

Large-scale physical activity data reveal worldwide activity inequality

Tim Althoff¹, Rok Sosič¹, Jennifer L. Hicks², Abby C. King^{3,4}, Scott L. Delp^{2,5} & Jure Leskovec^{1,6}

To be able to curb the global pandemic of physical inactivity^{1–7} and the associated 5.3 million deaths per year², we need to understand the basic principles that govern physical activity. However, there is a lack of large-scale measurements of physical activity patterns across free-living populations worldwide^{1,6}. Here we leverage the wide usage of smartphones with built-in accelerometry to measure physical activity at the global scale. We study a dataset consisting of 68 million days of physical activity for 717,527 people, giving us a window into activity in 111 countries across the globe. We find inequality in how activity is distributed within countries and that this inequality is a better predictor of obesity prevalence in the population than average activity volume. Reduced activity in females contributes to a large portion of the observed activity inequality. Aspects of the built environment, such as the walkability of a city, are associated with a smaller gender gap in activity and lower activity inequality. In more walkable cities, activity is greater throughout the day and throughout the week, across age, gender, and body mass index (BMI) groups, with the greatest increases in activity found for females. Our findings have implications for global public health policy and urban planning and highlight the role of activity inequality and the built environment in improving physical activity and health.

Physical activity improves musculoskeletal health and function, prevents cognitive decline, reduces symptoms of depression and anxiety, and helps individuals to maintain a healthy weight^{4,7}. Although prior surveillance and population studies have revealed that physical activity levels vary widely between countries¹, more information is needed about how activity levels vary within countries and the relationships between physical activity disparities, health outcomes (such as obesity levels), and modifiable factors such as the built environment. For example, while much is known about how both intrinsic factors (such as gender, age, and weight) and extrinsic factors (for example, public transportation density) are related to activity levels, evidence about how these factors interact (such as the influence of environmental factors on older or obese individuals) is more limited⁸. Understanding these interactions is important for developing public policy^{9,10}, planning cities¹¹, and designing behaviour-change interventions^{12,13}.

The majority of physical activity studies are based on information that is either self-reported, with attendant biases¹⁴, or is measured via wearable sensors, but limited in the number of subjects, observation period, and geographic range¹⁵. Mobile phones are a powerful tool with which to study large-scale population dynamics and health on a global scale^{12,16}, revealing the basic patterns of human movement¹⁷, mood rhythms¹⁸, the dynamics of the spread of diseases such as malaria¹⁹,

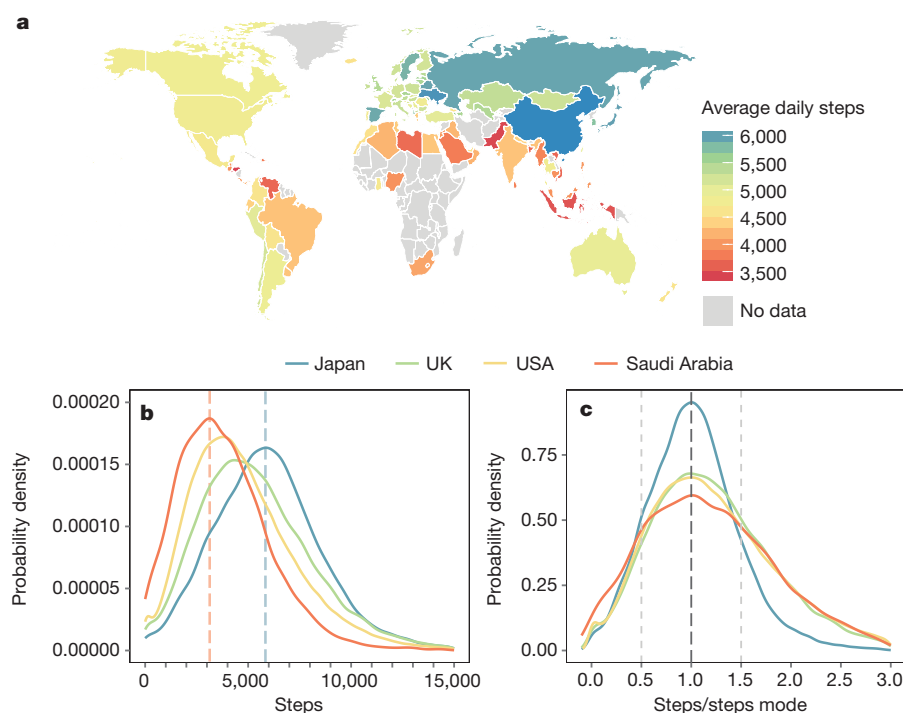


Figure 1 | Smartphone data from over 68 million days of activity by 717,527 individuals reveal variability in physical activity across the world.

a, World map showing variation in activity (mean daily steps) between countries measured through smartphone data from 111 countries with at least 100 users. Cool colours correspond to high activity (for example, Japan in blue) and warm colours indicate low levels of activity (for example, Saudi Arabia in orange). **b**, Typical activity levels (distribution mode) differ between countries. Curves show distribution of steps across the population in four representative countries as a normalized probability density (high to low activity: Japan, UK, USA, Saudi Arabia). Vertical dashed lines indicate the mode of activity for Japan (blue) and Saudi Arabia (orange). **c**, The variance of activity around the population mode differs between countries. Curves show distribution of steps across the population relative to the population mode. In Japan, the activity of 76% of the population falls within 50% of the mode (that is, between the light grey dashed lines), whereas in Saudi Arabia this fraction is only 62%. The UK and USA lie between these two extremes for average activity level and variance. This map is based on CIA World Data Bank II data, publicly available through the R package *mapdata* (<https://www.r-project.org/>).

¹Computer Science Department, Stanford University, Stanford, California, USA. ²Department of Bioengineering, Stanford University, Stanford, California, USA. ³Department of Health Research and Policy, Stanford University School of Medicine, Stanford, California, USA. ⁴Stanford Prevention Research Center, Department of Medicine, Stanford University School of Medicine, Stanford, California, USA. ⁵Department of Mechanical Engineering, Stanford University, Stanford, California, USA. ⁶Chan Zuckerberg Biohub, San Francisco, California, USA.

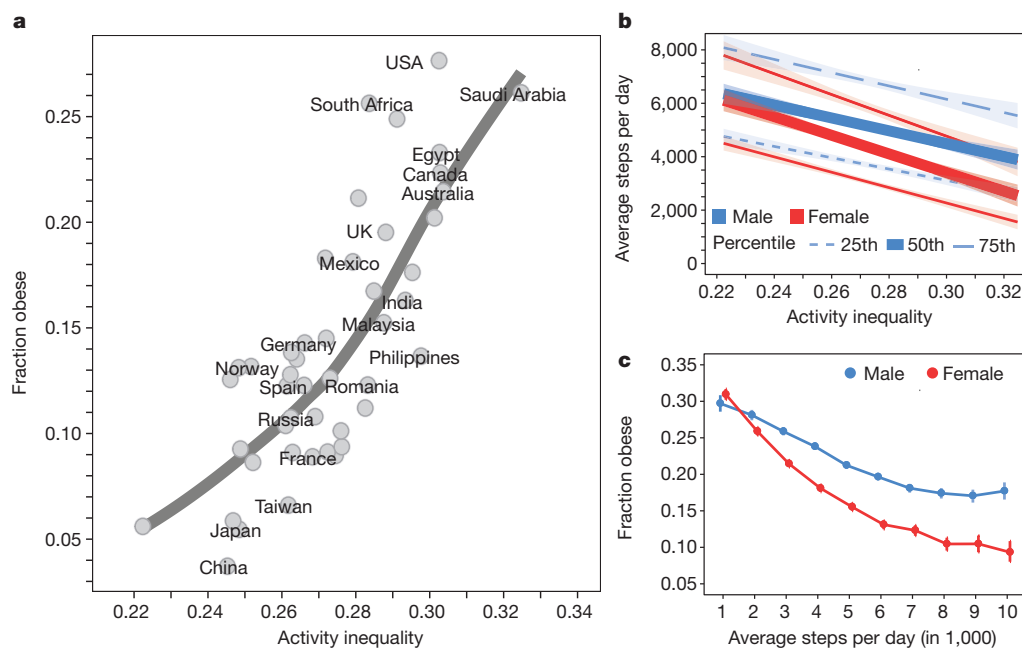


Figure 2 | Activity inequality is associated with obesity and increasing gender gaps in activity. **a**, Activity inequality predicts obesity (LOESS fit; $R^2 = 0.64$). Individuals in the five countries with highest activity inequality are 196% more likely to be obese than individuals from the five countries with the lowest activity inequality. **b**, Activity inequality is associated with reduced activity, particularly in females. The figure shows the 25th, 50th, and 75th percentiles of daily steps within each country along with 95% confidence intervals (shaded) as a linear function of activity inequality. As activity inequality increases, median activity (50th percentile) decreases by

39% for males (blue) and by 58% for females (red). **c**, The obesity–activity relationship differs between males and females and between high- and low-activity individuals. The plot shows the prevalence of obesity as a function of daily number of steps across all subjects in all countries (with 95% confidence intervals). For both males (blue) and females (red), a larger number of steps recorded is associated with lower obesity, but for females, the prevalence of obesity increases more rapidly as step volume decreases (232% obesity increase for females versus 67% increase for males; comparing lowest versus highest activity).

and socioeconomic status in developing countries²⁰. Smartphones are now being used globally, with the adoption rate among adults at 69% in developed countries and 46% in developing economies and growing rapidly²¹. With onboard accelerometers for automatic recording of activity throughout the day, smartphones provide a scalable tool with which to measure physical activity worldwide. Here, we use a large-scale physical activity dataset to quantify disparities in the distribution of physical activity in countries around the world, identify the relationship between activity disparities and obesity, and explore the role of the built environment, in particular its walkability, in creating a more equal distribution of activity across populations.

