needs to be justified more clearly. It is maybe also worth noting that the lack of population weights in the main analysis is unlikely to matter as LSOAs are roughly the same size. (11) Motivating the study It would have been nice to see more discussion in the introduction on why vaccination uptake is lower among young people, socioeconomic disadvantaged, and ethnic minorities. Do the studies referenced in the introduction infer or hypothesise why this is the case? This is important for understanding why the impact of the policy may have worked / not worked amongst different intersections of the population. The authors reference studies that investigated the impact of different types of interventions to increase vaccination uptake in the introduction. Did these studies also investigate inequalities in impact? This should be discussed in the final paragraph of the introduction if they do. If not, I think that the authors perhaps undersell their contribution to the literature on inequalities in vaccination uptake. This is an important part of the study. In particular, their approach to investigate inequalities in impact between intersections of the population (including ethnicity) is novel. I would emphasise this in the introduction and the discussion in addition to their other main contributions. Are there studies that look at impact of mobile units used to administer other forms of preventative care (e.g. cancer screening)? The authors should consider referencing these studies in the introduction if they exist as the mechanisms through which these units work would be similar (e.g. reduce time/cost for the patient to increase uptake). This could then be discussed in the discussion section when talking about the policy implications of the study (i.e. do the results suggest anything about the use of mobile units in other areas). (12) Information on the intervention The authors could have included more information on the details of the intervention. I have suggested some of the questions that came into my head when reading the paper below. I don’t think that all of

	the questions need to be answered given the word limit of the journal but more information would have definitely been helpful. • How long on average were the mobile units situated in each location? Did this differ between units? What is the timeline of units being deployed between first and last unit (i.e. were the majority of units deployed at the same time or were they spread out)? • Did the intervention include other types of intervention in addition to the mobile units (e.g. outreach programs)? The authors mention about not being able to distinguish between the two as a limitation in the discussion but do not mention this aspect of the program in the intervention section. • What classifies as a static vaccination site? Were they also setup during COVID and during the study period? If yes, the introduction of static sites may have biased the results of the study and should be addressed in the design of the analysis. • How severe was the pandemic throughout the study period and did this differ during the study period? How many people had already been vaccinated prior to the intervention? Which ages were eligible for a vaccine and did this change throughout the study period? Did other national interventions/rule changes occur throughout this period? Were static sites also available as walk-ins? These questions are important in understanding how easy it was to get vaccination in order to explain differences in impact between subgroups and areas (i.e. were the mobile sites attractive because they offered walk-ins, were open to all ages). (12) Discussing the study I feel that the authors could include a paragraph summarising the strengths of the paper before the limitations section in the discussion. In particular, they could highlight their contributions to the literature on vaccination uptake: (1) investigating impact of mobile units, and (2) investigating inequalities in impact of a vaccine intervention. Individual level data on ethnicity is often unavailable to researchers wanting to investigate inequalities in the impact of interventions too. It would be nice to include some discussion on the external validity of the findings and their policy recommendations. Do the authors think that the results would be applicable to other types of vaccine or care and at different points in time (i.e. outside of a pandemic)? i.e. what are the mechanisms driving the change? They mention about the mobile vaccination unit may have increased uptake due to improved accessibility through the mobile vaccination unit itself and/or the awareness raising and publicity associated with visits of these units but are there also other mechanisms? E.g. did static units also offer walk-ins, were more people eligible to attend the mobile units? (13) Other minor comments Why did you choose to combine information on an individual's ethnicity from 17 to 5 categories? Was this necessary given the large sample size of the data used in the study? Is it possible to include the location of static sites as well in Figure 3?

	It would be helpful to explain the red dots in key/legend in Figure 3. Please report figures in Table 1 to more decimal places to allow for comparisons between the intervention and non-intervention areas. In the results section, the authors write: “This overall effect size estimates 3723 additional vaccinations over three weeks of follow-up across the study area, similar to the actual number of people vaccinated in the mobile units (n=3824). This suggests that at least in the short term most of the vaccinations in the mobile units were additional. In other words, among the mobile units’ users there were few people, who would have otherwise gone to static centres to get their vaccine.” This should be moved to the discussion when interpreting the results. It is an important interpretation. The authors should justify why they run a sensitivity test excluding ethnicity from the set of control variables or drop it from the analysis. I cannot see why this particular sensitivity test was needed. The first paragraph of the discussion is not 100% clear in how it is worded. It might help to first discuss the overall impact of the policy and in a separate paragraph to discuss the results by subgroup. The authors also note that interpretation of Figure 3 should be interpreted with caution due to small sample sizes within each of the 45 subgroups but do not provide figures on these sample sizes for the reader to inform their own judgement.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1 **Dr. Kayoko Shioda, Yale University School of Public Health**	**Recommendation**	This study evaluated the impact of mobile vaccination units (e.g., vaccine buses or pop-up clinics) on the weekly number of first-dose vaccines administered among people aged 18+ years in the Northwest of England in 2021. The authors conducted two main analyses (the synthetic control analysis and the weighted Poisson regression models), generating robust estimates of the intervention impact controlled for both community-level and individual-level characteristics. The impact of mobile vaccination units had not been thoroughly evaluated, and this study provides important evidence on this topic. The authors found that, while mobile vaccination units successfully	Thank you very much. We really appreciate the encouraging comments.

		increased vaccine uptake, the effect was smaller among older adults, Asian and Black ethnic groups, and most socioeconomically deprived populations, providing important policy implications.	
	Comments	1. I was wondering how the authors determined the follow-up period (three weeks). Was this because vaccine buses were usually implemented for three weeks on average? Sorry if I missed it but I could not find the information on how long the buses and pop-up clinics ran.	Thank you for pointing this out. There are two main reasons for our choice of the follow-up period: 1. Practicality, feasibility and resource constraints. Our priority in conducting this study was to provide a rapid response during the pandemic on the effectiveness of this approach to inform its commissioning in the region. The study therefore was resourced for a limited follow up period in order to produce results within a time frame that was useful for decision makers. This does reflect the reality of an urgent public health response and the best use of available data to inform policies and tactics. Also although we have made all the effort to collect as much data as possible, we struggled to gather sufficient data for July and August 2021 for most places. That was because places started prioritising the administration of the second-dose vaccines on mobile vaccination units then, and the number of people receiving the first-dose vaccines on mobile vaccination units had dropped and as such records were incomplete. Some external factors may have caused that, such as the availability of resources (including funding, personnel, and logistical support) among the local team to collect such data in a complete manner, when they were struggling to meet the urgent healthcare needs of communities. Longer follow-up periods

			would have required more resources and funding to collect the needed data, which unfortunately wasn't available back then. 2. The nature of the intervention and our study objectives. We aimed to measure the short-term impact of mobile vaccination units using weekly number of first-dose vaccines administered among adults as our outcome. Our observation period spanned from seven weeks (22nd February 2021) before the first mobile vaccination unit visit on 12th April 2021 to three weeks (19th July 2021) after the last mobile vaccination unit visit on 28th June 2021, during which the first-dose vaccines were prioritised. Since July 2021, many places gradually switched to prioritising the administration of the second-dose vaccines. We found that 3 weeks was sufficient to capture the relevant data needed to address our research question – measuring the potential immediate effect of the intervention upon the weekly number of first-dose vaccines administered among adults, without much spill-over effect of prioritising the second-dose vaccines since July 2021. To summarise, we aimed to measure the immediate or short-term effect of the mobile vaccination units, which was a substantial intervention (54 visits during our study period). We didn't strive to measure long-term outcomes; neither was that practical or feasible due to resource constraints. We therefore chose our endpoint and the follow-up period as such to facilitate our investigation. We have added such explanation as outlined above in Methods in Lines 17-21 on Page 7, and more details on the deployment of the mobile vaccination units in Lines 14-24 on

			Page 6.
		2. For reproducibility, it would be great if the authors could post their codes online.	Thank you for your advice. We have uploaded our R scripts at our Civic Data Cooperative GitHub public repository: https://github.com/civicdatacoop/mobileVaxUnit.
		3. For the synthetic control analysis, I usually recommend a sensitivity analysis where the authors remove the most heavily-weighted control from the model and evaluate the intervention impact again, which will allow us to see how robust these estimates are. However, as the authors have already done other alternative methods to support their findings (e.g., individual-level Poisson model), I do not think that is necessary.	Thank you for your thoughtful advice. We appreciate your valuable feedback. After careful consideration of feedback of all reviewers, we have decided to conducted a series of sensitivity tests: apart from the sensitivity test of excluding ethnicity (Appendix 8) and the two pop-up sites (Appendix 9) from the analysis to assess the potential impact of the missing data and two pop-up static sites on our results and conclusions respectively, we additionally tested the dose effect of distance between mobile vaccination units and population weighted centroids of LSOAs in constructing the intervention and non-intervention groups (Appendix 10), the potential spatial spill-over effect of the intervention (Appendix 11), and survival analyses (Appendix 12) to further check the robustness of results of the synthetic control method and the weighted Poisson model.
Reviewer 2 **Dr. Gary W Reinbold, University of Illinois at Springfield**	**Recommendation**	The authors have an interesting research question and a very good source of individual-level data and seem to be using appropriate methods and reaching appropriate conclusions. However, they need to explain their data and research methods a little more carefully and interpret their results a little more explicitly for their readers. In particular, it seems that only	Thank you for your thoughtful review and for your kind consideration of our study. We appreciate your valuable feedback and have carefully considered and incorporated your suggestion regarding additional explanation on the data, research methods the interpretation of the results – particularly the subgroup analysis, as well as reaching appropriate conclusions. We understand your concerns and have responded accordingly throughout the manuscript.

