Global maps of travel time to healthcare facilities

D. J. Weiss [1,2,3] ✉, A. Nelson [4], C. A. Vargas-Ruiz[1], K. Gligorić[5], S. Bavadekar[6], E. Gabrilovich [6], A. Bertozzi-Villa[1,7], J. Rozier[2], H. S. Gibson [1], T. Shekel[6], C. Kamath[6], A. Lieber[6], K. Schulman[8], Y. Shao[9], V. Qarkaxhija[10], A. K. Nandi [1], S. H. Keddie[2], S. Rumisha [1], P. Amratia[1], R. Arambepola[1], E. G. Chestnutt[1], J. J. Millar [1], T. L. Symons[1], E. Cameron[2,3], K. E. Battle[7], S. Bhatt [11] and P. W. Gething[2,3]

Access to healthcare is a requirement for human well-being that is constrained, in part, by the allocation of healthcare resources relative to the geographically dispersed human population[1–3]. Quantifying access to care globally is challenging due to the absence of a comprehensive database of healthcare facilities. We harness major data collection efforts underway by OpenStreetMap, Google Maps and academic researchers to compile the most complete collection of facility locations to date. Leveraging the geographically variable strengths of our facility datasets, we use an established methodology[4] to characterize travel time to healthcare facilities in unprecedented detail. We produce maps of travel time with and without access to motorized transport, thus characterizing travel time to healthcare for populations distributed across the wealth spectrum. We find that just 8.9% of the global population (646 million people) cannot reach healthcare within one hour if they have access to motorized transport, and that 43.3% (3.16 billion people) cannot reach a healthcare facility by foot within one hour. Our maps highlight an additional vulnerability faced by poorer individuals in remote areas and can help to estimate whether individuals will seek healthcare when it is needed, as well as providing an evidence base for efficiently distributing limited healthcare and transportation resources to underserved populations both now and in the future.

Access to healthcare is a measure of human well-being that is constrained by numerous geographically varying factors[1–3], the most immediate of which is the time it takes individuals to travel to a properly equipped and adequately staffed healthcare facility. Due to spatial clustering of healthcare facilities in densely populated areas, individuals living in rural regions often face increased travel times and thus cost when seeking healthcare. This situation can be exacerbated by poor transportation infrastructure and lack of motorized transport, which further increase the time required for travel and could disproportionally affect lower-income populations. As such, people facing long travel times to healthcare facilities are less likely to seek care when it is needed[5–9], and the consequences of failing to seek care include increased mortality and morbidity from treatable conditions[10,11].

Effectively characterizing travel time can assist in identifying communities that would most benefit from additional healthcare resources. These resources include outreach programs, community healthcare workers, subsidized transportation, telemedicine, transfer of trained personnel to short-staffed facilities and optimally placed future facilities. Providing medical care to communities located in rural or remote areas is critical for ensuring health equity for populations that are often underserved due to a lack of skilled healthcare workers[12]. Within this context, maps of travel time to healthcare facilities provide a valuable resource for policymakers tasked with allocating personnel and resources to improve health outcomes. The main findings, limitations and policy implications of this work are summarized in Table 1.

Previous studies on travel time to healthcare facilities[13–17] provided the conceptual foundation for the current study, but these could not extend beyond localized or regional analyses without a suitable global dataset of healthcare facilities. We overcame this obstacle by leveraging multiple continental-to-global-scale datasets, which also helped to mitigate challenges associated with inconsistent data quality between countries and between public and private health sector data sources. Geographic locations of hospitals and clinics were acquired from Google Maps (https://www.google.com/maps/), OpenStreetMap (OSM; https://www.openstreetmap.org/) and other published sources[18,19], which were combined to create a global map (with a spatial resolution of 1 km × 1 km) in which 376,231 pixels contained one or more facilities (Extended Data Fig. 1). Among the sources, Google and OSM were global in scope and included a mixture of public and private facilities, while the published datasets were continental-scale data for Australia[18] and Africa[19], containing primarily public facilities. We then used an established methodology[4,20] to create maps of travel time for individuals with and without access to motorized transportation. For this process, we first updated global 'friction surfaces' containing the estimated time required to traverse each pixel with and without using motorized transportation. The friction surfaces were derived from a variety of geospatial datasets, the most important of which were roads location data from OSM and Google and road speed data from OSM (Supplementary Table 1). Lastly, we used a least-cost-path algorithm to create maps of travel time to the nearest healthcare facility (Methods).

The global map of optimal travel time to a hospital or clinic (Fig. 1) illustrates the spatial heterogeneity of geographically constrained healthcare access experienced by individuals with access to motorized transportation.

[1]Malaria Atlas Project, Big Data Institute, Nuffield Department of Medicine, University of Oxford, Oxford, UK. [2]Telethon Kids Institute, Perth Children's Hospital, Nedlands, Western Australia, Australia. [3]Curtin University, Bentley, Western Australia, Australia. [4]Department of Natural Resources, ITC Faculty of Geo-Information Science and Earth Observation, University of Twente, Enschede, the Netherlands. [5]Swiss Federal Institute of Technology Lausanne (École Polytechnique Fédérale de Lausanne), Lausanne, Switzerland. [6]Google, Mountain View, CA, USA. [7]Institute for Disease Modeling, Bellevue, WA, USA. [8]Stanford University, Palo Alto, CA, USA. [9]Department of Geography, Virginia Polytechnic Institute and State University , Blacksburg, VA, USA. [10]Vaccitech, The Oxford Science Park, Oxford, UK. [11]Department of Infectious Disease Epidemiology, Imperial College London, London, UK. ✉e-mail: daniel.weiss@telethonkids.org.au

