Defining service catchment areas in low-resource settings

Peter M Macharia 1,2, Nicolas Ray 3,4, Emanuele Giorgi 1, Emelda A Okiro 2,5, Robert W Snow 2,5

INTRODUCTION

An accurate service catchment area—a geographical area delineated around a service point (such as a health facility or a school) describing the population that uses its services—is important for robust and reliable estimation of population denominator used in planning, estimating commodity needs and mapping applications. In the literature, approaches used to define a service catchment area range from simple to complex approaches and are often presented as population to service ratios, use of administrative boundaries, proximity-based metrics and models accounting for utilisation rates. However, the implementation and choice of an approach are often based on convenience, disregarding the implications on accuracy of the catchment population, complexities of service use and the likely implications for planning. In this commentary, we highlight the common approaches, key issues and best practices in defining a reliable and representative service catchment area. We hope to create a dialogue within the geospatial community and allied stakeholders to evolve best-practice standards in defining service catchment areas for efficient and equitable planning.

The commentary is structured into four sections. We discuss the need for recording and geolocating user addresses followed by a summary of approaches used to define service catchment areas and conclude with the key characteristics of an ideal service catchment, best practices and key issues that require addressing.

NEED FOR RECORDING AND GEOLOCATING SERVICE USER ADDRESSES

The definition of a reliable service catchment area for health, education, social care and other local public services depends on the availability of spatially positioned residential addresses of those who seek services. In high-income countries, residential addresses (such as zip code, house number and street name) are part of well-defined registers that can be used to define service catchment areas. Conversely, most of sub-Saharan Africa (SSA) do not have the luxury of well-defined addresses or routine logging and geocoding of villages/estates of those seeking services. In cases where such data are available in SSA, there is heterogeneity between and within countries. For example, within health and demographic surveillance sites, addresses of
An illustration of several approaches used to define service catchment area, administrative unit (ward or subcounty), straight-line distance (10 km buffer and Thiessen polygon) and travel time (area within 60 min of a hospital). Illustration is based on 622 COVID-19 vaccination posts (https://medical.unon.org/node/169) approved in Kenya. If vaccine allocation was based on one of these catchment areas, there would be mismatches between those who attend a facility and the estimated population denominator in 2021\textsuperscript{8} across methods due to lack of geolocated residential data of care seekers at a particular post. The top right corner shows the illustrated region within Samburu County, Kenya.

households are well maintained. The absence of reliably defined service catchment areas challenges robust disease mapping, equitable provision of vaccines (eg, childhood immunisations), other preventive health commodities (such as bed-nets), and acute care, school needs, social care and the accurate compilation of health data. This is because such applications rely on a catchment population denominator, which is in turn derived by a spatial overlay of population density maps and the geographical extent of a service catchment area.\textsuperscript{4-7}

Across SSA, population settlement and sites of service provision are often not mapped on a regular basis nor are always available at small administrative units for efficient planning. Disaggregating low-resolution census data to presumed high-resolution settlement patterns guided by satellite imagery and building footprints has increased in sophistication in recent years.\textsuperscript{8} The locations of schools, health facilities and other amenities have improved with crowd-sourced Global Positioning Systems initiatives, participatory mapping, improved national gazetteers, high-resolution satellite imagery and demand for such data sets.\textsuperscript{5} However, linking population to service delivery remains crude and ignores the complexities of service demand and/or supply mainly due to a lack of geocoded data on where the users of services reside in SSA.