		the aggregate effect is presented in absolute terms and all of the subgroup effects are presented only in relative terms compared to reference subgroups, which makes it difficult to understand the size of the absolute effect for each subgroup.	
	Comments	1. Page 9, line 9: Do the authors mean “causal” instead of “casual”? If they do, that seems a little strong for quasi-experimental analysis.	Apologies for the typos. We have corrected it accordingly in Line 4 on Page 3 of the manuscript. We understand your concerns regarding drawing definitive causal inferences from quasi-experimental analysis, with challenges arising from potential pre-existing differences among participants before the intervention, and the non-random assignment of intervention and control groups. Although we have carefully chosen the synthetic control method to mitigate the potential biases and confounding effects mentioned above to address limitations of quasi-experimental analysis and strengthen causal inferences, we acknowledge that there are other limitations, uncertainties, or alternative explanations that could be explored. We therefore have now reworded the sentence to capture a more balanced and scientifically responsible interpretation of our findings in Line 4 of Page 3 of the manuscript.
		2. Page 9, lines 17-21: Couldn’t exclusion of these individuals also lead to overestimation of the effect?	Our thinking was that the mobile vaccination units might be expected to have had a larger effect amongst this group (unregistered with the GP) as they would not have had access to alternative sources of vaccination, i.e., through the GP practice, through which most vaccinations were taking place. Excluding them would then reduce the overall intervention effect. To be clear we do not have data on these individuals so couldn’t include them. However, it is true that it is possible that it could also have led to

			an overestimation if the effect was lower in this group (unregistered with the GP). It is unlikely this would have had a large impact on our result as there were probably quite a small proportion of such people. This was indicated by our results with the overall effect size being 3723 additional vaccinations only 3% smaller than the recorded number of people vaccinated in the mobile units (n=3824). We have reworded this to make it clearer in Lines 10-12 on Page 3.
		3. Page 9, lines 22-26: Isn't that limitation true regardless of the length of the follow-up period? One can never know the unobserved counterfactual.	Thank you for raising an important point regarding the limitations of any follow-up period and the challenge of knowing the unobserved counterfactual. It is true that in any study, regardless of the follow-up period's duration, it is impossible to directly observe the unobserved counterfactual or the outcome that would have occurred had a different intervention or condition been applied. This inherent limitation is a fundamental challenge in quasi-experimental design or causal inference research. Our point, however, was not about being able to observe the counterfactual, but rather being able to distinguish between the intervention bringing forward vaccinations by a few weeks versus the intervention leading to new people being vaccinated that would have never been vaccinated. The length of the follow-up period does have implications for our ability to distinguish between these two types of effects. A longer follow up would have enabled assessment of whether the effect on vaccination rates was sustained over the long term. We have rephrased this sentence to clarify this point in Lines 13-14 on

			Page 3 to: “As our analysis covers a relatively short follow-up period, the results may not reflect the sustained impact of the intervention over longer periods of time”.
		4. Page 10, line 16: Omicron isn’t really an emerging strain anymore.	Thank you for pointing this out. Changed to reflect the current context in Line 11 on Page 4.
		5. Page 10, line 21: Again, wasn’t this strategy implemented at the end of 2021. That hardly seems “current.”	Thank you for pointing this out. We have changed the text accordingly in Lines 14-16 on Page 4.
		6. Page 10, lines 52-57: I would think there would be a very large body of literature analyzing interventions to increase various types of vaccinations. If I were the authors, I would limit my literature review to (1) interventions to increase COVID vaccination rates, and (2) the use of mobile vaccination units to increase vaccination rates for any vaccines (although there apparently are no such studies in this second category).	Thank you for raising an important point regarding our summary of the relevance evidence. We appreciate your insight, and we agree that there is a large body of literature analysing interventions to increase vaccine uptakes. Please allow us to elaborate on how we made our choices in conducting the literature review. Because of the quasi-experimental design of our study, we have decided to limit our literature review to quasi-experimental and randomised control trials of interventions aiming to increase immunisation rates for COVID-19, Influenza, and pneumococcal vaccination, as these target similar population groups. We have also included studies focusing specifically on mobile vaccination units for these diseases, regardless of study designs. We therefore believe that our choices of the literature are broadly in line with your recommendation on conducting the literature search. We have also updated the text to reflect the latest emerging evidence published since we first wrote the paper in Lines 1-4 on Page 5.
		7. Page 11, lines 7-13: The first two sentences of this paragraph seem unnecessary. The authors had just stated that there are	Thank you for your feedback regarding the first two sentences of the paragraph. We appreciate your attention to detail. Please allow us to explain the rationale behind including

		no prior studies of the impact of mobile vaccination units, so the effects of those units are necessarily unknown.	those sentences and address your concern. While it is true that we mentioned the absence of prior studies on quantifying the intervention effect of mobile vaccination units, we believe that the additional context provided by those sentences serves an important purpose. Those sentences are also in line with findings of prior studies on deploying COVID-19 vaccines through the mobile vaccination units in San Francisco (reference 18, Marquez C, Kerkhoff AD, Naso J, et al. A multi-component, community-based strategy to facilitate COVID-19 vaccine uptake among Latinx populations: From theory to practice. PLOS ONE 2021;16:e0257111. Doi:10/gn2bbs) and the Great Boston area (reference 19, Gupta PS, Mohareb AM, Valdes C, et al. Expanding COVID-19 vaccine access to underserved populations through implementation of mobile vaccination units. Prev Med 2022;163:107226. Doi:10.1016/j.ypmed.2022.107226). As a reflective analysis on how the UK vaccine rollout delivered success so far, one study mentioned that mobile vaccination units had been deployed to a small number of people living in remote, rural areas in the UK (reference 2, Baraniuk C. Covid-19: How the UK vaccine rollout delivered success, so far. BMJ 2021;372:n421. Doi:10.1136/bmj.n421). By explicitly stating that how the intervention effects of mobile vaccination units can be quantified is uncertain we emphasised the significance of our study in addressing this knowledge gap and generating new insights in the field. Furthermore, the inclusion of these sentences helps set the stage for the subsequent discussion of our research objectives and the importance of investigating the

			intervention effect of mobile vaccination units. It reinforces the need for empirical evidence to understand the potential benefits and challenges associated with their implementation. Mindful of the need for concision, we have revised these sentences in Lines 11-13 on Page 5 of the manuscript to maintain the necessary context as well as acknowledging evidence of existing knowledge, while ensuring the paragraph remains clear.
		8. Page 11, lines 26-30: This issue first comes up here in the body of the paper, so I'll comment on it here. The fact that the study population is limited to people who are registered with a GP seems potentially significant. What percentage of the population in this area is registered with a GP? Do they differ in any important way from the people who are not registered?	Thank you for raising this. There are no reliable data for numbers of UK residents not registered for primary care with the NHS, just sparse estimates from various surveys of hard-to-reach population and attempts to get them to register with a GP. However, the advantages of NHS registration mean that the vast majority of any UK population will be registered. Whilst we don't have access to data on differences between the registered and unregistered population, based on our knowledge we believe that it would be difficult for certain groups of people to register with a GP, such as undocumented immigrants, displaced people, and some travellers such as the Romani people. There were about 2.5 million usual residents in Cheshire and Merseyside based on the 2021 Census https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/populationandhouseholdestimatesenglandandwalescensus2021, whilst there were about 2.7 people registered with the GP in the same year https://www.cheshireandmerseyside.nhs.uk/. There are a few possible reasons, which aren't mutually exclusive https://commonslibrary.parliament.uk/

			population-estimates-gp-registers-why-the-difference/: over-counting in GP practice registers with typical delays in updating the record when people move addresses and change their registered GPs, under-counting in population estimates, and different definitions of who counts as 'resident' in the country. We have added a statement about this, and acknowledged this data limitation and how it affected the interpretation of our findings in Lines 8-10 of Page 13.
		9. Page 11, lines 26-46: How often are these data updated? The discussion later implies that the data are updated weekly, but if that is true, on which day of the week are they updated? Did the authors match up the dates of data updating with the exact intervention dates to clearly identify pre- and post-intervention data?	Thank you for raising an important point. These data were updated by the data providers irregularly during our study period. We were unable to match up the dates of data updating with the exact intervention dates. However, the time windows are sufficient for reasonably relating outcome to exposure. We have now added more information on this in Lines 19-20 on Page 13 of the manuscript.
		10. Page 12, lines 10-25: Can the authors please explain more about the timing and duration of these interventions? How long were the units at each site? How did you handle sites that were visited twice?	Thank you for your feedback. We have now added more information about the timing and duration of these interventions in Lines 14-24 on Page 6 of the manuscript. We chose to treat sites with more than one visits the same as the rest in our main analysis.
		11. Page 12, lines 27-31: Was this method also used to identify treatment and control groups for the individual analysis? If so, couldn't that result in some individuals in the treatment group living further from a mobile unit than some individuals in the control group?	Thank you for your feedback. Yes, we used the same approach for defining treatment and control groups for the individual analysis, as in the aggregate analysis. We used the distance between the population weighted centroids of each individual neighbourhood to the mobile vaccination units to identify if they were in the intervention group. They were assigned to the intervention group if this distance was equal to or less than 1km. We use the centroid of their neighbourhoods as we did not have access to the exact location of

			their residence for confidentially purposes. This could mean that a small number of individuals in the control group could live closer to a mobile unit –than those of the intervention group. Such cases are most likely to emerge in some adjacent intervention and non-intervention LSOAs, where the centroid of the non-intervention LSOA was farther than 1km to the nearest mobile vaccination unit, but there are some residences bordering the intervention LSOA at a distance within the distance radius of 1km to the nearest vaccination unit. The average distance, however, was very different – Figure A8 in Appendix 5 shows this – giving the weighted distribution of intervention and controls. This would have made our effect estimate more conservative. We have added a note to the effect in discussing the limitation of our research in Lines 23-24 on Page 13. As a robustness test of these potential spill over effects, we have conducted additional analysis excluding LSOAs with population weighted centroids located at a distance between 1km and 1.5km from the nearest mobile vaccination unit in Appendix 11. Although result of this sensitivity test is very similar to that of the main analysis, as we expected it is indeed slightly larger than that of the main analysis.
		12. Page 12, lines 43-45: Why did the authors decide to use seven weeks of preintervention data?	We have now added more explanation on this in Lines 12-17 on Page 7 of the manuscript, in addition to the new visual evidence presented in Appendix 2. We choose seven weeks of preintervention data, primarily because it's when we started to observe signs of slowing-down of growth in crude uptake overall and expanding variations in the uptake of the first dose across LSOAs (Figure A1 and A2 in