We study 68 million days of minute-by-minute step recordings from 717,527 anonymized users of the Argus smartphone application developed by Azumio. The dataset includes recordings of physical activity for free-living individuals from 111 countries (Fig. 1a). We focus on the 46 countries with at least 1,000 users (Supplementary Table 1); 90% of these users were from 32 high-income countries and 10% were from 14 middle-income countries (including five lower-middle-income countries; Methods). The average user recorded 4,961 steps per day (standard deviation $\sigma = 2,684$) over an average span of 14 h. We verified that the smartphone application data reproduce established relationships between age, gender, weight status, and activity (Extended Data Fig. 1), as well as country-level variations in activity and obesity levels determined from prior surveillance data and population studies (Extended Data Fig. 2). Recent research has further demonstrated that smartphones provide accurate step counts²² and reliable activity estimates in both laboratory and free-living settings²³. We perform complete-case analyses accompanied by sample correction, stratification, outlier, and balance testing to verify that our conclusions are robust to missing data (see Supplementary Table 1) and biases in age and gender, and that they hold for both high- and middle-income countries (see Methods and Extended Data Figs 3 and 4).

Our large-scale activity measurements enable the characterization of the full distribution of activity within a population beyond activity-level

averages and including the tails of the distribution (Fig. 1b). Consider two countries with divergent activity distributions, Japan and Saudi Arabia. In Japan, the mode of recorded steps is high (Fig. 1b, dashed blue line; 5,846 steps), while in Saudi Arabia it is low (Fig. 1b, dashed red line; 3,103 steps). In Saudi Arabia, the mode is low, but also the variance of recorded steps across the population is larger (Fig. 1c). This larger variance means that some individuals are highly active and others record very little activity even relative to the low country baseline.

We formally characterize these systematic differences in country-level activity distributions by measuring activity inequality, which we define as the Gini coefficient of the population activity distribution^{24,25} (Extended Data Fig. 5). We find that not only is there inequality in how steps are distributed within countries, but also that activity inequality is associated with higher obesity levels (Fig. 2a). For example, Saudi Arabia has a high obesity rate in comparison to Japan. At the same time Saudi Arabia has lower average activity (Fig. 1b) and a wider activity distribution (Fig. 1c), that is, a higher activity inequality. This finding is independent of gender and age biases (Extended Data Fig. 3) and independent of a country's income level (high versus middle; no lower-income countries were included in our dataset; Extended Data Fig. 4). In fact, a country's activity inequality is a better predictor of obesity prevalence than the average volume of steps recorded ($R^2 = 0.64$ versus 0.47; $P < 0.01$; Extended Data Fig. 6). For example, the USA and Mexico have similar average daily steps (4,774 versus 4,692), but the USA exhibits larger activity inequality (0.303 versus 0.279; 10th versus 7th decile of country activity inequality distribution) and higher obesity prevalence (27.7% versus 18.1%; 10th versus 8th decile of country obesity prevalence distribution) compared to Mexico (Supplementary Table 1).

We find that in countries with high activity inequality, activity in females is reduced disproportionately compared to males, across all quartiles of activity (Fig. 2b). In particular, 43% of activity inequality is explained by the gender gap in activity (Extended Data Fig. 7). Therefore the larger variances we observe (Fig. 1c) are due to reduced

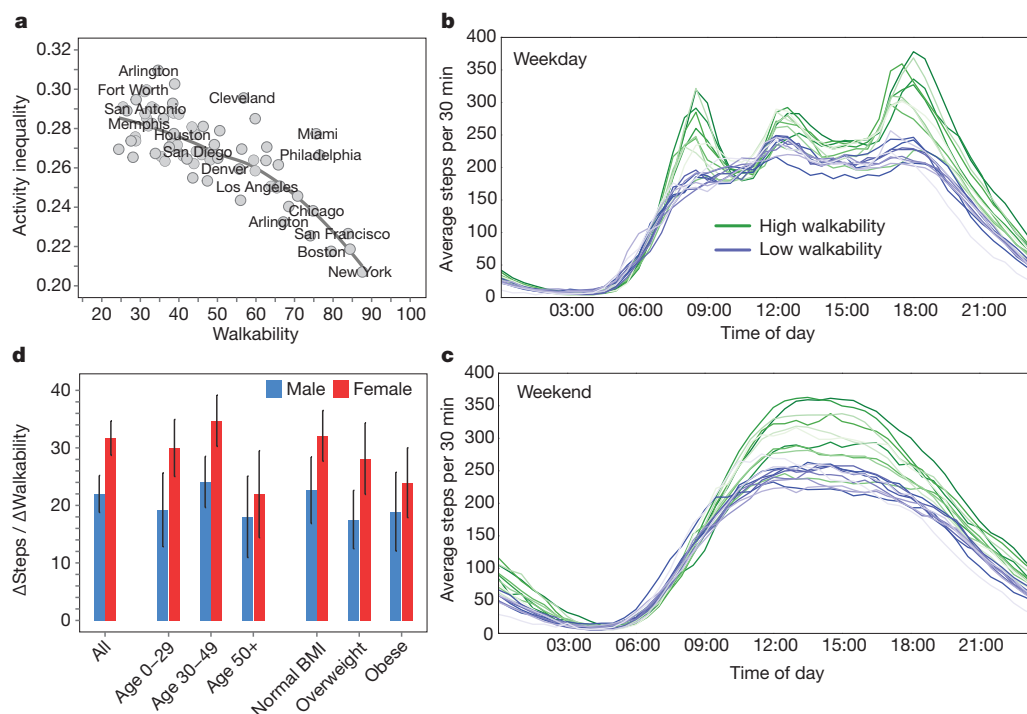


Figure 3 | Aspects of the built environment, such as walkability, may mitigate gender differences in activity and overall activity inequality.

a, Higher walkability scores are associated with lower activity inequality, based on data from 69 cities in the USA (LOESS fit; $R^2 = 0.61$). **b**, **c**, Walkability is linked to increased activity levels. Curves show average steps recorded throughout the day in cities in the USA with the top ten walkability scores (green) and bottom ten walkability scores (blue) in our dataset. **b**, On weekdays, walkable cities exhibit a spike in activity during morning commute (9:00), evening commute (18:00), and lunch times (12:00), while activity is relatively constant and lower overall in less walkable cities. **c**, On weekend days, people in more walkable cities take

more steps throughout the middle of the day, so walkability is associated with higher activity levels even at times when most people are not working or commuting. **d**, Higher walkability is associated with more daily steps across age, gender and BMI groups. Bars show the steps gained per day for each point increase in walkability score for 24 cities in the USA, including 95% confidence intervals (assuming linear model; Methods). Positive values across all bars reveal that, with increasing walkability, significantly more steps are taken by every subgroup (Student's t -tests, all $P < 10^{-6}$). The effect is significantly larger for females overall (left; Student's t -test $P < 10^{-4}$), with the greatest increases for women under 50 years (middle) and individuals with a BMI less than 30 (right).

activity for females in comparison to males and not just an increase in variance overall (Extended Data Fig. 7a). While lower physical activity in females has been reported in several countries^{1,26}, we discover that in countries with low activity and high activity inequality, the gender gap in activity is amplified (Extended Data Fig. 7b).

By quantifying the relationship between activity and obesity at the individual level (Fig. 2c), we were able to determine why a country's activity inequality is a better predictor of obesity than average activity level. We find that the prevalence of obesity increases more rapidly for females than males as activity decreases. And while lower activity is associated with a substantial increase in obesity prevalence for low-activity individuals, there is little change in obesity prevalence among high-activity individuals. So given two countries with identical average activity levels, the country with higher activity inequality will have a greater fraction of low-activity individuals (Fig. 1c), many of them female (Fig. 2b), leading to higher obesity than predicted from average activity levels alone. These findings are analogous to the phenomenon revealed in past studies of the effects of income inequality on health^{27,28}, whereby a relatively small change in income (in our case, activity) for an individual at the bottom of the distribution can lead to substantial improvements in health. On the basis of our model relating activity inequality to obesity prevalence (Fig. 2a), we also performed a simulation experiment which, assuming perfect information (Methods), suggests that interventions focused on reducing activity inequality could result in a reduction in obesity prevalence up to four times greater than in population-wide approaches (Extended Data Fig. 8).

We investigated the walkability of a city as a modifiable extrinsic factor that could increase activity levels⁸ and reduce activity inequality and the gender activity gap. Using data from 69 cities in the USA

(Supplementary Table 2), we find that higher walkability scores are associated with lower activity inequality (Fig. 3a) across all quartiles of median income (Extended Data Fig. 9). Examining San Francisco, San Jose, and Fremont—Californian cities in close geographic proximity—reveals that activity inequality is lowest in San Francisco, the city with the highest walkability (Supplementary Table 3), suggesting that the relationship between walkability and activity inequality holds even for geographically and socioeconomically similar cities. Furthermore, in more walkable cities, activity is higher on weekdays during morning and evening commute times and at lunch time and on weekends during the afternoon (Fig. 3b and c). This indicates that walkable environments increase physical activity during both work and leisure time.

