Supplementary Information for:

Taking connected mobile-health diagnostics of infectious diseases to the field

Christopher S. Wood[1,2,3,4,11], Michael R. Thomas[1,2,3,11], Jobie Budd[5], Tivani P. Mashamba-Thompson[6], Kobus Herbst[7], Deenan Pillay[7,8], Rosanna W. Peeling[9], Anne M. Johnson[10], Rachel A. McKendry[5] & Molly M. Stevens[1,2,3,4]*

[1]Department of Materials, Imperial College London, London, UK.
[2]Department of Bioengineering, Imperial College London, London, UK.
[3]Institute of Biomedical Engineering, Imperial College London, London, UK.
[4]Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden.
[5]London Centre for Nanotechnology, Division of Medicine, University College London, London, UK.
[6]Discipline of Public Health Medicine, School of Nursing and Public Health, University of KwaZulu-Natal, Durban, South Africa.
[7]Africa Health Research Institute, Durban, KwaZulu-Natal, South Africa.
[8]Division of Infection and Immunity, University College London, London, UK.
[9]Department of Clinical Research, London School of Hygiene and Tropical Medicine, London, UK.
[10]Institute for Global Health, University College London, London, UK.
[11]These authors contributed equally: Christopher S. Wood, Michael R. Thomas.
*e-mail: m.stevens@imperial.ac.uk

Raw and Intermediate data to recreate figures 1 and 4 can be found at DOI:10.5281/zenodo.1320937

This supplementary information document is divided into three sections:

A. Data sources and processing methods relating to Figure 1: Smartphone Adoption and Capability 2007-2020.

B. Data sources and processing methods relating to Figure 4: Access to Health Facilities and Mobile Networks in Uganda.

C. Data tables relating to both figures:

   Table 1. Smartphone adoption percentages, global and sub-Saharan Africa
   Table 2. Smartphone specifications, 'High End' and 'Low Cost'
   Table 3. Calculated Ugandan populations in range of cellular networks and health facilities

   Cell tower and health facility data points are not included in this document but can be found in the 'Raw Data' folder at DOI:10.5281/zenodo.1320937

D. References

A. Figure 1: Smartphone Adoption and Capability 2007-2020

Data Sources:

1. Smartphone Adoption

Global smartphone adoption data was obtained from the 2015 GSMA Mobile Economy report[1], with future projections maintained by the 2017 report[2]. Sub-Saharan Africa data was obtained from GSMA Mobile Economy Sub Saharan Africa 2014[3].

GSMA (GSM Association) is a global trade body that represents the interests of mobile network operators. It collates annual regional and global reports of the mobile ecosystem, including estimates for smartphone adoption calculated from active SIM subscriptions, and forecasted percentages informed by regional pricing and connectivity[4].

All data used for plotting is shown in **Table 1 in Section C** of this document.

2. Smartphone Hardware Specifications

'Low cost' smartphone manufacturers were identified from GSMA mobile trends report 2017[5] that states that "lower cost smartphones from local manufacturers such as *Huawei, Oppo, OnePlus* and *Xiaomi* in China, *Micromax* in India, and now *AfriOne* in Nigeria, are helping to address the affordability barrier". Phones from *Apple* and *Samsung*, traditionally best-selling in developed markets[6], were chosen for comparison as more expensive 'high end' smartphone model ranges.

All smartphone specifications were accessed from GSMArena[7] between July 2017 and April 2018.

GSMArena is an advertising-funded comparison website which collates manufacturer-stated hardware specifications and runs independent hardware comparison tests.

All data used for plotting is shown in **Table 2** in **Section C** of this document.

Processing Method:

The specifications of smartphone ranges from *Samsung, Apple, Huawei, Vivo, Xiaomi* and *Micromax* were found on GSMArena (No *AfriOne* phones were listed).
Phones from *Apple* and *Samsung*, traditionally best-selling and more expensive model ranges, were compared to models from ranges from *Huawei, Vivo* (same parent company as *Oppo*, but aimed at emerging markets), and *Micromax* that had models listed under €160 on gsmarena.com. No *OnePlus* models were listed below this price.