			Appendix 2), but with widening gaps among ethnic and socio-economic groups in the local population (Figure A3-A6 in Appendix 2). This choice was also motivated by the rollout of the vaccine prioritised by age groups, which expanded to all people aged 65 and over in the second half of February 2021 (see Appendix 1) whilst the vast majority of the adult population aged below 65. The 7 week's pre-intervention period commencing on 22nd February 2021 could therefore allow us to capture trends for the majority of the population eligible for the vaccine, whilst avoiding potential influences from prior changes in the severity of the pandemic and policy changes in COVID-19 local restriction measures and national lockdowns that inter-linked with each other and could have influenced the outcome in intricate ways that were not easily controlled for correctly. The choice of the 7-week pre-intervention period is valid with the successful calibration of the synthetic control weights in controlling for the pre-intervention trends in the outcome variable between the intervention and synthetic control, as evidenced by Appendix 4.
		13. Page 12, lines 45-47: Why were outcomes measured up to three weeks after the intervention? Wouldn't any vaccinations administered by the units have been recorded in the first week?	Thank you for raising an important point. Dates when mobile vaccination units visited areas were collected and provided to us by public health teams of different local authorities between 17th May and 10th November 2021. We chose the follow-up period of three weeks for a number of reasons below: 1. We are interested in the overall impact of the intervention on overall vaccine rates, not just the numbers vaccinated in the mobile units. It is possible that people could have just switched to the mobile unit

			instead of a conventional static site that required in-advance booking. That would have led to no increase overall in vaccine uptake. We needed to allow sufficient time to compare rates between intervention and control groups to estimate this. 2. The intervention which also included promotion in the local community that could have influenced vaccine uptake – not just in the mobile units but also at conventional static sites, this effect would have been more diffuse. 3. The units visited sites on multiple occasions over the time period, so effects would not necessarily be isolated to the first week. We have added such explanation as outlined above in Methods on in Lines 17-21 on Page 7.
		14. Page 12, lines 55-57: I understand that the authors matched on the weekly number of new first doses for the seven weeks prior to the intervention. Did they also match on the total number of pre-intervention first doses? If so, why?	Thank you so very much for raising this question. We didn't directly match on the total number of pre-intervention first doses at the LSOA level. Please allow us to explain it more explicitly. For the intervention and control groups, at the LSOA level we matched the cumulative number of the first dose at the start of the seven-week pre-intervention period (as an input variable; note this variable remains constant for each LSOA during the whole study period), and the weekly number of the first dose administered during the seven-week pre-intervention period (as the outcome variable). If we have understood you correctly, the total number of pre-intervention first doses that you referred to, is the total number of the first dose administered at the end of the seven-week pre-intervention period at the LSOA level, which is a natural product of our algorithm of weight calibration, by adding the LSOA-level cumulative

			number of the first dose at the start of the seven-week pre-intervention period and all the seven data points of the LSOA-level weekly number of the first dose administered during the seven-week pre-intervention period. We therefore don't need to additionally match the total number of pre-intervention first doses, and it would not be possible as it is co-linear with the variables the weighting algorithm as outlined above. Apart from adding a new appendix (Appendix 4) to illustrate this, we have more explanation on this in Lines 8-13 on Page 8 of the manuscript.
		15. Page 13, lines 3-12: Robbins et al. developed their approach for a situation with multiple outcomes and cautioned that "failure to include a robust set of outcomes may result in omitted variable biases." Is this issue a potential concern here? Did the authors consider alternate approaches designed for single outcomes that construct a separate synthetic control for each treated unit?	In our main analysis we used the synthetic control method for micro data developed by Robbins et al. As noted, Robins et al highlighted how this method can accommodate multiple outcomes, which has advantaged over other synthetic control approaches that can only account for single outcomes. As is highlighted in the example, Robbins et al used the reason that not including a robust set of outcomes could lead to bias, because the un-accounted prior trends of these "outcomes" could be confounders. In our study, however, the only relevant outcome is the vaccination rate, therefore it is not clear what other outcomes could be included as confounder. Not including other time varying confounders (which is what Robbins et al are referring to) would bias our results, a point we make in the paper in Lines 24-27 on Page 13. However, we have accounted for a large number of potential confounders (total GP registered population size, the proportion of Black/Black British people, the proportion of Asian/Asian British people, the proportion of people of mixed ethnicity, mean age of residents, population density, the proportion of women, the IMD score, the proportion of households living in overcrowded housing, the average travel time by car to the nearest static

			vaccine site, and the cumulative number of first dose administered at the pre-intervention start time). The issue raised by Robbins et al, is not addressed by other approaches that only allow for single outcomes. Those approaches would be affected by the same “omitted variable” or” unobserved confounding” bias, in fact this is the point, we think, that Robbins et al is making in highlighting the advantages of their approach.
		16. Page 13, lines 14-17: Do the authors believe that they were able to match on all important predictors of the outcome? Some studies indicate that other variables such as education, income, employment status, health status, and the presence of children in the household were important predictors of COVID vaccination uptake at the individual level. Did the authors assess the predictive power of their predictors on the preintervention outcomes?	Thank you very much for this comment. We agree with you that other variables such as education, income, employment status, health status, and the presence of children in the household, are important predictors of COVID vaccination uptake at the individual level. We are confident that we have effectively controlled for significant predictors of the outcome to the best of our knowledge through our matching process, as the Index of Multiple Deprivation (IMD) that we used in our model has included measures of these variables. IMD is a composite measure of deprivation based on a total of 37 separate indicators of seven domains, including income (such as the proportion of all children aged 0 to 15 living in income deprived families), employment, education, health, crime, barriers to housing & services, and living environment https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019. Furthermore, in our study, it would not be possible to include these variables as separate indicators, because we didn’t have reliable individual-level data on people’s education, income, employment status, or the presence of children in the household. They are also highly correlated with each other meaning that they might not be appropriate to include (especially

			income, education, employment, and deprivation). We therefore are unable to include these variables separately in the analysis. We, however, did include chronic health conditions diagnosed in primary care in our individual-level analysis (including chronic obstructive pulmonary disease, chronic heart disease, diabetes, chronic kidney disease asthma, cancer, obesity, depression, and stroke/transient ischaemic attack). We have now added more explanation about the IMD score in Lines 1-3 on Page 6. We have also included an appendix to assess the predictive power of our predictors on the preintervention outcome (Appendix 4). It further proves that with our carefully selected set of matching variables the weighted average of outcome variable was identical in the synthetic control to those in the intervention areas during the preintervention period, as the matching algorithm achieved an exact match.
		17. Page 13, lines 43-48: Was the Poisson regression analysis also limited to people registered with a GP?	Yes, that's correct. We have revised the manuscript accordingly to further clarify this in Lines 24-25 on Page 8.
		18. Page 13, lines 43-60: Did the authors feel that they were able to include all important potential confounders in the Poisson regression analysis? (I assumed that the reason they started with a synthetic control analysis instead of a multivariate regression analysis is because individual-level data weren't available for all necessary control variables.) Did they consider using a difference-in-differences analysis as a robustness check in addition to or instead of the Poisson	Thank you very much for this comment. We acknowledge the concerns raised and appreciate the opportunity to address them. Regarding the inclusion of important potential confounders in the Poisson regression analysis, our choices of variables are based on the propensity to get vaccinated, drawing on learning from the former research and how it can be applied within the context of our study. We agree with the reviewer's assumption that individual-level data were not available for all necessary control variables. Due to this data limitation, we opted to use a synthetic control analysis to estimate causal effects on

		regression analysis?	the LSOA level. Although we have cautiously selected matching covariates to the best of our knowledge, we do acknowledge that the synthetic control method may not capture the full range of confounders in Lines 23-26 on Page 8. For that reason, we supplemented this with a regression model, weighted using the weights from the LSOA synthetic control model and including additional individual level control variables. This model adjusts for both the area-based differences between intervention and control neighbourhoods *and* individual differences. We agree that difference-in-differences (DiD) analysis of the individual-level data is useful and that is essential what we have done. The synthetic control method developed by Robbins et. A. is a generalisation of DiD, weighted for differences between intervention and control group (traditional DiD would treat intervention and control units as having equal weights). Our Poisson regression analysis is a difference-in-differences (DiD) analysis at the individual level, weighted with the synthetic control weights. After careful examination, we have chosen to implement survival analyses of individual-level data to further support the robustness of our results. Results of the additional survival analyses are in line with those of the synthetic control method and the Poisson regression analysis overall, enhancing the robustness and comprehensiveness of our study. We have now added it in Appendix 12 and referred it in the manuscript accordingly.
		19. Page 16, lines 30-39: The authors should better explain Table A2. Why is the point	Thanks a lot for your comments. We would like to clarify. Table A4 in Appendix 6 presents the model