Table 1 | Policy implications

|	
Background	The time required to travel to healthcare facilities influences whether individuals seek care when it is needed. This relationship is most problematic in low-income settings, where long travel times are also associated with higher relative transport costs.
Main findings and limitations	We produce the first high-resolution global maps of travel time to hospitals and clinics. The results show that 91.1% of the world's population can reach a hospital or clinic within an hour if they have access to motorized transportation, but only 56.7% can do so by walking. The maps were made by combining healthcare facility location data from world-leading sources and recently published research, but the accuracy and completeness of the facility list remain potential limitations.
Policy implications	Policymakers may benefit from maps showing travel time to healthcare facilities as these highlight areas most in need of additional personnel and resources. By increasing the efficiency of resource allocation, the maps could help increase health equity without requiring additional resources. Critically, by freely providing the tools to make custom maps of travel time, we also enable public health professionals to characterize accessibility to specialized services such as emergency care.
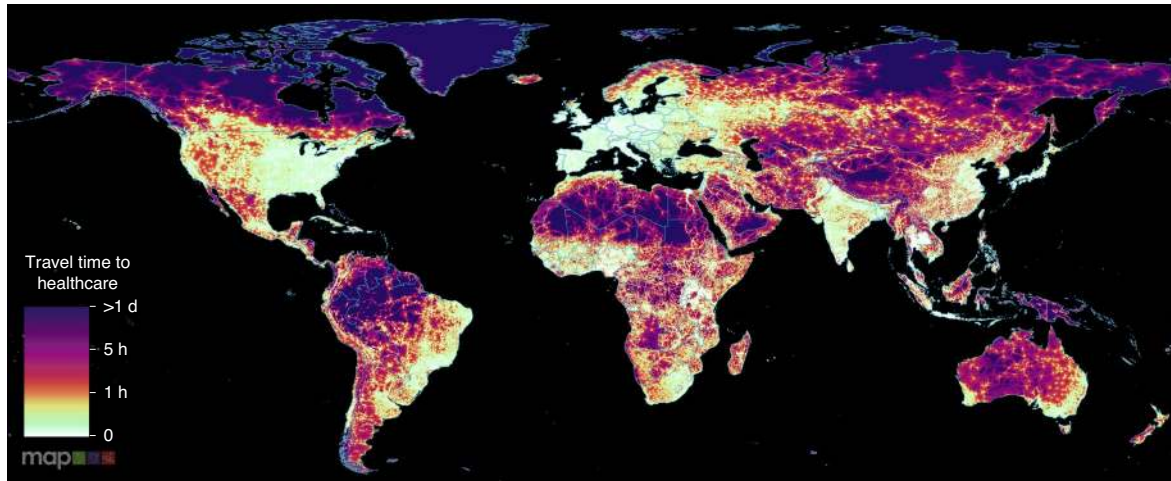
Fig. 1 | The global map of optimal travel time to healthcare with access to motorized transport. Color-coded logarithmic timescale from minutes (yellow) to 24 h (dark purple).

The walking-only map of travel time to healthcare facilities (Fig. 2) shows similar spatial patterns but with far longer travel times for rural areas, thus highlighting a stark reality for rural populations in need of healthcare but without access to motorized transportation. Hospitals and clinics are typically located in densely populated areas, thus supplying the local demand for healthcare. As such, while the broad patterns evident in our maps are similar to those within the map of travel time to cities[4], the maps of travel time to healthcare are much more nuanced due to the comparatively large number and wide spatial distribution of healthcare facilities.

By intersecting the maps of travel time with a gridded population surface, we quantified the proportional travel times to hospitals and clinics for the global population relative to each facility location dataset, as well as for the combined facility dataset (Fig. 3). With access to motorized transportation, 60.3% of the Earth's population lives within 10 min of a hospital or clinic, while 82.6% and 91.1% live within 30 and 60 min, respectively. These percentages correspond to 4.39, 6.02 and 6.64 billion people, respectively. These results suggest that the spatial coverage of healthcare facilities is very good in most regions for individuals with access to the fastest means of transportation. By contrast, the walking-only results show that just 14.2%, 39.8% and 56.9% (1.19, 2.86 and 4.13 billion) of people live within 10, 30 and 60 min, respectively, of a healthcare facility when traveling by foot. The disparity between motorized and nonmotorized travel times illuminates a potential causal association between individuals' decisions on whether to seek healthcare and their available modes of transportation. Furthermore, because a lack of financial resources could constrain an individual's transportation options, as well as their ability to pay

for healthcare, wealth and travel time to a healthcare facility are intrinsically linked and combine to influence whether individuals seek healthcare when it is needed[2,3,21]. While we present a dichotomy of people either moving at the optimal speed or having to walk to healthcare facilities, in reality large numbers of people will fall between these estimates if, for example, their journeys combine walking and public transport. Furthermore, the urgency of individual healthcare needs will affect their choice of transportation mode and being unwell could increase travel times.

The completeness and accuracy of the healthcare facility dataset vary by country (Supplementary Table 2) and are important limitations of this analysis that should be considered when interpreting the maps. Despite using the best available data, the constituent healthcare facility datasets, and hence the combined dataset, were susceptible to errors of omission, could be outdated as a consequence of facilities opening or closing and might contain incorrect facility locations. These concerns provide the rationale for only including datasets that were published in 2019 and, in the cases of Google Maps and OSM data, are frequently updated products with robust quality controls. Although the quality of the healthcare facility dataset could not be fully ascertained, technical validation showed that the travel time estimates derived from the friction surface were generally accurate[4,22], with the caveat that they did not account for daily or seasonal variability in travel times or factors such as delays in journeys incurred while waiting for public transportation. Similarly, the facility location data are not temporally dynamic and thus do not reflect when facilities are open but understaffed, closed for nights, weekends, holidays or closed seasonally or permanently.
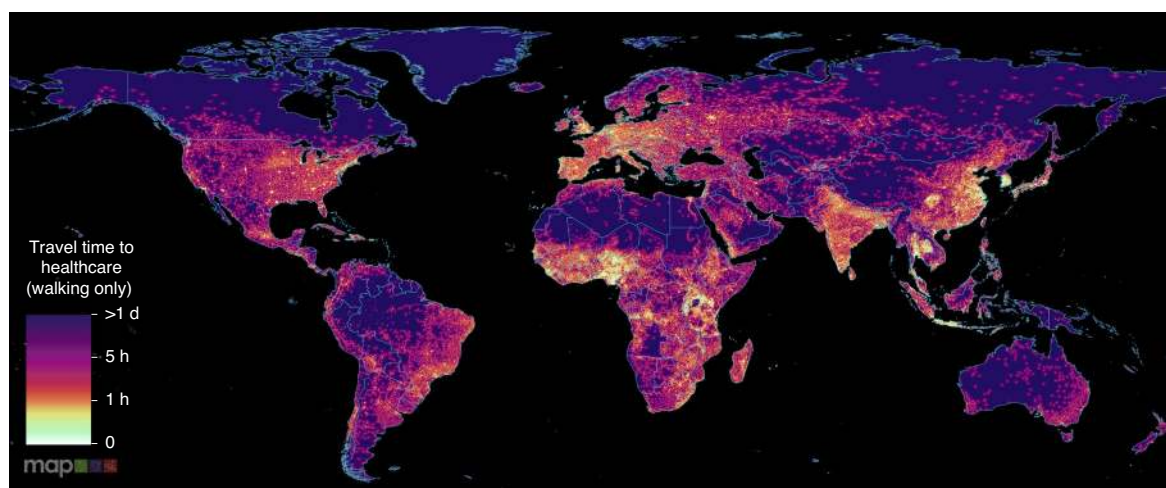
Fig. 2 | The walking-only map of travel time to healthcare without access to motorized transport. Color-coded logarithmic timescale from minutes (yellow) to 24 h (dark purple).
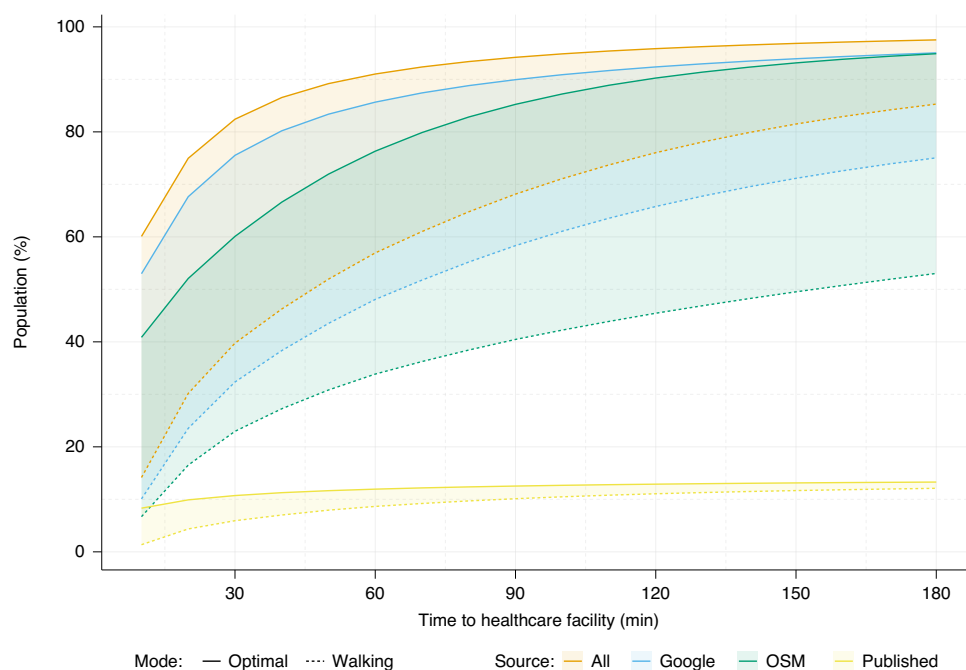

