CrossMark

# Comparing alternative methods to measuring pedestrian access to community pharmacies

**Miguel Padeiro[1]**

**Abstract** The aim of this paper is to compare several methods for measuring geographical accessibility to community pharmacies in the Lisbon Metropolitan Area (LMA). Twelve measures of pedestrian distance between spatial units and the closest community pharmacy were computed based on the combination of 4 parameters: type of distance, location, centroid definition, and level of spatial unit. For this, the Google Maps Application Programming Interface was used for calculating network pedestrian distances, using a list of 801 community pharmacies and population data from the Census 2011. Correlations between every method were performed, and the variations of the estimated number of inhabitants served at an 800-m distance were analyzed. Local errors were assessed comparing every combination to the most accurate one. The results show that the number of people served ranges from 70 to 89% of the total population, depending upon the method used. The use of pedestrian network distances decreases by more than 10 points the population coverage, compared to crow-fly distances. The finest parameters of population assignment are more inclusive than coarser ones. This research demonstrates the influence of several measurement methods on coverage estimations. Empirical evidence indicates that both measurement and policies should be called into question in order to improve actual coverage.

✉ Miguel Padeiro
   jmpadeiro@campus.ul.pt

[1] IGOT (Institute of Geography and Spatial Planning), University of Lisbon, Rua Branca Edmée Marques, 1600-274 Lisbon, Portugal

 Springer

# 1 Introduction

With the rise of equity issues, accessibility to healthcare services has been under growing scrutiny (Gulliford and Morgan 2013). Equal access to healthcare is a key objective of national policies and international strategies (NHS 2008; US DHHS 2011; Evans et al. 2013; EC 2016), leading to significant improvements since the end of the 1990s. Among many definitions including the affordability and availability dimensions (Penchansky and Thomas 1981; Gulliford et al. 2002), geographical accessibility is referred to as the ease with which an individual can reach a given destination (Jones 1981). It is thus a measure of potential access.

Geographical accessibility can be threatened by many factors related to local barriers and poor local environment, physical impairment, cultural and linguistic difficulties, scarce financial resources and the lack of social capital (Ensor and Cooper 2004). The lack of geographical accessibility may induce lower rates of real healthcare utilization, as a distance decay effect has been frequently observed (Hassell et al. 2000; Lin et al. 2002; Allard et al. 2003; LaVela et al. 2004; Arcury et al. 2005; Hiscock et al. 2008; Turnbull et al. 2008; Saramunee et al. 2014), resulting in increased morbidity (Aakvik and Holmås 2006; Blalock et al. 2013) and raising health equity concerns (Korda et al. 2007). Vulnerable social groups tend indeed to experience more mobility issues than others, defined either in medical terms, as the expression of difficulties in walking (Simonsick et al. 2008), or in geographical terms relating to transport disadvantage (Markovich and Lucas 2011). As a consequence, concerns about the most efficient and inclusive location of and access to healthcare have arisen, necessitating the development of accurate measurement methods (Guagliardo 2004; Bell et al. 2013).

In the last decade progress has been made based on the growing use of geographical information systems (GIS) that allow for better accessibility measures (Higgs 2004). A wide variety of methods have been developed, ranging from catchment areas approaches (Schuurman et al. 2006) to more complex measures based on travel behavior, gravity models and utility-based methods (Talen and Anselin 1998; Baradaran and Ramjerdi 2001; Geurs and Ritsema Van Eck 2001). As a result, and despite the many improvements made, debate continues about the best methods for estimating population access to healthcare. Beyond the choice of the method for measuring accessibility, several parameters must be defined and may significantly influence the accessibility measurements (Fortney et al. 2000; Hewko et al. 2002; Apparicio et al. 2008):

- Location georeferencing: exact address location is not always available due to confidentiality issues, incomplete data, or even time consumption. Alternatively, zipcodes have been increasingly used as a proxy (Gatrell et al. 1991), introducing some well-accepted uncertainty, as zipcodes do not exactly match addresses. However, the resulting bias remains generally unclear (Martin and Higgs 1997). This may not be relevant when analyzing driving distances, but it can create inaccuracies in pedestrian-based accessibility measures (Karimi and Kasemsuppakorn 2013).
- Distance measurement: many studies still use crow-fly distance for determining healthcare services catchment areas or for estimating rates of population served (Delamater et al. 2012). Public authorities generally use it as a decision tool (Schuurman et al. 2006) despite some doubts about its reliability. Network-based real distance or time may be more accurate, mainly in regions and cities where roads and street patterns exhibit high degrees of sinuosity and of community severance due to infrastructure-based or topographical barriers.

- Aggregation method, including two different parameters:

    - Level of spatial unit, such as counties, municipalities, census tracts and census blocks;
    - Population assignment: reducing a given spatial unit to its geometric centroid is one of the most frequent and practical approaches. A variant is to get a population-weighted centroid when the available lower-level spatial units allow weighting. Other methods may be used, such as dasymetric techniques and pro rata population distribution (Langford and Higgs 2006), the latter relying on the assumption that population is homogeneously distributed there. Assignment methods are a major source of bias in final results (Hewko et al. 2002; Langford and Higgs 2006).

The main aim of this study is to investigate the differences between several ways to measure geographical accessibility, based on the four parameters aforementioned, and to determine which one is best suited for pedestrian access measurement. The paper focuses on community pharmacies in the Lisbon Metropolitan Area (LMA) and provides an estimation of the total population served within an 800-m radius.

The next section will present a brief overview of the study area, data collection and study procedure. Then, the results of the several methods computed are analyzed. Finally, the main findings will be discussed, mentioning their strengths and limitations, and highlighting the scientific and policy implications.

## 2 Study design

### 2.1 Study area

The study focuses on accessibility to community pharmacies in the Lisbon Metropolitan Area (LMA). Two reasons led to the choice of this specific type of health care service. First, community pharmacies are one of the most important proximity services (Anderson 2000; Hassell et al. 2000; McMillan et al. 2013), with more than 90% of the entire population utilizing one at least once a year (Anderson 2000). Pharmacies are particularly sought after by unhealthy and/or elderly people who frequently experience more mobility issues than other social groups (Lin 2004). Second, as a proximity service, pedestrian access is particularly relevant. When pedestrian distances are concerned, the precision of geographical accessibility measures is even more relevant, as a simple 50-m difference can induce an error as high as 10% for a 10-min walk, or more when slow-walking elderly are considered. The level of precision is therefore sensitive and worth analyzing. It might affect the decision to include or exclude a spatial unit from the group of reachable ones.