		estimate for the mobile vaccination units 1.68 as compared with 1.23 in Table 2? Which groups actually have increased vaccination rates with the mobile vaccination units? The authors say that they found lower impacts for some groups, but were the impacts still positive? The authors might specifically interpret the 0.64 number for the interaction term for 30-65 year olds with mobile vaccination units and explain how that number can be used to derive the absolute effect of the intervention on that age group (which seems more important than the relative effect of the intervention on that age group as compared with the effect on the 18-30 year old age group).	including interactions, between the intervention and age group, ethnicity, and deprivation. The parameters associated with the intervention can only be interpreted in the context of these interactions. So, 1.68 is the relative risk (RR) to the comparison group – i.e., 18–30–year–old, white people in the least deprived areas. 1.23 is the RR in a model without interactions – i.e., the average effect across the whole population. The effect in each subgroup, can be calculated using the combination of the effect on the comparison group and the interaction terms. 0.64 is the interaction term for 30–65–year–olds. So the effect in 30-65 year olds, white people in the least deprived areas would be 1.68*0.64 =1.07. We have derived these from the model for each subgroup and present them in Figure 3. Therefore for those where the relative risk is above 1 in Figure 3, the intervention was estimated to have a positive effect, whilst those where it is below 1 the effect was negative, although we only highlighted differences that were statistically significant at the 5% level in the narrative. We don't feel producing an additional plot with the absolute rather than relative effects would be beneficial, as the relative measures are consistent with the presentation of the rest of the results and provide a better basis for comparing between groups with differing baseline levels of vaccination (see visual evidence in Appendix 2). We have now added more explanation in Appendix 6 and Lines 1-11 on Page 12 of the manuscript.
		20. Page 16, lines 43-59: Similarly, the authors should better explain Table 3. They might specifically interpret the first number in the table as an	We assumed that you meant Figure 3. We would like to clarify that it's not our intention to give the impression that the table shows increased uptakes for all groups in areas with

		example. The authors state that the table shows increased uptake for all groups in areas with the mobile vaccination units, but that isn't obvious to me from the table.	the mobile vaccination units. The figure shows the mobile vaccination units increased uptakes for people aged 18-30 from all socio-economic backgrounds and all ethnic groups except for Black/Black British, White/White British people aged above 30 from least and intermediate deprived areas, people aged 30-65 of mixed and other ethnic groups from the intermediate deprived areas, and people aged above 65 of mixed and other ethnic groups from least and intermediate deprived areas. In addition to explaining the coefficient for the reference group (1.68 at the left bottom), we have now added other texts clarify this in Lines 1-11 on Page 12 of the manuscript.
Reviewer 3 **Dr. Philip Britteon, The University of Mancheste r**	**Recommenda tion**	The authors use the synthetic control method to assess the impact of mobile vaccination units on vaccination uptake during the COVID-19 pandemic, by exploiting data on the geographic deployment of the mobile units in the North-West of England. They estimate the impact of mobile vaccination units on the vaccination rates of the population in immediate area around each mobile unit and assess how this effect differed between different subgroups within the targeted population (by age, income deprivation, ethnicity). The results suggest mobile units increased vaccination rates in the targeted areas. However, this effect differed across subgroups and intersections of the population in each area. In particular, the impact of the intervention was smaller for Asian and Black ethnic groups, the older population, and in the most deprived targeted areas.	Thank you for your review and positive feedback on the paper. We appreciate your comments and suggestions, which provide valuable insights for improving the policy recommendations, strengthening the model's assumptions, and refining the motivation and discussion. We acknowledge the importance of addressing the limitations identified in your comments. By doing so, we aim to enhance the robustness and applicability of the study's findings and policy implications. We have carefully considered your feedback and made the necessary revisions to address the limitations you highlighted, as detailed below. Thanks again for your thorough review and constructive feedback!

		The paper is well written and addresses a well-motivated research question using a policy evaluation method that is appropriate for the study. My comments highlight limitations that could be addressed to improve the policy recommendations that can be drawn from the study, to strengthen confidence in the underlying assumptions of the model, and to refine the motivation and discussion of the paper.	
	Comments	1. Exploiting distance from mobile vaccination unit The authors use a 1km radius around each mobile unit to define the intervention areas. This allows the authors to estimate the targeted effect of the intervention on those living closest to each mobile vaccination unit. However, I think it is important to also investigate whether there was a "dose" effect of the policy as the distance from mobile unit increased in order to strengthen the findings of the study. As a simple test, this could be done reestimating separate synthetic control models on two or more treatment zones (e.g. less than 1km, 1km-3km) and using the other wider regions as controls. This test would serve three purposes. First, it would allow the authors to check whether they were underestimating the true impact of the reform by restricting the intervention group to a small surrounding area. Second, it would allow the author to test the hypothesis that distance matters (i.e. did the effect	Thank you for your feedback. Yes, we used the distance between the population weighted centroids to mobile vaccination units to identify the intervention and control group for the individual analysis, by setting the threshold as 1km. This is to make sure even for the individual-level analysis the LSOA-level observed differences were accounted for between the intervention and control group with the weights calibrated from the synthetic control method, which is essential for our quasi-experimental study design and drawing any possible causal associations. We chose the threshold of 1km after conducting extensive experiments of the "dose" effect. We have now added more information on this as the sensitivity test of using different distance thresholds (0.5, 1.5, 2 and 3km), and reported the results in Appendix 10. Results of this sensitivity test show that our analysis is sensitive to the choice of distance in constructing the intervention and non-intervention areas. We would like to highlight that this is a plausible consideration that we had incorporated in our study design, given the spatial nature of the data. The distance between the mobile vaccination units and population

		dimmish and over what distance). Third, it could also be interpreted as a robustness check when testing the assumptions of the model (see point on spillovers below). Depending on how the treatment areas are defined, the authors should also test for spillover effects by excluding surrounding/adjacent areas to the intervention areas from the pool of potential control areas in a sensitivity test. Spillover effects from intervention to non-intervention areas in the synthetic control group could bias the estimates of the model (most likely downwards, if the mobile units also attracted people outside the 1km boundary). The authors could also test whether their results were sensitive how they defined distance in the analysis. The authors identify distance from each mobile unit based on geographical distance (less than 1km), but travel time could be a better measure in this instance if data on this is available. As the authors have data on travel time to nearest static clinic they would then be able to define the full intervention population as being equal to the population whose travel time to the mobile clinic was less than their travel time to static clinic. This would allow the authors to capture the full impact of the intervention (dependent on data availability, of course!).	weighted centroids of LSOAs has a significant effect on the number of people that the mobile vaccination units could reach, which is critical in evaluating the true impact of the mobile vaccination units. We have also now added the sensitivity test of the spatial spill over in Appendix 11. Having accounting for the spatial spill over effect between 1 and 1.5 km, the results of model 1e are well in line with the main model (model 1), further confirming the robustness of our main analysis. We have already used the average travel time by car from each person's address to the nearest conventional static vaccination centre (conventional static vaccination centres that require booking in advance, not the two pop-in static sites that we included as mobile vaccination units in our LSOA level analysis; we have now clarified on that throughout the manuscript to avoid confusion) by LSOA in our LSOA-level synthetic control analysis. To further mitigate the potential impact of our research decision on the individual-level analysis, we then included the travel time by car from each person's address to the nearest conventional static vaccination centre in our weighted Poisson regression model. By doing this, we could not only account for systematic area-based differences, but also control for individual-level geographical differences in accessing the main stream vaccination site. As we don't have access to the address of individual participants for data security and ethical reasons, we couldn't calculate the direct distance, or the travel time to specific mobile vaccination units for them. We don't have access to data on the location or operational dates of conventional

			static sites within our study area and period either. The travel time by car from each person's address to the nearest conventional static vaccination centre, admittedly the second-best variable compared to the direct distance, or travel time to mobile vaccination units for individuals, was the only available information provided to us to measure the geographical accessibility of vaccinations. We believe that we did the best we could within our power and capacity to make our individual-level analysis as robust as possible.
		2. Extending the follow-up period I was also wondering why the authors chose to restrict the follow-up period to only 3 weeks. This should be justified in the methods section. If data is available for a longer follow-up period, then this could then be used to exploit the "switch off" of the treatment effect to begin to unpick whether the mobile units attracted people who would otherwise not have been vaccinated - a current limitation of the study and an important policy question. The question would then be: did the vaccination rate return to the same or a lower level than the control group after the intervention? However, this would require information on how long each unit was deployed in each area which may not be available to the authors. Extending the follow-up period until the time that the mobile unit was removed would also allow the authors	Thank you for pointing this out. There are two main reasons for our choice of the follow-up period: 1. Practicality, feasibility and resource constraints. Our priority in conducting this study was to provide a rapid response during the pandemic on the effectiveness of this approach to inform its commissioning in the region. The study therefore was resourced for a limited follow up period in order to produce results within a time frame that was useful for decision makers. This does reflect the reality of an urgent public health response and the best use of available data to inform policies and tactics. Also although we have made all the effort to collect as much data as possible, we struggled to gather sufficient data for July and August 2021 for most places. That was because places started prioritising the administration of the second-dose vaccines on mobile vaccination units then, and the number of people receiving the first-dose vaccines on mobile vaccination units had dropped and as such records

		to test whether vaccinations diminished during their tenancy, allowing the authors to infer policy recommendations on the length of deployment.	were incomplete. Some external factors may have caused that, such as the availability of resources (including funding, personnel, and logistical support) among the local team to collect such data in a complete manner, when they were struggling to meet the urgent healthcare needs of communities. Longer follow-up periods would have required more resources and funding to collect the needed data, which unfortunately wasn't available back then. 2. The nature of the intervention and our study objectives. We aimed to measure the short-term impact of mobile vaccination units using weekly number of first-dose vaccines administered among adults as our outcome. Our observation period spanned from seven weeks (22nd February 2021) before the first mobile vaccination unit visit on 12th April 2021 to three weeks (19th July 2021) after the last mobile vaccination unit visit on 28th June 2021, during which the first-dose vaccines were prioritised. Since July 2021, many places gradually switched to prioritising the administration of the second-dose vaccines. We found that 3 weeks was sufficient to capture the relevant data needed to address our research question – measuring the potential immediate effect of the intervention upon the weekly number of first-dose vaccines administered among adults, without much spill-over effect of prioritising the second-dose vaccines since July 2021. To summarise, we aimed to measure the immediate or short-term effect of the mobile vaccination units, which was a substantial intervention (54 visits during our study period). We

