Human movement can inform the spatial scale of interventions against COVID-19 transmission

Authors: Hamish Gibbs[1*], Emily Nightingale[2], Yang Liu[1], James Cheshire[3], Leon Danon[4,5,6], Liam Smeeth[7], Carl AB Pearson[1], Chris Grundy[1], LSHTM CMMID COVID-19 working group[1], Adam J Kucharski[1] and Rosalind M Eggo[1*].

Affiliations:

[1]Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine, Keppel Street, London. WC1E 7HT. UK

[2]Department of Global Health and Development, London School of Hygiene & Tropical Medicine, Keppel Street, London. WC1E 7HT. UK

[3]Department of Geography, University College London, Gower St, Bloomsbury, London. WC1E 6BT. UK

[4]Department of Computer Science, University of Exeter, Stocker Rd, Exeter. EX4 4PY. UK

[5]The Alan Turing Institute, British Library, 96 Euston Rd, London NW1 2DB, UK.

[6]Population Health Sciences, University of Bristol, BS8, UK

[7]Faculty of Epidemiology and Population Health, London School of Hygiene & Tropical Medicine. UK

Corresponding author: [*] joint corresponding authors: hamish.gibbs@lshtm.ac.uk r.eggo@lshtm.ac.uk

Classification: Biological Sciences; Population Biology

Keywords: COVID-19; mobility; interventions; networks; community structure

Abstract

The UK enacted an intensive, nationwide lockdown on March 23 2020 to mitigate transmission of COVID-19. As restrictions began to ease, resurgence in transmission has been targeted by geographically-limited interventions of various stringencies. Determining the optimal spatial scale for local interventions is critical to ensure interventions reach the most at risk areas without unnecessarily restricting areas at low risk of resurgence. Here we use detailed human mobility data from Facebook to determine the spatially-explicit network community structure of the UK before and during the lockdown period, and how that has changed in response to the easing of restrictions and to locally-targeted interventions. We found that the mobility network became more sparse and the number of mobility communities decreased under the national lockdown. During this period, there was no evidence of re-routing in the network. Communities in which locally-targeted interventions have happened following resurgence did not show reorganization but did show small decreases in measurable mobility effects in the Facebook dataset. We propose that geographic communities detected in Facebook or other mobility data be part of decision making for determining the spatial extent or boundaries of interventions in the UK. These data are available in near real-time, and allow quantification of changes in the distribution of the population across the UK, as well as people's travel patterns to give data-driven metrics for geographically-targeted interventions.

Significance Statement

Large-scale intensive interventions in response to the COVID-19 pandemic have affected human movement patterns. Mobility data show spatially-explicit network structure, but it is not clear if that structure changed in response to national or locally-targeted interventions. We used daily Facebook for Good mobility data to quantify changes in the travel network in the UK during the national lockdown, and in response to local interventions. The network community structure inherent in these networks can help quantify which areas are at risk of resurgence, or the extent of locally-targeted interventions aiming to suppress transmission. We showed that spatial mobility data available in real-time can give information on connectivity that can be used to optimise the scale of geographically-targeted interventions.

Main Text

Introduction

Fine-scale geographic monitoring of large populations provides a valuable resource for increasing the accuracy and responsiveness of epidemiological modelling, outbreak response, and intervention planning in response to public health emergencies like the COVID-19 pandemic (1–5). Population and mobility datasets collected from the movement of individuals' mobile phones provide empirical, near-real time metrics of population movement between different geographic regions.(6) The COVID-19 pandemic response could potentially benefit from the availability of new data sources for measuring human movement, aggregated from mobile devices by network providers and popular applications including Google Maps, Apple Maps, Citymapper, and Facebook.(7)

Travel and movement behavior during epidemics may change in response to imposed interventions, perceived risk, and due to seasonal activities such as vacations (8,9). During the COVID-19 pandemic, mobility data has been used to assess adherence to movement restrictions (10,11), the impact of movement restrictions on the transmission dynamics of COVID-19 (12–14), the socioeconomic impacts of large scale movement restrictions (15,16).

These are typically retrospective, describing past movement patterns to understand their impact, although the use of movement datasets to assist in developing policy responses to target key populations at risk during a disease outbreak is increasing (17). Following the relaxation of nationwide restrictions in May 2020, the United Kingdom adopted a policy of targeted local interventions, aimed at reducing transmission in areas with resurgences, to avoid reimposing national restrictions (18). However, the effectiveness of these measures will depend, amongst other things, on how the interconnections between areas change over time, and how 'local' areas are defined. We used mobility data from Facebook to investigate changes in travel behaviour in response to national and local policy changes. Large-scale movement datasets quantify these interconnections and could inform the appropriate geographic extent of control measures.

In this analysis, we used Facebook Data for Good UK movement and population data (March 19 to October 16 2020), which records approximately 15 million daily locations of 4.8 million users (19), and UK census population, age, ethnicity, and socioeconomic deprivation data to understand the changes of travel behavior in response to initially stringent movement restrictions and subsequent easing between March and September 2020 (20). Using network analytic methods to understand the structure of interconnected communities in the movement network, and traced the evolution of these geographic communities through time, comparing them to intervention measures implemented in response to local outbreaks.

We also analysed temporal changes in population movements, identifying outflows from population centers preceding the implementation of movement restrictions, as well as patterns of increased movement most likely resulting from holiday travel. We determined limitations of the dataset for quantifying movement patterns, and discuss the implications of the identified movement patterns on future policy responses.

Results

Data representativity

We used anonymised, aggregated mobility and population data from Facebook (19). These data record the number of users who have opted-in to Location History data sharing with the Facebook mobile app in a ~5km^2 grid cell, giving the population size and the number of Facebook users travelling between cells, in eight-hour windows from March 19th to September 10th 2020 (Supplemental Figure 1). Before being shared with the research team, data were aggregated by Facebook from individual adult user geolocation trajectories into between-cell flux data, assigning users to a grid cell by their modal location in a given time period. Movement flux between squares that involved fewer than 10 users were not shared to preserve privacy (Supplemental Figure 2).

To understand how representative Facebook data were of the general population, we explored the size of the population of Facebook users included in the movement dataset, and compared this population to 2011 UK Census estimates. In the dataset, there were an average of 15.06 million daily movements, ranging from 15.86 million on April 7th, to 14.27 million on June 28th (Figure 1a).

The percentage of Facebook users per cell was comparable in the four nations of the UK (Fig 1b) and was fairly homogeneous across the entire study area (Fig 1d). The ratio of number of Facebook users and total population was relatively constant across cells (Fig 1c). There were no strong associations between the percentage of Facebook users and the average age, percent minority ethnic, population density, or index of multiple deprivation of each cell (Supplemental Figures 3-5).