The LMA extends over 3015 km², on the shores of the Atlantic Ocean and on both banks of the River Tagus. It has 2.8 million inhabitants, including the Portuguese capital with its 505,000 inhabitants. It is divided into 18 municipalities, with all of them subdivided into 211 parishes,[1] 4521 census tracts (Portuguese 'statistical sections'), and 34,937 census blocks ('statistical subsections').

---

[1] A territorial reform conducted in 2011–2013 reduced their number to 183, with most of the merged parishes being located in Lisbon. Here, the former parishes are used, due to (i) data issues and (ii) the better level of precision they allow.

## 2.2 Data

Three data sources were used. The 2011 population census provides the most recent data on residents at the census blocks and tracts levels. Community pharmacies were exhaustively collected in November 2016 using the National Health System database. The list was updated using a web directory and making several phone calls for confirmation, resulting in the collection of 801 pharmacies within the LMA, 260 of which were in Lisbon city. Zipcodes were gathered using the Geopostcodes dataset. Finally, pedestrian distances between spatial units and the closest community pharmacy (with a maximum crow-fly distance of 1 km) were calculated using the Google Maps Directions API service, resulting in a total of 33,965 origin–destination pedestrian trips measured. The use of the Google Maps API service is rather recent. It provides several advantages, among which is the permanent access to dynamically updated network data (Zhang et al. 2010; Wang and Xu 2011). Using the Google Maps API service only requires the preparation of a dataset containing the coordinates of all points, as well as a Windows macro for calling the API. This method thus provides a useful way to easily measure local accessibility to services that can then be used for specific population groups or other transportation modes.

## 2.3 Procedure

Three main issues were analyzed. The choice of parameters leads to a set of combinations resulting in 12 different estimates (Fig. 1):

- Precision of pharmacy location: exact address versus zipcode. In the dataset, the Portuguese 7-number zipcode refers to the linear centroid of the street(s) included in the zipcode.
- Type of distance: Euclidean versus shortest network distance. Other types of distance such as Manhattan distance and the shortest network time are not addressed here. The former is actually rarely used for accessibility analyses, and the latter is not required for pedestrian trips if distance is addressed.
- Aggregation method:

    - Level of spatial units: census tracts versus census blocks. The municipality is in many countries the privileged spatial unit for demographic and geographic criteria. In Portugal, a community pharmacy can open only when the resulting ratio at the municipal scale is at least 3500 inhabitants per pharmacy. However, census tracts and blocks may provide better measures, since municipalities may comprise well-served communities and others that are not served at all. Throughout this paper, 'census tracts' will refer to the Portuguese 'statistical sections', and 'census blocks' will refer to 'subsections'.
    - Population assignment: geometric versus population-weighted centroids. At the census block level, centroids are necessarily geometric, as there are no lower-level units allowing for population-weighted methods. At the census tract level, geometric and population-weighted centroids are compared.

After obtaining distances via Google Maps API, the combinations of parameters are compared to the most accurate of them (network-based real pedestrian distance of census blocks, BNAG) in terms of: (i) origin–destination distances; (ii) global errors in total
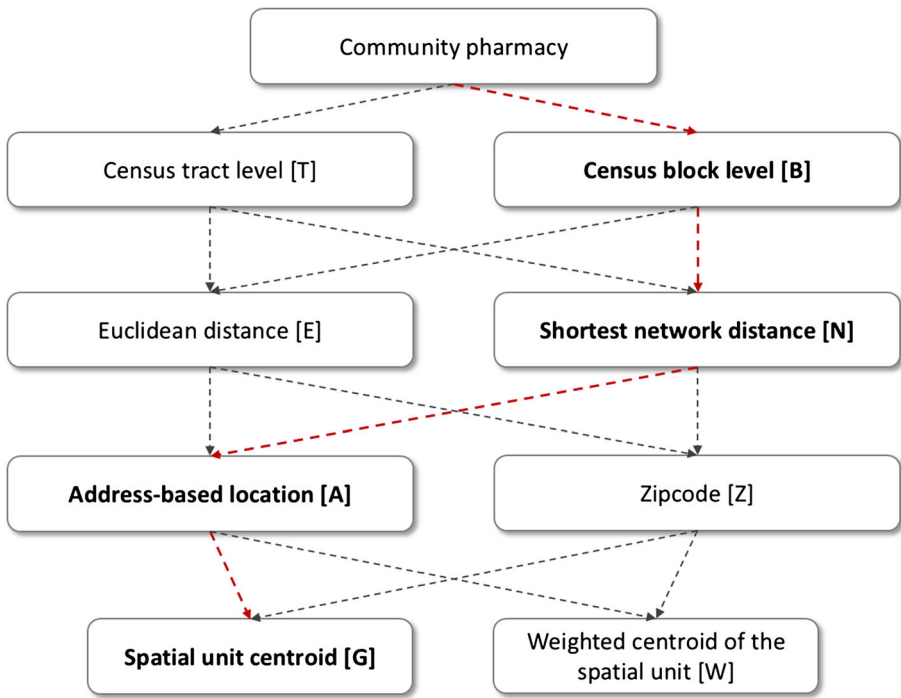
Letters within brackets define the code for each set of parameters (TEAG: census tract level, measured with Euclidean distance, at the pharmacy's exact address and using the geometric centroid). Bold: most accurate combination. [W] is not possible for census block level

**Fig. 1** Combinations of parameters defining the 12 methods tested

population served at different distances from 500 to 1000 m, although only the 800-m distance is presented; (iii) local errors (catchment areas).