			didn't strive to measure long-term outcomes; neither was that practical or feasible due to resource constraints. We therefore chose our endpoint and the follow-up period as such to facilitate our investigation. Because of that, we are unable to infer policy recommendations on the length of deployment through extending the follow-up period. We have added such explanation as outlined above in Methods in Lines 17-21 on Page 7, and more details on the deployment of the mobile vaccination units in Lines 14-24 on Page 6.
		3. Exploiting the timing of deployment of each mobile unit The authors could also consider whether the timing of the deployment of each mobile unit affected the treatment effect by stratifying the analysis based on whether the unit was deployed in the first or second half of their sample. This would help to understand whether differences in the severity of the pandemic and the stage of the rollout of vaccination program affected the results to help draw conclusions on the applicability of their findings in different settings. It was also unclear how the authors dealt with the fact that mobile units appeared to have been deployed to the same areas more than once – "visited 37 sites on 54 occasions". Does the estimated effect capture the impact on first time deployments only or does it include the impact of second, third time visits as well? If the	Thank you for raising an important point. We acknowledge the value of stratifying the analysis based on the deployment timing of the units, such as the first or second half of the sample. However, there are limitations and constraints that prevent us from implementing this stratification in our analysis. Firstly, our dataset does not provide sufficient quality or data structure to accurately divide the sample into distinct halves based on deployment timing. The timings of the deployment were selected by the local public health team according to the availability of their resources during our study period. The visits were therefore distributed far from evenly. These data were then collected by the data providers and provided/updated to us sporadically during our study period. Although our study period was sufficient for reasonably relating exposure to outcome, stratifying the analysis into two equal halves based on deployment timing requires a reasonably evenly distributed

		latter, then this could be explored further in the paper by looking at whether the impact of the intervention reduced in subsequent visits.	deployment dates throughout the study period with substantial number of observations in each stratum to ensure statistical robustness and reliable comparisons. For instance, the majority of the visits were conducted in June 2021 (34 out of the total 54 visits). Our dataset therefore unfortunately doesn't allow us to do that. Secondly, we have limited information on how the severity of the pandemic and the stage of the vaccination program rollout across different settings had impacted the deployment strategies in each local site. To conduct meaningful stratification, it would require comprehensive and reliable data on these factors for each deployment unit, which isn't feasible or accessible to us. Based on anecdotal sources in a few sites, the deployment strategies were in fact very much influenced by the availability of medical staff and free-parking public open spaces in local neighbourhoods, if not more so than the severity of the pandemic and the stage of the vaccination program rollout across different settings at the macro level. But again, we didn't have comprehensive or sufficient information for all the sites. Furthermore, stratifying the analysis based on deployment timing could introduce additional complexities and potential confounding factors. The severity of the pandemic and

			the vaccination program rollout are interconnected and influenced by numerous contextual variables, making it challenging to isolate their independent effects in the regression models. In the framework of the synthetic control methods, however, accounting for all the relevant variables and their interactions in a stratified analysis would require extensive data and more sophisticated statistical modelling techniques. Even if we were able to split the sample into two exact halves based on deployment timing somehow, they would have to be assigned with either unavoidably overlapping or tremendously shortened (for instance, 3.5 and 1.5 week's respectively) pre-intervention and post-intervention periods. Either would be extremely difficult, if not impossible at all, to estimate with our quasi-experimental design in the framework of the synthetic control method. Another widely used way to stratify the analysis based on deployment timing is to simply include a dummy variable denoting the first and second half of the period. But then again, this isn't supported by our unevenly distributed data. While stratifying the analysis based on deployment timing could provide some insights into the effects of pandemic severity and vaccination program stage, it is important to consider the trade-offs between the complexity of the analysis and the availability and reliability of the data. In our study, we have focused on analysing the overall effect of the

			intervention while acknowledging that variations in pandemic severity and vaccination program stage may have co-existed across settings. This approach allows us to draw broader conclusions about the applicability of our findings in different contexts. We have now added more information about the timing and duration of these interventions in Lines 14-18 on Page 6 and in the final paragraph of Discussion where we discuss the limitations of our study (Lines 20-22 on Page 13 of the manuscript). We chose to treat sites with more than one visits the same as the rest in our analysis. Please allow us to explain why we made this decision: Firstly, considering each site as a separate entity would introduce complexity and potential bias due to the variation in the number of visits across sites. Treating them equally allows for a simpler and more straightforward analysis. Secondly, by treating all sites uniformly, we aim to minimise the potential influence of site-specific factors that could confound the analysis. If we were to differentiate sites based on the number of visits, it might introduce site-level characteristics or biases that could obscure the true effects of the intervention. We also don't have sufficient information about site-level characteristics for all the site to support us to do that. Thirdly, treating all sites equally provides a conservative approach that avoids potentially

			overemphasising the impact of sites with multiple visits. By not giving undue weight to these sites, we ensure that the overall analysis remains balanced and representative of the entire sample. While it is possible that sites with multiple visits may have different dynamics or characteristics, our decision to treat them the same as sites with one visit is a simplifying assumption that allows us to focus on the overall effect of the intervention rather than site-specific variations. This approach enables clearer interpretation and generalisability of the results. We have also now added more explanation on this in Lines 18-24 on Page 6 of the manuscript.
		4. Defining the synthetic control group in the inequalities analysis The individual level analysis allows the authors to easily stratify their analysis by different population characteristics. This inequalities analysis is a key part of the paper. However, I have questions about how the authors define the control group in these subgroup analyses. Do the authors use the same synthetic control group as in the main analysis? I think it would be more informative to identify new synthetic control groups for each subgroup so that you are comparing like-for-like across the intervention and non-intervention areas. This dawned on me when looking at the results in Figure 3 where the authors find an adverse impact of mobile vaccination units on certain subgroups despite there	Thank you very much for raising this point. We would like to take this opportunity and clarify that it's not plausible to construct the exact same synthetic control of the individual analysis with that of the LSOA-level analysis. We therefore don't have a control group in the same manner as that of the LSOA-level synthetic control, although the intervention and synthetic control LSOA areas remain the same in the individual-level analysis. In the individual level analysis, we used a weighted Poisson regression model with robust sandwich variance estimators to assess the association between the intervention and outcome, whilst controlling for area-level differences by weighing with synthetic control weights of the main analysis and additional potential confounding individual-level variables, including age, sex, ethnicity, health conditions (as defined above), whether people were in contact with social care, whether they were paid carers and the travel time by car from each person's address to the nearest conventional static vaccination

		being no theoretical reason why mobile units should decrease vaccination uptake. These negative findings suggest that other unobserved differences over time might be influencing the results (e.g. 30-65 year old, black men and women, in the most deprived intervention areas had different underlying trends to the same demographic in the synthetic control group from the main analysis). Estimating a new synthetic control group for each subgroup would provide an arguably more meaningful comparison (providing that there is sufficient overlap to estimate the synthetic control method for each subgroup). If not, then the authors should at least present the trends for each sub-group in the intervention and non-intervention areas to allow the reader to assess whether they believe that the control group is a good comparison.	centre. The control group in the individual-level analysis consists of people who had been registered for primary care with the National Health Service, aged over 18, living in the non-intervention LSOAs (as defined in the synthetic control method), but hadn't been vaccinated before the intervention or visited by the mobile vaccination unit in their neighbourhood. It is not appropriate here to use the same synthetic control group as in the main analysis to include all the eligible adult population at the aggregated level of LSOAs as the control population: among all eligible adults of the non-intervention areas, those who had received their first-dose vaccine before the introduction of the intervention should be excluded. That's because they hadn't been subject to any influence from the mobile vaccination units to take up the vaccine. For the same reason, in the intervention group we have also excluded those individuals of the intervention area who had received their first-dose vaccine before the introduction of the intervention. To summarise, in the analysis of individual level, we couldn't stick with the exact same synthetic control group of individuals as those of the LSOA-level analysis. We would like to clarify about the results in Figure 3. Due to the small sample sizes of these 45 subgroups (see Appendix 7), these results lack statistical power and are inaccurate and unstable estimates that should be treated with caution. They are only indicative of the likely pattern of effects of the intervention across these groups. We have now added Appendix 7 and other texts clarify this in Line 12 on Page 12 of the manuscript. It is also not beneficial here for us to construct new control groups by

			conducting new synthetic control analyses respectively for each subgroup either. It would drastically reduce our sample size (See Appendix 7 for more information) and thus reduce the explanatory power of our analysis. It would also be technically difficult, if not impossible at all, to accommodate the main and interaction effects simultaneously using separated and fragmented synthetic control weights generated by the synthetic control algorithm in the same weighted Poisson regression model. Not to mention the additional biases that would be introduced in the model. We have now added more explanation on the subgroup analysis in both the main manuscript (Lines 1-11 on Page 12) and the appendices (Appendix 6 and 7). We have also provided more description on the intervention and non-intervention areas (Appendix 2 and 3).
		5. Investigating differences in deployment strategies The authors could also investigate whether there were differences in treatment effects between local authorities given that the decision to deploy units was made at this level. However, I would only suggest testing this if it is possible for the authors to meaningfully interpret the output from these results (i.e. if they are able to distinguish between the different approaches taken by each locality when deciding where to deploy each mobile facility). Alternatively, it might be possible to infer information on the deployment strategies of each mobile unit from the data (i.e. look at differences	Thank you for this suggestion. It would have been interesting to explore this aspect. However, we have limited information on how the deployment strategies were made in each local site. From sporadic anecdotal sources in a few sites, the deployment strategies were influenced by the availability of medical staff and free-parking public open spaces in local neighbourhoods, as well as the severity of the pandemic and the stage of the vaccination program rollout across different settings at the aggregated level. To conduct stratification and interpret the results meaningfully, we would require comprehensive and reliable data on these factors for each deployment unit, which unfortunately isn't accessible to us. We therefore are not able to draw meaningful policy recommendations for each local authority on where decision makers