Fig. 3 | Proportional distribution of the human population relative to healthcare facilities present within each dataset. Note that the published sources only contain data for Africa and Australia, whereas the others have a global extent. Solid lines represent travel to healthcare facilities at the optimal speed, while dashed lines represent walking-only travel.

Additional limitations of the maps of travel time relate to whether individuals receive care at their closest facility. First among these considerations is that an individual who is capable of physically reaching a healthcare facility cannot necessarily receive care at that facility if, for example, the cost of care is prohibitive for the patient[23]. Second, all hospitals and facilities do not offer the same services, such as emergency care, and travel times to more specialized services could be much greater than our maps indicate. As such, the provided maps are not a universally applicable assessment of healthcare access in all circumstances, but are indicative of access only to basic services, such as care provided by a general practitioner or trained nurse. Third, our maps of travel time to healthcare provide estimates of potential rather than actual travel times to healthcare, because although individuals have a propensity for

choosing the closest facility[24], they could choose to seek healthcare from facilities farther from home. Addressing the disparity between potential and actual travel time to healthcare is a priority for future research. Such research is now feasible due to advances in the quality and availability of empirical human movement data such as user locations logged by GPS-equipped smartphones and collated by technology companies. Finally, our research focuses exclusively on the location of geographically fixed facilities and ignores the potential use of mobile or temporary clinics for providing healthcare in remote areas.

Despite these limitations, the maps of global travel time to hospitals and clinics highlight areas most in need of additional healthcare facilities, and thus could be used to assess the adequacy of healthcare provision, optimally position new facilities and allocate

funding toward subsidized transportation or telemedicine to communities most in need. Our results could also benefit travelers, as specialists in some countries use access to healthcare to determine whether to issue standby emergency treatments to individuals who will be in areas far from facilities. Furthermore, we freely provide the friction surfaces and the travel time mapping code; thus, our approach can readily be used to produce custom maps of travel time, such as those for specific healthcare services like pediatric emergency care[25]. Additionally, the underlying approach we developed can be adapted for new applications such as assessing disease outbreak preparedness[26] or helping to define healthcare facility catchment areas. Our maps are also useful for estimating healthcare treatment-seeking behavior of individuals, which declines with increased travel time and is also affected by access to motorized transport (Extended Data Fig. 2). Improved characterization of healthcare seeking is critical because individuals failing to seek care face an increased risk of death from treatable conditions, and thus this metric is a much-needed predictor in datasets for modeling mortality[27]. Lastly, by creating the walking-only map, we provide a means of highlighting regions in which individuals lacking resources for motorized transport are particularly disadvantaged in their ability to seek healthcare.

To our knowledge, the maps of travel time to hospitals and clinics are the first, global-scale, high-resolution maps of this critical healthcare accessibility metric yet published. As such, the maps have the potential to make substantial contributions to the public health and epidemiological research communities by providing estimates for the time required to reach healthcare from every inhabited region on Earth. By combining the strengths of multiple healthcare facility data sources, this project demonstrates the power of data amalgamation for creating outputs that were, until recently, prohibitively expensive and time consuming to make. However, because the results of this work are contingent upon the accuracy of hospital and clinic locations, this project also represents a call for action to improve the quality, consistency, completeness and availability of healthcare facility data globally.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41591-020-1059-1.