# 3 Results

## 3.1 Differences in distance and location

### 3.1.1 Euclidean versus network distance

Differences between Euclidean and network-based distances can be assessed using a sinuosity index $S_i$ (shortest network distance/Euclidean distance). $S_i$ ranges from 1, indicating no sinuosity, to $+\infty$, indicating a high degree of sinuosity. The maximum sinuosity varies across the methods from 16 to 47. However, for all methods the median is around 1.46, with the first and third quartiles close to 1.2 and 2.0 respectively. No effect of distance from the center has been detected, as the sinuosity index is rather homogeneously distributed across space, contrary to the hypothesis that Euclidean distances are less precise in suburban than in central areas (Apparicio et al. 2008). This might be due to the fact that

only a small proportion of census tracts (< 1 crow-fly km from the closest pharmacy) was previously selected for further analysis. Including all the LMA census tracts might therefore show increased indexes in the peripheral areas.

### 3.1.2 Zipcodes versus exact address

Calculating the locational difference between each pharmacy's zipcode and exact address provides a first insight on the reliability of using the former as a proxy to the latter. The distance between them ranges from 0 to 739 m (mean 62 m, median 42 m, standard deviation 62 m). A total of 22 pharmacies out of the 801 collected are located more than 200 m from their zipcode point, 109 over 100 m, and 412 (more than half) over 50 m. Such a spatial offsetting between the real and zipcode-based location might be relevant for pedestrian trips, especially for people walking slowly, as do the elderly and people with disabilities. Differences could possibly be greater in peripheral areas where zipcodes are expected to define larger areas; however, the spatial distribution of each class of differences is rather homogeneous across the LMA (Fig. 2a). This finding is in accordance with the findings of Fortney et al. (2000), who did not find any significant difference between urban and rural counties.

### 3.1.3 Aggregation methods

The spatial offsetting between geometric and population-weighted centroids is even greater, as the distances range from 0 to 6.67 km, with means of 170.2 m, standard
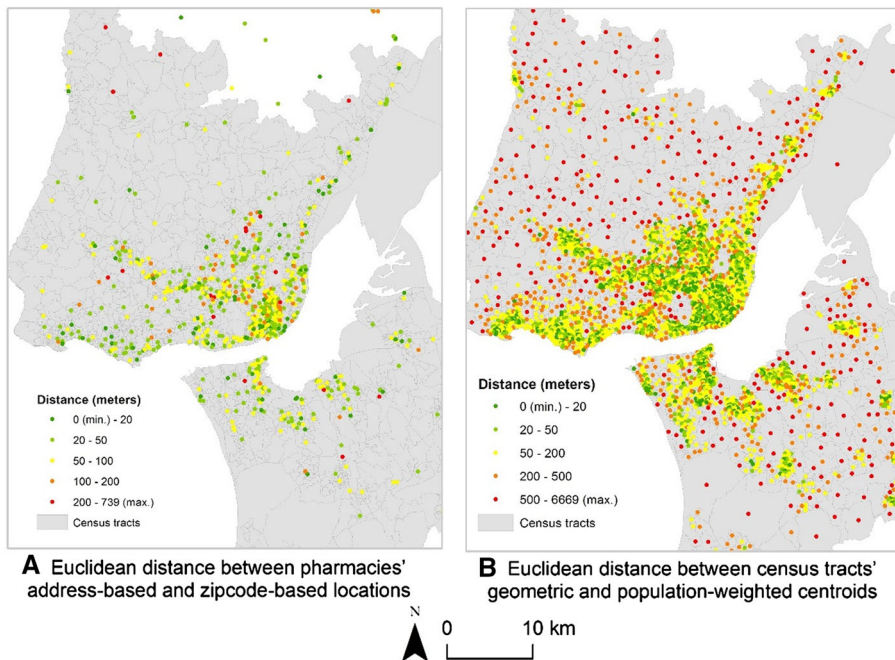


**Fig. 2** Differences across location parameters: real versus zipcode locations of pharmacies, geometric versus population-weighted centroids

deviation equal to 347.2 m, and median equal to 70.1 m. Contrary to the pharmacy locations, there is a clear discrepancy between densely-populated areas and peripheral ones (Fig. 2b). This was highly expectable, as the size of both census tracts and blocks increases with distance from the center.

## 3.2 Global errors: how do differences influence estimates of the population served?

Until now, the analysis has focused on distances themselves. The resulting level of accessibility might be influenced by differences across methods; however, it may also not produce such a bias that would justify the use of time-consuming methods instead of quicker and less exhaustive ones. In order to address this question, the total population living within 800 m from the nearest community pharmacy was calculated using every possible combination, and compared to the most accurate one (BNAG). Having its centroid included in an 800-m distance or not therefore determines the inclusion or exclusion of a spatial unit from the reachable group. Other distances were also tested, from 500 to 1000 m. The choice of presenting the results for a distance of 800 m is because this is an intermediary one, being relatively close to the 15 min walk that many elderly people consider as a reasonable distance (Padeiro 2017).

Regardless of the global result, the geography of (in)accessible residents may significantly change with the parameters set. To address this question, two series of association measures were performed (Table 1).

First, Pearson's coefficients were estimated between all pairs of methods based on the distance from census tracts and blocks to the closest pharmacy (below diagonal in Table 1). Among many more possibilities (Hubálek 1982; Zhang and Srihari 2003), Pearson's correlation is a popular, simple and reliable tool for evaluating presumably linear association between continuous variables. The correlations were estimated taking account of population of each spatial unit. Correlations between tracts-based and blocks-based distances were also made possible by attributing to each census block the same distance as the census tract the block belongs to. Sensitivity to outliers was assessed by using Cook's distance (CD). These allowed to detect between 5 and 9% of potential outliers (always more than 300 spatial units with $CD > 4/n$). However, a visual inspection of scatter plots showed that the outliers did not change the results nor the assumptions. The regression line is the same in both cases, with an approximately 0.1-point increase in Pearson's coefficients when dropping them. As a consequence, there was no reason to remove them.

Second, three similarity measures were computed: Jaccard-Needham, Rogers-Tanimoto, and Sokal-Michener were tested (Zhang and Srihari 2003), and provided rather similar results. Table 1 presents the latter one above diagonal. Similarity measures were performed based on binary variables, where value '1' means "the spatial unit is located less than 800 m from the nearest pharmacy" and value '0' means "not".