		in the area level characteristics of the areas where each unit). The authors could then stratify the analysis accordingly to draw further policy recommendations on where decision makers might want to choose to deploy vaccination units in the future to achieve different aims (e.g. increase total vaccination rate, reduce vaccination inequalities). This is subtly different from stratifying the analysis based on individual level characteristics. However, I recognise that this could form a new paper in itself.	might want to choose to deploy vaccination units in the future to achieve different aims at the neighbourhood level (e.g. increase total vaccination rate, reduce vaccination inequalities). To better illustrate the context, we have now provided more description by different local authorities (Figure A7 in Appendix 2).
		6. Defining the outcome variable It would be helpful if the authors clearly stated the outcome variable used in the study in the data/methods section and in the tables/figures (i.e. how they define and assess “vaccination uptake”). This was not initially clear when reading through the paper.	We appreciate the reviewer's feedback regarding the clarity of our outcome variable, specifically the definition and assessment of "vaccination uptake." We apologise for any confusion caused by the initial presentation of this information. To address this concern, we have made sure to explicitly explain the outcome variable at its first occurrence in the abstract. We have also reiterated its definition in the Data and Methods section, and the tables and figures where appropriate. We hope this could help the readers easily identify and understand the specific metric used in our analysis.
		7. Adjusting for ceiling effects The authors should also justify why they chose to estimate the impact of the intervention on weekly first dose vaccinations rather than the total first dose vaccinations. There are advantages to using the cumulative total vaccination rate when constructing the	Thank you so very much for raising this question. We have considered the ceiling effect and that’s why we didn’t directly match on the total number of pre-intervention first doses or the cumulative total vaccination rate at the LSOA level. Please allow us to explain it more explicitly. For the intervention and control groups, at the LSOA level we matched the cumulative number of the first dose at the start of the seven-week pre-

		synthetic control group. Namely, using total vaccinations allows one to construct a control group that is more likely to follow the same path in the absence of the intervention when ceiling effects come into play (e.g. if people are less and less forthcoming to be vaccinated as the vaccination rate increases). This potential issue is highlighted in Table 1, which highlights that the first dose vaccine uptake rate was higher in the non-intervention areas (69%) than in the intervention areas (53%). It could therefore be the case that the slight decline in the weekly vaccination rate in the control group (shown in Figure 2) may have been driven by a ceiling effect that would not have occurred in the intervention areas, raising concerns that the observed treatment effect is in fact an overestimate of the true effect of the intervention. Using total first rate vaccinations as the outcome to construct the synthetic control group would overcome this limitation. There is also the issue that when comparing effects across subgroups in Figure 3 that differences in the impact of the intervention could also be driven differences in the underlying vaccination rate and ceiling effects. It would be helpful to include descriptive statistics (either trends over time in a figure or averages in a table) of the total vaccination rate by subgroup.	intervention period (as an input variable; note this variable remains constant for each LSOA during the whole study period), and the weekly number of the first dose administered during the seven-week pre-intervention period (as the outcome variable). If we have understood you correctly, the cumulative total vaccination rate that you referred to, is the percentage of the total number of the first dose administered for adults among the total adult population at the end of the seven-week pre-intervention period at the LSOA level, which is a natural product of our algorithm of weight calibration, by adding the LSOA-level cumulative number of the first dose at the start of the seven-week pre-intervention period and all the seven data points of the LSOA-level weekly number of the first dose administered during the seven-week pre-intervention period. We therefore don't need to additionally match the total number of pre-intervention first doses. Additionally matching the total number of pre-intervention first doses will introduce the issue of perfect collinearity in the model and cause it to collapse. We chose to do that is because the weekly number of the first dose administered is a more sensitive measurement towards our research question. We also chose to include the cumulative number of the first dose administered at the beginning of the intervention, which is beneficial for measuring the baselines as well as accounting for the ceiling effect. Apart from adding a new appendix (Appendix 4) to illustrate this, we have more explanation on this in Lines 8-13 on Page 8 of the manuscript. We have also included descriptive figures and tables in Appendix 1, 2 and 3 to further assist with the understanding of our analysis.

		8. Including additional descriptive statistics I would have also liked to see descriptive statistics on the full non-intervention population, not just those included in the synthetic control group in Table 1. This would help the reader to infer information on how the intervention areas differed from the full population (i.e. how local authorities may have made decisions on where to deploy the mobile units).	Thank you for raising this. These have now been included in Appendix 3 and referred to in the main manuscript where appropriate.
		9. Defining the start date The authors should also justify their choice of start date (7 weeks prior to the deployment of the mobile unit). The synthetic control method performs better with a longer pre-policy period so unless there is a reason for using this cut-off point (which I suspect there might be) then a longer pre-policy period should be used.	We have now added more explanation on this in Lines 12-17 on Page 7 of the manuscript, in addition to the new visual evidence presented in Appendix 2. We choose seven weeks of preintervention data, primarily because it's when we started to observe signs of slowing-down of growth in crude uptake overall and expanding variations in the uptake of the first dose across LSOAs (Figure A1 and A2 in Appendix 2), but with widening gaps among ethnic and socio-economic groups in the local population (Figure A3-A6 in Appendix 2). This choice was also motivated by the rollout of the vaccine prioritised by age groups, which only expanded to people aged 65 and over in the second half of February 2021 (see Appendix 1) whilst the vast majority of the adult population aged below 65. The 7 week's pre-intervention period commencing on 22nd February 2021 could therefore allow us to capture trends for the majority of the population eligible for the vaccine, whilst avoiding potential influences from prior changes in the severity of the pandemic and policy changes in COVID-19 local restriction measures and national lockdowns that inter-linked with each other and could

			have influenced the outcome in intricate ways that were not easily controlled for correctly. The choice of the 7-week pre-intervention period is valid with the successful calibration of the synthetic control weights in controlling for the pre-intervention trends in the outcome variable between the intervention and synthetic control, as evidenced by Appendix 4.
		10. Weighting by population size The authors do not weight by population size in model 1. I was confused by their explanation as to why this was the case: “a weighted regression model to estimate this effect would potentially underestimate standard errors and p-values as this would not account for complex aspects of the process used to generate the synthetic control”. This needs to be justified more clearly. It is maybe also worth noting that the lack of population weights in the main analysis is unlikely to matter as LSOAs are roughly the same size.	Thank you very much for raising this important point and giving us the opportunity to explain it. We didn’t weigh by population size (LSOA-level measures of total GP registered population size to be more specific) separately in model 1 (the synthetic control method) as we have already included it as the matching variable (Lines 19-26 on Page7), and the essence of the synthetic control method is to calibrate the weights to construct a synthetic control group that is comparable to the intervention group regarding the selected matching variables. Population size is therefore already controlled and weighed in our model 1, with our calibrated synthetic control weights. We don’t need to additionally weigh the model specifically and separately by population size in model 1, the LSOA-level synthetic control analysis. We also want to clarify that we chose to weigh the individual-level weighted Poisson regressions by the calibrated synthetic control weights of model 1 (rather than by population size alone) to account for aggregate-level confounders. This is to make sure that in addition to LSOA-level measures of total GP registered population size, the intervention and synthetic control groups are also comparable in Poisson regressions for other LSOA-level matching variables (the proportion of

			Black/Black British people, the proportion of Asian/Asian British people, the proportion of people of mixed ethnicity, mean age of residents, population density, the proportion of women, the IMD score, the proportion of households living in overcrowded housing, the average travel time to the nearest conventional static vaccine site by car, the cumulative number of first dose administered at the intervention start time, and the number of new first dose vaccinations administered per week). “a weighted regression model to estimate this effect would potentially underestimate standard errors and p-values as this would not account for complex aspects of the process used to generate the synthetic control” refers to “individual-level analysis using weighted Poisson regressions”. We have now moved the sentences to the next paragraph of the individual-level analysis and revised the text in Lines 30-34 on Page 8 to clarify that.
		11. Motivating the study It would have been nice to see more discussion in the introduction on why vaccination uptake is lower among young people, socioeconomic disadvantaged, and ethnic minorities. Do the studies referenced in the introduction infer or hypothesise why this is the case? This is important for understanding why the impact of the policy may have worked / not worked amongst different intersections of the population. The authors reference studies that investigated the impact of different types of interventions to increase	Thank you, we have expanded the discussion accordingly. Vaccination uptake for the first dose of COVID-19 vaccine is lower among young people, socioeconomic disadvantaged, and ethnic minorities for serval reasons such as vaccine eligibility and vaccine hesitancy. We have now added more information on this in Lines 6-9 on Page 1 of the manuscript. Vaccination uptake is lower among ethnic minority groups, and those with lower education, income and social-economic status tends to be related to the relatively higher level of vaccine hesitancy, issues with convenience and access, health disparities and inequalities, language and cultural barriers, trust in healthcare systems, communication and messaging, and