MODEST APPROACHES USED TO DEFINE SERVICE CATCHMENT AREAS

Population-to-service ratios are the most common metrics, ranging from wide subnational units to national estimates for international ranking. These metrics merely require an estimate of population size from national censuses and the number of facilities within a census-defined polygon. Smaller, subnational polygons can be used, for example, subcounties, wards, municipalities or parishes (figure 1), where service provision is decided by local health and education authorities. However, people residing along administrative boundaries will likely access the most proximal service which might be in the neighbouring administrative unit. Similarly, people in need of a particular service far away from the border might be attracted to a particular service point in a neighbouring unit due to preferred, desirable characteristics. For example, women bypassed the nearest facility in Tanzania,\textsuperscript{9} Ghana,\textsuperscript{10} Kenya,\textsuperscript{11} Mozambique, India and Pakistan.\textsuperscript{12} As a result, it is common practice to have childhood immunisation coverage from routine data greater than 100% in administrative units that attract many people from neighbouring administrative units.\textsuperscript{13,14} This is likely due to a challenge of assigning population denominators to an administrative area instead of a well-defined service catchment area based on the actual users of a particular service.

Where both the population (whether actual or modelled) and amenity locations (geocoded with a longitude and latitude) are available at high resolution and accuracy,\textsuperscript{8,15} service catchment areas can be empirically modelled. The most common and rudimentary approach is the use of an arbitrary radius (buffer) around a service point, for example, a 10 km buffer. The radius of the buffer can be variable depending on the capacity of the provider, type of services offered and the mode of transport used to access the service (figure 1). Related to the buffer is Thiessen polygon, where each catchment area incorporates all points that are closer to a given service provider than any other provider (figure 1). These two approaches rely on straight-line distance, which does not reflect actual distance and/or time travelled, which are in turn impacted by topography, land use/cover, road network, transport modes and associated transport speeds. Furthermore, service points are also weighted equally despite differences in the capacity and size.

IMPROVED APPROACHES TO DEFINING SERVICE CATCHMENT AREAS

The complexity of physical access, choice and demand characteristics can be tackled with more complex models of service use. Gravity models assume that the flow of people from residential areas to service providers is proportional to the demand for services and inversely proportional to physical access.\textsuperscript{16} The demand for services is defined by morbidity, age, and social structure and the capacity of a facility.\textsuperscript{16} To be more realistic, this
physical access can be estimated by modelling the travel time to the nearest provider, by accounting for travel factors and barriers mainly through cost distance algorithms and network analysis. The travel time is then binned to define a service catchment area based on either an arbitrary time threshold (figure 1) or known cut-offs, where available (eg, 2 hours for obstetrical complications). A range of software including AccessMod, ArcMap (Esri, Redlands, CA, USA, 2021) and QGIS (QGIS.org, 2021. QGIS Association. QGIS Geographic Information System) can be used to estimate travel times and corresponding catchment areas.

The use of travel time is appealing, but it requires expert knowledge to implement. It also assumes travel time to the nearest facility, which is not always the case and often does not account for competition and user preferences. Furthermore, catchment boundaries might be highly permeable when arbitrarily defined thresholds are applied. The issue of arbitrarily defined time threshold can be minimised by combining travel time with utilisation rates typically from sample household survey data, to inform a reasonable threshold via the inflection point of a decay curve. The gravity models can be adapted with additional information and sophisticated models to account for competition between service points, variation in capacity, available equipment, quality and affordability of the service provided.

**IDEAL SERVICE CATCHMENT AREA AND BEST PRACTICES**

An ideal service catchment area should meet a minimum set of conditions from both the supply and demand side: (a) capture a significant proportion of its activities, (b) exclude areas whose activity contribution is due to random variation (eg, service users who are not usual residents such as visitors), (c) account for geography (physical access and its barriers), (d) differentiate across specialty types (eg, vaccination sites, tuberculosis detection, snake antivenom availability, or primary and tertiary educational levels) and capacities (affordability, size, staffing and equipment), (f) competition from the network of neighbouring providers and (g) seasonal changes due to weather or population flows.

The use of models to define a catchment area can, however, be bypassed if data needed to define an ideal catchment were available, enabling the use of simpler geospatial information systems techniques. Therefore, defining a representative catchment is premised on the availability of geocoded data at high spatial and/or temporal resolution, reflecting residential location of service users, service points, utilisation rates, service use and factors that affect travel and behavioural data on service choice. Future ambitions to define service catchment areas should be aligned to several key considerations listed in table 1.