In general terms, the tested methods correlate with each other less than expected, with a median coefficient of 0.679. Only 13 pairs of methods (23%) show a correlation coefficient above 0.8. High correlations are mostly related with pairs of close methods where only one parameter differs. This is the case, for instance, of TEAW/TEAG, TEZG/TEAG, TEZW/TEAW, TEZW/TEZG, and TNZG/TNAG. The majority of pairs with at least two divergent parameters (for example between TNZG and TNAW) show low to moderate correlation, below 0.7.

Within census tracts (T), methods based on euclidean distances show the highest correlation coefficients between each other. Those based on network distances only show high

**Table 1** Measures of association (Peasrson's correlations and Sokal–Michener similarity measure) between alternative types of distance. *Source:* INE 2011 Population Census, CAOP 2013, National Health System database. Google Maps, author's calucations

| Combinations of parameters | TEAG | TEAW | TEZG | TEZW | TNAG | TNAW | TNZG | TNZW | BEAG | BEZG | BNAG | BNZG |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Exact address and geom. centroid (TEAG) | 1 | 0.968 | 0.991 | 0.965 | 0.853 | 0.886 | 0.851 | 0.879 | 0.873 | 0.872 | 0.778 | 0.776 |
| Exact address and weighted centroid (TEAW) | 0.837 | 1 | 0.966 | 0.995 | 0.833 | 0.888 | 0.834 | 0.883 | 0.935 | 0.934 | 0.789 | 0.786 |
| Zipcode and geom. centroid (TEZG) | 0.974 | 0.820 | 1 | 0.967 | 0.853 | 0.888 | 0.853 | 0.882 | 0.871 | 0.874 | 0.778 | 0.778 |
| Zipcode and wieghted centroid (TNAG) | 0.813 | 0.972 | 0.837 | 1 | 0.834 | 0.887 | 0.835 | 0.883 | 0.911 | 0.913 | 0.772 | 0.771 |
| Exact address and geom. centroid (TNAW) | 0.732 | 0.590 | 0.713 | 0.573 | 1 | 0.895 | 0.954 | 0.880 | 0.717 | 0.717 | 0.805 | 0.797 |
| Exact address and weighted centroid (TNZG) | 0.639 | 0.804 | 0.624 | 0.777 | 0.665 | 1 | 0.884 | 0.961 | 0.811 | 0.812 | 0.859 | 0.846 |
| Zipcode and geom. centroid (TNZW) | 0.684 | 0.560 | 0.708 | 0.579 | 0.910 | 0.598 | 1 | 0.896 | 0.719 | 0.722 | 0.799 | 0.810 |
| Zipcode and weighted centroid (TNZW) | 0.596 | 0.749 | 0.621 | 0.784 | 0.596 | 0.884 | 0.661 | 1 | 0.787 | 0.789 | 0.826 | 0.842 |
| Exact address and geom. centroid (BEAG) | 0.763 | 0.834 | 0.745 | 0.818 | 0.546 | 0.679 | 0.518 | 0.643 | 1 | 0.983 | 0.816 | 0.809 |
| Zipcode and geom. centroid (BEZG) | 0.739 | 0.812 | 0.761 | 0.835 | 0.530 | 0.659 | 0.532 | 0.658 | 0.972 | 1 | 0.813 | 0.811 |
| Exact address and geom. centroid (BNAG) | 0.583 | 0.644 | 0.565 | 0.625 | 0.570 | 0.710 | 0.509 | 0.625 | 0.779 | 0.753 | 1 | 0.951 |
| Zipcode and geom. centroid (BNZG) | 0.551 | 0.619 | 0.575 | 0.643 | 0.513 | 0.643 | 0.555 | 0.702 | 0.731 | 0.758 | 0.882 | 1 |

Below diagonal Pearson's correlation between distances calculation methods (all significant at 1% level). Above, Sokal–Michener similarity measure between binary values of inclusion (= 1) and exclusion (= 0) of spatial units. All observations included, population-weighted, distance for similarity measures 800 m

correlation when the divergent parameter is the centroid location (population-weighted vs geometric centroid) (TNAW vs TNZW; TNAG vs TNAW). The clearest divide line separates euclidean and network-based methods, with lower correlation coefficients between methods belonging to different groups. Finally, similarity measures show the same commonalities and dividing lines between the tested methods, and a fairly high level of similarity between them, with a median of 0.853.

Regarding the most accurate method (BNAG), the most striking result is that there is only one correlation coefficient above 0.8 (with BNZG—blocks-network-zipcode). The BNAG method, which is the most accurate one, is thus also the one with the lowest coefficients, suggesting that there may be important divergences between the other methods and the reference (BNAG).

Table 2 shows the differences in population coverage between every method tested and the most accurate one (BNAG: census blocks, network-based, precise address location). Three distances are included. At a 800-m distance from the pharmacy location point (real address, coded A, or zipcode-based location, coded Z), the coverage ranges from 70.7 to 88.7% of the population, with Euclidean distance being more inclusive than the real pedestrian one. This clear dividing line was rather expected, particularly in a European country where street layouts are generally more irregular than in other continents,

**Table 2** Under- and overestimation of population coverage resulting from different combinations of parameters, against the most accurate one (BNAG), at different distances. *Sources*: INE 2011 Population Census, CAOP 2013, National Health System database, Google Maps API, author's calculations

| Combinations of parameters | | | % Pop. cov. (800 m) | Differences against BNAG in pop coverage (%) at distance | | |
|---|---|---|---|---|---|---|
| | | | | 500 m | 800 m | 1000 m |
| Tracts | Euclidean | Exact address, and geom. centroid (TEAG) | 84.6 | 28.5 | 10.8 | 8.7 |
| | | Exact address, and weighted centroid (TEAW) | 88.7 | 38.2 | 16.2 | 12.8 |
| | | Zipcode, and geom. centroid (TEZG) | 84.4 | 27.9 | 10.6 | 8.6 |
| | | Zipcode, and weighted centroid (TEZW) | 86.7 | 35.4 | 13.5 | 10.0 |
| | Network | Exact address, and geom. centroid (TNAG) | 70.7 | − 9.1 | − 7.4 | − 4.7 |
| | | Exact address, and weighted centroid (TNAW) | 77.4 | 3.7 | 1.4 | 3.6 |
| | | Zipcode, and geom. centroid (TNZG) | 70.7 | − 9.3 | − 7.4 | − 4.8 |
| | | Zipcode, and weighted centroid (TNZW) | 75.0 | − 0.7 | − 1.8 | 0.9 |
| Blocks | Euclidean | Exact address, and geom. centroid (BEAG) | 88.1 | 34.5 | 15.4 | 9.5 |
| | | Zipcode, and geom. centroid (BEZG) | 88.0 | 34.7 | 13.1 | 9.5 |
| | Network | Zipcode, and geom. centroid (BNZG) | 76.1 | − 1.6 | − 2.3 | − 0.1 |