References

1. Fullman, N. et al. Measuring performance on the healthcare access and quality index for 195 countries and territories and selected subnational locations: a systematic analysis from the Global Burden of Disease Study 2016. *Lancet* **391**, 2236–2271 (2018).
2. Wang, F. & Luo, W. Assessing spatial and nonspatial factors for healthcare access: towards an integrated approach to defining health professional shortage areas. *Health Place* **11**, 131–146 (2005).
3. Buzza, C. et al. Distance is relative: unpacking a principal barrier in rural healthcare. *J. Gen. Intern. Med.* **26**, 648–654 (2011).
4. Weiss, D. J. et al. A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature* **553**, 333–336 (2018).
5. Battle, K. E. et al. Treatment-seeking rates in malaria endemic countries. *Malar. J.* **15**, 20 (2016).
6. Stock, R. Distance and the utilization of health facilities in rural Nigeria. *Soc. Sci. Med.* **17**, 563–570 (1983).
7. Alegana, V. A., Wright, J., Pezzulo, C., Tatem, A. J. & Atkinson, P. M. Treatment-seeking behaviour in low-and middle-income countries estimated using a Bayesian model. *BMC Med. Res. Methodol.* **17**, 67 (2017).
8. Nemet, G. F. & Bailey, A. J. Distance and healthcare utilization among the rural elderly. *Soc. Sci. Med.* **50**, 1197–1208 (2000).
9. Syed, S. T., Gerber, B. S. & Sharp, L. K. Traveling towards disease: transportation barriers to healthcare access. *J. Community Health* **38**, 976–993 (2013).
10. Thaddeus, S. & Maine, D. Too far to walk: maternal mortality in context. *Soc. Sci. Med.* **38**, 1091–1110 (1994).
11. Manongi, R. et al. Inpatient child mortality by travel time to hospital in a rural area of Tanzania. *Trop. Med. Int. Health* **19**, 555–562 (2014).
12. World Health Organization. Increasing access to health workers in remote and rural areas through improved retention: global policy recommendations. https://apps.who.int/iris/bitstream/handle/10665/44369/9789241564014_eng.pdf (WHO, 2010).
13. Juran, S. et al. Geospatial mapping of access to timely essential surgery in sub-Saharan Africa. *BMJ Glob. Health* **3**, e000875 (2018).
14. Noor, A., Zurovac, D., Hay, S., Ochola, S. & Snow, R. Defining equity in physical access to clinical services using geographical information systems as part of malaria planning and monitoring in Kenya. *Trop. Med. Int. Health* **8**, 917–926 (2003).
15. Perry, B. & Gesler, W. Physical access to primary healthcare in Andean Bolivia. *Soc. Sci. Med.* **50**, 1177–1188 (2000).
16. Tanser, F., Gijsbertsen, B. & Herbst, K. Modelling and understanding primary healthcare accessibility and utilization in rural South Africa: an exploration using a geographical information system. *Soc. Sci. Med.* **63**, 691–705 (2006).
17. Alegana, V. A. et al. Spatial modelling of healthcare utilisation for treatment of fever in Namibia. *Int. J. Health Geogr.* **11**, 6 (2012).
18. Barbieri, S. & Jorm, L. Travel times to hospitals in Australia. *Sci. Data* **6**, 248 (2019).
19. Maina, J. et al. A spatial database of health facilities managed by the public health sector in sub-Saharan Africa. *Sci. Data* **6**, 134 (2019).
20. Nelson, A. Travel time to major cities: a global map of accessibility. https://forobs.jrc.ec.europa.eu/products/gam/ (Global Environment Monitoring Unit, Joint Research Unit of the European Commission, 2008).
21. Wong, K. L. M., Brady, O. J., Campbell, O. M. R., Banke-Thomas, A. & Benova, L. Too poor or too far? Partitioning the variability of hospital-based childbirth by poverty and travel time in Kenya, Malawi, Nigeria and Tanzania. *Int. J. Equity Health* **19**, 15 (2020).
22. Nelson, A. et al. A suite of global accessibility indicators. *Sci. Data* **6**, 266 (2019).
23. Gulliford, M. et al. What does 'access to healthcare' mean? *J. Health Serv. Res. Policy* **7**, 186–188 (2002).
24. Haynes, R., Lovett, A. & Sünnenberg, G. Potential accessibility, travel time and consumer choice: geographical variations in general medical practice registrations in Eastern England. *Environ. Plan. A* **35**, 1733–1750 (2003).
25. Johansson, E. W. et al. Accessibility of basic paediatric emergency care in Malawi: analysis of a national facility census. *BMC Public Health* **20**, 992 (2020)
26. Hulland, E. N. et al. Travel time to health facilities in areas of outbreak potential: maps for guiding local preparedness and response. *BMC Med.* **17**, 232 (2019).
27. Weiss, D. J. et al. Mapping the global prevalence, incidence and mortality of Plasmodium falciparum, 2000-17: a spatial and temporal modelling study. *Lancet* **394**, 322–331 (2019).

Methods

Facility data. Until recently, large and comparable datasets of healthcare facility locations have been rare, and even the most comprehensive sources continue to have geographically heterogeneous levels of data quality and completeness. To mitigate these limitations, we used healthcare facility location data from two of the largest global databases: (1) OSM data that was collated and made available for download at https://www.healthsites.io/[28] and (2) data from Google Maps that was extracted specifically for this project. We augmented the global datasets with continental-scale facility locations that were recently published for Africa[19] and Australia[18]. To align temporally, each of these datasets was extracted in mid-2019. To facilitate comparisons between data sources, we used only facilities defined as hospitals and clinics, which matched the definition used in the published datasets[18,19] and was distinguishable within the OSM data by using the facility-type attribute. In contrast, Google Maps data were filtered to isolate hospitals and clinics from the full set of geolocated entities. For each dataset, multiple points found within the same ~1 km × 1 km pixel were merged to match the resolution of the subsequent maps of travel time. The resulting counts of pixels containing one or more hospitals or clinics, by data source, were 201,799 from Google Maps, 148,522 from OSM and 87,402 from the published sources[18,19]. Maps of travel time were prepared for each set of healthcare facility pixels and then combined to form the final travel time map, in which a total of 376,231 pixels contained one or more hospitals or clinics. A sensitivity analysis confirmed that slight spatial misalignments were not responsible for the differences between the facility databases. As such, the lack of agreement between the data sources (Extended Data Fig. 1) suggests that they vary considerably in their completeness from country to country, or they capture fundamentally different segments of healthcare systems such as public versus private facilities. For example, the published data for Africa[19] contains primarily public facilities, while OSM and Google are a mixture of both data types. Finally, because the published data are not global in scale, the implication of Extended Data Fig. 1 is that the Google data lack many of the hospitals and clinics in the focal areas of the published results (Africa in particular).

The healthcare facility location points used for this research were obtained from Google Maps, OSM and recently published healthcare facility datasets[18,19]. The OSM dataset is a global extraction of healthcare facilities collated and distributed by https://www.healthsites.io/. OSM is a collaborative project designed to create a free and editable geospatial database of the whole world. OSM is one of the most successful examples of a volunteered geographic information project[29]. OSM is built by a large user community that uses aerial imagery, GPS devices and low-tech field maps to verify that OSM is accurate and up to date. By February 2020, the project had more than 6 million registered users, while the number of contributors was 1.4 million. OSM follows the peer production model that created Wikipedia; intrinsic within this model are three approaches to ensure data quality: crowdsourcing, social and geographic[30]. A considerable number of scientific studies have evaluated different data quality elements of OSM data[31–37], including completeness, logical consistency, positional accuracy, temporal accuracy and thematic accuracy[38]. Overall, these studies agree that different quality elements show heterogeneous patterns across space, and that dense areas with a high number of OSM contributors assessed have higher accuracy and completeness than rural areas. However, these studies do not attempt to define minimum acceptable levels of quality for geographic data, suggesting that OSM data may have different degrees of suitability for specific purposes. In the case of geospatial features representing health facilities, the Global Healthsites Mapping Project (https://www.healthsites.io/)[28] provides a domain-specific view of OSM data, focusing on health facility data. The healthsites.io platform integrates OSM data with several databases from trusted partners. To enhance the reliability of health facilities data, healthsites.io performs manual and automated processes to assess the quality of data through a validity index, for example, calling the facilities by telephone and comparing the street address associated with the facilities with the corresponding address returned by external geocoding services, respectively.