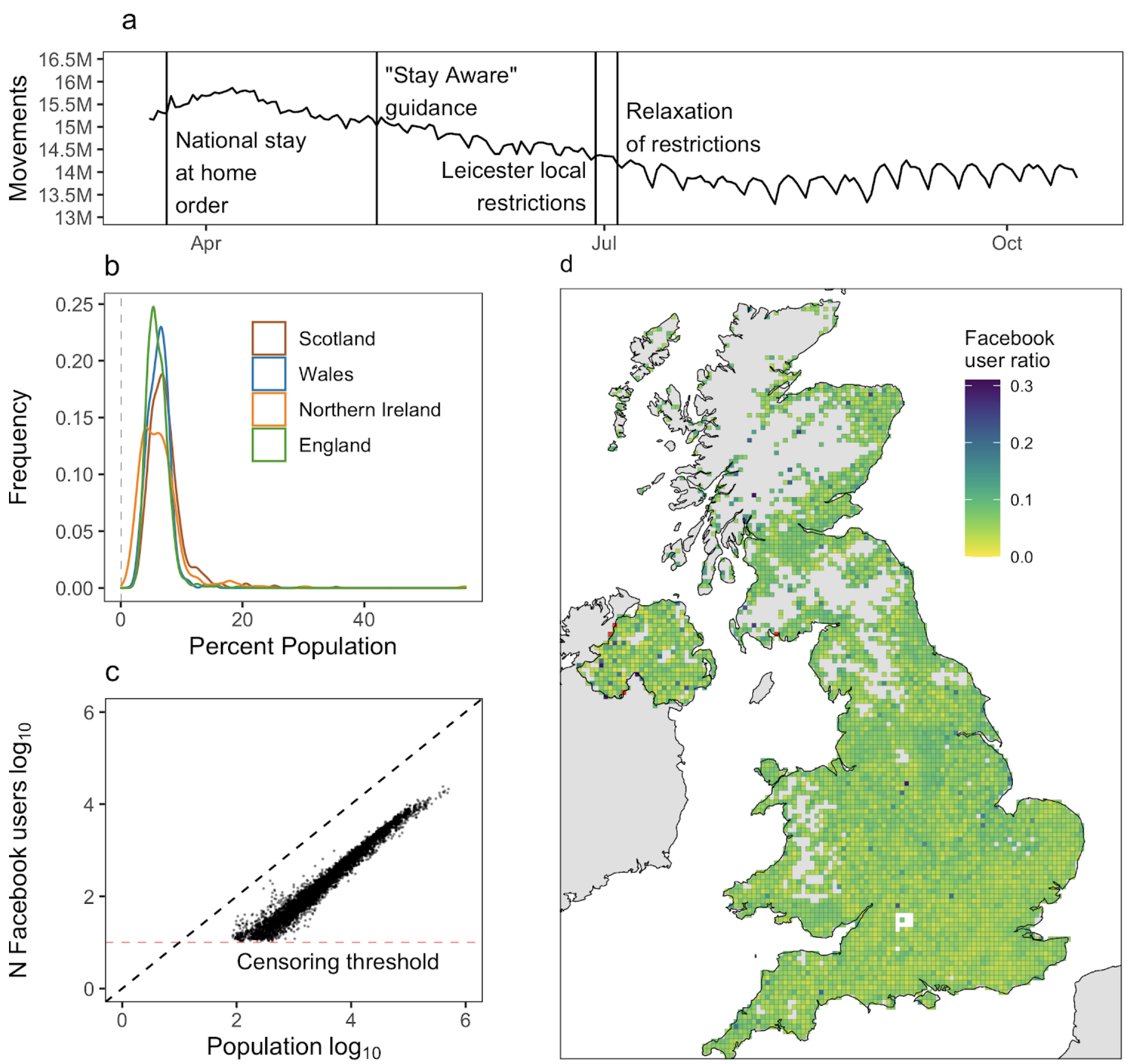

Figure 1. Characteristics of the between-cell Facebook mobility dataset. a) The daily total number of movements recorded. b) The probability density functions of the percentage of Facebook users in the census population for cells by country (the median of all countries, 6.09%, is displayed with a dashed line). c) The relationship between the number of Facebook users and census population for all cells. d) The spatial distribution of the percentage of Facebook users in the census population of each cell across the UK. Grey cells are generally missing because of low

numbers, which are suppressed for privacy reasons. Note however, 12 cells around the town of Swindon are missing due to a data processing error prior to data sharing (shown in white).

Network Structure

To quantify how the structure of the overall network changed during the first wave of the epidemic, we computed the edge betweenness centrality of connections between cells, a measure of the relative importance of a given connection in the network. There was a large difference in the centrality of journeys on days with high volumes of travel, with the majority of the UK travel network forming a single connected component, with clear links between major urban centres (Figure 2a). This contrasts with the centrality of connections on days of low travel volume, where central journeys only serve as connections between distinct network components (Figure 2b-c and Supplemental Figures 6-8).

To understand the changes in movement volume across the study period and identify similarities between travel networks of different dates we compared the network on each day using Canberra distance (21), which measures the difference between two matrices, where a smaller value means greater similarity. This comparison of network pairs shows the increasing dissimilarity of the network caused by decreased movement as lockdown measures were implemented in late March. The regeneration of the weekend difference in mobility after lockdown is clearly visible (Figure 2d). Similarities between weekend flows are observed throughout the time series with greater similarity between weekends and mid-lockdown travel, as overall volume of travel decreased during lockdown.