BNAG is used as a reference (pop. coverage: 2,154,247 residents, or 76.3% of total population at 800 m; 1,575,638 at 500 m—55.8%; 2,297,198—81.4%—at 1000 m). Total LMA population: 2,821,876 inhabitants

especially in North America, thus leading to greater differences in distance between crow-fly and real distances.

Regarding the difference between census tracts and blocks, two opposite effects were possibly competing: on the one hand, as census blocks are more precise, their use could significantly reduce the area and population covered. On the other hand, less precise census tracts could contribute to exclude too large areas, and consequently the use of census blocks could correct the bias. This latter effect has proven to predominate, although it is likely that both effects coexist across the LMA. Finally, between tracts-based combinations population-weighted centroids are much more inclusive than geometric ones, with an inclusion increase ranging from 2.3 to 6.7 points. The use of exact address locations is marginally more inclusive than zipcode-based locations, with the difference ranging from 0 to 2.4 points.

The same table shows how different methods may under- or overestimate population coverage, compared with the most accurate method and setting as references three different distances (500, 800, and 1000 m). Based on the BNAG method, 2.154 million individuals live at a 800-m distance from the closest pharmacy. At this distance, the use of other methods leads to an overestimation ranging from 30,000 to 350,000 residents, and to an underestimation nearly between 40,000 and 160,000 residents. Maximum overestimation occurs at the 500-m threshold (600,000 residents according to TEAW method), maximum underestimation at the 800-m one (160,000 people based on TNZG method). The use of Euclidean distance is clearly the main parameter leading to the highest deviations from the BNAG method. With network-based distances the resulting population coverage systematically becomes more similar to BNAG-based estimations. Using population-weighted centroids instead of geometric ones improves accuracy, with less deviations from BNAG-based measures. Finally, the difference between zipcodes and precise address locations is more variable, and not always favourable to the address-based measure.

## 3.3 Local errors and variability

Based on the correlation analysis and on counting the population living less than 800 m (and 500, and 1000 m) from a community pharmacy, global errors have shown how some methods are more acceptable than others. Locally, the question refers to which spatial units are included in one method and excluded from another one, and to the level of disagreement of every method compared to the reference one (BNAG). It may therefore be useful to assess the spatial distribution of these differences between different methods and the most accurate one, since the spatial units and population covered through a given method might not be the same as with another one. Figure 3 exhibits two examples of this non-matching effect. In both of them, the reference for comparisons is the BNAG method (blocks-network-address).

Map A (Fig. 3) shows how the most contrasting method (TEAW—tracts-euclidean-address) differs from BNAG, the most accurate one: at a 500-m distance, TEAW overestimates population served in more than 600,000 people; at 800 m, the overestimation is still 350,000 residents. Many pockets are highly visible in the Lisbon city center: in these areas, the accurate BNAG method excludes a much larger portion of the Lisbon territory, while the TEAW method includes it. This observation can be extented to the entire LMA, as purple areas (those located near a pharmacy according to the TEAW method, but not according to the BNAG method) are relatively numerous. Methods based on euclidean distances and census tracts are frequently utilized in planning, due to data availability and to the ease with which distances are estimated. Such a difference between both methods
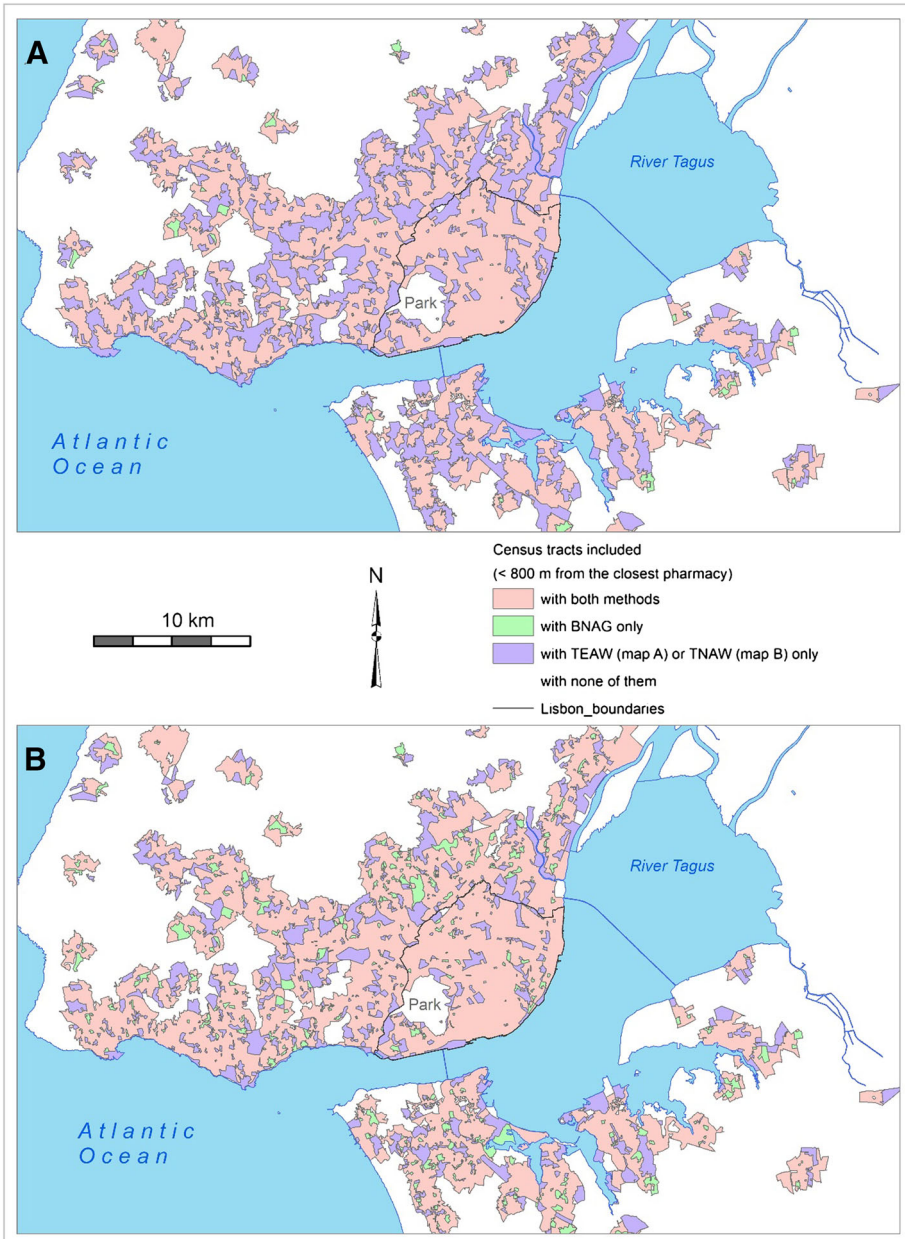
**Fig. 3** Spatial distribution of census tracts within 800 m from the nearest community pharmacy, according to 3 methods (BNAG vs TEAW, BNAG vs TNAW)