Hospitals and clinics that were open in August 2019 were extracted from the full set of geolocated places within Google Maps. The set of hospitals and clinics was derived based on a proprietary ontology of geographic entities and a coarse classification relying on estimates produced by machine-learning models. Facilities were retained if they were broadly categorized as hospitals, clinics or facilities providing urgent or emergency medical care with an entry subtype indicating they were a hospital or clinic. The underlying sources of the geolocated entities within Google Maps included publicly available data, licensed third-party data and data contributed by users[39]. Publicly available and third-party data may be associated with dataset-specific metadata that describe their accuracy and completeness, and, as with OSM, users of Google Maps can flag and report potential errors.

The published dataset for Africa[19] contained healthcare facilities created through a multiyear project that collated publicly available data (for example, from online repositories) and data obtained via personal communication with researchers and officials from numerous countries. All facilities without coordinates were manually name matched to known geographic locations by, for example, searching for place names within Google Maps. The Africa dataset was limited to facilities defined as hospitals or clinics and, with the exception of Botswana, included only public facilities. The published hospitals and clinics

dataset for Australia[18] consisted of a mixture of public and private hospitals from an inventory distributed by the Australian Institute of Health and Welfare.

The set of facility locations from Google Maps was used to generate source-specific maps of travel time, which were later merged (using a minimum function) with similar maps of travel time made from the OSM and published facility locations. The OSM and published facility points were downloaded from their repositories and only marginally modified from their original forms. The only changes consisted of removing any points that lacked spatial coordinates, had coordinates of 0° latitude and 0° longitude, were otherwise erroneously located in the ocean or were located outside the analysis window (for example, two points in Antarctica). In cases where multiple facilities were found within single pixels, or where a facility was in multiple datasets, the points were effectively merged such that the pixel was considered a single source location of healthcare.

Travel time mapping. The first method for modeling travel time globally was established by Nelson[20] and substantially improved upon by Weiss et al. [4]. This approach was predicated on the creation of friction surfaces that quantified the time required to traverse each pixel within a ~1 km × 1 km gridded representation of the Earth's surface. The movement rates encapsulated within the friction surfaces were based on landscape characteristics and transportation infrastructure, and no individual movement data were used to create them. The friction surfaces were made by combining geographic datasets that each characterized a layer-specific cost of moving through the landscape[4]. These datasets consisted of roads, railways, navigable waterways, bodies of water, land cover types, topographic conditions (elevation and slope angle) and national borders. Descriptions of these datasets and their associated processing steps are provided in Weiss et al. [4]. Note, however, that we did not consider air travel in this analysis.

As people typically take the path of least resistance when traveling, the route calculated to the most accessible healthcare facility typically relied on road networks for a large portion of the journey distance. As such, the road layers used to derive the friction surfaces were the most critical inputs for overall model accuracy. The roads datasets used for this project were the OSM and Google Roads datasets. Firstly, OSM roads data were extracted in October 2019 and converted into a rasterized surface of road types whereby the fastest road type in each pixel took precedence. In a second processing step, the OSM data were analyzed to determine the median speed limit value for each road type in each country (Supplementary Table 1). These values were subsequently used to attribute road speeds to specific pixels using a lookup table. In cases where a road type was present in a country, but associated speed limit information was unavailable, a global average speed limit for that road type was used instead. The Google roads data consisted of a distance to road layer, which effectively became a rasterized roads surface by defining all pixels with distance values of less than 500 m as roads. Because the OSM data were split by road type for which we had speed limit information, these roads data took precedence when both datasets indicated the existence of a road within a pixel. In cases where only Google roads data indicated the presence of a road, those pixels were assigned the generic 'road' classification for the purposes of assigning speed limits via the lookup tables. While the approach for making the friction surface was identical to the one used in earlier research[4], reanalysis was warranted because substantial additions have since been made to OSM. These changes included the addition of many previously missing roads, particularly in developing countries, and far more comprehensive speed limit information to support country-specific road speed attribution.

Travel times were calculated to the nearest geolocated hospital or clinic. A critical assumption of the travel time calculations, however, was that people moved through the landscape at an optimal speed. In reality, factors such as wealth and the availability of public transportation affect individuals' decisions on when and how to travel to obtain healthcare, thereby preventing segments of the population from moving this quickly. To account for this aspect of human movement, we also created a walking-only variant of the friction surface. In the optimal friction surface, we assumed that humans traveling along roadways, railways and on water moved at the motorized movement rates assigned to each type. In the walking-only friction surface, we assumed that all road travel was limited to a walking speed of 5 km per hour, and all water travel was reduced to the rate for traversing a wetland or swimming of 1 km per hour. The walking-only friction provided an upper bound of movement rates for individuals who lack access to or cannot use motorized transport. Examples of factors preventing the use of motorized transport include roads being impassible due to natural disasters or the presence of travel restrictions such as those put in place in response to the 2020 coronavirus pandemic. It should be noted, however, that in scenarios where emergency healthcare is required (for example, when someone is injured and incapacitated), even the walking-only estimates may be too optimistic unless transportation assistance such as an ambulance service is available. Lastly, only optimal and walking-only maps were produced because only they could be reasonably parameterized at a global scale. In reality, individuals travel using a variety of modes of transportation that move at a range of speeds, and even those able to travel by motorized vehicle may not be able to do so immediately. As such, understanding the prevalence of varying modes of transportation, as well as their speeds, is critical for contextualizing local travel times within the upper and lower bounds provided by our maps. Likewise, the modes of transportation and routes

of transport that individuals use can shift over time as people respond to changing conditions, such as seasonal flooding. Each of these topics warrants additional study, albeit for smaller geographic areas where highly localized conditions, dynamics and adaptations can be adequately characterized.