Prices listed on GSMArena refer to cost at the time of data entry, which may be relevant to purchase of model at time of release, or market price of older models. Here they are only used for comparison and to identify the most affordable model ranges.

Models chosen from these manufacturers were those that had the longest continuous release range, e.g: *Samsung Galaxy S*, *Samsung Galaxy S2*….*Samsung Galaxy S8* or *Xiaomi Redmi 1-5*. Since smartphone manufacturers release different model lines with different audiences and therefore different specs, following the evolution of one model line over time shows the improvement in manufacturing capability and reductions in manufacturing costs. In the case that many similar models were released on the same range (e.g. *MicroMax Canvas*), models of similar price and released more than six months after the previous model were chosen to plot.

The specifications collected from GSMArena are: release date, primary camera megapixels, CPU GHz, CPU cores, network technology (highest generation, GSM = 2G, CDMA/UMTS = 3G, LTE = 4G)

The release date was decimalised for ease of plotting (June 2016 became 2016.5).

The CPU clock frequency was calculated using Amdahl's Law[8] at 50% efficiency to take into account CPU frequency and the number of cores:

Increase in speed =1/((1-parallelisation)+(parallelisation/number of cores))

Multiple CPU cores effectively increase processor speed, but only when computations are parallelisable. Different smartphone operating systems and apps allow for different parallelisation efficiencies, but extra cores allow multiple apps to be run at the same time. An arbitrary value of 50% parallelisation was chosen for all smartphone models.

Data taken from gsmarena and calculations saved as phone specs for plotting.xlsx with plotted values highlighted in yellow.

In the case that CPU cores of two different clock speeds are used in same device, the mean is used for calculation.

Decimalised release date (x), calculated CPU frequency (y) and camera MP (z) were plotted using Matlab's scatter function scatter(x,y,z) where the z value is shown as the size of the plotted point (larger z value, larger radius of point). The camera MP values were multiplied by a constant of 50 so that point sizes were easier to see. A straight line fit for all data was plotted using Matlab's curve fitting toolbox (polynomial fit with 1 degree) Graph was imported into Adobe Illustrator where divisions between data points based on network compatibility were made manually. Vector circles were drawn over data points to reduce pixelation.

Smartphone penetration data was plotted on the same axis using Matlab's bar() function. Vector squares were drawn over bars in Adobe Illustrator, where the colour was changed for predicted values.

B. Figure 4: Access to Health Facilities and Mobile Networks in Uganda

Usage Note:

This figure is only intended to estimate proximity to cellular network and health infrastructure for the illustration of the 'digital divide' in Uganda. It is relevant to the date of each dataset included, and is not intended for use in planning public interventions.

Data Sources:

1. Population distribution grid data for Uganda

Population distribution data was downloaded from WorldPop[9] (Creative Commons Attribution 4.0 International License) and has a resolution of 100x100 metres per pixel.
This data is an estimate for 2015, projected from data available up to 2010, and normalised to UN predictions of population. This data was generated by Linard *et. al*[10]**,** collating government census population data and land use data derived from satellite imagery[11].

This data is accessible as a .tif image file in raster data format listed as PopulationDensityB.tif in 'Raw Data'.

2. Location of health facilities

Health facility data was downloaded on 14/08/17 from the Uganda Bureau of Statistics[12] where it was uploaded by UBOS-GIS, the mapping administration of the Uganda Bureau of Statistics. This list contains 3876 facilities.

This dataset is a collation of studies from Makerere University Department of Geography in 2003 and census information collected by the Uganda Bureau of Statistics up to 2011. Some health facilities are listed as 'under construction' but since this data comes from a survey from 2011, they are assumed to be open at time of processing. Location coordinates of nearby health facilities were collected by field workers using GPS during census collection, however some facilities are listed as more than 1 km from the reported location and so not all locations are exact. This dataset was chosen as being the only available official health facility inventory for the whole of Uganda with location coordinates at the time of writing. The United Nations Office for the Coordination of Humanitarian Affairs also surveyed health facility locations in the North of the country in 2011[13].