means that the use of a less precise one may lead to serious misconceptions about how many residents live near a pharmacy. The second map (Fig. 3b) shows one of the most similar combinations compared to BNAG: TNAW (tracts-network-address-weighted centroid). At a 800-m distance, the difference between both methods is reduced (only 1.4%

in population coverage—see Table 2), with an overestimation of 30,000 residents. Despite this, their spatial distribution is fairly divergent, with the persistence of numerous pockets (purple areas in the map) located near a pharmacy only according to the TNAW method. Again, this may raise concerns about the impacts of less precise measurement methodologies on local accessibility to healthcare.

## 4 Discussion

The evidence from this study reinforces the need for more awareness and better accuracy in accessibility measurement. Modulating four parameters for testing twelve different combinations induced several behaviors. One of them (BNAG) is clearly the most accurate method, as it is based on the most precise spatial units level, on real address locations and on the network. It was thus set as the reference for comparisons. Given the number of individuals concerned, some methods are clearly to be rejected if the objective is to provide a good estimation of the population coverage. Methods based on Euclidean distances have proven much more error-prone, with very high numbers of people mistakenly included within a given distance from a pharmacy. As a consequence, of great importance is the gain provided by the use of the real pedestrian distance instead of the classical Euclidean distance. Euclidean-based measurements should be avoided whenever possible.

This result contrasts with some studies that concluded that more than 90% of the variance of real distance was explained by Euclidean distance, leading to a relatively low measurement error (Phibbs and Luft 1995; Fortney et al. 2000; Apparicio et al. 2008). Such analyses might end up qualifying the interest of time-consuming road networks based distance, and legitimised the use of Euclidean distances as a fairly good proxy, while encouraging caution. This might be true in certain cases, for example when distances are used as an independent variable in regression models and when, as a result, their internal variability is more relevant than their accuracy against real distances. However, when the objective consists of estimating the number of people inside and outside reachable areas for healthcare location policies and monitoring purposes, variance and explaining power do not matter anymore. This is particularly true in the case of community pharmacies, which need a high pedestrian accessibility level by elderly people. That is why, even in the case of zipcode versus address-based location which gave very similar results, the most precise methods should be adopted, since a substantially different spatial distribution of population served is noticeable. However, these methods may be adopted when available data do not enable the use of a BNAG combination: using census tracts instead of blocks, and zipcodes instead of exact address locations, would not jeopardize the estimations at a metropolitan or regional level.

Fortunately, not only is accessibility measurement becoming easier than ever, thanks to the development of many softwares and modules and to the increased availability of network datasets, but even easily used online resources are now available, consuming less time and allowing calculations that in many countries would be inaccessible due to the lack of data. The use of a Google Maps API for calculating pedestrian distances has proven to be useful here and may be viewed as a response to difficulties arising from unavailable or incomplete data and from time-consuming and skill-demanding operations. Not all territories and transportation modes are covered at this time, but much progress has been made, and using the Google Maps Directions or Matrix API may increasingly provide powerful ways to overcome such difficulties.

Consequently, the now widely used geographical information systems (GIS) and the increasingly easy way that distances and durations can be calculated for accessibility measurements should lead to more accurate policies. This study reinforces the argument that Euclidean-based catchment areas should be replaced by network-based distances, especially when individuals with mobility issues are taken into account. Similarly, the best level for action and criteria in healthcare location policies should be called into question. This is particularly relevant at a time when local communities and resources are becoming a major topic in urban sustainability and social inclusion purposes. In ageing societies, the time and distance to community pharmacies and other healthcare services are growing in relevance, requiring a better match between location policies and demand for services. Regarding pharmacies, many countries have established a municipal-based number of pharmacies per capita, in a struggle to avoid uneven spatial distributions. Other countries require that the majority of the population should live at a 1-km distance from the nearest pharmacy. In the former case, a simple number of providers per capita at the municipal level does not ensure that all of the residents are actually at an acceptable distance. Regarding the 1-km buffer, crow-fly distances are frequently used, potentially leading to some misunderstanding and underestimating the number of people actually unserved.