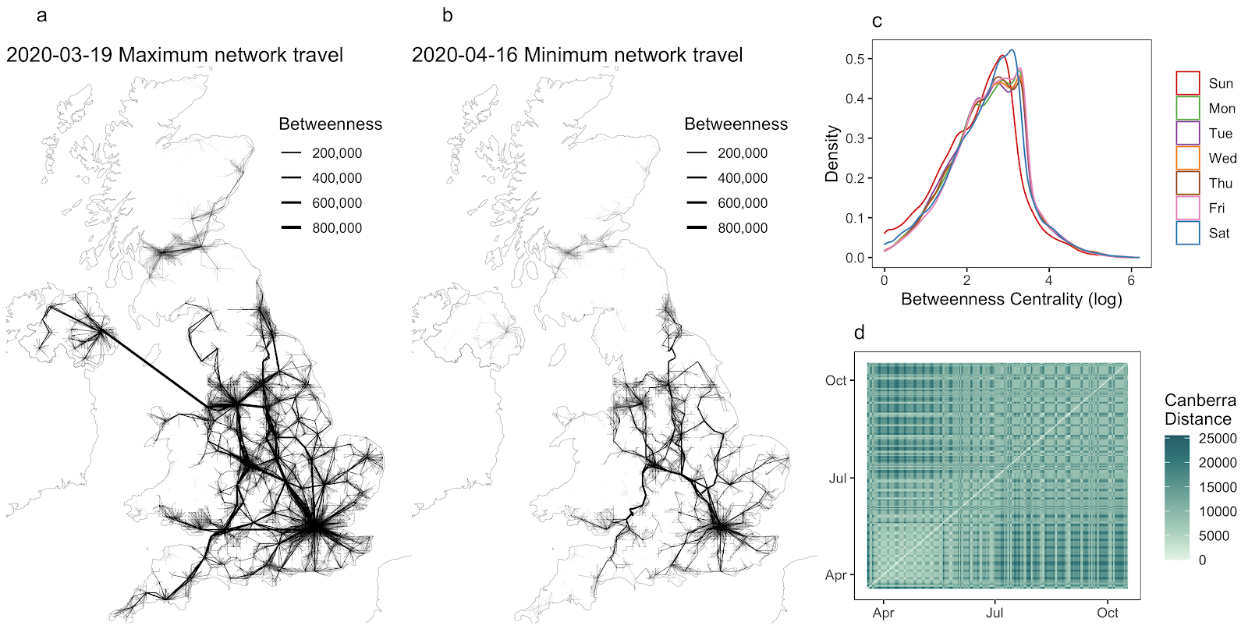

Figure 2. Measures of network structure. All network connections are displayed on Thursday March 19th, the day of maximum network travel (a) and April 16th, the Thursday closest to the day of minimum network travel (b). The thickness of edges corresponds to their betweenness centrality. c) The density of betweenness centrality for all journeys by weekday, showing a weekly pattern of lower betweenness on weekend days. d) Canberra distance, a measure of matrix similarity, between pairs of daily weighted, directed network adjacency matrices.

Community Detection

We used movement between cells to identify geographically-explicit "communities of interaction" in the weighted, directed network of user movements between cells, where the weights are the frequency of cell-to-cell movements. These communities partition the cells in the movement network into groups with relatively higher within-group connectivity than between-group connectivity.

To determine daily community structure we used the InfoMap algorithm (22), a flow-based community detection method demonstrated to have high accuracy and relatively low computational complexity for large networks (23). We conducted sensitivity analyses using the Leiden algorithm (24), finding close agreement between the spatial intersection and community boundaries identified by both methods (see Methods), where Leiden identified larger communities of which InfoMap communities were usually a subset (Supplemental Figures 9-10).

The geographic distribution of communities before the nationwide UK lockdown on March 23rd became fragmented during lockdown as the cell-level network became more sparse because fewer cells were reported by Facebook due to lower numbers. The geographic extent of communities was smaller (Figure 3a and Supplemental Figure 11). As restrictions on travel in the UK were eased in early July, the geographic community structure began returning to pre-lockdown number, geographic extent, and location of communities around population centers.

The community structure of the movement data changed on weekends, showing an increase in the number and a decrease in the size of communities, due to reduced travel on weekends and bank holidays (Figure 3b, Supplemental Figures 7-8). This fragmentation of the movement network on weekends reflects more local patterns of movement, and an absence of weekend commuting travel. This weekend effect was particularly pronounced in London, which on weekdays exhibited strong interconnectedness across the entire metropolitan area, but on weekends consistently fractured into four separate communities (Supplemental Figure 12).

The network structure during lockdown shares similarities in the number of communities and the volume of between tile-travel compared to typical weekends, where commuting travel is reduced, and overall travel between cells decreases. We did not observe reorganisation of the network around population centres through rerouting of connections, or other responses to the nationwide or local interventions.

We tracked the temporal evolution of communities using a heuristic approach (25), transferring community labels between time steps based on the proportion of community members shared between time steps. This approach allowed identification of the spatial extent of mobility communities extent and persistence through time. We found that the most persistent communities existed around residential areas, with the exception of London which split into more communities on weekends (Supplemental Figure 12-13). We also identified the communities which overlapped certain administrative geographies (Figure 4a,b), allowing us to combine these communities with data on the date and extent of local area interventions.

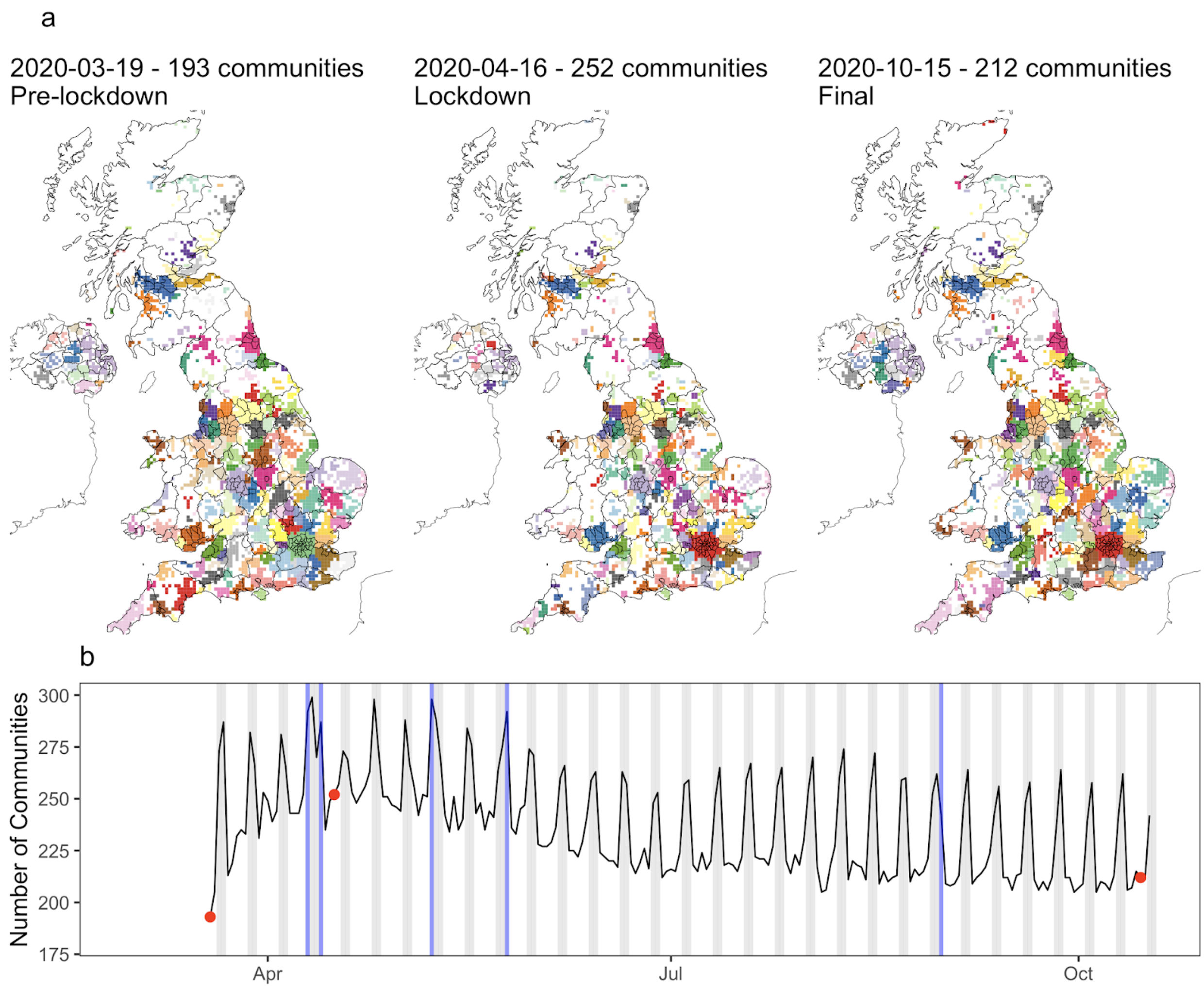

Figure 3. Community detection using the InfoMap algorithm. Communities detected on a) March 19th, the date of maximum network travel, April 16th, the same week day during nationwide lockdown, and October 15th. Missing tiles recorded fewer than 10 people moving and have been censored for privacy and appear white. All cells with data are assigned to a single community. b) The number of communities on each day, with weekends (grey lines) and bank holidays (blue lines), and the dates of snapshots in 1a (red points).