Since this review, the Uganda Bureau of Statistics has released a data portal showing the same dataset updated in 2014. This dataset contains 3715 health facilities[14]. The Ugandan Ministry of Health also released an updated Health Facilities Inventory in 2017, although without location coordinates[15].

This data is accessible as a .csv file with health facility location and metadata listed as HealthCentrePointsB.csv in 'Raw Data'.

Only the health facility location was used for this review.

3. Cell tower data

Cell tower data including location coordinates, network generation (2G, 3G, 4G), and range (distance of measurement from cell tower in metres) were downloaded from OpencelliD[16].

OpenCelliD is an open-source community project with the original intention of listing cell tower locations as an alternative to satellite navigation systems to provide cell phone users with location information. Cell tower location data is contributed by smartphone users with tracking apps, commercial tracking devices, and corporations who have donated their data.
As tracking relies on the contributor passing within range of a cell tower, some cell towers in sparsely populated areas may be missed and therefore cellular network coverage underestimated from this dataset. Likewise, the distance of the signal measured does not necessarily correlate to the furthest distance from the cell tower that the signal is detectable, further underestimating the calculated signal coverage.

Conversely, the location of measurement is not stated and so a uniform radial range of signal up to the measured distance is assumed, likely overestimating network coverage as variations in elevation, landscape and building height and density can all affect signal strength and range.
Detected signals in the dataset from 2G cell towers over the theoretical limit of ~35 km are assumed to be from a boosted signal, although again this is unlikely to be uniformly radial from the original signal source, and likely to cause overestimations in calculated signal coverage.

cell_towers.csv, as noted in[17] was downloaded 08/08/17 with a temporary API key (licensed under a Creative Commons Attribution-ShareAlike 4.0 International License). This data is updated daily and so is relevant up to download date. The cell towers located in Uganda were extracted from the global file, by including only those with mobile country code 641, as listed in the 'mcc' column. This data was then separated into GSM (2G), UMTS (3G), and LTE (4G)-transmitting cell towers as listed in the 'radio' column.

This data is accessible as three .csv files with cell tower location and metadata, listed as 2G_cell_towers_uganda.csv, 3G_cell_towers_uganda.csv and 4G_cell_towers_uganda.csv in 'Raw Data'.

Only the cell tower location, range, and network type were used in this review.

4. Uganda national border

Downloaded from Stanford University Earthworks[18]. This data was sourced from the Uganda Bureau of Statistics in 2012 and references the national border claimed by Uganda as of 2006.

This data is accessible as .shp and .shx files as UgandaBorder.shp, UgandaBorder.shx in 'Raw Data'.

This data is not included in the final figure and only used for intermediate calculations.

All location data are saved in the following Coordinate Reference System (CRS): EPSG: 32636 Arc 1960 / UMT Zone 36N. Apart from cell tower location raw data which is saved in EPSG: 4326, WGS 84.

Processing Method:

Mapping was performed using an open source geographical information system application QGIS[19]. Version 2.18 was used to perform the mapping used in Figure 4

1. The population density geo-referenced image file was uploaded into QGIS mapping software and recoloured in the 'style' menu to show population density into four discrete intervals of population per pixel (0-1, 1-3, 3-5, >5), denoting the population in a 100 m x 100 m area.