We find that higher walkability is associated with significantly more daily steps across all age, gender, and body mass index (BMI) groups (Fig. 3d; Student's t -tests, all $P < 10^{-6}$). The relationship between walkability and activity is significantly stronger for females (Student's t -test on gender difference, $P < 10^{-4}$), whose activity was also disproportionately reduced with higher activity inequality, with the greatest increases for women under 50 years old. For example, our linear model shows that for 40-year-old women, a 25-point increase in walkability (such as from Sacramento, California, to Oakland, California) is associated with 868 more steps per day, while for men, this 25-point increase is associated with only 622 additional daily steps. While walkability was associated with the greatest increases in recorded steps among normal-weight individuals, even overweight and obese individuals in more walkable cities record more steps.

There are limitations in the instrument we used to collect daily physical activity. For example, our sample is cross-sectional and potentially biased towards individuals of higher socioeconomic status, particularly

in lower-income countries, and towards people interested in their activity and health. However, we find that activity inequality predicts obesity in both middle- and high-income countries (Extended Data Fig. 4) and that walkability predicts activity inequality across four quartiles of median income in cities in the USA (Extended Data Fig. 9), suggesting that our findings are robust to variation in socioeconomic status. The majority of adults in developed countries already own a smartphone and the number of smartphone connections worldwide is expected to increase by 50% by 2020²¹, so we expect any biases to diminish in the future. Walking is the most popular aerobic physical activity²⁹, but our dataset may fail to capture time spent in activities where it is impractical to carry a phone (for example, playing soccer or swimming) or steps are not a major component of the activity (for example, bicycling), and, because users must carry their phone for steps to be recorded, there may exist systematic differences in wear time based on gender and age. However, analysis of our dataset reproduces previously established relationships between activity across geographic locations, gender, and age (Extended Data Figs 1 and 2). We also find that, across countries, the span of time over which steps were recorded is uncorrelated with the number of steps (Extended Data Fig. 10), and thus systematic wear time differences are unlikely to affect our country-level comparisons. Together, these results provide confidence that our dataset is able to identify activity differences between countries, gender, and age groups.

This study presents a new paradigm for population activity studies by demonstrating that smartphones can deliver new insights about key health behaviours. We examine the distribution of activity in 46 countries around the world, including rarely studied countries such as Saudi Arabia and Mexico. Our findings highlight activity inequality as an important indicator of activity disparities in the population and identify 'activity poor' subpopulations, such as women, who could benefit most from interventions to promote physical activity. We further find that higher walkability is associated with reduced activity inequality and greater activity across age, gender, and BMI groups, which indicates the importance of the built environment to global activity levels and health. Our findings can help us to understand the prevalence, spread, and effects of inactivity and obesity within and across countries and subpopulations and to design communities, policies, and interventions that promote greater physical activity.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 1 March; accepted 6 June 2017.

Published online 10 July 2017.

- Hallal, P. C. *et al.* Global physical activity levels: surveillance progress, pitfalls, and prospects. *Lancet* **380**, 247–257 (2012).
- Lee, I.-M. *et al.* Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *Lancet* **380**, 219–229 (2012).
- UN Secretary General. Prevention and control of non-communicable diseases. <http://www.who.int/nmh/publications/2011-report-of-SG-to-UNGA.pdf> (Regional Office for South-East Asia, World Health Organisation, 2011); accessed 21 April 2016.
- World Health Organization (WHO). *Global Recommendations On Physical Activity For Health*. <http://www.who.int/dietphysicalactivity/publications/9789241599979/en/> (WHO, 2010); accessed 21 April 2016.
- Kohl, H. W. *et al.* The pandemic of physical inactivity: global action for public health. *Lancet* **380**, 294–305 (2012).
- Tudor-Locke, C., Hatano, Y., Pangrazi, R. P. & Kang, M. Revisiting “how many steps are enough?”. *Med. Sci. Sports Exerc.* **40**, S537–S543 (2008).
- Sallis, J. F. *et al.* Progress in physical activity over the Olympic quadrennium. *Lancet* **388**, 1325–1336 (2016).

- Bauman, A. E. *et al.* Correlates of physical activity: why are some people physically active and others not? *Lancet* **380**, 258–271 (2012).
- Physical Activity Guidelines Advisory Committee. *Physical Activity Guidelines Report*. <https://health.gov/paguidelines/report/pdf/CommitteeReport.pdf> (US Department of Health and Human Services, 2008); accessed 21 April 2016.
- Chokshi, D. A. & Farley, T. A. Changing behaviors to prevent noncommunicable diseases. *Science* **345**, 1243–1244 (2014).
- Sallis, J. F. *et al.* Physical activity in relation to urban environments in 14 cities worldwide: a cross-sectional study. *Lancet* **387**, 2207–2217 (2016).
- Servick, K. Mind the phone. *Science* **350**, 1306–1309 (2015).
- Reis, R. S. *et al.* Scaling up physical activity interventions worldwide: stepping up to larger and smarter approaches to get people moving. *Lancet* **388**, 1337–1348 (2016).
- Prince, S. A. *et al.* A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. *Int. J. Behav. Nutr. Phys. Act.* **5**, 56 (2008).
- Van Dyck, D. *et al.* International study of objectively measured physical activity and sedentary time with body mass index and obesity: IPEN adult study. *Int. J. Obes.* **39**, 199–207 (2015).
- Walch, O. J., Cochran, A. & Forger, D. B. A global quantification of “normal” sleep schedules using smartphone data. *Sci. Adv.* **2**, e1501705 (2016).
- González, M. C., Hidalgo, C. A. & Barabási, A.-L. Understanding individual human mobility patterns. *Nature* **453**, 779–782 (2008); addendum **458**, 238 (2009).
- Golder, S. A. & Macy, M. W. Diurnal and seasonal mood vary with work, sleep, and daylight across diverse cultures. *Science* **333**, 1878–1881 (2011).
- Wesolowski, A. *et al.* Quantifying the impact of human mobility on malaria. *Science* **338**, 267–270 (2012).
- Blumenstock, J., Cadamuro, G. & On, R. Predicting poverty and wealth from mobile phone metadata. *Science* **350**, 1073–1076 (2015).
- Anthes, E. Mental health: there's an app for that. *Nature* **532**, 20–23 (2016).
- Case, M. A., Burwick, H. A., Volpp, K. G. & Patel, M. S. Accuracy of smartphone applications and wearable devices for tracking physical activity data. *J. Am. Med. Assoc.* **313**, 625–626 (2015).
- Hekler, E. B. *et al.* Validation of physical activity tracking via android smartphones compared to ActiGraph accelerometer: laboratory-based and free-living validation studies. *JMIR mHealth uHealth* **3**, e36 (2015).
- Atkinson, A. B. On the measurement of inequality. *J. Econ. Theory* **2**, 244–263 (1970).
- Allison, P. D. Measures of inequality. *Am. Sociol. Rev.* **43**, 865–880 (1978).
- Brown, W. J., Mielke, G. I. & Kolbe-Alexander, T. L. Gender equality in sport for improved public health. *Lancet* **388**, 1257–1258 (2016).
- Wagstaff, A. & Van Doorslaer, E. Income inequality and health: what does the literature tell us? *Annu. Rev. Public Health* **21**, 543–567 (2000).
- Lynch, J. W., Smith, G. D., Kaplan, G. A. & House, J. S. Income inequality and mortality: importance to health of individual income, psychosocial environment, or material conditions. *Br. Med. J.* **320**, 1200 (2000).
- Centers for Disease Control and Prevention (CDC) *Vital Signs: More People Walk to Better Health*. <http://www.cdc.gov/vitalsigns/walking/> (2012); accessed 3 November 2016.

Supplementary Information is available in the online version of the paper.

Acknowledgements Further information and data are available at <http://activityinequality.stanford.edu>. We thank Azumio for donating the data for independent research, and T. Uchida and W. Hamilton for comments and discussions. T.A., R.S., J.L.H., A.C.K., S.L.D. and J.L. were supported by a National Institutes of Health (NIH) grant (U54 EB020405, Mobilize Center, NIH Big Data to Knowledge Center of Excellence). T.A. was supported by the SAP Stanford Graduate Fellowship. J.L.H. and S.L.D. were supported by grants R24 HD065690 and P2C HD065690 (NIH National Center for Simulation in Rehabilitation Research). J.L. and R.S. were supported by NSF grant IIS-1149837 and the Stanford Data Science Initiative. J.L. is a Chan Zuckerberg Biohub investigator.

Author Contributions T.A. performed the statistical analysis. T.A., R.S., J.L.H., A.C.K., S.L.D. and J.L. jointly analysed the results and wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to J.L. (jure@cs.stanford.edu).