Local Lockdown Extents

Motivated by the need to identify communities associated with epidemic resurgences and responses to reactive interventions, we compared the extent and date of local interventions with the spatial extent and temporal persistence of network communities. We examined movement changes before, during, and after the first "local lockdown" in the UK, which happened around the city of Leicester on June 29th. The local intervention area was much smaller than the community containing Leicester (Figure 4a). We observed a decrease in internal movement in the Leicester community following the local intervention lasting approximately two weeks (Figure

4b, c). Other mobility datasets, such as that from Google (26), can give detailed information on the setting of interventions, and read in concert with Facebook mobility data, as well as case data, allow fine-scaled quantification of responses to local interventions (Figure 4d & e).

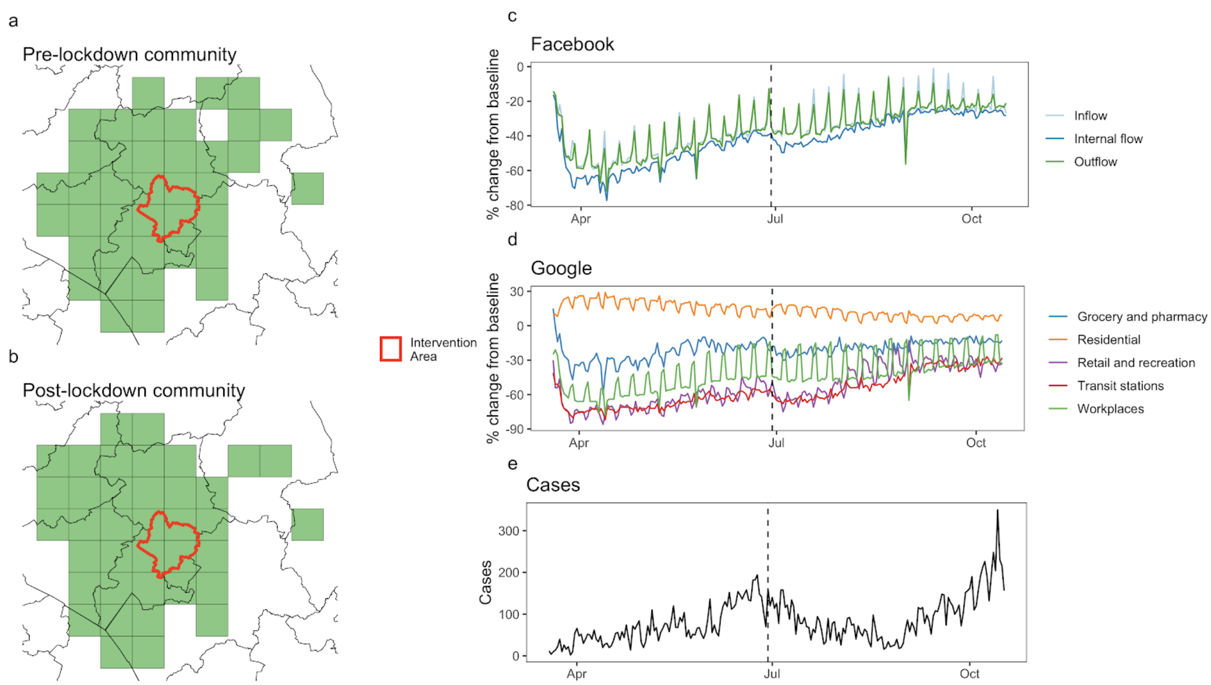

Figure 4. Changes in the Leicester community around the local intervention on June 29th. a) Geographic extent of the Leicester community (green) in the 2 weeks before local interventions, and the area of the interventions (red line). b) Same as a showing 2 weeks after local interventions. c) Movement recorded entering, leaving, or within Leicester in the Facebook mobility data. d) Movement recorded in different settings by Google mobility reports (26), and e) Confirmed COVID-19 cases in Leicester local authority. The date of local interventions is shown with a dashed line.

In the UK, Leicester was the first area to have a local intervention, and other areas followed with different rules and stringency. There was closer agreement between movement communities and the geographic extent of local area interventions from July onwards, particularly in Manchester (Supplemental Figure 14). Some interventions also spanned multiple movement communities, as in the North West and North East of England (Supplemental Figures 15-16). Early local intervention measures at limited spatial extents may not have fully encompassed the area of transmission resurgence, however UK policy has changed over time, and has begun to enforce collaborative local area interventions comprising multiple Local Authorities, which include constituent parts of mobility communities detected here. Additionally, movement communities evolve over time, and have the potential to shift following local area interventions, requiring an understanding of real-time patterns of movement to monitor the appropriateness of a given measure.

Changes in Facebook population distribution

The major holiday season in the UK occurred between July and September, and we used Facebook population data, collected in higher resolution cells than mobility data, to determine if there was a reorganization of population distribution during this period. We found a decrease in the number of Facebook users in some areas preceding the enforcement of lockdown restrictions in March (Fig 5a & b), which were generally urban areas (Supplemental figure 17). We also found a sizable summer increase in Facebook population in holiday destinations, most notably in Cornwall (Figure 5).

The same pattern of high volumes of travel to tourist sites can be observed when analysing the maximum deviations from baseline movement during weekends (Friday, Saturday, Sunday) recorded by the Facebook movement dataset. Those journeys which experience greater than 100% deviation from baseline on Fridays, Saturdays, or Sundays identify major holiday destinations, Cornwall, Brighton, and Blackpool, and North Berwick (Supplemental Figure 18) (27).