In conclusion, service choice will always depend on a complex array of social factors, quality (perceived or actual), previous experience, affordability and physical accessibility encompassing both the supply and the demand side. Such factors dictate how often lower level health facilities are bypassed, or the choice of an educational institution or social care programmes. The availability of well-curated and regularly updated service utilisation data and locations of service users will be important in accounting for such unique characteristics.

<table>
<thead>
<tr>
<th>Category</th>
<th>Ambitions and best practices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data collection</td>
<td>Improving collection and geocoding of residential address (village/estate) data from service users by healthcare providers, educational institutions, local governments and national statistical agencies. This will enhance the definition of service catchment areas for effective planning. High-resolution population density maps and databases of road network, land use/cover, travel barriers, care-seeking behaviour, modes of transport and travel speeds also need a careful curation.</td>
</tr>
<tr>
<td>Data and software sharing</td>
<td>Important data sets to define service catchment areas should no longer be kept in silos. The culture of making open-access data analytical models and software should improve across researchers and organisations in SSA. With increasing model sophistication, there is a need for software that can easily be used to define realistic service areas especially for planners.</td>
</tr>
<tr>
<td>Community</td>
<td>Building a community of researchers, sharing best practices, identifying difference between services, different diseases, service interruptions (eg, COVID-19 or natural disasters), ecological contexts and demography will be useful.</td>
</tr>
<tr>
<td>Service use</td>
<td>With a growth in availability of geocoded data and spatial epidemiologists across SSA, there is a need for an increased investment in research aimed at mapping higher resolution data on service use. Studies should also consider different forms of service access such as vaccination, healthcare, education and social care.</td>
</tr>
<tr>
<td>Disease mapping</td>
<td>The use of spatial statistics to map diseases, health outcomes, and demographic and socioeconomic indicators has witnessed huge advancements. However, the use of data from routine services (such as disease registries) together with reliably defined catchment areas requires more attention and quantifying the role played by different approaches and their impact on the mapped quantities.</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>Where modelling must be conducted due to inadequate data, authors should test the sensitivity and uncertainty of several models that are used to define a service area. Comparisons will tease out if there any gains in using complex geospatial approaches in lieu of simpler approaches (more accessible to non-experts) to define service areas. There is also a need to recognise limitations such as bypassing the nearest provider due to personal preferences, quality and capacity when results are being interpreted.</td>
</tr>
</tbody>
</table>

SSA, sub-Saharan Africa.
and defining reliable service catchment areas going forward.

Twitter Peter M Macharia @Pete_M_M

Contributors PMM conceived and designed the commentary and was responsible for the production of the first draft. RWS contributed to the second draft. All authors provided scientific input and contributed to the writing of the commentary.

Funding PMM is supported under the Royal Society Newton International Fellowship (NIF/R1/201418), NR acknowledges support from the Swiss National Science Foundation (project number 315130_176271). PMM, EAO and RWS are grateful to the support of the Wellcome Trust to the Kenya Major Overseas Programme (number 203077). EAO is supported as a Wellcome Trust Intermediate Fellow (number 201866). RWS is supported by a Fellowship (NIF/R1/201418). NR acknowledges support from the Swiss National Science Foundation (project number 315130_176271). EG is supported by a Royal Society Newton International Fellowship (https://royalsociety.org/). Use of mobile data for population mapping using random forests with remotely-sensed and ancillary data. PLoS One 2015;10:e0107042.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made: See: https://creativecommons.org/licenses/by/4.0/.

ORCID iDs Peter M Macharia http://orcid.org/0000-0003-3410-1881
Nicolas Ray http://orcid.org/0000-0002-4696-5313
Emanuele Giorgi http://orcid.org/0000-0003-0640-181X
Emelda A Okiro http://orcid.org/0000-0001-9543-8360
Robert W Snow http://orcid.org/0000-0003-3725-6089

REFERENCES

22. Gilmour SJ. Identification of hospital catchment areas using clustering: an example from the NHS. Health Serv Res 2010;45:497–513.