Siting of HIV/AIDS diagnostic equipment in South Africa: a case study in locational analysis

Honora K. Smith[1*], Jon P. Smith[2], Deborah K. Glencross[3, 4], Naseem Cassim[3,4], Lindi M. Coetzee[3,4], Sergio Carmona[3, 4], Wendy Stevens[3, 4]

[1] *University of Southampton, Southampton, UK;* [2] *Selective Analytics Ltd, UK;* [3] *National Health Laboratory Service and* [4] *the Department of Molecular Medicine and Haematology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa*

*Correspondence: Honora Smith, University of Southampton, Highfield, Southampton, SO17 1BJ, UK
E-mail: honora.smith@soton.ac.uk
Telephone: +44 (0)23 8059 3700

Abstract

This paper describes a practical application of locational analysis to the siting of HIV/AIDS diagnostic equipment in laboratories across South Africa. Classical location analytical techniques were extended to ensure that laboratories are sited as close as possible to major centres of demand from hospitals and clinics. A particular advantage of the modified set covering algorithm developed is that choices between laboratory sites are made in a transparent manner. In order to find appropriate numbers and ideal placement of CD4 laboratories, runs were undertaken for various scenarios based on maximum travel time from health facilities to laboratory sites. Results demonstrated to decision makers showed close comparisons with

pilot review projects undertaken in four health districts of South Africa. The research has potential to impact health care delivery to HIV sufferers in the poorest rural regions of the country.

_____

1. Introduction

The National Health Laboratory Service (NHLS) of South Africa has been considering ways to streamline service delivery at their current testing facilities, especially for HIV/AIDS and TB (Hamers et al., 2012; Nahid et al., 2012; Glencross et al., 2014; Cassim et al., 2016). This paper describes the heuristic methodology used to suggest locations for laboratories that match local service demands. The study focuses on tests carried out to monitor CD4 blood levels (see Glossary for this and other medical terms), tests frequently administered for those who are HIV-positive to indicate disease progression and eligibility for antiretroviral treatment. Even with a large number of CD4 testing facilities (60 at the time of this case study), there are still districts in South Africa that do not have a CD4 testing facility within reasonable distance from the referring health facilities. While some areas are over-provided for, others require that long distances be travelled to the nearest laboratory, increasing costs and decreasing sample integrity (Glencross et al., 2014; Cassim et al., 2016).

For the CD4 test, samples given by patients at hospitals and other health facilities have to be transported to a laboratory site with the requisite testing platform. The objective of this study was to find an "appropriate" number of CD4 test laboratories, with "efficient" locations. In this study, we explore what

"appropriate" and "efficient" mean in comparison to current numbers and locations for laboratories using different scenarios for travel times. A further criterion was that laboratories should be sited close to health facilities with high demand for CD4 tests. Additionally, to maintain service in hard-to-reach areas, locations for placement of point-of-care devices were also to be recommended.

In this extension of classical location analytical techniques, a methodology is offered to decision makers for re-siting and possible reduction of numbers of laboratories testing CD4. The process is transparent in that where there is a choice of laboratory sites for retention or removal of CD4 testing, it is those closest to health facilities with greatest demand for those tests that are recommended for retention. Taking a modified set-covering approach, the relational-algebraic capacitated location (RACL) algorithm was developed to enable clearly justified choices to be made when allocating demand to CD4 laboratories.

Results for several different scenarios were presented to decision makers considering the possible relocation of NHLS services. The scenarios represented a range of maximum travel times and service needs. Further comparisons (supplementary to this methodology) were made in terms of running costs and quality of service.

To describe this case-oriented study, this paper proceeds as follows. Firstly, the background to the problem is given in Section 2, with a description of the South African laboratory service for HIV/AIDS diagnostics. An outline follows in Section 3 of literature regarding locational analysis. The methodology used to approach the problem is presented in detail in Section 4. Results obtained are presented in Section 5; limitations of the study are discussed and conclusions are drawn in the final sections.