This study obviously has several limitations. There are many more methods and parameters available for calculating accessibility to services than those tested here (Fortney et al. 2000; El-Geneidy and Levinson 2006; Cheng et al. 2012; Karou and Hull 2012). This was just an attempt to explore some of the most commonly used parameters requiring a choice that can potentially affect the results. Regarding the parameters analyzed here, some types of distance were not considered although they could be approached when the car-based distance is more relevant, for example in remote areas where community pharmacies are sparsely distributed (Law et al. 2011; Norris et al. 2014). Another limitation lies on the potential bias arising from discrepancies in the years of data (2011 for population data, 2016 for pharmacy data). More recent data are not available at the census tracts and blocks levels in Portugal (estimations are provided annually at the municipal level by Statistics Portugal). However it is unlikely that such a bias produce significant changes in the results obtained. Despite these shortcomings, this study sheds light on several of the most frequently used parameters and might provide orientations for researchers and planners who have to make difficult choices that may significantly affect their workload, as they calculate accessibility to healthcare services or determine catchment areas. The results are also transposable to other fields necessitating accessibility measurements.

## 5 Conclusion

This paper set out to compare several methods of measuring pedestrian accessibility to community pharmacies. For this purpose, 11 possible combinations of 4 parameters (location precision, type of distance, spatial units level and centroid definition) were explored and compared to a reference one in order to estimate the distance between spatial units and the closest community pharmacy, to calculate the resulting number of people served at several distances (500, 800, and 1000 m), and to measure the level of disagreement against the most accurate method.

The empirical results indicate that the number of population served varies considerably according to the chosen combination. The proportion of the LMA population covered ranges from 70.7 to 88.7%. The bottom line is that using Euclidean distances instead of

real pedestrian distances is the most influential choice, as there is a clear dividing line between the methods based on these two distances. The use of the finest spatial unit appears to be more inclusive, as well as the use of population-weighted centroids instead of simple geometric ones. Using zipcodes instead of the real address-based location has not proven to be a limiting factor, as they gave fairly similar results. All comparisons of methods have demonstrated that there are a significant number of spatial units that can be included in or excluded from the 800-m buffer, depending on the method. As a result, the use of the most accurate parameters (census blocks, network, real address location) is recommended, although zipcodes and census tracts may be used as a good alternative to addresses and census blocks when data are not available. In contrast, Euclidean distance should be avoided whenever possible. The development and availability of network data are a major advance in this respect.

Of course, the relationship between accessibility and outcome was not analyzed here. Wealthy residents and communities are less likely to experience a lack of accessibility than low-income groups and minorities, and in some cases there might be no decay effect. Such an approach requires further research, as several authors have called for recently (Higgs 2004; Arcury et al. 2005; Hiscock et al. 2008; Higgs 2009).

**Compliance with ethical standards**

**Ethical approval** This article does not contain any studies with human participants or animals performed by any of the authors.

# References

Aakvik, A., Holmås, T.H.: Access to primary health care and health outcomes: the relationships between GP characteristics and mortality rates. J. Health Econ **25**, 1139–1153 (2006)

Allard, S.W., Tolman, R.M., Rosen, D.: Proximity to service providers and service utilization among welfare recipients: the interaction of place and race. J. Pol. Anal. Manag. **22**, 599–613 (2003)

Anderson, C.: Health promotion in community pharmacy: the UK situation. Patient Educ. Couns. **39**, 285–291 (2000)

Apparicio, P., Abdelmajid, M., Riva, M., Shearmur, R.: Comparing alternative approaches to measuring the geographical accessibility of urban health services: distance types and aggregation-error issues. Int. J. Health Geogr. **7**, 1–14 (2008)

Arcury, T.A., Gesler, W.M., Preisser, J.S., Sherman, J., Spencer, J., Perin, J.: The effects of geography and spatial behavior on health care utilization among the residents of a rural region. Health Serv. Res. **40**, 135–156 (2005)

Baradaran, S., Ramjerdi, F.: Performance of accessibility measures in Europe. J. Transp. Stat. **4**, 31–48 (2001)

Bell, S., Wilson, K., Bissonnette, L., Shah, T.: Access to primary health care: does neighborhood of residence matter? Ann. Assoc. Am. Geogr. **103**, 85–105 (2013)

Blalock, S.J., Roberts, A.W., Lauffenburger, J.C., Thompson, T., O'Connor, S.K.: The effect of community pharmacy-based interventions on patient health outcomes a systematic review. Med. Care Res. Rev **70**, 235–266 (2013)

Cheng, J., Bertolini, L., Le Clercq, F., Kapoen, L.: Understanding urban networks. Cities **31**, 16–25 (2012)

Delamater, P.L., Messina, J.P., Shortridge, A.M., Grady, S.C.: Measuring geographic access to health care: raster and network-based methods. Int. J. Health Geogr. **11**, 15 (2012)

EC 2016: Expert Panel on effective ways of investing in Health (EXPH). Report on Access to Health Services in the European Union Brussels, European Commission

El-Geneidy, A., Levinson, D.: Access to destinations: development of accessibility measures. University of Minnesota, Minneapolis (2006)

Ensor, T., Cooper, S.: Overcoming barriers to health service access: influencing the demand side. Health Pol. Plan. **19**, 69–79 (2004)

Evans, D.B., Hsu, J., Boerma, T.: Universal health coverage and universal access. Bull. World Health Org. **91**, 546 (2013)

Fortney, J., Rost, K., Warren, J.: Comparing alternative methods of measuring geographic access to health services. Health Serv. Outcomes Res. Method. **1**, 173–184 (2000)

Gatrell, A.C., Dunn, C.E., Boyle, P.J.: The relative utility of the central postcode directory and pinpoint address code in applications of geographical information systems. Environ. Plan. A **23**, 1447–1458 (1991)

Geurs, K.T., Ritsema van Eck, J.R.: Accessibility Measures: Review and Applications. Bilthoven, RIVM (2001)

Guagliardo, M.F.: Spatial accessibility of primary care: concepts, methods and challenges. Int. J. Health Geogr. **3**, 3 (2004)

Gulliford, M., Figueroa-Munoz, J., Morgan, M., Hughes, D., Gibson, B., Beech, R., Hudson, M.: What does 'access to health care' mean? J. Health Serv. Res. Pol. **7**, 186–188 (2002)

Gulliford, M., Morgan, M.: Access to Health Care. Routledge, Abingdon (2013)