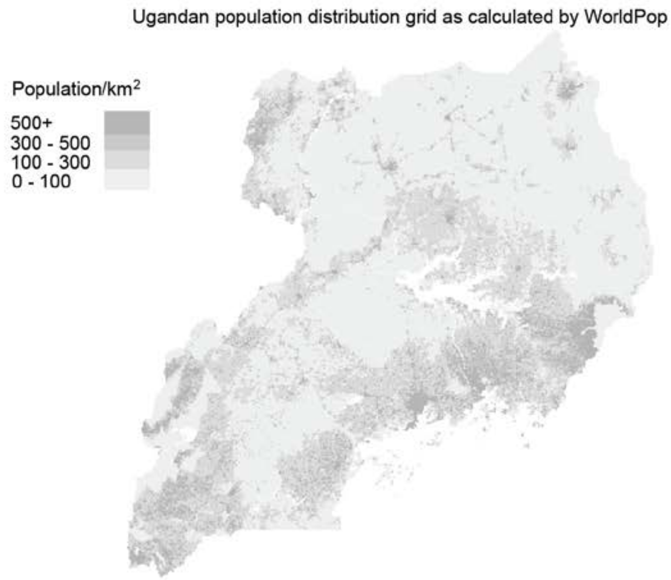

Figure 1: *Population geo-reference image from WorldPop. Recoloured to show population density in discrete intervals, where darker greys show higher population density.*

2. Cell tower points were separated and saved into 2G, 3G and 4G files, (denoted by GSM, UMTS and LTE respectively in the 'radio' column) before the three .csv files were loaded into QGIS mapping software.

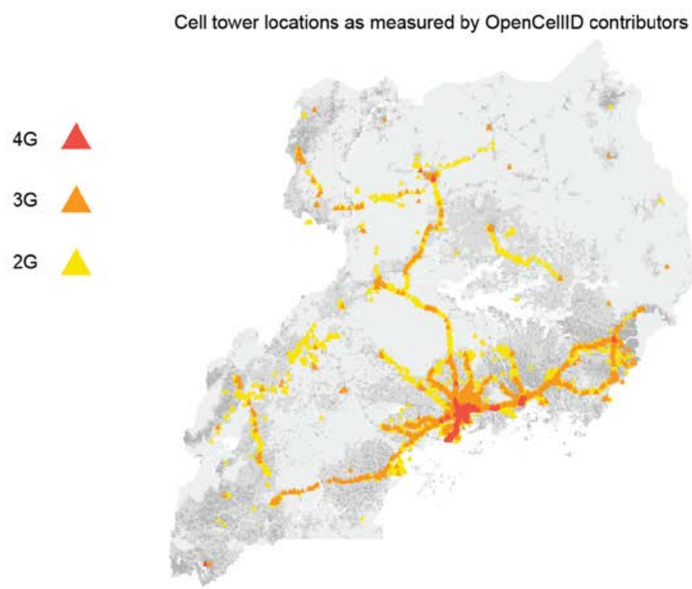

Figure 2: *Cell tower locations from OpenCelliD were uploaded into the QGIS mapping software, overlaying the population raster. Here they are shown as triangles and coloured by network generation, where 4G cell towers are shown in red, 3G in orange, and 2G in yellow.*

3. The ranges of the cell towers from the OpenCelliD dataset were then visualised:

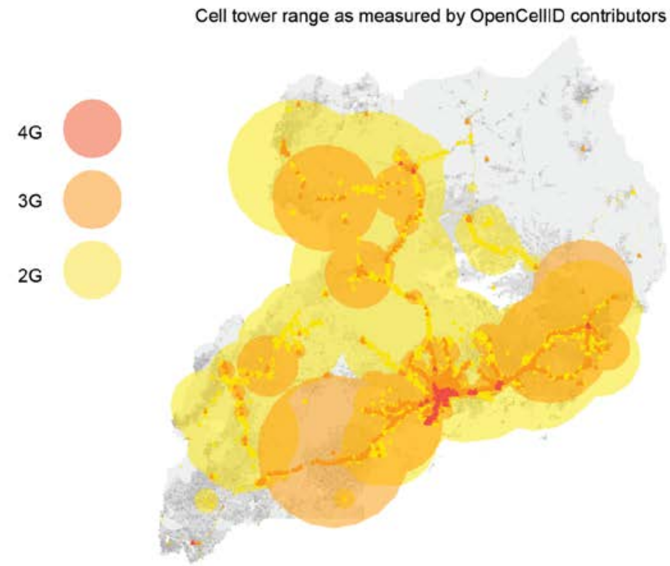

Figure 3: *Cell tower ranges visualised from the OpenCelliD dataset. The measured ranges are shown in red for 4G cell towers, orange for 3G, yellow for 2G. Measured ranges can be seen to vary in size. Ranges are assumed to be uniformly radial for the purpose of this estimation.*