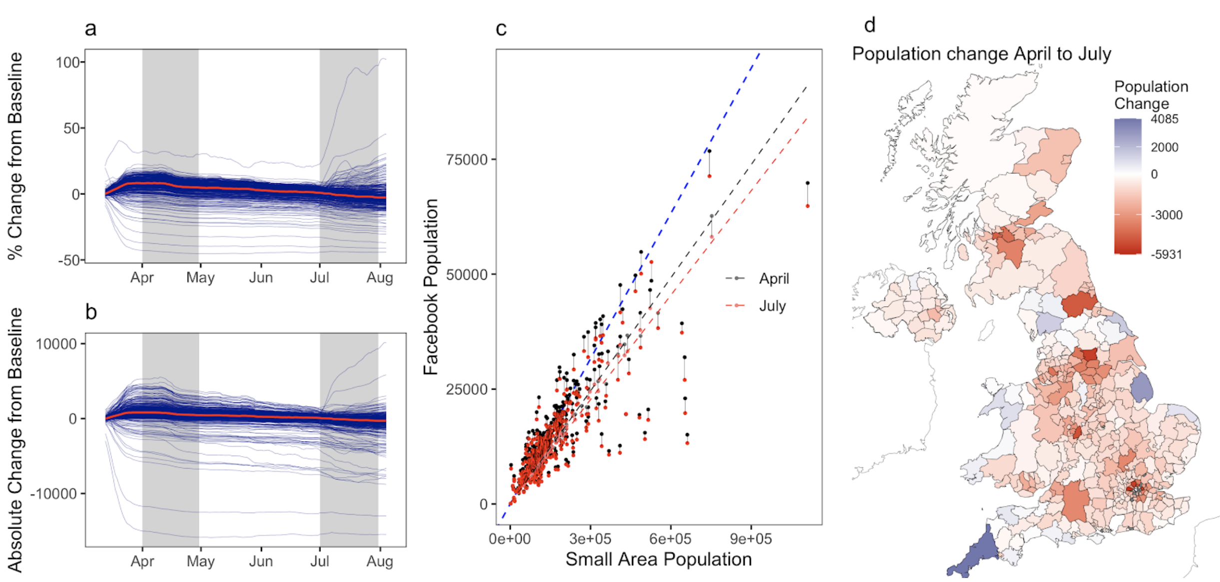

Figure 5. Population changes in the UK. The percent (a) and absolute (b) population change from baseline for each of 406 local authorities through time, the largest increase in population is in Cornwall, with the largest decreases in London boroughs. Red lines show the median daily population change for all local authorities and grey windows mark April and July. c) The relationship between Facebook user population and census small area population

estimates for April and July, shows the average proportion of Facebook users compared to census small area population. Blue dashed line shows the relationship over all months, and blak and red show relationship in April and July. d) A map of the absolute difference between population estimates in April and July. Note that the overall number of movements recorded by Facebook also decreased between April and July (Figure 1).

Discussion

This study used a large, anonymized movement dataset to detect geographically-explicit community structure in the mobility network, and assessed how those communities were affected by interventions. Using these communities we can delineate local areas of high mobility that can be used to determine the spatial scale of locally-targeted interventions to mitigate disease resurgences during the COVID-19 pandemic. We also explored the structure of the UK travel network through the pandemic, identifying variations in the central connections between population centres and changes in the structure of closely connected communities.

Gridded mobility datasets such as those made available by Facebook provide granular, near real-time information about the movement patterns of a large sample of the population. While these datasets could usefully inform epidemic responses (15,28–30), there remain questions about the generalizability of the movement recorded in these datasets to the overall population movements in the UK (31,32). The privacy preserving structure of the Facebook movement dataset means that low frequency journeys are not recorded, precise locations are replaced by modal locations, and data is provided in a grid so cells vary in population size. By comparing to Google mobility data, we showed that this can obscure changes in mobility occurring at a finer spatial scale, and thus limits the granularity at which local patterns can be identified from one data source alone. Ideally, multiple mobility data can be analysed in concert to assess changes in response to interventions.

Using a network analysis of mobility data, we identified geographically distinct communities with strong interconnections , which are a potentially more relevant target area for coordinated responses than traditional administrative boundaries. We found that these communities were stable around population centres between March and October (Supplemental Figure 13), but can be affected by public health interventions such as national travel restrictions. If restrictions are changed in the UK, the boundaries and size of communities could change, especially if home-working decreases and commuting behaviour returns to previous levels.

Following COVID-19 resurgence in a particular area, determining the geographic limit of reactive interventions should be driven by areas at risk of increased transmission, which may not intersect with administrative boundaries. We found that while communities tended to stabilise around populous areas, there was disagreement between the extent of these communities and the boundaries at which local area interventions have been introduced in the UK. This demonstrates the potential value of the approach for the formulation of policies to contain outbreaks or resurgence.

In response to the nationwide lockdown starting March 23rd, the UK mobility network changed drastically. The movement community structure over time became more sparse and the patterns appeared similar to those present on weekends prior to the lockdown. We did not find evidence of large scale restructuring or rerouting, either in communities or in the betweenness-structure of edges in the network during this time. We found variations in the size and interconnection of communities in relation to changes in overall travel and small decreases in travel following local area interventions. The summer period brought large changes in Facebook population sizes in the UK, with a decrease in population around major population centers and an increased population in some rural areas (most notably in Cornwall) that tend to be holiday destinations. Quantifying these changes in real-time is critical for informed decision-making about the scale of interventions and areas at risk of introductions.

There are several caveats to the methods of community detection used in this study, as the extent of communities could be influenced by the level of aggregation of the Facebook mobility data, and cells were assigned to a single community each day. We conducted a sensitivity analysis using two methods for identifying communities but there are a wide variety of community detection algorithms which emphasize different aspects of network structure. Questions also remain about the general reliability of community detection methods, developed on well understood network structures, applied to real-world networks (33). The effect of local area interventions on travel depends on the specifics of each intervention and their stringency. Additionally, interventions occur at multiple spatial scales, and across overlapping time periods. For example, in the UK, national interventions coincide with local interventions, and each may contribute differently to changes in movement behaviour. Finally, we assessed the representativeness of the Facebook data at a geographic level, limited by the available levels of aggregation of both the Facebook and census data sets.

Conclusion

Data-driven approaches using mobility data can provide evidence for defining the spatial scale of geographically-targeted public health interventions.

Methods

Facebook Data

Data provided by the Facebook Data for Good partner program (19) makes use of aggregated and anonymized user data to create a number of different data products. In this study, we used the movement and population data which is computed by Facebook using the geolocation of users with location services actively enabled. Data used in this study were provided as cell-to-cell movements between Bing Maps cells in England, Wales, Scotland, and Northern Ireland, aggregated in 8-hour time windows from March 10th to September 10th 2020. We used the number of users making journeys within and between map cells. A user's location is defined as their modal location in a map cell in sequential 8 hour periods, which defines the beginning and end points of journeys within or between cells. Journeys with fewer than 10 travellers were removed by Facebook to preserve privacy. Any cell that does not record any journeys with greater than 10 travellers in a given time point is omitted from the dataset, regardless of whether that cell recorded an internal number of users greater than 10. Facebook population data records the number of active users in each cell during a certain 8 hour period. In our network analysis, nodes are cells, edges exist when there are movements from one cell to another, and weights are the frequency of travellers along each edge.