Reviewer Information Nature thanks N. Christakis, P. Hallal, J. Han and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

Statistics. No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

Data set description. We analysed anonymized, retrospective data collected between July 2013 and December 2014 from Apple iPhone smartphone users of the Azumio Argus app, a free application for tracking physical activity and other health behaviours. Data are available at <http://activityinequality.stanford.edu>. We define a step as a unit of activity as determined through iPhone accelerometers and Apple's proprietary algorithms for step counting. The app records step measurements on a minute-by-minute basis. We considered only users with at least ten days of steps data. The dataset contains 111 countries with 100 users or more (717,527 users; 68 million days of data; Fig. 1a). We restricted further analyses to the 46 countries with at least 1,000 or more users (693,806 users; 66 million days of data). We aggregated data from all of these users to the country level. A user's country was assigned based on the most common country identified through the user's IP addresses. In the USA, users were assigned to a city based on the most commonly occurring location of weather updates in the user's activity feed. Weather updates are automatically added to the feed of each user according to the nearest cell phone tower. The user enters gender, age, height and weight in the app settings, and can change these values at any time; we used the most recent recorded values. 28.9% of users report multiple values for their weight; among these users, weight changed by 0.24 kg on average between the first and last recorded weight. Users had on average 95 days with recorded steps, although variation was large (standard deviation $\sigma = 313$ days). Subjects were excluded from a particular analysis if information was unreported (for example, subjects with no reported height or weight were excluded from the analysis of Fig. 2a). The amount of data for each country can be found in Supplementary Table 1. To verify that subjects with missing data on gender, age, or BMI are not different from those who report data, we computed the standardized mean difference in age, gender, BMI, and average steps per day between groups with and without missing data. Across all combinations of missing variables (age, gender, BMI) and outcomes (age, gender, BMI, daily steps), the groups were balanced³⁰, with all standard mean differences lower than 0.25. Data handling and analysis was conducted in accordance with the guidelines of the Stanford University Institutional Review Board.

Verifying established physical activity trends. To determine the ability of our dataset to identify relationships between physical activity and gender, age, BMI, and geographic location, we confirmed that the activity measure (daily steps) in our dataset reproduces trends established in prior work. We find that activity decreased with increasing age^{1,8,31,32} and BMI^{8,15,32}, and is lower in females than in males^{1,8,31–33}, which is consistent with previous reports (Extended Data Fig. 1). We compared our physical activity estimates to physical activity data aggregated by the World Health Organization (WHO)³⁴. The comparison between recorded steps in our dataset and the WHO data are limited for the following reasons. The WHO's dataset is based on self-reports instead of accelerometer-defined measures as in our dataset. It contains the percentage of the population meeting the WHO guidelines for moderate to vigorous physical activity rather than recorded steps, and there is no published direct correspondence between the WHO data and daily steps. Furthermore, the confidence intervals in the WHO dataset are often very large and make a comparison complicated (see for example, Japan: 28–89% meeting guidelines³⁴). Yet we do observe moderate correlation between the two measures (Pearson's correlation coefficient $r = 0.3194$; $P = 0.0393$, Extended Data Fig. 2a). Similarly, we determined the correlation between obesity prevalence in a country in our dataset and comparable WHO estimates from 2014³⁵ ($r = 0.691$; $P < 10^{-6}$, Extended Data Fig. 2b). In addition, we find a significant correlation between the gender gap in activity in our dataset and that reported by the WHO ($r = 0.52$, $P < 10^{-3}$, Extended Data Fig. 2c). For these analyses we used the 46 countries with 1,000 users in our dataset that also had WHO data^{34,35} (that excludes Hong Kong and Taiwan).

Daily recorded steps and wear time. We define a proxy for wear time of the activity-tracking smartphone as daily span of recorded activity; that is, the time between the first and the last recorded step each day. We find that users have an average wear time of 14.0 h per day. To verify that differences in recorded steps between countries are not confounded by differences in wear time from country to country, we compared the average wear time in each country versus the average number of daily steps (Extended Data Fig. 10). We find no significant correlation ($r = -0.086$, $P = 0.57$). Across the 46 countries, males have a 30 min longer average wear time than women (14.2 h versus 13.7 h), which is consistent with longer average sleep duration of females^{16,36}.

Defining activity inequality. We used the Gini coefficient^{24,25} to compute activity inequality, as it is the most commonly used measure with which to quantify inequality and statistical dispersion³⁷. The Gini coefficient is based on the Lorenz

curve, which plots the share of the population's total average daily steps that is cumulatively recorded by the bottom $X\%$ of the population (Extended Data Fig. 5). The Gini coefficient is the ratio of the area that lies between the line of equality and the Lorenz curve (marked A in the diagram) to the total area under the line of equality (marked A and B in the diagram): Gini coefficient = $A/(A + B)$. The Gini coefficient ranges from 0 (complete equality) to 1 (complete inequality), since step counts cannot be negative. Several other measures have been used to quantify inequality and statistical dispersion including the coefficient of variation^{24,25}, decile ratio³⁸, and others^{37,38}; we find that these measures are all highly correlated with the Gini coefficient ($r = 0.96$ or higher) when applied to step counts within countries.

Correlation between activity inequality and obesity. We computed the Pearson correlation coefficient of activity inequality and the prevalence of obesity in a country (Fig. 2a; $r = 0.79$; $P < 10^{-10}$; $R^2 = 0.64$) using local polynomial regression fitting (LOESS; R statistical software package with a re-descending M estimator and Tukey's biweight function). We included all subjects with reported height and weight. We additionally correlated obesity with average daily steps for users in a country and compared the Pearson correlation coefficient for average daily steps with that for activity inequality ($r = -0.62$; $P < 10^{-5}$; $R^2 = 0.47$; Extended Data Fig. 6). Steiger's Z -test³⁹ shows that activity inequality is more strongly correlated with obesity than the average volume of steps recorded in a country ($r = 0.79$ versus -0.62 ; $N = 46$; $t = 2.86$; $P < 0.01$). For example, the UK exhibits higher obesity prevalence than Germany and France (19.5% versus 14.3% and 8.9%, respectively), even though the UK has higher average daily steps (5,444 versus 5,205 and 5,141, respectively). However, the high obesity levels in the UK are matched to their high activity inequality compared to Germany and France (0.288 versus 0.266 and 0.268, respectively).

Robustness of correlation between activity inequality and obesity. While for some countries the gender ratio in our sample closely matched official estimates (for example, the USA, Canada and Australia), our sample is more biased in other countries (for example, Japan, Germany and India; Supplementary Table 1). There is also a bias towards younger subjects in many countries (for example, population median age⁴⁰ for the USA is 34 years versus 38 years; UK is 33 versus 41; Japan is 38 years versus 47 years; Brazil is 33 years versus 32 years). Our sample further includes both middle- and high-income countries, as classified by the World Bank⁴⁰. To verify the robustness of our results, we calculated gender-unbiased estimates for activity inequality and obesity prevalence for each country by reweighting males and females in our sample to match exactly World Bank estimates⁴¹ using a bootstrap⁴² with 500 replications. In addition, we computed activity inequality separately for males and females in each country and then correlated the activity inequality for each gender with obesity prevalence for that gender. We also computed the correlation between obesity and activity inequality for specific age groups in our dataset: [10,20], [20,30], [30,40], [40,50] and [50,100], again using only subjects with a reported age. In addition, we stratified countries by middle- versus high-income status. In all cases, activity inequality remains a strong predictor of obesity (Extended Data Figs 3 and 4), which makes our findings independent of the exact age and gender distributions in our sample and suggests that our results are not confounded by the middle- versus high-income status of countries or isolated to high-income countries. Note that the results of these robustness analyses also show that our findings are not explained by patterns of missing data in our sample. We find similar results in analyses that include all subjects (Fig. 2a) or only those that report gender (Extended Data Fig. 3a, b) or age (Extended Data Fig. 3c). We further verified that the relationship between activity inequality and obesity is not unduly driven by outliers. We removed the potential outliers of Indonesia, Malaysia and the Philippines from our dataset and found that activity inequality was still a better predictor of obesity than average volume of steps recorded (R^2 was 0.69 for activity inequality versus 0.56 for average steps).

Gender gaps in activity and obesity. To determine how activity varies with increasing activity inequality across countries, we calculated the 25th, 50th and 75th percentile of daily steps in each country, with separate calculations for males and females. We then fitted a linear model based on each country's activity inequality to each percentile/gender group, along with 95% confidence intervals (Fig. 2b). We determined the relationship between obesity prevalence and average daily steps for males and females in our sample by measuring the fraction of obese subjects who recorded a certain amount of activity (1,000–2,000 daily steps, 2,000–3,000 daily steps, ..., 10,000–11,000 daily steps) and then computing bootstrapped 95% confidence intervals (Fig. 2c). This analysis included all subjects in the dataset who reported height and weight ($n = 297,268$). We computed the proportion of variability explained by the gender gap in activity using the R^2 measure (Extended Data Fig. 7b).

City walkability analysis. Walkability scores were obtained from Walk Score⁴³. Scores are on a scale of 1 to 100 (100 = most walkable) and are based on amenities (such as shops and parks) within a 0.25 to 1.5 mile radius (a decay function

penalizes more distant amenities) and measures of friendliness to pedestrians, such as city block length and intersection density. At a city level, the score shows good correlation with gold-standard measures of walkability from geographic information systems⁴⁴. For the 69 cities in the USA with at least 200 Azumio users (Supplementary Table 2), we correlated walkability scores with the activity inequality on the city level (that is, using the within-city distribution of average daily step counts). We verified that correlations between walkability and activity inequality are similar when controlling for the median income level of the city by grouping the 69 cities used in Fig. 3a into quartiles based on median household income data from the 2015 American Community Survey^{45,46}. We find that walkable environments are associated with lower levels of activity inequality for all four median income groups (Extended Data Fig. 9). We next analysed activity in our dataset throughout the day on weekdays and weekend days in the ten cities with the highest walkability scores, and the ten cities with the lowest walkability score. We considered only cities with at least 20,000 weekdays of tracked steps across all users for this analysis. We aggregated steps taken over time within each city to the average number of steps per 30-min interval. We considered only days (1) with at least 60 min with non-zero steps, (2) with first and last recorded step at least 8 h apart, and (3) with recorded total steps between 500 and 100,000. We examined a subset of similar cities in close geographic proximity to show that our results cannot be explained by simple differences in geographic variation or city populations (Supplementary Table 3).