		vaccination uptake in the introduction. Did these studies also investigate inequalities in impact? This should be discussed in the final paragraph of the introduction if they do. If not, I think that the authors perhaps undersell their contribution to the literature on inequalities in vaccination uptake. This is an important part of the study. In particular, their approach to investigate inequalities in impact between intersections of the population (including ethnicity) is novel. I would emphasise this in the introduction and the discussion in addition to their other main contributions. Are there studies that look at impact of mobile units used to administer other forms of preventative care (e.g. cancer screening)? The authors should consider referencing these studies in the introduction if they exist as the mechanisms through which these units work would be similar (e.g. reduce time/cost for the patient to increase uptake). This could then be discussed in the discussion section when talking about the policy implications of the study (i.e. do the results suggest anything about the use of mobile units in other areas).	information disparities among those groups. As with the lower uptake among young people during our study period, it's mainly because of the vaccine eligibility. In the early stages of the vaccination rollout, priority was given to the elderly, healthcare workers, and other vulnerable groups. Young people, who generally have a lower risk of severe COVID-19 outcomes, were not initially eligible to receive the vaccine, leading to a lower uptake among this age group. Moreover, young people, especially those without underlying health conditions, may perceive themselves as having a lower risk of severe COVID-19 outcomes. This perception could lead to a belief that vaccination is unnecessary, impacting their motivation to get vaccinated. Regarding the referenced studies that investigated the impact of different types of interventions to increase vaccination uptake, we have now added more information on investigating inequalities in impact in Lines 1-4 and 11-13 on Page 2 of the Introduction. We thank the review for advising on adding references that look at impact of mobile units used to administer other forms of preventative care (e.g. cancer screening). We appreciate your insight, and we agree that there is a large body of literature analysing impact of mobile units used to administer other forms of preventative care. Please allow us to elaborate on how we made our choices in conducting the literature review. Because of the quasi-experimental design of our study, we have decided to limit our literature review to quasi-experimental and randomised control trials of interventions aiming to increase immunisation rates for COVID-19, Influenza, and pneumococcal

			vaccination, as these target similar population groups. We have also included studies focusing specifically on mobile vaccination units for these diseases, regardless of study designs. We therefore believe that our choices of the literature are broadly in line with your recommendation on conducting the literature search. We have now added discussion on using mobile units to administer other forms of preventative care with study designs of quasi-experimental and randomised control trials in Lines 27-28 on Page 1 of the Introduction and Lines 28-30 on Page 12 of the Discussion in the manuscript. The following are the relevant references: 12 Williams EM, Vessey MP. Randomised trial of two strategies offering women mobile screening for breast cancer. *BMJ* 1989;**299**:158–9. 13 Greenwald ZR, El-Zein M, Bouten S, *et al.* Mobile Screening Units for the Early Detection of Cancer: A Systematic Review. *Cancer Epidemiol Biomarkers Prev* 2017;**26**:1679–94. doi:10.1158/1055-9965.EPI-17-0454 14 Reuben DB, Bassett LW, Hirsch SH, *et al.* A Randomized Clinical Trial to Assess the Benefit of Offering On-Site Mobile Mammography in Addition to Health Education for Older Women. *Am J Roentgenol* 2002;**179**:1509–14. doi:10.2214/ajr.179.6.1791509 15 Shima A, Tanaka H, Okamura T, *et al.* Offering on-site mammography in workplaces improved screening rates: Cluster randomized controlled trial. *J*

			Occup Health 2023;**65**:e12389. doi:10.1002/1348-9585.12389
		12. Information on the intervention The authors could have included more information on the details of the intervention. I have suggested some of the questions that came into my head when reading the paper below. I don't think that all of the questions need to be answered given the word limit of the journal but more information would have definitely been helpful. (1) How long on average were the mobile units situated in each location? Did this differ between units? What is the timeline of units being deployed between first and last unit (i.e. were the majority of units deployed at the same time or were they spread out)? (2) Did the intervention include other types of intervention in addition to the mobile units (e.g. outreach programs)? The authors mention about not being able to distinguish between the two as a limitation in the discussion but do not mention this aspect of the program in the intervention section. (3) What classifies as a static vaccination site? Were they also setup during COVID and during the study period? If yes, the introduction of static	Thank you, we have added these details accordingly throughout the manuscript where appropriate. We have also provided a summary here for your reference: (1) The mobile units situated in each location typically from 9am to 4.30pm or 5pm on the date of the visit, with four visits operating from 8am to 6pm, 9am to 3pm, 9.15am to 2.30pm and 10am to 2pm respectively. The first and last visits were deployed on 12th April and 28th June 2021 respectively in our study. The visits were distributed unevenly during our study period, with the majority of the visits conducted in June 2021 (34 out of 54). We have now added more explanation on this in Lines 14-18 on Page 6 of the manuscript. (2) The deployment of the mobile vaccination units did involve outreach programmes (e.g., leaflet campaigns to advertise the mobile vaccination units and explain the benefits of vaccination). We have now added that information in Lines 23-24 on Page 6 of the manuscript. (3) The two visits from the two pop-up static vaccination clinics that offered walk-in services during our study period were included in our study, as they essentially offered the same drop-in services without the need to make an appointment in advance. They also involved the same outreach programmes as visits of mobile units. Our advisors of the local public health teams suggested us to include them in the evaluation. We

		sites may have biased the results of the study and should be addressed in the design of the analysis. (4) How severe was the pandemic throughout the study period and did this differ during the study period? How many people had already been vaccinated prior to the intervention? Which ages were eligible for a vaccine and did this change throughout the study period? Did other national interventions/rule changes occur throughout this period? Were static sites also available as walk-ins? These questions are important in understanding how easy it was to get vaccination in order to explain differences in impact between subgroups and areas (i.e. were the mobile sites attractive because they offered walk-ins, were open to all ages).	however don't have any other information on how they were set up, or whether they were set up specifically during our study period only or throughout the rollout of COVID-19 vaccination program. Apart from more clarification on this throughout the manuscript, we have also provided a sensitivity test by excluding them from our analysis in Appendix 9 and we found our results are robust. (4) The severity of the pandemic didn't change that much during our study period according to the government statistics on the number of cases by specimen date https://coronavirus.data.gov.uk/details/cases?areaType=region&areaName=North%20West. Other than the relaxing of restrictions on 17th May 2021 to allow indoor hospitality to reopen with groups of up to six people, we are not aware of other national interventions/rule changes during our study period. Such relatively stability in the severity of the pandemic and COVID-19 restriction policies facilitates the choice of our pre-intervention period and is therefore mentioned in Lines 15-17 on Page 7. There were 1282550 people that had received the first dose of the COVID-19 vaccine prior to the intervention within Cheshire and Merseyside. We have also provided the breakdown of the accumulated vaccination rate by intervention and non-intervention areas in Table 1 of the manuscript and Appendix 2. Conventional static vaccination clinics normally required users to book an appointment in advance. We are neither aware of nor have data on other static vaccination clinics offering walk-in services other than the two

			pop-up ones that we have already included in our analysis. We have now provided more clarification on this throughout the manuscript. The age eligibility for the first dose of the COVID-19 vaccine changed during our study period as detailed below (we have now provided the full timeline for the first dose of the COVID-19 vaccine in England in Appendix 1 and referred to it in the manuscript where appropriate: • 1 March, all aged 60 and over; • 6 March, all aged 56 and over; • 17 March, all aged 50 and over; • 13 April, all aged 45 and over; • 26 April, the programme was rolled out to those aged 44; • 27 April, people aged 42 and over were invited to book their first COVID vaccine; • 30 April, people aged 40 and over were invited to book their first COVID vaccination; • 3 May, people in England aged 38 and 39 became eligible for their first COVID vaccine; • 18 May, People in England

			aged 37 became eligible for their first COVID vaccine; • 20 May, People aged 34 and 35 were invited to book their first COVID vaccine; • 8 June, England's vaccination programme was extended to people aged 25–29. • 15 June, People aged 23 and 24 became eligible to book their first COVID vaccination; • 16 June, the vaccine rollout opened to those aged 21 and 22 in England; • 18 June, people aged 18, 19 and 20 became eligible for their first COVID vaccination.
		13. Other minor comments (1) Why did you choose to combine information on an individual's ethnicity from 17 to 5 categories? Was this necessary given the large sample size of the data used in the study? (2) Is it possible to include the location of static sites as well in	(1) The 5 categories represent the groups with distinctive lived experiences and public health characteristics relevant to practical steps that the local public health teams wished to take in an urgent public health context. Therefore the 5 aggregated categories align well with our study's goal to examine broad racial and ethnic groupings while still capturing meaningful variations in the data. By using the broader 5

		Figure 3? (3) It would be helpful to explain the red dots in key/legend in Figure 3. (4) Please report figures in Table 1 to more decimal places to allow for comparisons between the intervention and non-intervention areas. (5) In the results section, the authors write: "This overall effect size estimates 3723 additional vaccinations over three weeks of follow-up across the study area, similar to the actual number of people vaccinated in the mobile units (n=3824). This suggests that at least in the short term most of the vaccinations in the mobile units were additional. In other words, among the mobile units' users there were few people, who would have otherwise gone to static centres to get their vaccine." This should be moved to the discussion when interpreting the results. It is an important interpretation. (6) The authors should justify why they run a sensitivity test excluding ethnicity from the set of control variables or drop it from the analysis. I cannot see why this particular sensitivity test was needed. (7) The first paragraph of the discussion is not 100% clear in how it is worded. It might help to first discuss the overall impact of	categories, our study can also facilitate comparison with other studies that have adopted similar aggregated classifications. This enhances the ability to benchmark and compare findings across different research projects. Among the 17 categories, combining categories is necessary to avoid small sizes of some groups, which can lead to unreliable estimates and statistical instability. These small groups may also raise ethical concerns related to privacy and confidentiality, as we are also looking at other characteristics such as age, sex, deprivation, and their interaction terms (see Appendix 6 and 7 for more details). For those categories with very low representation in the dataset, combining them with similar categories facilitates data analysis and interpretation. So it was necessary to group ethnicity into 5 categories in our study. (2) We think you were referring to Figure 1. We have now added the location for both mobile (red) and the two pop-up static (cyan) sites in the maps (Figure 1 in the manuscript and A8 in Appendix 4). We, however, don't have access to data on the addresses of the conventional static sites. (3) We think you were referring to Figure 1. We have now added the legend for the location for both mobile (red) and the two pop-up static (cyan) sites in the maps (Figure 1 in the manuscript and A8 in Appendix 4). (4) Done. (5) Done. (6) We performed this sensitivity test in response to the missing ethnicity data, which accounted for 18.62% of the total dataset. By conducting this test, we aimed to assess the potential impact of the missing data on our results

		the policy and in a separate paragraph to discuss the results by subgroup. (8) The authors also note that interpretation of Figure 3 should be interpreted with caution due to small sample sizes within each of the 45 subgroups but do not provide figures on these sample sizes for the reader to inform their own judgement.	and conclusions. It allowed us to understand the robustness of our findings and evaluate the extent to which the missing data may have influenced our analysis. This sensitivity test helps ensure the validity and reliability of our study by addressing the potential bias introduced by the missing ethnicity data. We have now edited the main text in Lines 10-12 on Page 9 to clarify this. (7) Done. We have now changed the manuscript in Lines 23-30 on Page 13 to reflect this. (8) Done. We have now provided the sample sizes for the reader in Appendix 7 and referred to it in the manuscript where appropriate.