Bing Maps Tile System

The Bing Maps Tile System is a standard geospatial reference used primarily for serving web maps and cells of geospatial raster data at varying zoom levels (34). The system is divided into 23 zoom levels ranging from global level 1, (map scale: 1:295,829,355.45) to detailed level 23, (map scale: 1:70.53). Each Bing Map cell is referenced with a “quadkey”, or unique identifier of the zoom level and pixel coordinates of an individual cell. In this analysis, all mobility, population, and other census datasets were referenced to Bing Maps cells using a unique quadkey identifier. Facebook mobility datasets were referenced to Bing Maps cell zoom level 12 (approximately 4.8 to 6.2 km^2 in the UK - measured at 60.77° and 50.59° respectively). Facebook population datasets were referenced to Bing Maps cells at zoom level 13 (approximately 2.4 to 3.1 km^2 in the UK - measured at 60.77° and 50.59° respectively). The ground resolution of Bing Maps cells varies with latitude, with cells at higher latitudes covering a smaller ground area than those at lower latitudes. This distortion results from the distortion inherent in the Web Mercator projection (EPSG:3857) used by the Bing Maps Tile System.

Demographic information

We compared the age distribution, population, ethnicity, and socioeconomic deprivation of each cell to the population of Facebook users to determine if the percentage of users varied by these demographic factors. We extracted these variables from national statistics agencies (Office for National Statistics, Northern Ireland Statistics and Research Agency, Scottish Government, and Welsh Government) and aggregated to Bing Tile level (cells). Census variables age, ethnicity, and socioeconomic deprivation data (Index of Multiple Deprivation in England & Wales) were referenced to different statistical units by country. In Northern Ireland, census variables were referenced to Super Output areas (SOAs), in England and Wales, Lower Super Output areas (LSOAs), and in Scotland, 2011 Data Zones (DZs). Detailed population data was also collected from national statistics agencies, providing a measure of population for Small Areas (Northern Ireland), Output Areas (OA; England and Wales), and Data Zones (Scotland).

Census variables referenced to different national statistical areas were aggregated to align with mobility datasets at Bing Tile level 12. First, we combined the 2011 population weighted centroids of each OA (or equivalent) from the UK Census with 2020 mid-year population estimates in each UK country. We then assigned each OA centroid to the Bing Maps level 12 cell it falls within. We then joined 2011-derived census variables (Age, Ethnicity and Socioeconomic Deprivation) to the OA centroids and computed an average of each census variable for each Bing Maps cell, weighted by the OA population estimates. For socioeconomic deprivation data (recorded as ranks) we ranked the weighted average values to create a rank of cells by their population weighted deprivation. OAs are much more granular than the Bing cells and therefore nested within them in the majority of cases, minimising the risk of the cells detrimentally intersecting OAs during the demographic assignment.

To assess the correlation between census variables and the proportion of Facebook users in each cell, we computed the Pearson correlation coefficient and two-sided p-values between the proportion of Facebook users in a cell and each census variable.

Temporal aggregation

Both Facebook movement and population datasets include a baseline prediction, estimating the volume of travellers along a given journey (or within a given cell) considering “normal travel patterns”. Both the baseline and observed number of users is recorded in 8 hour intervals. These data display strong and consistent intraday and intraweek patterns. To isolate changes in daily mobility, data collected in 8 hour periods were aggregated to daily periods by taking the sum of the observed and expected number of travellers along a journey for all periods within a day.

The baseline volume of travel was defined by Facebook as the mean volume of travel between two cells during the same 8-hour period on each day of the week computed using the 45 days prior to the creation of a data collection. For example, the baseline value for journeys from Cell A to Cell B on Tuesdays between 00:00 to 08:00 is the mean of the number of users who travelled between cell A and cell B on Tuesdays between 00:00 to 08:00 during the 45 day reference period. Any journeys that do not appear in the baseline period are not included in the movement dataset and outliers are winsorized by Facebook prior to data sharing. This method of computing baseline values was originally developed for rapid responses to natural disasters, but has limitations for long time series as the definition of a baseline reference period can be arbitrary based on the date of a data collection's creation. For Facebook mobility data in the UK, baseline travel values were computed by Facebook using data from January 29th to March 9th 2020.

Facebook Population

The population of Facebook users, recorded in level 13 Bing cells, was aggregated to the boundaries of Local Authorities. We computed the spatial intersection of local authorities and Bing level 13 cell centroids. Population data intersecting a given cell were then summed to compute the population of Facebook users in each Local Authority boundary.

Community Detection

Community detection methods are algorithms for identifying groups of meaningfully connected vertices. Many methods exist, with various tradeoffs on computational performance, resolution, or other characteristics (23,33,35,36). In this study, we employed two different algorithms, InfoMap and Leiden. InfoMap assesses the movement of a random walker around a network, identifying the partition of the network that minimizes the description length required to describe the movements of the walker (22). The Leiden algorithm maximizes the modularity of different node partitions, identifying the partition for which communities possess stronger connections to community members than to other nodes (24). We used the Leiden implementation in the *leidenalg* Python package with resolution optimised by the algorithm (37).

We tested the effect of the community detection algorithm, and found that they aligned hierarchically, where the Leiden algorithm identified geographically larger communities. If the communities detected by a one method are largely a superset of the communities detected by another, with shared boundaries between the defined communities, this likely represents a differing hierarchical structure, compared to a different interpretation of community structure. We assessed the agreement between community detection methods to understand the stability of detected communities by comparing the proportion of nodes in each community detected using InfoMap with all communities determined using Leiden, and vice versa (Supplemental Figure 9). This comparison allows for the computation of the proportion of shared nodes between both algorithms. The maximum and mean overlap of communities in each algorithm helped to identify the agreement between each method of community detection. In general we found that Leiden detected larger communities, for which the InfoMap communities were (for the most part) sub-communities.

Community Label Inheritance

The community detection methods used in this study identified communities each day. To track the evolution of communities over the study period, we employed a heuristic approach, assigning the label of a given community identified in a certain time step to that community with the highest number of shared nodes in the following time step (25). When multiple communities in a certain time step “claim” the same community in the following timestep, the community with the closest size to the community in the following timestep “wins” the right to pass its own label to the following time step. This situation typically occurs when a large community incorporates members from smaller communities. When a community divides into a collection of smaller communities, the community with the largest proportion of shared nodes inherits the original community label while other communities are assigned new unique identifiers.

Canberra Distance

To assess the differences in network structure between travel matrices, we used Canberra distance, a distance metric for comparing the similarity between pairs of matrices. This metric describes the distance between two vectors in n-dimensional space. To construct symmetric matrices with identical members through time (including censored edges), all cells which recorded travel at any point in the time series were transformed into an empty symmetric matrix. Matrix values were then inserted from the movement data on each day. Canberra distance was computed for each pair combination of movement matrices, resulting in a Canberra distance value comparing each pair of dates in the time series.

COVID-19 Data

We used confirmed COVID-19 cases from the UK Pillar 1 and Pillar 2 testing schemes. Data were available with data of specimens at the Lower Tier Local Authority level (38).

Google Mobility Data

We used data from Google mobility reports (26), a freely available service giving information on trends in movements to locations of 6 broad types: Retail and Recreation, Supermarket and Pharmacy, Parks, Public Transport, Workplaces, and Residential. We excluded Parks from the analysis, because they are likely low risk for transmission.

Code and Data Availability

All code is available on Github at https://github.com/hamishgibbs/facebook_mobility_uk

Ethics Approval

This research was approved by the LSHTM Observational Research Ethics Committee (ref 16834-1).