Impact of walkability on daily steps. We computed the relationship between walkability and average daily steps for several subgroups of our sample. We used data from cities in the USA that had at least 25 Azumio users in each subgroup (age 0–29, age 30–49, age 50+, normal BMI, overweight, obese, and all; for both males and females). There are 24 such cities in the dataset (Supplementary Table 4). The number of subjects for each group and city is shown in Supplementary Table 4. For each group, we ran independent linear regressions of steps on walkability at the per-subject level. The models include an intercept coefficient. We determined the estimated coefficient of walkability (that is, the increase in daily steps for each one-point increase in walkability of a city) along with 95% confidence intervals (based on Student's *t*-distribution) for each subgroup (Fig. 3d). We refer to the set of these coefficients as our linear model in the main text. We performed Student's *t*-tests on the regression model coefficients, which establish that higher walkability is associated with significantly more daily steps across all age, gender, and BMI groups (all $P < 10^{-6}$) and that the relationship between walkability and activity is significantly stronger for females than males (Student's *t*-test, $P < 10^{-4}$).

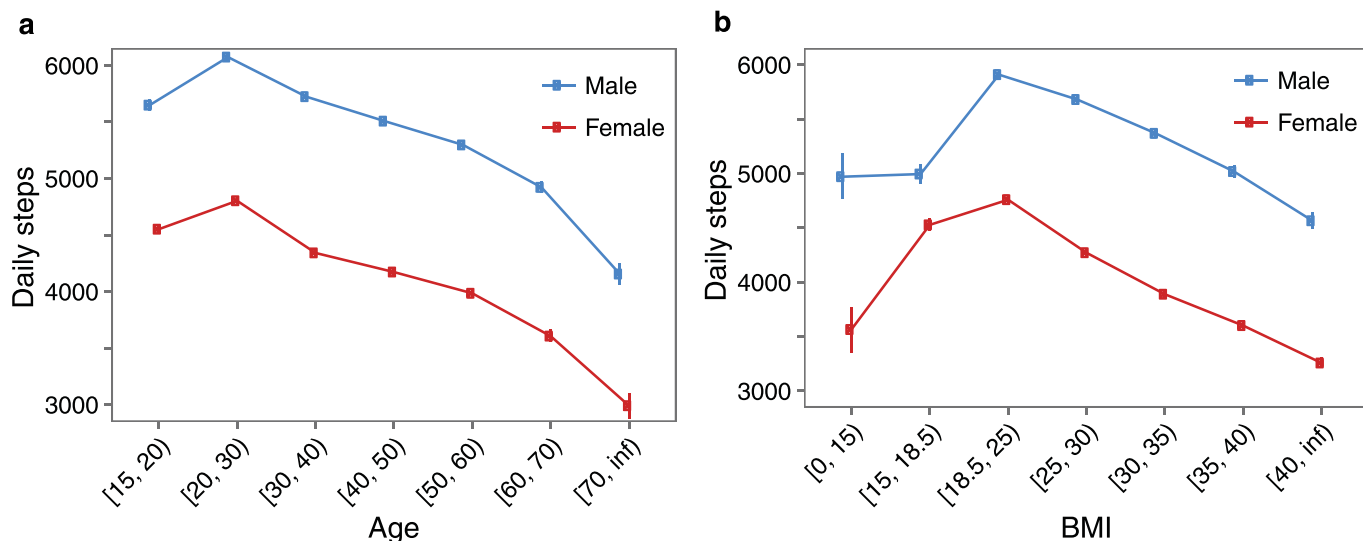
Simulating population-level changes in activity. We used our model relating activity inequality to obesity prevalence to simulate how changes in activity might affect a country's obesity prevalence. We consider an activity budget of 100 additional daily steps per person in a country to distribute across the population (we found similar results for different activity budgets). We compared two strategies for distributing the steps—a population-wide distribution and an inequality-centric distribution. Both strategies result in the same shift in the average activity level of a country. For the population-wide distribution strategy, we increased each individual's daily activity by 100 steps. We then recomputed the country's activity inequality after the redistribution and estimated the country's new obesity prevalence based on our inequality-obesity model (LOESS fit in Fig. 2a). We next tested a strategy focused on activity inequality, in which we distributed the activity budget equally among the activity-poorest $X\%$ of the population (for example, the bottom 20% of the population would increase their daily steps by 500). For the

inequality-centric strategy, we computed the optimal fraction X for each country that results in the greatest reduction in the country's activity inequality. Optimal values for X across all countries ranged from 5% to 9%. Further, assuming a fixed X (for example, $X = 10\%$) yielded similar results. Our simulation assumes perfect knowledge of population activity levels and perfect compliance; that is, any user targeted in this simulation would increase their activity levels according to the available budget. We also assume that other factors affecting weight would be held constant when activity levels change. In our simulations, the inequality-centric intervention resulted in reductions in obesity prevalence of up to 8.3% (median 4.0%; Extended Data Fig. 8), whereas the population-wide approach led to reductions of up to 2.3% (median 1.0%). Thus, activity inequality-centric interventions could result in a reduction in obesity prevalence with a median up to four times greater than in the population-wide approach.

Data availability. Data are available at <http://activityinequality.stanford.edu/>.

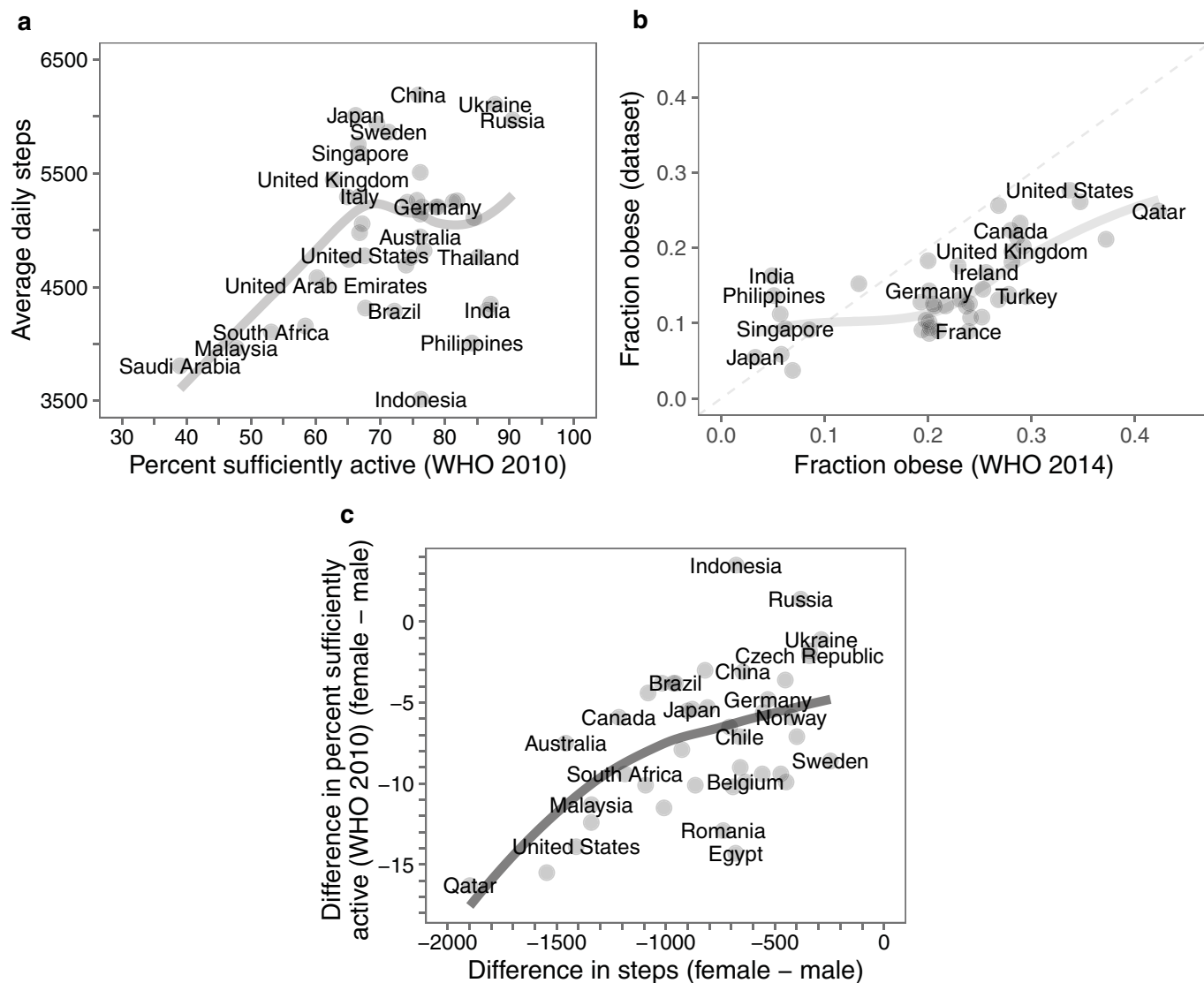
Code availability. Code is available at <http://activityinequality.stanford.edu/>.