A least-cost-path algorithm was used to calculate the travel time from every pixel to a hospital or clinic accessible via the shortest journey. The least-cost-path algorithm was developed by Dijkstra[40] for use in graphs (that is, networks of connected nodes) and has since been adapted within several software packages for use with gridded datasets. For this work, we used the versions of Dijkstra's algorithm coded into Google Earth Engine[41] (as the cumulativeCost() function) and the gDistance package[42] in R (as the accCost() function). The underlying premise of this approach was to effectively test all possible routes from each 1 km × 1 km pixel to every healthcare facility and then display the minimum travel time for each pixel within the map of travel time. The maps extend from 56° S to 83° N and span the full circumference of the Earth (that is, −180° to +180°). This extent includes virtually all land areas on Earth other than Antarctica. The map resolution is 30 × 30 arc-seconds, which is roughly equivalent to 1 km × 1 km at the equator. Ocean areas were masked in the final maps of travel time and set to a no-data value of −9999. The units of the map are in minutes of travel time to the most accessible facility, which need not be the closest facility in terms of geographic distance due to transportation infrastructure patterns.

Model validation. A thorough validation of the underlying method used to calculate travel time was presented by Nelson et al. [22]. This validation compared results generated through the least-cost-path algorithm applied to an earlier friction surface[4] with travel time estimates for the same journeys as those calculated using the Google Maps API. Unlike our raster-based approach, the Google Maps API estimated travel time using the vector-based road network. These validation results indicated that the travel time estimates from the friction surface were, on average, within ±15.8 min of those from the alternative source. There was, however, spatial variability in the model accuracy, with some areas being more prone to underestimates, while other areas tended to have overestimates[22]. Cursory analysis of the new, optimal friction surface suggested that it produced similar accuracies.

No validation was performed on the input healthcare facility locations as these data were published elsewhere by research groups that have their own data quality protocols. A caveat for the maps of travel time to healthcare facilities is that definitions for hospitals and clinics are likely to vary from country to country and, in the case of data originating from OSM, between the users entering the data. As such, countries that appear to have few facilities relative to the population may reflect a stricter definition of what constitutes a facility, and vice versa. Supplementary Table 2 enables users of our maps of travel time to better assess the completeness of the facility dataset for individual countries, as well as the percentage of the population living beyond 10, 30 and 60 min from a healthcare facility. In cases where these data (Supplementary Table 2) suggest that our facility set is incomplete (for example, the Democratic People's Republic of Korea has only 26 facility pixels in our dataset), our code and friction surface provide a ready means of creating updated maps of travel time if improved facility location data become available.

The most important limitation of this work, however, is the level of completeness of the healthcare facility inventory, which varies by country. This challenge provided the rationale for using multiple data sources, as they had different strengths in different areas. For example, the Google dataset provided the best source of information on facility location in Asia, OSM was the best source for data on facilities in remote locations such as Pacific Islands and high latitudes, and the published data provided the majority of the facility information in Africa. The implications of incomplete or inaccurate facility locations were a notable concern of this work. While we had no means of discovering facilities that were erroneously attributed to locations many kilometers from their true positions, we did conduct a sensitivity analysis to test whether there were slight spatial discrepancies between the input datasets (for example, whether a hospital attributed to one pixel in Google was likely to be found within a neighboring pixel within OSM). This analysis consisted of aggregating facility location pixels to lower spatial resolutions, calculating the proportion of pixels containing healthcare facility pixels that were merged and then comparing results derived within and between the facility data sources. These results showed that facility pixels from the same dataset were more likely to be merged than those from different datasets (Supplementary Table 3), which suggests that minor discrepancies in facility geolocations between sources did not inflate the combined facility count or cause noteworthy inaccuracies in the final maps. Lastly, because the comprehensiveness and accuracy of the healthcare facility datasets will always be speculative, and new or updated datasets will continue to improve data quality, we have made our code and friction surfaces freely available to support (1) the production of custom maps in cases where users have their own facility location data and (2) frequent updates to our global maps of travel time to hospitals and clinics as new facility locations become available.

Extended analysis. To provide additional context with which to interpret the optimal and walking-only maps of travel time, we explored the association between travel time to a healthcare facility and healthcare utilization. For this, we analyzed treatment seeking among children under five years of age with a fever using household survey data collected by the Demographic and Health Survey Program (DHS). Treatment-seeking data were extracted for geolocated household clusters from 150 surveys collected across 56 countries between 2000 and 2017. Only households within a 2-hour walk of a hospital or clinic were included to retain large sample sizes for households both with and without motorized transportation within each travel time threshold range. The resulting dataset consisted of 257,416 children under the age of five with a fever, 159,863 of whom were taken for medical care.

The results of the extended treatment-seeking analysis (Extended Data Fig. 2) provided additional empirical evidence that the time it takes individuals to reach a hospital or clinic affects the likelihood that they seek medical attention, but the comparison was imperfect as some of the DHS survey responses may have predated the building of facilities contained within our database. A more robust finding, which was consistent whether the optimal or walking-only maps of travel time were used as the basis for the Extended Data Fig. 2, was a strong association between household ownership of motorized transportation (including motorcycles) and care seeking for children with fevers. Interestingly, this relationship was apparent even in cases where a healthcare facility was less than a 10-min walk from the household, which highlights the importance of household wealth as a determinant of healthcare seeking[43–45] even when travel time is a minor consideration.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The code, friction surfaces and maps of travel time to hospitals and clinics are available for download in GEOTIFF format (.tif) at https://malariaatlas.org/research-project/accessibility_to_healthcare/ and for analysis within Google Earth Engine[41]. For users with limited computational capacity, we have also created an online tool for freely creating accessibility maps with limited spatial extents at https://access-mapper.appspot.com/. The OSM roads and healthcare facilities are available through https://www.openstreetmap.org/ and https://healthsites.io/, respectively. Google roads and healthcare facilities can be viewed within Google Maps. Published healthcare facility datasets are available through repositories associated with those publications[18,19].