Author Contributions

RME conceived the project, and planned the methods and analytical plan with HG, CABP, EN, and YL. JC and LS contributed to analysis of census variables. CABP and LD contributed to community detection methods and analysis of the methods. JC and CG contributed to analysis of spatial data. YL and EN contributed to statistical analysis of the mobility, census, and case data. AJK acquired the data and contributed to the analysis and understanding of Facebook data. HG wrote and implemented the majority of the code, and made the figures, with input from all authors. HG and RME wrote the first draft, to which all authors contributed. All authors interpreted the findings and were involved in the analysis and decision to publish.
Facebook had no role in study design, the analysis, or decision to publish.

Acknowledgements

The following funding sources are acknowledged as providing funding for the named authors. This research was partly funded by the Bill & Melinda Gates Foundation (INV-003174: YL; NTD Modelling Consortium OPP1184344: CABP; OPP1183986: ESN). DFID/Wellcome Trust (Epidemic Preparedness Coronavirus research programme 221303/Z/20/Z: CABP). This project has received funding from the European Union's Horizon 2020 research and innovation programme - project EpiPose (101003688: YL). HDR UK (MR/S003975/1: RME). UK DHSC/UK Aid/NIHR (ITCRZ 03010: HPG). UK MRC (MC_PC_19065: RME, YL). Wellcome Trust (206250/Z/17/Z: AJK), UK MRC (MC_PC_19067, LD), The Alan Turing Institute under the EPSRC (grant EP/N510129/1: LD). This research was partly funded by the National Institute for Health Research (NIHR) using UK Aid from the UK Government to support global health research. The views expressed in this publication are those of the author(s) and not necessarily those of the NIHR or the UK Department of Health and Social Care (16/137/109: YL; NIHR200908: RME).

The following authors were part of the Centre for Mathematical Modelling of Infectious Disease COVID-19 working group. Each contributed in processing, cleaning and interpretation of data, interpreted findings, contributed to the manuscript, and approved the work for publication: Sam Abbott, Kaja Abbas, Kiesha Prem, Sebastian Funk, Jon C Emery, Georgia R Gore-Langton, Fiona Yueqian Sun, Arminder K Deol, Alicia Showering, Nicholas G. Davies, Nikos I Bosse, Samuel Clifford, Anna M Foss, Graham Medley, C Julian Villabona-Arenas, Timothy W Russell, Amy Gimma, W John Edmunds, Gwenan M Knight, Yung-Wai Desmond Chan, Yalda Jafari, Quentin J Leclerc, Rein M G J Houben, Akira Endo, Sophie R Meakin, Petra Klepac, Joel Hellewell, Naomi R Waterlow, Kevin van Zandvoort, Christopher I Jarvis, Rachel Lowe, Matthew Quaife, Charlie Diamond, Megan Auzenbergs, Simon R Procter, Rosanna C Barnard, Oliver Brady, Katherine E. Atkins, Katharine Sherratt, Thibaut Jombart, Stéphane Hué, Kathleen O'Reilly, Jack Williams, David Simons, Stefan Flasche, Mark Jit, James D Munday, Billy J Quilty, Frank G Sandmann, Damien C Tully, James W Rudge, Alicia Rosello.

The following funding sources are acknowledged as providing funding for the working group authors. Alan Turing Institute (AE). BBSRC LIDP (BB/M009513/1: DS). This research was partly funded by the Bill & Melinda Gates Foundation (INV-001754: MQ; INV-003174: KP, MJ; NTD Modelling Consortium OPP1184344: GFM; OPP1180644: SRP; OPP1191821: KO'R, MA). BMGF (OPP1157270: KA). DFID/Wellcome Trust (Epidemic Preparedness Coronavirus research programme 221303/Z/20/Z: KvZ). DTRA (HDTRA1-18-1-0051: JWR). Elrha R2HC/UK DFID/Wellcome Trust/This research was partly funded by the National Institute for Health Research (NIHR) using UK aid from the UK Government to support global health research. The views expressed in this publication are those of the author(s) and not necessarily those of the NIHR or the UK Department of Health and Social Care (KvZ). ERC Starting Grant (#757699: JCE, MQ, RMGJH). This project has received funding from the European Union's Horizon 2020 research and innovation programme - project EpiPose (101003688: KP, MJ, PK, RCB, WJE). This research was partly funded by the Global Challenges Research Fund (GCRF) project 'RECAP' managed through RCUK and ESRC (ES/P010873/1: AG, CIJ, TJ). MRC (MR/N013638/1: NRW). Nakajima Foundation (AE). NIHR (16/136/46: BJQ; 16/137/109: BJQ, CD, FYS, MJ; Health Protection Research Unit for Immunisation NIHR200929: NGD; Health Protection Research Unit for Modelling Methodology HPRU-2012-10096: TJ; NIHR200929: FGS, MJ; PR-OD-1017-20002: AR, WJE). Royal Society (Dorothy Hodgkin Fellowship: RL; RP\EA\180004: PK). UK MRC (LID DTP MR/N013638/1: GRGL, QJL; MC_PC_19065: AG, NGD, SC, TJ, WJE; MR/P014658/1: GMK). Authors of this research receive funding from UK Public Health Rapid Support Team funded by the United Kingdom Department of Health and Social Care (TJ). Wellcome Trust (206250/Z/17/Z: TWR; 206471/Z/17/Z: OJB; 208812/Z/17/Z: SC, S Flasche; 210758/Z/18/Z: JDM, JH, KS, NIB, SA, SFunk, SRM). No funding (AKD, AMF, AS, CJVA, DCT, JW, KEA, SH, YJ, YWDC).