30. Stuart, E. A. Matching methods for causal inference: a review and a look forward. *Stat. Sci.* **25**, 1–21 (2010).
31. Bassett, D. R., Wyatt, H. R., Thompson, H., Peters, J. C. & Hill, J. O. Pedometer-measured physical activity and health behaviors in U.S. adults. *Med. Sci. Sports Exerc.* **42**, 1819–1825 (2010).
32. Troiano, R. P. et al. Physical activity in the United States measured by accelerometer. *Med. Sci. Sports Exerc.* **40**, 181–188 (2008).
33. Tudor-Locke, C., Johnson, W. D. & Katzmarzyk, P. T. Accelerometer-determined steps per day in US adults. *Med. Sci. Sports Exerc.* **41**, 1384–1391 (2009).
34. World Health Organization. *Prevalence of Insufficient Physical Activity among Adults: Data by Country*. <http://apps.who.int/gho/data/node.main.A893?lang=en> (Global Health Observatory data repository, WHO, accessed 19 May 2016).
35. World Health Organization. *Obesity (Body Mass Index ≥ 30) (Age-Standardized Estimate): Estimates by Country*. <http://apps.who.int/gho/data/node.main.A900A?lang=en> (Global Health Observatory data repository, WHO, accessed 19 May 2016).
36. Basner, M. et al. American time use survey: sleep time and its relationship to waking activities. *Sleep* **30**, 1085–1095 (2007).
37. De Maio, F. G. Income inequality measures. *J. Epidemiol. Community Health* **61**, 849–852 (2007).
38. Kawachi, I. & Kennedy, B. P. The relationship of income inequality to mortality: does the choice of indicator matter? *Soc. Sci. Med.* **45**, 1121–1127 (1997).
39. Steiger, J. H. Tests for comparing elements of a correlation matrix. *Psychol. Bull.* **87**, 245–251 (1980).
40. CIA World Factbook. *Field Listing: Median Age* <https://www.cia.gov/library/publications/the-world-factbook/fields/2177.html> (CIA, accessed 22 June 2017).
41. World Bank. *World Bank Country and Lending Groups*. <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups> (World Bank, accessed 5 October 2016).
42. World Bank. *Population, Female (% of Total)*. <http://data.worldbank.org/indicator/SP.POP.TOTL.FE.ZS> (World Bank, accessed 10 May 2016).
43. Efron, B. & Tibshirani, R. J. *An Introduction to the Bootstrap* (CRC Press, 1994).
44. Walk Score. <https://www.walkscore.com/cities-and-neighborhoods/> (Walk Score, accessed 17 May 2016).
45. Duncan, D. T., Aldstadt, J., Whalen, J., Melly, S. J. & Gortmaker, S. L. Validation of Walk Score for estimating neighborhood walkability: an analysis of four US metropolitan areas. *Int. J. Environ. Res. Public Health* **8**, 4160–4179 (2011).
46. United States Census Bureau (USCB). *American Community Survey*. <http://www.census.gov/programs-surveys/acs/> (USCB, accessed 5 October 2016).
47. United States Census Bureau (USCB). *Bay Area Census. 2010 Census and American Community Survey 2006–2010*. <http://www.bayareacensus.ca.gov/cities/cities.htm> (accessed 5 July 2016).



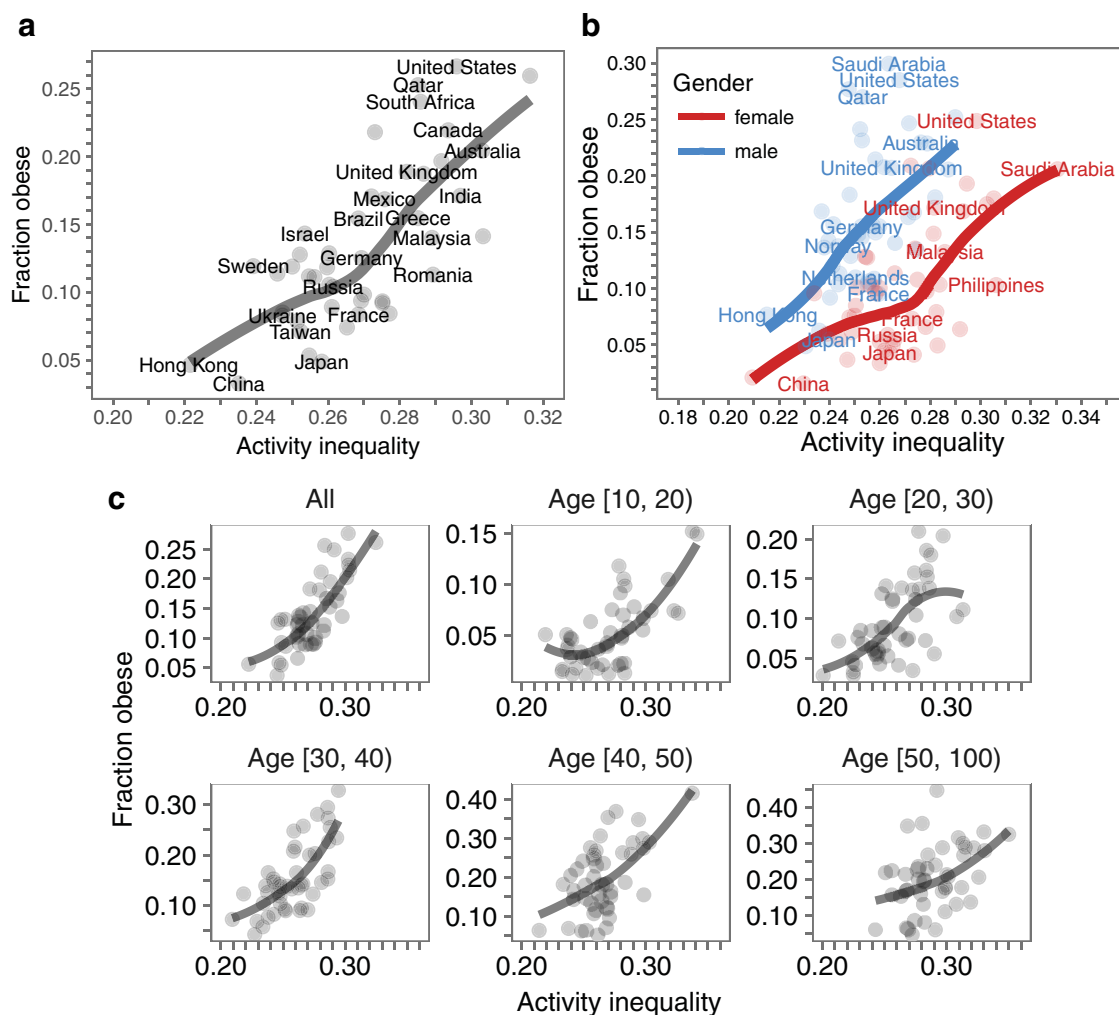
Extended Data Figure 1 | Activity and obesity data gathered with smartphones exhibit well established trends. Daily step counts are shown across age (a) and BMI groups (b) for all users. Error bars correspond to

bootstrapped 95% confidence intervals. Observed trends in the dataset are consistent with previous findings; that is, activity decreases with increasing age^{1,8,31,32} and BMI^{8,15,31}, and is lower in females than in males^{1,8,31–33}.



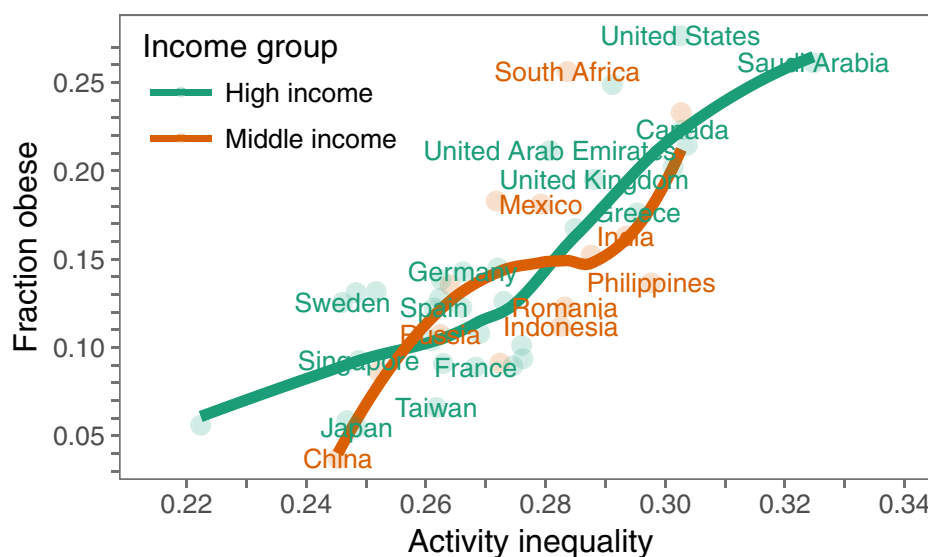
Extended Data Figure 2 | Activity and obesity data gathered with smartphones are significantly correlated with previously reported estimates based on self-report. **a**, WHO physical activity measure³⁴ versus smartphone activity measure (LOESS fit). The WHO measure corresponds to the percentage of the population meeting the WHO guidelines for moderate to vigorous physical activity based on self-report. The smartphone activity measure is based on accelerometer-defined average daily steps. We find a correlation of $r = 0.3194$ between the two measures ($P < 0.05$). Note that this comparison is limited because there is no direct correspondence between the two measures—values of self-report and accelerometer-defined activity can differ¹⁴, and the WHO

confidence intervals are very large for many countries (Methods). **b**, WHO obesity estimates³⁵, based on self-reports to survey conductors, versus obesity estimates in our dataset, based on height and weight reported to the activity-tracking app (LOESS fit). We find a significant correlation of $r = 0.691$ between the two estimates ($P < 10^{-6}$). **c**, Gender gap in activity estimated from smartphones is strongly correlated with previously reported estimates based on self-report (LOESS fit). We find that the difference in average steps per day between females and males is strongly correlated to the difference in the fraction of each gender who report being sufficiently active according to the WHO ($r = 0.52$, $P < 10^{-3}$).