2. Background to the problem: the NHLS laboratory service for HIV/AIDS testing

In 2010, the South African National Department of Health (NDOH) began a renewed and significant HIV counselling and testing (HCT) campaign to extend the take-up of antiretroviral therapy (ART) (NDOH, 2010). Those infected with HIV/AIDS were estimated to number 5.3 million or 11% of the population (Rehle et al., 2010). The NHLS made it a priority to support NDOH's HCT campaign with improved accessibility throughout the country to HIV/AIDS diagnostic and monitoring services (NHLS, 2013a; Cassim et al., 2016). A review of services was therefore initiated with the aim of possible rationalisation of the NHLS laboratories, in view of anticipated increased demand on services.

=================== insert Figure 1 here =========================

The NHLS operates laboratories across nine provinces in South Africa, predominantly serving the public health sector (Cassim et al., 2016). The laboratories operate all levels of pathology service, including routine diagnostic chemistry, haematology and microbiology services, with testing for TB and for HIV/AIDS, such as CD4 and HIV Viral Load. To simplify this analysis of service, scenario planning was carried out for the locations of CD4 testing. Figure 1 shows the locations of the 60 NHLS laboratories nationwide carrying out CD4 testing in 2011, with straight lines indicating links from health facilities to laboratories.

The different regions of South Africa present contrasts in terms of population and accessibility to diagnostic CD4 testing facilities. In sparsely populated rural areas, transportation of CD4 blood samples from health facilities to the nearest laboratory can be lengthy and over difficult terrain.

Sample quality can deteriorate during transit, particularly due to shaking of blood samples when transported over rough terrain and in high temperatures (Rehle et al., 2005). Contrastingly, some densely populated eastern and central regions are over-serviced with respect to CD4 testing. Figure 1 illustrates travel times of up to eleven hours needed to transport samples from some rural health facilities to nearest laboratories. Shorter travel times of less than two hours prevail towards the east of the country.

The focus of this research was finding "ideal" locations for CD4 testing laboratories based on local service demand and proximity to referring health facilities, for various scenarios. Large volumes of CD4 tests are performed on behalf of HIV/AIDS patients: these tests are used to determine whether a patient who is HIV-positive is at an appropriate stage to commence ART, and to monitor patients on ART (NDOH, 2013). The NHLS carried out over 3.8 million CD4 tests in the fiscal year 2011/2012.

This research considered 3,266 state-run hospitals, community health centres and clinics in South Africa which send blood tests for analysis by the NHLS. Such health facilities range from the largest (academic) hospitals to one-roomed clinics in rural areas. The volume of demand for tests at each health facility was not generally available, but estimates could be made on the basis of the HIV/AIDS prevalence of the surrounding population.

It is now possible to install point-of-care (PoC) equipment in small rural health facilities so that CD4 testing can be carried out on-site: this option is currently being investigated by NHLS and NDOH (Glencross et al., 2012, 2014). A particular objective of this study was therefore to identify those rural regions of the country where it would be advantageous to install PoCs. PoC devices do not need to be operated within a laboratory and therefore placing them directly in health facilities can avoid the transportation difficulties that can reduce sample quality. Moreover, use of such devices can speed up the return of

test results to patients and can help maintain contact with people who have travelled long distances to receive service.

This analysis highlights PoC placement in health facilities that were further from any potential laboratory location than a given distance chosen by the decision maker, and with suitably small demand requirements. In comparison with tests carried out at laboratories, the running costs per PoC test are high, and capital installation costs have to be considered including costs of a data link (Cassim et al., 2014). Furthermore, CD4 test results obtained from PoCs may be of lower accuracy than laboratory instruments if care is not taken in blood collection (Glencross et al., 2014). Recommendations for locations for PoC devices were therefore made only at sites from which transportation of samples to laboratories could not easily be made, and at which demand was not more than the capacity of PoC devices.

To summarise, the aims for optimised CD4 testing were as follows:

- Equity of cover: all health facilities should be either covered by a laboratory within a certain time *'T'* (chosen by the decision maker) or by a PoC device on-site (if demand is sufficiently low and higher costs could be offset by clinical advantage);
- Efficiency: the efficiency of the CD4 test locations should be improved by using less sites to achieve coverage whilst also shortening travel times needed to transport samples;
- Capacity: laboratory sizes (from a minimum, up to a maximum capacity) should be determined as appropriate for the daily volume of tests.

Runs of the RACL algorithm were undertaken for various scenarios meaningful to decision makers, based on maximum travel times from health facilities to laboratory testing sites, estimated from Euclidean distances.