References

1. The duration of travel impacts the spatial dynamics of infectious diseases | PNAS [Internet]. [cited 2020 Sep 14]. Available from: https://www.pnas.org/content/117/36/22572
2. Quantifying seasonal population fluxes driving rubella transmission dynamics using mobile phone data | PNAS [Internet]. [cited 2020 Sep 17]. Available from: https://www.pnas.org/content/112/35/11114.short
3. Kramer AM, Pulliam JT, Alexander LW, Park AW, Rohani P, Drake JM. Spatial spread of the West Africa Ebola epidemic. R Soc Open Sci. 3(8):160294.
4. Wesolowski A, Eagle N, Tatem AJ, Smith DL, Noor AM, Snow RW, et al. Quantifying the Impact of Human Mobility on Malaria. Science. 2012 Oct 12;338(6104):267–70.
5. Balcan D, Gonçalves B, Hu H, Ramasco JJ, Colizza V, Vespignani A. Modeling the spatial spread of infectious diseases: The GLobal Epidemic and Mobility computational model. J Comput Sci. 2010 Aug 1;1(3):132–45.
6. González MC, Hidalgo CA, Barabási A-L. Understanding individual human mobility patterns. Nature. 2008 Jun;453(7196):779–82.
7. Oliver N, Lepri B, Sterly H, Lambiotte R, Deletaille S, Nadai MD, et al. Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle. Sci Adv. 2020 Jun 1;6(23):eabc0764.
8. Funk S, Salathé M, Jansen VAA. Modelling the influence of human behaviour on the spread of infectious diseases: a review. J R Soc Interface. 2010 Sep 6;7(50):1247–56.
9. Meloni S, Perra N, Arenas A, Gómez S, Moreno Y, Vespignani A. Modeling human mobility responses to the large-scale spreading of infectious diseases. Sci Rep. 2011 Aug 12;1(1):62.
10. Pullano G, Valdano E, Scarpa N, Rubrichi S, Colizza V. Population mobility reductions during COVID-19 epidemic in France under lockdown. medRxiv. 2020 Jun 1;2020.05.29.20097097.
11. Pepe E, Bajardi P, Gauvin L, Privitera F, Lake B, Cattuto C, et al. COVID-19 outbreak response: a first assessment of mobility changes in Italy following national lockdown. medRxiv. 2020 Apr 7;2020.03.22.20039933.
12. Anderson RM, Heesterbeek H, Klinkenberg D, Hollingsworth TD. How will country-based mitigation measures influence the course of the COVID-19 epidemic? The Lancet. 2020 Mar 21;395(10228):931–4.
13. Chinazzi M, Davis JT, Ajelli M, Gioannini C, Litvinova M, Merler S, et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. Science. 2020 Apr 24;368(6489):395–400.
14. Ferguson N, Laydon D, Nedjati Gilani G, Imai N, Ainslie K, Baguelin M, et al. Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand [Internet]. Imperial College London; 2020 Mar [cited 2020 Sep 17]. Available from: http://spiral.imperial.ac.uk/handle/10044/1/77482
15. Economic and social consequences of human mobility restrictions under COVID-19 | PNAS [Internet]. [cited 2020 Sep 17]. Available from: https://www.pnas.org/content/117/27/15530.short
16. Reductions in commuting mobility correlate with geographic differences in SARS-CoV-2 prevalence in New York City | Nature Communications [Internet]. [cited 2020 Sep 17]. Available from: https://www.nature.com/articles/s41467-020-18271-5

17. Buckee CO, Balsari S, Chan J, Crosas M, Dominici F, Gasser U, et al. Aggregated mobility data could help fight COVID-19. Science. 2020 Apr 10;368(6487):145–6.
18. Coronavirus: local lockdowns [Internet]. The Institute for Government. 2020 [cited 2020 Sep 17]. Available from: https://www.instituteforgovernment.org.uk/explainers/coronavirus-local-lockdowns
19. Disease Prevention Maps [Internet]. Facebook Data for Good. [cited 2020 Sep 17]. Available from: https://dataforgood.fb.com/tools/disease-prevention-maps/
20. COVID-19 policy tracker [Internet]. The Health Foundation. [cited 2020 Sep 23]. Available from: https://www.health.org.uk/news-and-comment/charts-and-infographics/covid-19-policy-tracker
21. Lance GN, Williams WT. Computer Programs for Hierarchical Polythetic Classification (“Similarity Analyses”). Comput J. 1966 May 1;9(1):60–4.
22. Bohlin L, Edler D, Lancichinetti A, Rosvall M. Community Detection and Visualization of Networks with the Map Equation Framework. In: Ding Y, Rousseau R, Wolfram D, editors. Measuring Scholarly Impact [Internet]. Cham: Springer International Publishing; 2014 [cited 2020 Sep 11]. p. 3–34. Available from: http://link.springer.com/10.1007/978-3-319-10377-8_1
23. Yang Z, Algesheimer R, Tessone CJ. A Comparative Analysis of Community Detection Algorithms on Artificial Networks. Sci Rep. 2016 Aug 1;6(1):30750.
24. Traag VA, Waltman L, van Eck NJ. From Louvain to Leiden: guaranteeing well-connected communities. Sci Rep. 2019 Dec;9(1):5233.
25. Gao Y, Zhu Z, Kali R, Riccaboni M. Community evolution in patent networks: technological change and network dynamics. Appl Netw Sci. 2018 Aug 13;3(1):26.
26. COVID-19 Community Mobility Reports [Internet]. [cited 2020 Oct 22]. Available from: https://www.google.com/covid19/mobility/
27. Coastal towns in England and Wales - Office for National Statistics [Internet]. [cited 2020 Oct 6]. Available from: https://www.ons.gov.uk/businessindustryandtrade/tourismindustry/articles/coastaltownsinenglandandwales/2020-10-06
28. Effects of human mobility restrictions on the spread of COVID-19 in Shenzhen, China: a modelling study using mobile phone data - The Lancet Digital Health [Internet]. [cited 2020 Oct 6]. Available from: https://www.thelancet.com/journals/landig/article/PIIS2589-7500(20)30165-5/fulltext
29. Weill JA, Stigler M, Deschenes O, Springborn MR. Social distancing responses to COVID-19 emergency declarations strongly differentiated by income. Proc Natl Acad Sci. 2020 Aug 18;117(33):19658–60.
30. Gibbs H, Liu Y, Pearson CA, Jarvis CI, Grundy C, Quilty BJ, et al. Changing travel patterns in China during the early stages of the COVID-19 pandemic [Internet]. Public and Global Health; 2020 May [cited 2020 Aug 27]. Available from: http://medrxiv.org/lookup/doi/10.1101/2020.05.14.20101824
31. Jeffrey B, Walters CE, Ainslie KEC, Eales O, Ciavarella C, Bhatia S, et al. Anonymised and aggregated crowd level mobility data from mobile phones suggests that initial compliance with COVID-19 social distancing interventions was high and geographically consistent across the UK. Wellcome Open Res. 2020 Jul 17;5:170.
32. The Limited Value of Non-Replicable Field Experiments in Contexts With Low Temporal Validity - Kevin Munger, 2019 [Internet]. [cited 2020 Oct 6]. Available from: https://journals.sagepub.com/doi/full/10.1177/2056305119859294

33. Riolo MA, Newman MEJ. Consistency of community structure in complex networks. Phys Rev E. 2020 May 8;101(5):052306.
34. rbrundritt. Bing Maps Tile System - Bing Maps [Internet]. [cited 2020 Sep 17]. Available from: https://docs.microsoft.com/en-us/bingmaps/articles/bing-maps-tile-system
35. Good BH, de Montjoye Y-A, Clauset A. The performance of modularity maximization in practical contexts. Phys Rev E. 2010 Apr 15;81(4):046106.
36. Danon L, Díaz-Guilera A, Duch J, Arenas A. Comparing community structure identification. J Stat Mech Theory Exp. 2005 Sep 26;2005(09):P09008–P09008.
37. Traag VA. leidenalg: Leiden is a general algorithm for methods of community detection in large networks. [Internet]. [cited 2020 Oct 23]. Available from: https://github.com/vtraag/leidenalg
38. Coronavirus (COVID-19) in the UK: Cases [Internet]. [cited 2020 Oct 22]. Available from: https://coronavirus.data.gov.uk/cases?areaType=ltla&areaName=Northumberland