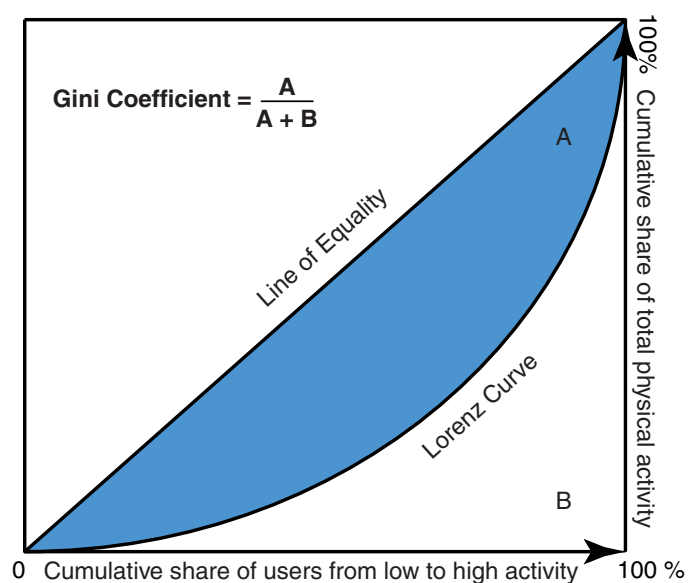
Extended Data Figure 3 | Activity inequality remains a strong predictor of obesity levels across countries when reweighting the sample based on officially reported gender distributions and when stratifying by gender or age. a, Obesity versus activity inequality on a country level where subjects are reweighted to accurately reflect the official gender distribution in each country (LOESS fit; Methods). The gender-unbiased estimates are very similar to estimates using all data ($r = 0.953$ for activity inequality and

$r = 0.986$ for obesity). **b**, Obesity versus activity inequality on a country level for males and females. Activity inequality predicts obesity for both genders (LOESS fit). **c**, Obesity versus activity inequality on a country level across different age groups. We find that associations between activity inequality and obesity persist within every age group (LOESS fits). These results indicate that our main result—activity inequality predicts obesity—is independent of any potential gender and age bias in our sample.

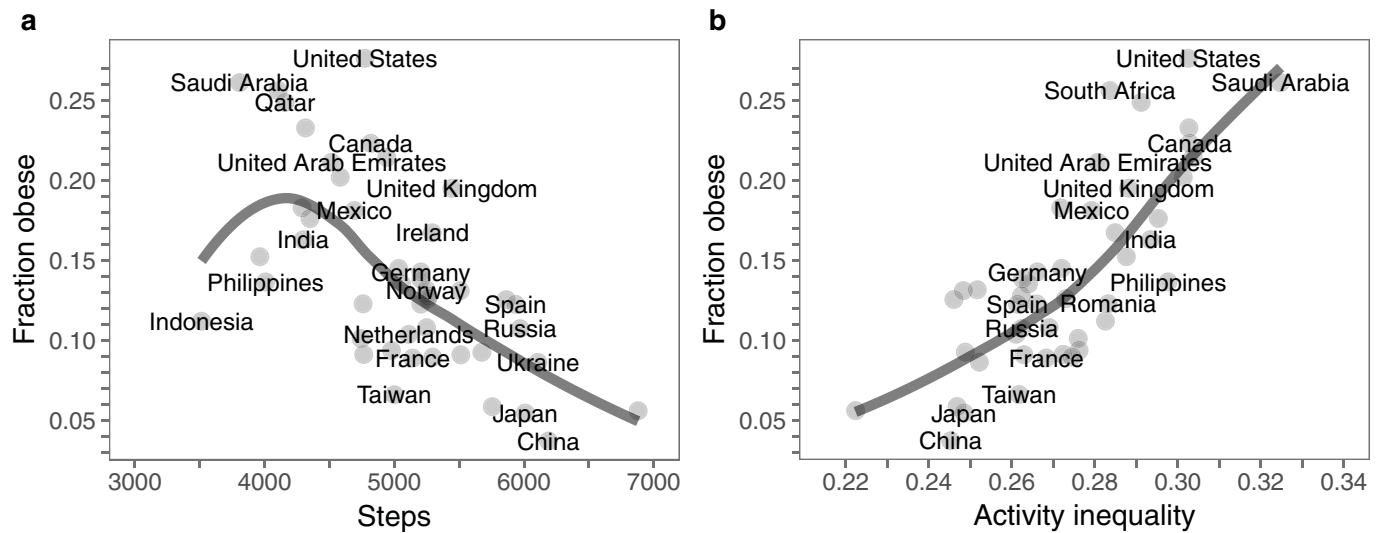


Extended Data Figure 4 | Relationship between activity inequality and obesity holds within countries of similar income. Of the 46 countries included in Fig. 2a, we have 32 high-income (green) and 14 middle-income (orange) countries according to the current World Bank classification⁴⁰. We find that activity inequality is a strong predictor of obesity levels in both high-income countries and middle-income countries

(LOESS fit). Although in middle-income countries, iPhone users might belong to the wealthiest in the population, in high-income countries iPhones are used by a larger proportion of the population. That we find a strong relationship between activity inequality and obesity in both groups of countries suggests that our findings are robust to differences in wealth in our sample.

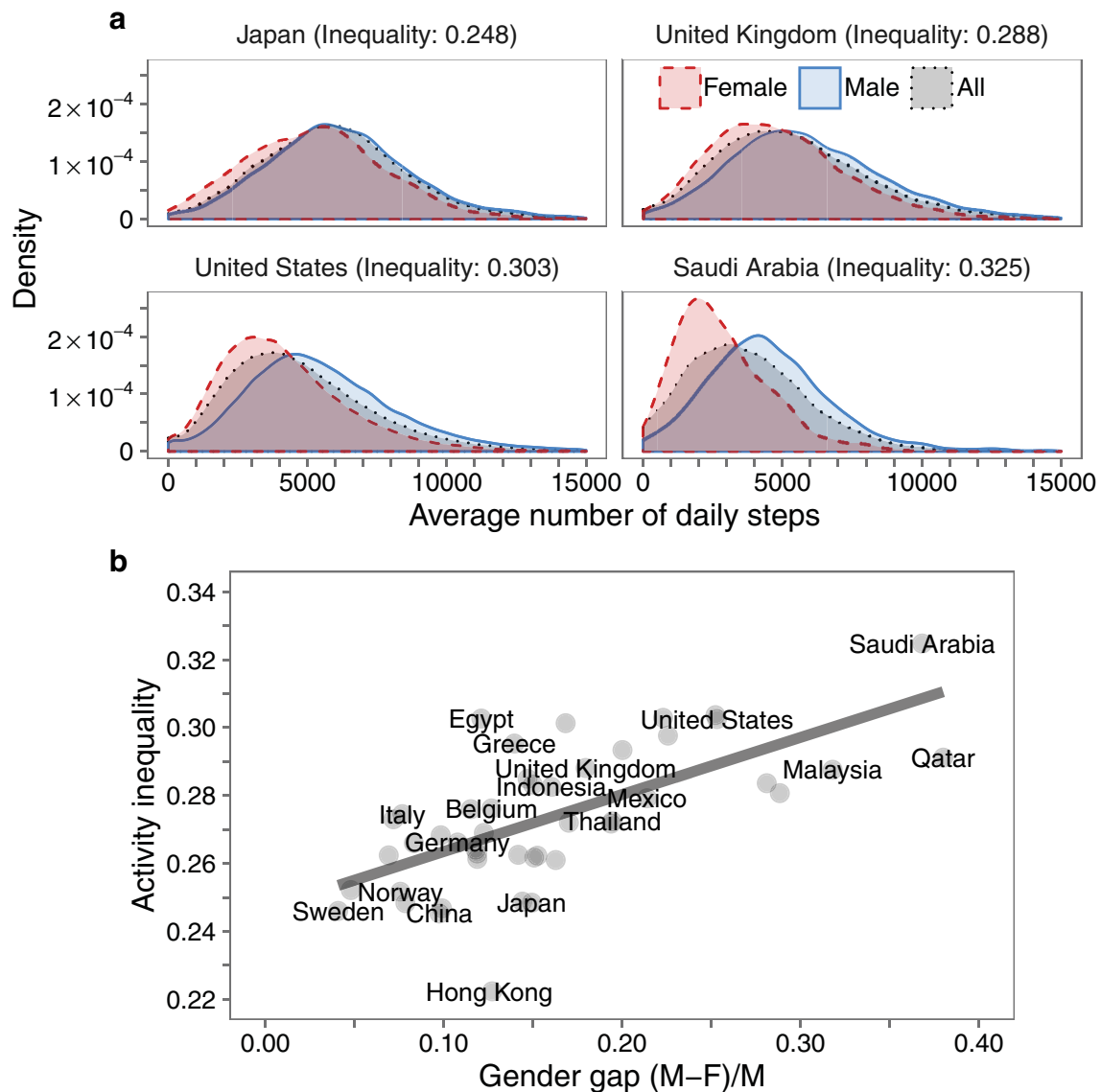


Extended Data Figure 5 | Graphical definition of activity inequality measure using the Gini coefficient. The Lorenz curve plots the share of total physical activity of the population on the y axis that is cumulatively performed by the bottom X% of the population, ordered by physical activity level. The diagonal line at 45 degrees represents perfect equality of physical activity (that is, everyone in the population is equally active). The Gini coefficient is defined as the ratio of the area that lies between the line of equality and the Lorenz curve (marked A in the diagram) to the total area under the line of equality (marked A and B in the diagram). The Gini coefficient for physical activity can range from 0 (complete equality) to 1 (complete inequality).