3. Approach to problem and literature

Several classical locational analytical approaches are applicable to a situation of finding optimal service locations from amongst a finite set of candidate locations (i.e. a discrete problem) (Smith et al., 2009). We consider here the applicability of three such classical problems: the p-median model (Hakimi, 1964, 1965; ReVelle and Swain, 1970), the location set covering model (Toregas et al., 1971) and the maximal covering location problem of Church and ReVelle (1974).

Much research has been based on the classical p-median problem (Hakimi, 1964, 1965), which minimises the total demand-weighted distance travelled to the nearest of a number (p) of service locations. Numerous variations on the stated problem and a variety of solution methods have been developed for this NP-hard problem: however, few real-world applications of the problem have been published, in comparison with theoretical contributions (Marianov and Serra, 2011). The p-median problem is suited to situations where the use of a service is a necessity for an entire population or region. Pizzolato (1994) applies this methodology to evaluating locations of public schools in Rio de Janeiro. Higgins and Laredo (2006) use a capacitated p-median model to present scenarios for cost cutting within a sugar value chain in a region of north-east Queensland, Australia. Smith et al. (2013) present hierarchical p-median-type models with bicriteria efficiency/equity objectives for health facility location. Carling et al. (2012), in applying the p-median model to locations of emergency hospitals in a rural region of Sweden, find, as expected, that solutions are sensitive to distance measures, either Euclidean or travel time, and to the uneven distribution of population in the region.

Covering models in classical locational analysis are based on the concept of demand that is 'covered' within a certain distance or time by a particular service point. Based on Hakimi's (1965) introduction of set covering, Toregas et al. (1971) formulate the location set covering problem (LSCP), which finds the minimum number of locations needed to provide cover for all demand. Church

and ReVelle (1974) follow with the maximal covering location problem (MCLP) which maximises the demand covered by a given number of service points. MCLP modelling admits the possibility that some demand can be outside cover distance from any facility, given the number of service points for which funds are available. In contrast, the LSCP pursues total cover with no constraint on numbers needed. A capacitated version of the MCLP (Chung et al., 1983), allows a maximum capacity to be defined at each candidate facility location.

Applications of the LSCP and MCLP are possible in a wide variety of arenas (Chung, 1986; Almiñana and Pastor, 1994). Commonly, these covering models have been applied to location of emergency service vehicles: a review of such modelling and extensions is provided by Li et al. (2011). Shariat-Mohaymany et al. (2012) extend the LSCP for ambulance location in the rural outskirts of Tehran. Maximal covering is applied to ambulance location by Erkut et al. (2009) and also to police patrol areas (Curtin et al., 2010). Branas and ReVelle (2001) develop a heuristic for maximal cover location of air ambulance bases and trauma centres. With a contrasting application, Hudak (1994) incorporates the LSCP in modelling designed to find minimum numbers of wells able to capture plumes of contaminated water. Ye et al. (2011) apply the LSCP to finding locations for recycling centres in Taiwan, and Bell et al. (2011) apply it to aircraft alert sites for protection of strategically important areas. It should be noted that both the latter two applications used LSCP to find the minimal number of locations with a second stage of applying the p-median model to minimise aggregated distances.

Burkey et al (2012) apply both *p*-median and maximum cover models to state-wide provision of healthcare in the United States, addressing general efficiency, service availability and equity of access. Using similar performance measures, Shuklaa et al (2015) evaluate different scenarios for location of radiotherapy centres in New South Wales, Australia.

The LSCP principle has been chosen in this study for the algorithmic approach to reduction of laboratory numbers since all demand in the country must be covered. A modification is made to retain laboratories closest to clustered health facilities, thus reducing travel times as a secondary objective. Additionally, a flexible approach is required, to utilise existing CD4 laboratory sites preferentially, with use of other laboratory sites secondarily.

4. Methodology

This section provides a description of data collection and preparation, the problem formulation, the heuristic methodology and its implementation, with a discussion of the system design.