For each three sets of cell tower points, the QGIS 'variable distance buffer' tool was used with cell tower location coordinates and the distance of measurement ('range' column in dataset, described as 'Estimate of cell range, in meters') to show measured range of each cell tower. Here a radius of this range is drawn around each respective cell tower, and the combined area that these radii cover is taken as the area in range of that generation of cell towers. Data saved as .shp and .shx files in 'network ranges (overlapping)'.
In order to visualise the ranges, the coordinate reference system was converted from WGS 84 (global) to Arc 1960/UMT Zone 36N (local) so that distances could be processed in metres (the cell tower range is given in metres).

4. Overlapping cell tower ranges were removed, and the population image beneath each network range area segmented:

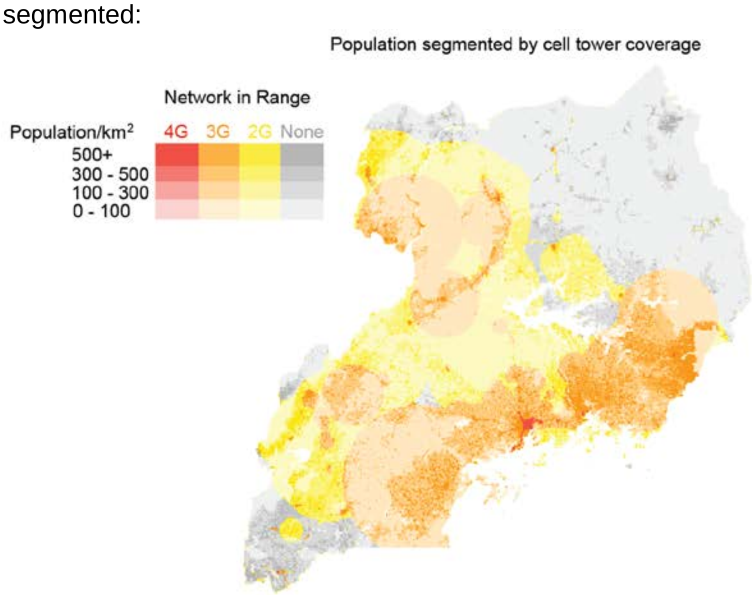

Figure 4: *Populations with connectivity. Clipped using national border and land outline (pop data).*

Some ranges overlapped (2G and 3G, 3G and 4G) so areas with only 2G and 3G-but-no-4G were created to be used in addition to the 4G range area and area with no network in range. Here it was assumed that when more than one network generation is available, the best network in range will be used, so in overlapping areas only the best network is shown to be in range. This also means that populations living under these areas will add up to the total population. These new areas were created using the previous set of areas with the QGIS vector clipping tool. Data saved as .shp and .shx files in 'network ranges (clipped)'.
The area of Uganda not in range of any network was calculated by using the border of Uganda, and subtracting the areas of 2,3,4 G ranges from it using the QGIS vector clipping tool.

Population rasters more than 5 km from health centres were displayed and coloured in the same way, with population density split into four discrete ranges of population per pixel (0-1, 1-3, 3-5, >5). These ranges had approximately equal amounts (calculated automatically using 'equal interval') of people when whole population raster was processed, and so these ranges were rounded for simplicity of reading the map key, and kept for all rasters for consistency.
The 5 km distance was chosen as it is a target set by the Ugandan government.[20]

The population density for the map key was calculated converting the pixel value (equal to population in 100 square metres) to people per square kilometre.

5. Health facility points from the Uganda Bureau of Statistics were uploaded into QGIS software. This data was then converted to local coordinates (Arc 1960/UMT Zone 36N), so that the range could be calculated in metres. A 5 km radius around all health centres was drawn using QGIS 'fixed distance buffer' tool. Data saved as as 'HC5' .shp and .shx files in 'Health Centre Ranges'.