VERSION 2 – REVIEW

REVIEWER	Gary W Reinbold University of Illinois at Springfield
REVIEW RETURNED	30-Aug-2023

GENERAL COMMENTS	The authors made an excellent effort to respond to the reviewers' comments and I am satisfied on all of the substantive issues. There are still a few places, however, where I think their language could be a little more precise. P.2, line 9: The treatment group could be described a little more precisely here. Not everyone in the treatment group "lived" within 1 km of a mobile site. P.2, lines 17-20: Saying the effect was "smaller" implies that there was still an effect for these groups, but Figure 3 suggests that there was no effect for many of these groups. By the way, thank you to the authors for adding Figure 3, which I found to be very helpful. P.3, lines 10-12: If I understand correctly, the authors weren't able to include any individuals who had not been registered with a GP at the time of the study, not just individuals who used the mobile units. P.7, line 5: If I understand correctly, the 1112 non-intervention LSOAs also did not have population centroids within 1 km of a mobile unit. P. 8, lines 6-14: I understand that some of this language was added to respond to my prior comment, but I think the authors misunderstood my comment, which probably means that I didn't

	state it clearly enough. It is obvious that the sum of (1) the cumulative number of first doses prior to week t - 7 and (2) the number of new first doses for weeks t – 7 through t - 1 would equal (3) the cumulative number of first doses after week t - 1. I only wanted the authors to be clearer on whether they are matching on just (2), on both (1) and (2), or on both (2) and (3). I understand now that they are matching on both (1) and (2), although the document still doesn't quite state that here. The authors could simply state that more clearly here and, if they would like, remove the added language about collinearity. P. 9, lines 2-4: Table A4 indicates that IMD tercile was included in these regressions, but it isn't included in the list of variables here. P. 11, lines 21-25. Again, saying that there was a "lower impact" for these groups implies that there was still an impact, but Figure 3 suggests otherwise. P. 12, lines 4-6. It could state more clearly here that the 1.68 indicate that the mobile units increased uptake by 68% for this group relative to the reference group. Appendix 10 seems like it merits more discussion in the text. It seems that the results were quite sensitive to the threshold choice, with only a 1 to 1.5 km threshold producing significant results. The authors might discuss possible reasons for those results.

REVIEWER	Philip Britteon The University of Manchester
REVIEW RETURNED	08-Aug-2023

GENERAL COMMENTS	Thank you for your clarifications and thorough response to my comments.

VERSION 2 – AUTHOR RESPONSE

		Comment/ Recommendation details	**Responses**
Reviewer 2 **Dr. Gary W Reinbold, University of Illinois at Springfield**	**Recommendation**	The authors made an excellent effort to respond to the reviewers' comments and I am satisfied on all of the substantive issues. There are still a few places, however, where I think their language could be a little more precise.	We sincerely appreciate your thorough review of our manuscript and your positive feedback on our efforts to address the reviewers' comments. Your constructive feedback is invaluable to us, and we are committed to further improving the precision of our language. We have carefully reviewed your comments regarding areas where our language could be more precise, and we have taken actions accordingly to address your concerns, which are detailed below point-by-point. Additionally, we have conducted a thorough proofreading and editing process to address any remaining issues related to precision

			throughout the manuscript. We hope that these revisions better align the manuscript with your expectations.
	Comments	1. P.2, line 9: The treatment group could be described a little more precisely here. Not everyone in the treatment group "lived" within 1 km of a mobile site.	We have now reworded to "We defined intervention neighbourhoods as having their population weighted centroid located within 1km of mobile vaccination sites (338,006 individuals)" in Line 9-10 on Page 2.
		2. P.2, lines 17-20: Saying the effect was "smaller" implies that there was still an effect for these groups, but Figure 3 suggests that there was no effect for many of these groups. By the way, thank you to the authors for adding Figure 3, which I found to be very helpful.	We agree, as there are interactions with age, deprivation, and ethnicity, describing the subgroup effects simply is challenging. You are right in many of these subgroups there is no effect. We have reworded this to try and reflect that in Lines 17-19 on Page 2: "Interaction analyses showed smaller or no effect amongst older age groups, Asian and Black ethnic groups, and the most socioeconomically deprived populations."
		3. P.3, lines 10-12: If I understand correctly, the authors weren't able to include any individuals who had not been registered with a GP at the time of the study, not just individuals who used the mobile units.	Thanks a lot for pointing it out. This has now been reworded to "we were not able to include individuals that had not been registered with the GP at time of our study" in Lines 10-11 on Page 3.
		4. P.7, line 5: If I understand correctly, the 1112 non-intervention LSOAs also did not have population centroids within 1 km of a mobile unit.	This now have been reworded to "this left 1112 non-intervention LSOAs within the eight participating local authorities that did not have population weighted centroids within 1km of a mobile unit" in Lines 5-6 on Page 7.
		5. P. 8, lines 6-14: I understand that some of this language was added to respond to my prior comment, but I think the authors misunderstood my comment, which probably means that I	Thanks a lot for explaining this. We have now explicitly pointed out what we matched on and removed the discussion about collinearity in Lines 4-8 on Page 8. The current text reads "The weights were calculated using the raking method [34] so that the weighted averages for all the variables outlined above in the synthetic

		didn't state it clearly enough. It is obvious that the sum of (1) the cumulative number of first doses prior to week t - 7 and (2) the number of new first doses for weeks t – 7 through t - 1 would equal (3) the cumulative number of first doses after week t - 1. I only wanted the authors to be clearer on whether they are matching on just (2), on both (1) and (2), or on both (2) and (3). I understand now that they are matching on both (1) and (2), although the document still doesn't quite state that here. The authors could simply state that more clearly here and, if they would like, remove the added language about collinearity.	control group were the same as for the intervention. This included the cumulative number of first vaccine doses administered prior to the pre-intervention period, the number of first dose vaccines administered in each of the seven weeks prior to the intervention (Appendix 4) and each local area characteristic, outlined above".
		6. P. 9, lines 2-4: Table A4 indicates that IMD tercile was included in these regressions, but it isn't included in the list of variables here.	Thank you for pointing this out. We'd like to clarify that we focused on comparing the intervention and non-intervention areas on the aggregated level in this paragraph, so we presented the original measurement of IMD in Table 1. We have now made it clearer in Lines 3-5 on Page 10.
		7. P. 11, lines 21-25. Again, saying that there was a "lower impact" for these groups implies that there was still an impact, but Figure 3 suggests otherwise.	Thank you we have tried to rephrase this to make it clearer in Lines 18-25 on Page 11: "We found a significant interaction between deprivation and the intervention indicating reduced effectiveness with increased deprivation($p<0.001$), a significant interaction with ethnicity indicating reduced effectiveness for Asian/Asian British ($p=0.006$), Black/Black British ($p=0.005$) or other ethnic groups ($p=0.010$), compared to White/White British people, and a significant interaction with age indicating reduced effectiveness in older age groups. The combination of these interaction effects means that in many groups defined by age,

			ethnicity, and deprivation there was no evidence of effectiveness (see Figure 3)."
		8. P. 12, lines 4-6. It could state more clearly here that the 1.68 indicate that the mobile units increased uptake by 68% for this group relative to the reference group.	We'd like to clarify we refer to 18–30–year–olds of white ethnicity living in the least deprived neighbourhoods that were visited by the mobile vaccination unit as the reference group in Figure 3. We have now reworded Lines 6-9 on Page 12 to make it clearer.
		9. Appendix 10 seems like it merits more discussion in the text. It seems that the results were quite sensitive to the threshold choice, with only a 1 to 1.5 km threshold producing significant results. The authors might discuss possible reasons for those results.	We sincerely appreciate your insightful feedback and your attention to the sensitivity of our results to the threshold choice in our study. We acknowledge that the choice of a 1 to 1.5 km threshold was observed to produce significant results and recognise that this sensitivity merits careful consideration. In response to your suggestion, we have now added more explanation on possible reasons for the observed sensitivity to threshold choice in our revised manuscript in Lines 28-32 on Page 13.
Reviewer 3 **Dr. Philip Britteon, The University of Manchester**	**Recommendation**	Thank you for your clarifications and thorough response to my comments.	You're very welcome, and we appreciate your kind words. We are glad we could address your comments and provide the necessary clarifications. Your feedback has been invaluable in improving our work, and we are thankful for your time and expertise in reviewing it.

VERSION 3 – REVIEW

REVIEWER	Gary W Reinbold University of Illinois at Springfield
REVIEW RETURNED	19-Sep-2023

GENERAL COMMENTS	I have no further comments. Thank you for your thoughtful responses to my prior comments.