Hassell, K., Rogers, A., Noyce, P.: Community pharmacy as a primary health and self-care resource: a framework for understanding pharmacy utilization. Health Soc. Care Commun. **8**, 40–49 (2000)

Hewko, J., Smoyer-Tomic, K.E., Hodgson, M.J.: Measuring neighbourhood spatial accessibility to urban amenities: does aggregation error matter? Environ. Plan. A **34**, 1185–1206 (2002)

Higgs, G.: A literature review of the use of GIS-based measures of access to health care services. Health Serv. Outcomes Res. Method **5**, 119–139 (2004)

Higgs, G.: The role of GIS for health utilization studies: literature review. Health Serv. Outcomes Res. Method **9**, 84–99 (2009)

Hiscock, R., Pearce, J., Blakely, T., Witten, K.: Is Neighborhood Access to Health Care Provision Associated with Individual-Level Utilization and Satisfaction? Health Serv. Res. **43**, 2183–2200 (2008)

Hubálek, Z.: Coefficients of association and similarity, based on binary (presence-absence) data: an evaluation. Biol. Rev. **57**, 669–689 (1982)

Jones, S.: Accessibility measures: a literature review, TRRL Report 967. Crowthorne, Berkshire (1981)

Karimi, H.A., Kasemsuppakorn, P.: Pedestrian network map generation approaches and recommendation. Int. J. Geogr. Inf. Sci. **27**, 947–962 (2013)

Karou, S. & Hull, A.: Accessibility measures and instruments. In: Hull, A., Silva, C. & Bertolini, L. (eds.) Accessibility Instruments for Planning Practice. COST Office (2012)

Korda, R.J., Butler, J.R., Clements, M.S., Kunitz, S.J.: Differential impacts of health care in Australia: trend analysis of socioeconomic inequalities in avoidable mortality. Int. J. Epidemiol. **36**, 157–165 (2007)

Langford, M., Higgs, G.: Measuring potential access to primary healthcare services: the influence of alternative spatial representations of population. Prof. Geogr. **58**, 294–306 (2006)

Lavela, S.L., Smith, B., Weaver, F.M., Miskevics, S.A.: Geographical proximity and health care utilization in veterans with SCI&D in the USA. Soc. Sci. Med. **59**, 2387–2399 (2004)

Law, M.R., Dijkstra, A., Douillard, J.A., Morgan, S.G.: Geographic accessibility of community pharmacies in ontario. Healthc. Pol. **6**, 36–46 (2011)

Lin, G., Allan, D.E., Penning, M.J.: Examining distance effects on hospitalizations using GIS: a study of three health regions in British Columbia, Canada. Environ. Plan. A **34**, 2037–2053 (2002)

Lin, S.-J.: Access to community pharmacies by the elderly in Illinois: a geographic information systems analysis. J. Med. Syst. **28**, 301–309 (2004)

Markovich, J. & Lucas, K.: The social and distributional impacts of transport: a literature review, Oxford, University of Oxford, Working Paper no. 1055 (2011)

Martin, D., Higgs, G.: Population georeferencing in England and Wales: basic spatial units reconsidered. Environ. Plan. A **29**, 333–347 (1997)

McMillan, S.S., Wheeler, A.J., Sav, A., King, M.A., Whitty, J.A., Kendall, E., Kelly, F.: Community pharmacy in Australia: a health hub destination of the future. Res. Soc. Admin. Pharm. **9**, 863–875 (2013)

Nhs, S.: Better Health, Better Care: Action Plan. The Scottish Government, Edinburgh (2008)

Norris, P., Horsburgh, S., Sides, G., Ram, S., Fraser, J.: Geographical access to community pharmacies in New Zealand. Health Place **29**, 140–145 (2014)

Padeiro, M.: Geographical accessibility to community pharmacies by the elderly in metropolitan Lisbon. Res. Social Adm. Pharm. (2017). https://doi.org/10.1016/j.sapharm.2017.07.014

Penchansky, R., Thomas, J.W.: The concept of access. Med. Care **19**, 127–140 (1981)

Phibbs, C.S., Luft, H.S.: correlation of travel time on roads versus straight line distance. Med. Care Res. Rev. **52**, 532–542 (1995)

Saramunee, K., Krska, J., Mackridge, A., Richards, J., Suttajit, S., Phillips-Howard, P.: How to enhance public health service utilization in community pharmacy? General public and health providers' perspectives. Res. Soc. Admin. Pharm. **10**, 272–284 (2014)

Schuurman, N., Fiedler, R.S., Grzybowski, S.C., Grund, D.: Defining rational hospital catchments for non-urban areas based on travel-time. Int. J. Health Geogr. **5**, 43 (2006)

Simonsick, E.M., Newman, A.B., Visser, M., Goodpaster, B., Kritchevsky, S.B., Rubin, S., Nevitt, M.C., Harris, T.B., FOR THE HEALTH, A. & STUDY, B. C.: Mobility limitation in self-described well-functioning older adults: importance of endurance walk testing. J. Gerontol. Ser. A Biol. Sci. Med. Sci. **63**, 841–847 (2008)

Talen, E., Anselin, L.: Assessing spatial equity: an evaluation of measures of accessibility to public playgrounds. Environ. Plan. A **30**, 595–613 (1998)

Turnbull, J., Martin, D., Lattimer, V., Pope, C., Culliford, D.: Does distance matter? Geographical variation in GP out-of-hours service use: an observational study. Br. J. Gen. Pract. **58**, 471–477 (2008)

Us, D.H.H.S.: Healthy People 2020. US Department of Health and Human Services, Washington, DC (2011)

Wang, F., Xu, Y.: Estimating O-D travel time matrix by Google Maps API: implementation, advantages, and implications. Ann. GIS **17**, 199–209 (2011)

Zhang, B. & Srihari, S.N.: Properties of binary vector dissimilarity measures. In: Proceedings of the JCIS International Conference on Computer Vision, Pattern Recognition, and Image Processing (2003)

Zhang, H., Li, M., Chen, Z., Bao, Z., Huang, Q. & Cai, D.: Land use information release system based on Google Maps API and XML. In: 2010 18th International Conference on Geoinformatics, IEEE, 1–4 (2010)