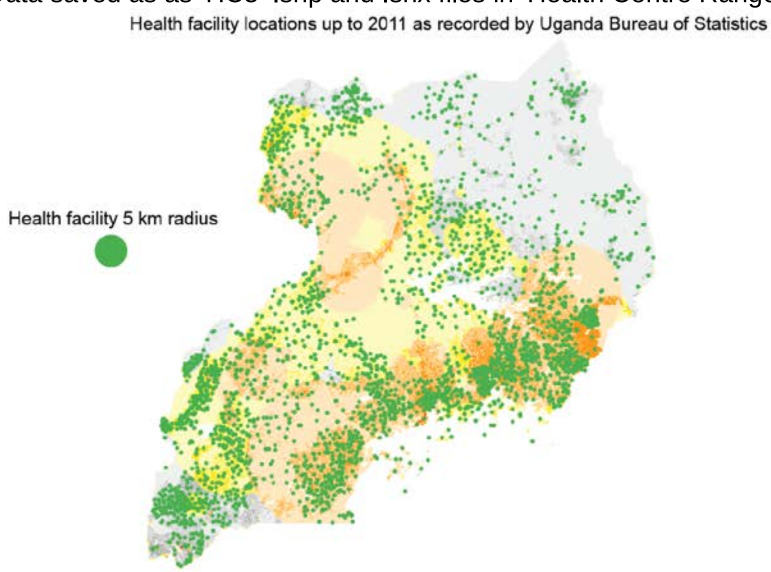

Figure 5: *Health centres and their 5km range, dividing the map beneath into the population within 5 km of health facilities, and the population more than 5 km from health facilities.*

The negative of this area was created using it and the border of Uganda (with QGIS vector difference tool), to show areas >5 km from a health centre. Data saved as '_HC5' .shp and .shx files in 'Health Centre Ranges'.

Health centres within each type of network range (2G, 3G, 4G, not network) were found using QGIS vector clipping tool, with the network calculated ranges and the health centre points. Each set of points were coloured according to the network in range and set to a width equal to 5 km.
Processing:

The QGIS vector clipping tool was used to find overlapping areas between the network ranges and areas more or less than 5 km from health centres, giving 8 areas in total (clipping the 2G area within the 5 km health centre range gave an invalid geometries error and so is not used. i.e. files 2Gonly_HC5_w were not generated).

These areas are saved as .shp and .shx files in ‘Network Ranges (clipped) clipped by Health Centre Ranges’. Here _w in the file name denotes the area within 5 km from health centres and _o the area out of this range.

The population density raster was cropped by 7 of the 8 areas using the QGIS raster clipper tool, creating 7 new rasters. Data saved as .tif files in ‘Clipped Population Rasters’. (Clipping by the 2Gonly_HC5_w area gave an error, so population within this area is calculated by the difference of the sum of all other areas and the total population). Here _w in the file name denotes the area within 5 km from health centres and _o the area out of this range.

The total population in each raster was calculated using the QGIS ‘Raster layer statistics tool’ (where the ‘sum’ output showed the population contained in that area), giving the number of people in range of networks and within range of health centres or not. In order to ensure the populations in and out of range of health facilities add up to the necessary total – population rasters clipped by only connectivity are created (the three used are also saved as .tif files in ‘Clipped Population Rasters’).

The total population for the original raster file was also calculated, giving the total population. The above populations were normalised against this total, and the proportions were rounded to two decimals places and used to create pie charts in Adobe Illustrator. (Population totals and proportions saved in excel file ‘UgandaChartStatistics2’ and in **Section C Table 3** of this document).

C. Data Tables

Table 1. *Smartphone adoption percentages, GSM Association*[1234].

Year	World	Sub Saharan Africa
2008	4	n/a
2009	6	n/a
2010	8	n/a
2011	13	6
2012	20	9
2013	28	13
2014	37	17
2015	44	23
2016	50	29
2017	55	35
2018	59	42
2019	62	49
2020	65	55

Table 2. *Smartphone specifications, GSMArena*[7].