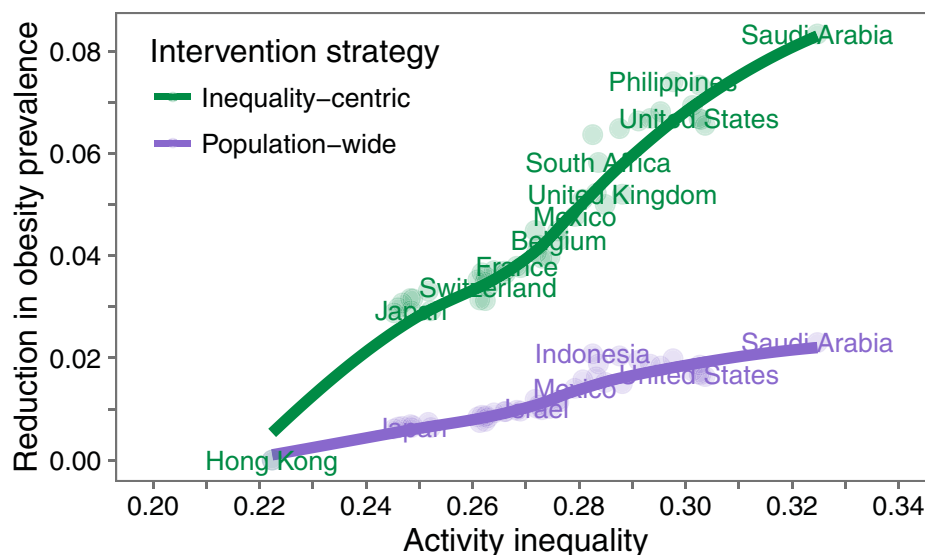
Extended Data Figure 6 | Activity inequality is a better predictor of obesity than the average activity level. **a**, Obesity is significantly correlated with the average number of daily steps in each country (LOESS fit; $R^2 = 0.47$). **b**, However, activity inequality is the better predictor of obesity (LOESS fit; $R^2 = 0.64$). The difference is significant according to Steiger's Z-test ($P < 0.01$; Methods). This shows that there is value

to measuring and modelling physical activity across countries beyond average activity levels. Activity inequality captures the variance of the distribution; that is, how many activity-rich and activity-poor people there are, allowing for better prediction of obesity levels. Panel **b** is repeated from Fig. 2a for comparison.



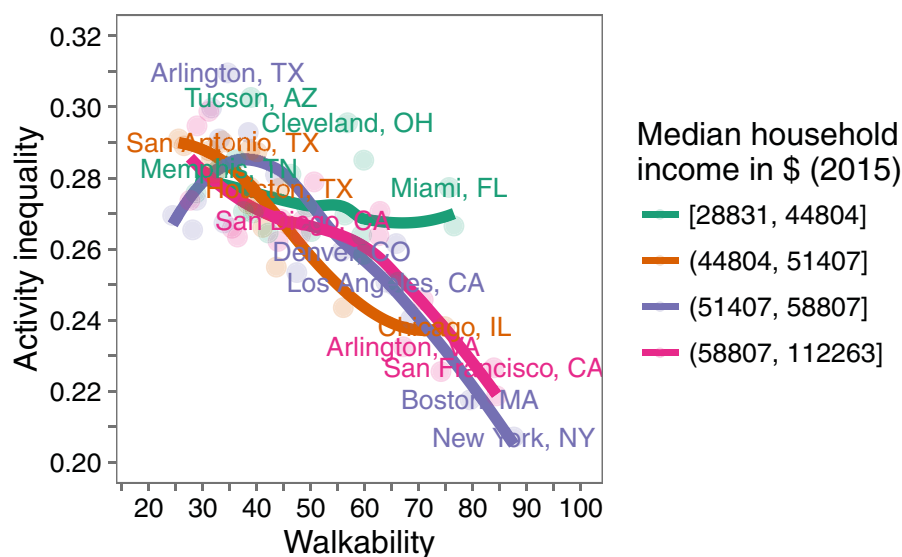
Extended Data Figure 7 | Female activity is reduced disproportionately in countries with high activity inequality. **a**, Distribution of daily steps for females, males, and all users in representative countries of increasing activity inequality (Japan, the UK, the USA and Saudi Arabia). While in countries with low activity inequality females and males have very similar levels of activity (for example, Japan), the distributions of female and male activity differ greatly for countries with high activity inequality (for example, Saudi Arabia and the USA). Activity distributions in these countries demonstrate that larger variances in activity (Fig. 1c) are due

to a disproportionate reduction in the activity of females and not just an increase in variance overall. **b**, Activity inequality increases with the relative activity gender gap on a country level (linear fit; Methods). We find that the relative gender gap ranges between 0.041 (Sweden) and 0.380 (Qatar). The average of daily steps for females is lower than for males in all 46 countries. The gender gap explains 43% of the observed variance in activity inequality (linear fit; $R^2 = 0.43$). This suggests that activity inequality could be greatly reduced through increases in female activity alone.



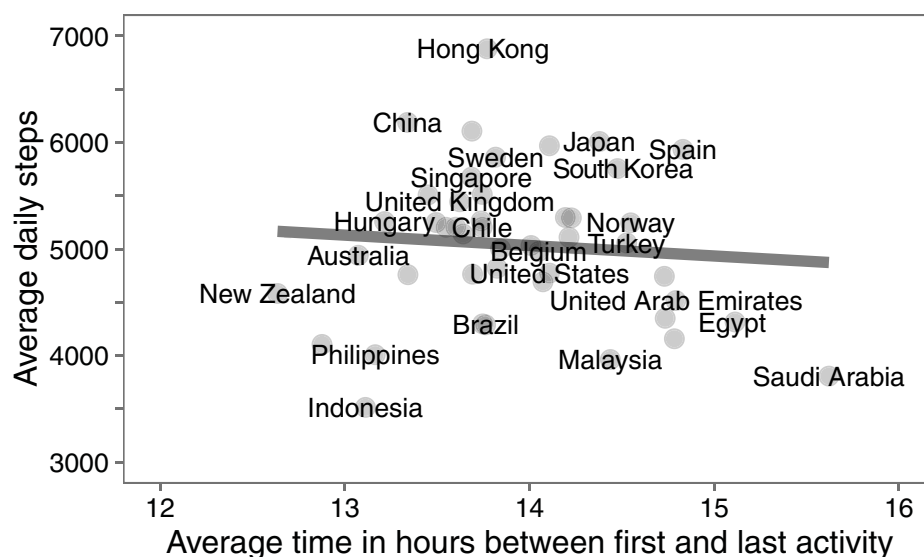
Extended Data Figure 8 | Interventions focused on reducing activity inequality could result in reductions in obesity prevalence up to four times greater than for population-wide approaches. Given a fixed activity budget (100 daily steps per individual) to distribute across the population, we compare an inequality-centric strategy that distributes this budget equally to minimize activity inequality (100/ X % daily steps increase for the activity-poorest X % where X minimizes the country's

resulting activity inequality; Methods) and a population-wide strategy which distributes the budget equally across the entire population (100 daily steps per individual; Methods). From our simulations, we find that the inequality-centric strategy would lead to predicted reductions in obesity prevalence of up to 8.3% (median 4.0%), whereas the population-wide approach would lead to predicted reductions of up to 2.3% (median 1.0%). Lines correspond to LOESS fits.



Extended Data Figure 9 | Relationship between walkability and activity inequality holds within cities in the USA of similar income. Walkable environments are associated with lower levels of activity inequality within socioeconomically similar groups of cities. We group the 69 cities into quartiles based on median household income (data from the 2015 American Community Survey⁴⁶). We find that walkable environments

are associated with lower levels of activity inequality for all four groups (LOESS fit). The effect appears to be attenuated for cities in the lowest median household income quartile. These results suggest that our main result—activity inequality predicts obesity and is mediated by factors of the physical environment—is independent of any potential socioeconomic bias in our sample.



Extended Data Figure 10 | Differences in country level daily steps are not explained by differences in estimated wear time. Users have an average span of 14.0h between the first and last recorded step, our proxy for daily wear time (Methods). While on an individual level, longer estimated wear time is associated with more daily steps ($r = 0.427$, $P < 10^{-10}$), on a country level, there is no significant association between

wear time and daily steps ($r = -0.086$, $P = 0.57$). The line shows the linear fit using the 46 countries with at least 1,000 users. This suggests that differences in recorded steps between countries are due to actual differences in physical activity behaviour and are not explained by differences in wear time.