References

28. Saameli, R., Kalubi, D., Herringer, M., Sutton, T. & de Roodenbeke, E. Healthsites.io: the Global Healthsites Mapping Project. In *Technologies for Development: From Innovation to Social Impact* (eds. Hostettler, S. et al.) 53–59 (Springer, 2018).
29. Goodchild, M. F. Citizens as sensors: the world of volunteered geography. *GeoJournal* **69**, 211–221 (2007).
30. Goodchild, M. F. & Li, L. Assuring the quality of volunteered geographic information. *Spat. Stat.* **1**, 110–120 (2012).
31. Brovelli, M. A., Minghini, M., Molinari, M. & Mooney, P. Towards an automated comparison of OpenStreetMap with authoritative road datasets. *Trans. GIS* **21**, 191–206 (2017).
32. Barrington-Leigh, C. & Millard-Ball, A. The world's user-generated road map is more than 80% complete. *PLoS ONE* **12**, e0180698 (2017).
33. Jackson, S. P. et al. Assessing completeness and spatial error of features in volunteered geographic information. *ISPRS Int. J. Geoinf.* **2**, 507–530 (2013).
34. Hecht, R., Kunze, C. & Hahmann, S. Measuring completeness of building footprints in OpenStreetMap over space and time. *ISPRS Int. J. Geoinf.* **2**, 1066–1091 (2013).
35. Canavosio-Zuzelski, R., Agouris, P. & Doucette, P. A photogrammetric approach for assessing positional accuracy of OpenStreetMap roads. *ISPRS Int. J. Geoinf.* **2**, 276–301 (2013).
36. Haklay, M. How good is volunteered geographical information? a comparative study of OpenStreetMap and Ordnance Survey datasets. *Environ. Plann. B Plann. Des.* **37**, 682–703 (2010).
37. Girres, J. F. & Touya, G. Quality assessment of the French OpenStreetMap dataset. *Trans. GIS* **14**, 435–459 (2010).
38. International Organization for Standardization. ISO 19157:2013. Geographic information—data quality (2013).
39. Google. How Google sources and uses information in local listings https://support.google.com/business/answer/9851099?hl=en/ (2020).
40. Dijkstra, E. W. A note on two problems in connexion with graphs. *Numer. Math.* **1**, 269–271 (1959).
41. Gorelick, N. et al. Google Earth Engine: planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **183**, 250–264 (2017).
42. van Etten, J. R package gdistance: distances and routes on geographical grids. *J. Stat. Softw.* **76**, 1–21 (2017).
43. Kassile, T., Lokina, R., Mujinja, P. & Mmbando, B. P. Determinants of delay in care seeking among children under five with fever in Dodoma region, central Tanzania: a cross-sectional study. *Malar. J.* **13**, 348 (2014).

44. Yaya, S., Bishwajit, G. & Shah, V. Wealth, education and urban–rural inequality and maternal healthcare service usage in Malawi. *BMJ Glob. Health* **1**, e000085 (2016).
45. Najnin, N., Bennett, C. M. & Luby, S. P. Inequalities in care-seeking for febrile illness of under-five children in urban Dhaka, Bangladesh. *J. Health Popul. Nutr.* **29**, 523–531 (2011).

Acknowledgements

Author contributions

D.J.W. designed and led the analysis and wrote the manuscript. A.N. developed the original modeling approach and contributed to the model validation. C.A.V.-R. extracted the OSM data. K.G., S. Bavadekar and E.G. extracted the Google Maps facility data and produced the Google-specific map of travel time. A.B.-V., H.S.G. and J.J.M. assisted with the analysis. J.R. produced the global maps. T.S. and A.L. provided administrative support at Google. P.G. and S. Bhatt assisted with the conceptual design of the project. All authors provided critical review and assisted in writing the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41591-020-1059-1.

**Supplementary information** is available for this paper at https://doi.org/10.1038/s41591-020-1059-1.

Correspondence and requests for materials should be addressed to D.J.W.