Name	Month Release	Year Release	2/3/4G Capability	CPU Clock Frequency (Mean)	CPU Cores	Amdahls Law 50%	Primary Camera Megapixels	Cost (EUR)	Link to page
LOW COST									
Xiaomi Redmi 5	12	2017	4	1.800	8	3.200	12.00	160	Link to page
Xiaomi Redmi 4 (China)	11	2016	4	1.400	8	2.489	13.00	130	Link to page
Xiaomi Redmi 3	1	2016	4	1.350	8	2.400	13.00	160	Link to page
Xiaomi Redmi 2	1	2015	4	1.200	4	1.920	8.00	130	Link to page
Xiaomi Redmi	8	2013	3	1.500	4	2.400	8.00	100	Link to page
Vivo Y53	3	2017	4	1.400	4	2.240	8.00	120	Link to page
Vivo Y31	9	2015	3	1.300	4	2.080	8.00	130	Link to page
Vivo Y28	12	2014	3	1.300	4	2.080	5.00	100	Link to page
Vivo Y15	6	2013	3	1.300	4	2.080	5.00	110	Link to page
Huawei Y3	9	2017	4	1.300	4	2.080	8.00	90	Link to page
Huawei Y3II	6	2016	4	1.000	4	1.600	5.00	80	Link to page
Huawei Y360	6	2015	3	1.200	4	1.920	5.00	90	Link to page
Huawei Ascend Y330	3	2014	3	1.300	2	1.733	3.15	70	Link to page
Huawei Ascend Y320	12	2013	3	1.300	2	1.733	2.00	100	Link to page
Huawei Ascend Y300	3	2013	3	1.000	2	1.333	5.00	90	Link to page
Huawei Ascend Y	10	2012	3	0.800	1	0.800	3.15	60	Link to page
Micromax Canvas Infinity	9	2017	4	1.400	4	2.240	13.00	123	Link to page
Micromax Canvas Amaze 2 E475	7	2016	4	1.200	8	2.133	13.00	100	Link to page
Micromax Canvas Pulse 4G E451	12	2015	4	1.300	8	2.311	13.00	130	Link to page
Micromax A108 Canvas L	8	2014	3	1.300	4	2.08	8.00	140	Link to page
Micromax A114 Canvas 2.2	12	2013	3	1.300	4	2.08	8.00	150	Link to page
Micromax A110 Canvas 2	11	2012	3	1.000	2	1.333	8.00	160	Link to page
HIGH END									
Apple iPhone X	10	2017	4	2.390	6	4.097	12.00	1000	Link to page
Apple iPhone 7	9	2016	4	2.340	4	3.744	12.00		Link to page
Apple iPhone 6S	9	2015	4	1.840	2	2.453	12.00		Link to page
Apple iPhone 6	9	2014	4	1.400	2	1.867	8.00		Link to page
Apple iPhone 5S	9	2013	4	1.300	2	1.733	8.00		Link to page
Apple iPhone 5	9	2012	4	1.300	2	1.733	8.00		Link to page
Apple iPhone 4S	11	2011	3	1.000	2	1.333	8.00		Link to page
Apple iPhone 4	6	2010	3	1.000	1	1.000	5.00		Link to page
Apple iPhone	6	2009	3	0.600	1	0.600	3.15		Link to page

3GS									
Apple iPhone 3G	6	2008	3	0.412	1	0.412	2.00		Link to page
Apple iPhone	6	2007	2	0.412	1	0.412	2.00		Link to page
Samsung Galaxy S9	3	2018	4	2.050	8	3.644	12.00	840	Link to page
Samsung Galaxy S8	4	2017	4	2.125	8	3.778	12.00		Link to page
Samsung Galaxy S7	3	2016	4	1.950	8	3.467	12.00		Link to page
Samsung Galaxy S6	4	2015	4	1.800	8	3.200	16.00		Link to page
Samsung Galaxy S5	4	2014	4	2.500	4	4.000	16.00		Link to page
Samsung I9500 Galaxy S4	4	2013	3	1.400	8	2.489	13.00		Link to page
Samsung I9300 Galaxy S III	5	2012	3	1.400	4	2.240	8.00		Link to page
Samsung I9100 Galaxy S II	4	2011	3	1.200	2	1.600	8.00		Link to page
Samsung I9000 Galaxy S	6	2010	3	1.000	1	1.000	5.00		Link to page

Table 3. *Calculated Ugandan populations in range of cellular networks and health facilities, used to plot the pie chart in Figure 4.*

	Total	>5km from Health Facility	<5km from Health Facility	Calculated Total
4G	2604405	2823	2601582	2604405
3G no 4G	19479330	4052019	15432285	19484304
2G only	9598106	1889345	7708761	9598106
No connection	7458223	2622501	4835722	7458223
Calc. total	39140064*	8566688	30578350	39145038*

*Difference in measured and calculated values due to overlap of border pixels of areas used for segmentation of the population raster. Normalised calculated population totals were used to plot the pie chart in Figure 4.

D. References

1 GSMA. *The Mobile Economy*, https://www.gsma.com/mobileeconomy/archive/GSMA_ME_2015.pdf (2015).

2 GSMA. *The Mobile Economy*, https://www.gsmaintelligence.com/research/?file=9e927fd6896724e7b26f33f61db5b9d5&download (2017).

3 GSMA. *Mobile Economy Sub-Saharan Africa* https://www.gsma.com/mobileeconomy/archive/GSMA_ME_SubSaharanAfrica_2014 (2014).

4 GSMA. *Smartphone forecasts and assumptions 2007–2020*. https://www.gsmaintelligence.com/research/2014/9/smartphone-forecasts-and-assumptions-20072020/443/ (2014).

5 GSMA. *Global Mobile Trends*, https://www.gsma.com/globalmobiletrends/index.html (2017).

6 Gartner, Global Smartphone Sales Report Fourth Quarter 2017, https://www.gartner.com/en/newsroom/press-releases/2018-02-22-gartner-says-worldwide-sales-of-smartphones-recorded-first-ever-decline-during-the-fourth-quarter-of-2017 (2018)
7 GSMArena. https://www.gsmarena.com/ (2017).
8 Wikipedia. *Amdahl's law*, https://en.wikipedia.org/wiki/Amdahl%27s_law (2017).
9 WorldPop. *Uganda 100m Population*, http://www.worldpop.org.uk/data/summary/?doi=10.5258/SOTON/WP00283 (2013).
10 Linard, C., Gilbert, M., Snow, R.W., Noor, A.M. and Tatem, A.J. Africa population count data, PLoS ONE, 7(2): e31743, (2012).
11 WorldPop. *Methods*, http://www.worldpop.org.uk/data/methods/ (2017)
12 Uganda Bureau of Statistics. *Health Centres Uganda*, http://maps.data.ug/layers/geonode%3Ahealth_centres_ubos_and_others_merged (Accessed 08/17)
13 United Nations Office for the Coordination of Humanitarian Affairs. *Uganda - Health*, https://data.humdata.org/dataset/uganda-health (2011).
14 Uganda Bureau of Statistics. *Uganda Health Facilities*. http://ubos.geosolutions.it/layers/geonode:health_facilities (2014).
15 Uganga Ministry of Health. *National Health Facility Master List 2017*, http://library.health.go.ug/publications/health-facility-inventory/national-health-facility-master-list-2017 (2017).
16 OpenCelliD. https://www.opencellid.org (2017)
17 OpenCelliD. *Cells database:*, http://wiki.opencellid.org/wiki/Menu_map_view#database (2018).
18 Uganda Bureau of Statistics. *National Boundary: Uganda, 2006*, https://earthworks.stanford.edu/catalog/stanford-ks911mt6862 (2006).
19 QGIS. https://www.qgis.org/ (2017)
20 Uganda Bureau of Statistics. *Uganda National Household Survey Report Chapter 5.0: Health*, https://www.ubos.org/wp-content/uploads/publications/03_20182016_UNHS_FINAL_REPORT.pdf (2016).