Peer review information Jennifer Sargent was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Reprints and permissions information is available at www.nature.com/reprints.
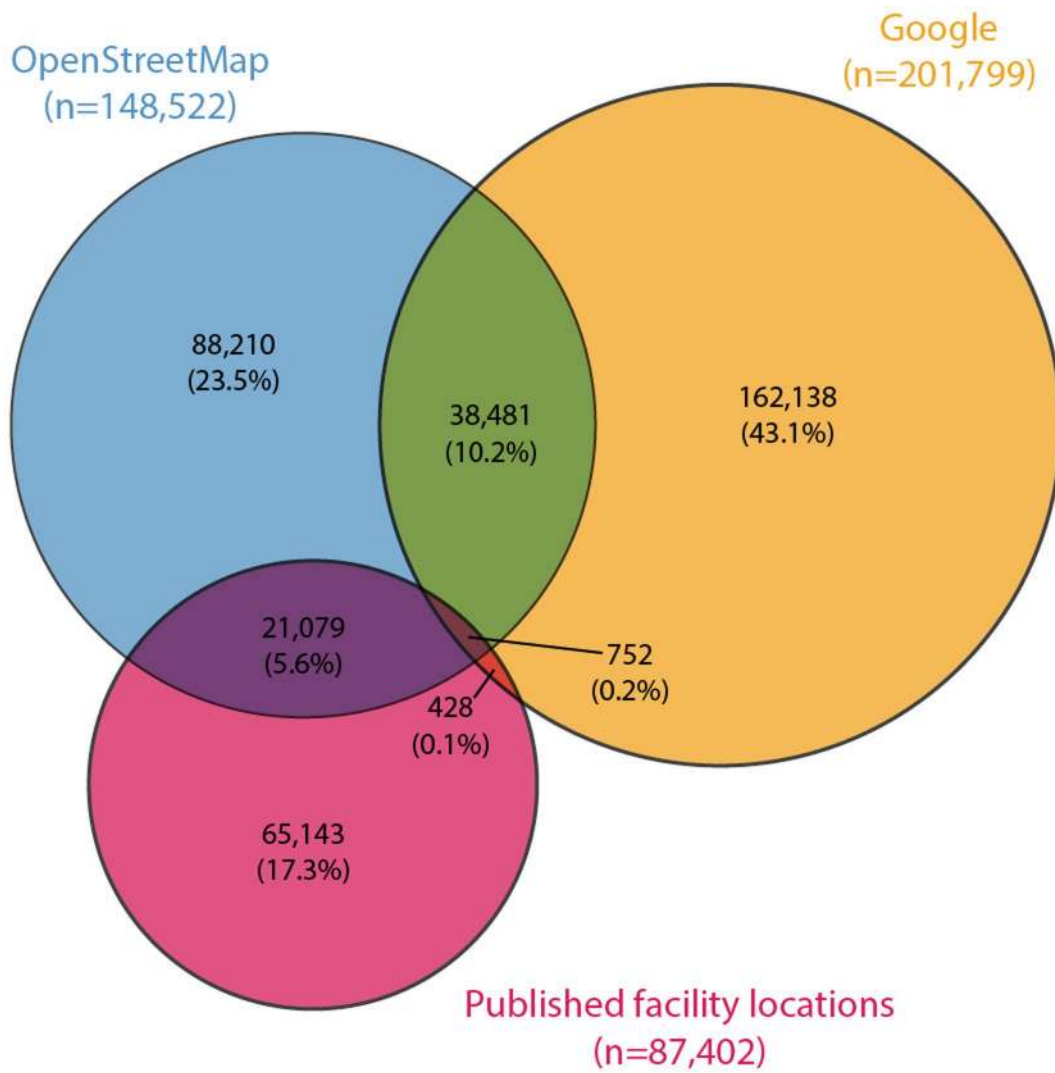
Extended data Fig. 1 | Venn diagram of the shared pixels counts of hospital and clinic locations contained within each of the facility datasets. Percentages in parentheses are relative to the combined set of facility pixels (n = 376,231).
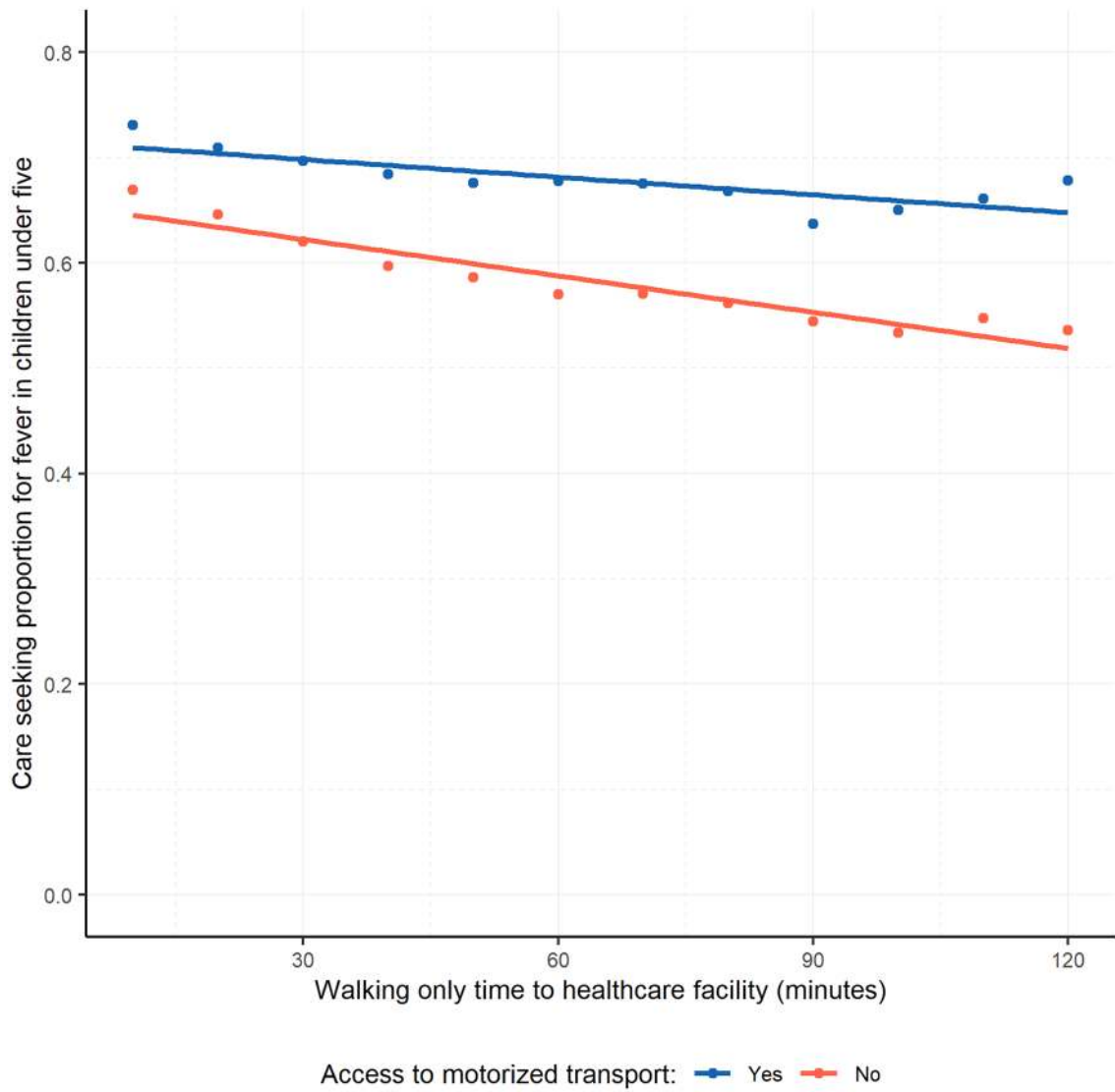
Extended data Fig. 2 | Proportions of healthcare seeking and linear trendlines for fevers in children under five from households with and without motorized transportation. The geolocated survey data were intersected with the walking-only travel time map and aggregated into ten-minute intervals.

Corresponding author(s): Daniel J. Weiss

Last updated by author(s): Mar 9, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☒	☐	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No data was collected for this research. All data was acquired from external sources and referenced within the text.
Data analysis	The friction surface was made in Google Earth Engine (https://earthengine.google.com/) and the travel time maps were created within R version 3.2.3 using the gdistance package (https://cran.r-project.org/web/packages/gdistance/index.html). Source code and datasets required for making the accessibility maps are available at https://malariaatlas.org/research-project/accessibility_to_healthcare/.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The code, friction surfaces, and maps of travel time to hospitals and clinics are available for download in GEOTIFF format (.tif) at https://malariaatlas.org/research-project/accessibility_to_healthcare/ and for analysis within Google Earth Engine. For users with limited computational capacity, we have also created an online tool for freely creating accessibility maps with limited spatial extents at https://access-mapper.appspot.com/. The OSM roads and healthcare facilities are available, respectively, through https://www.openstreetmap.org/ and https://healthsites.io/. Google roads and healthcare facilities can be viewed within Google Maps. Published healthcare facility datasets are available through repositories associated with those publications.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No samples were used in this analysis. All healthcare facilities other than the exclusions listed below were used in the map. The ancillary analysis of household survey data consisted only of descriptive statistics and included all available data points.
Data exclusions	The only facility data excluded were those without corresponding geographic coordinates and those labeled as a facility other than a hospital or clinic.
Replication	No replication was done as no statistical tests were done for this analysis.
Randomization	No randimization was done as no statistical tests were done for this analysis.
Blinding	No binding was done as no statistical tests were done for this analysis.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒ ☐	Antibodies
☒ ☐	Eukaryotic cell lines
☒ ☐	Palaeontology
☒ ☐	Animals and other organisms
☒ ☐	Human research participants
☒ ☐	Clinical data

Methods

n/a	Involved in the study
☒ ☐	ChIP-seq
☒ ☐	Flow cytometry
☒ ☐	MRI-based neuroimaging