4.1 Data collection and preparation

Either travel distances or times were applicable to the methodology used, but, in this case study, travel time was considered critical by decision makers. However, the full matrix of drive times could not be obtained for this case study, e.g. via the Google® Maps Directions API Web Service, because of licence costs. Euclidean distances between health facilities and laboratories were obtained from SQL server, with calculations taking into account the curvature of the earth's surface. The Web Service was used to obtain a number of estimations of drive time between a health facility and a laboratory, and these were compared with Euclidean distances. A straight line was fitted to this data (see Figure 2) and this was used to convert between Euclidean distance and travel time. However in rural areas with a limited road network this could underestimate travel time by up to 40%.

==================insert Figure 2 here ====================

The numbers of tests requested by health facilities for CD4 tests were available in only limited numbers of cases, due to time and cost constraints for this project. Estimations were therefore made of demand for tests from individual facilities. Estimates for disease prevalence were provided for each district, and also the population served at each facility. Knowing when the tests are likely to be needed in the progression of the disease, the likelihood of a particular test could thus be estimated, and hence demand from facilities.

Comparisons of estimated demand with actual figures, where available, showed the estimation approach to overestimate demand in regions of high demand and underestimate in areas of low demand.

4.2 Problem formulation

The core goal is to cover all health facilities reachable within a given travel time T with a minimal number of laboratories; a secondary objective is to reduce travel times from facilities to laboratories. Alongside the travel time constraint, demand at any one laboratory is limited to D_{max}. Any uncovered facilities are then deemed to need PoC instruments to provide local testing.

Input parameters are travel times from clustered health facilities to laboratory sites, and estimated demand from facilities. Decision variables are the locations for CD4 testing and the allocations of demand points to laboratory sites.

The candidate locations available for CD4 testing include the full set of existing CD4 test sites, augmented by the set of all other laboratory sites. It is desirable to use the existing CD4 test sites in preference to other sites, in order to minimise costs, and disruption to staff and services.

4.3 Heuristic methodology

We describe here the different elements of the heuristic methodology used to suggest sites for different levels of CD4 diagnostic facilities. Firstly, geographical clustering of health facilities was employed to find areas of high demand for CD4 tests. Drive times from clusters to nearest equipment sites were then estimated, as were levels of demand from health facilities, if these were unavailable. Next, the RACL algorithm was applied to find lowest numbers of CD4 laboratories needed for a given maximal cover distance, while reducing distances travelled.

4.3.1 Clustering of facilities

With demand occurring at more than 3,200 health facilities, it was found that heuristic performance was improved by clustering of facilities, i.e. the combining of facilities that were within, say, distance *'C'* of each other. Not only was there some improvement in speed of the heuristic, but, more importantly, there was better resulting location of laboratories. It is better to locate a laboratory, for example, within a cluster of five close-by health facilities, each producing 10 samples per day, than to locate near to a more remote health facility producing 30 samples per day.

The recursive QuadTree clustering approach (Finkel and Bentley, 1974) was implemented, as specified in Appendix A, making use of some of the spatial commands available in the SQL server database.

The QuadTree method provided a list of clusters and all the health facilities in each cluster. The resulting number of clusters was 1,800. The centre of each cluster was defined to be the health facility closest to average of the latitude and longitude of all the facilities in the cluster. So that clusters around these

facilities could be visualised on a map, geographic shapes were produced which were the union of circles with radius '*C*' drawn round each separate facility, as exemplified in Figure 3.

================== insert Figure 3 here =====================

4.3.2 The Relational-Algebraic Capacitated Location (RACL) algorithm

The Relational-Algebraic Capacitated Location (RACL) algorithm employs a modified set covering approach to find the required set of laboratories for this problem. Relational algebraic operations are used in the SQL database queries made (see Section 4.4).

RACL was applied to the 1,800 clustered demand points (see Section 4.3.1) with approximately 230 candidate (possible) service locations (NHLS laboratories). In order to give preference to the 60 existing CD4 laboratories in this case study, first runs for each scenario were made just with existing CD4 locations as candidate sites: second runs were then made with existing sites removed from the set of candidate laboratory sites along with the demand covered by them.

Details of the RACL algorithm follow, where *I* is the set of health facilities where blood samples are collected, i.e. demand nodes, and *J* is the set of candidate laboratory location sites. The total demand at each laboratory is computed from demand at facilities within a given travel time *T*. Where total demand exceeds level D_{max}, this is reduced by de-allocating the most distant facilities. Laboratories are then 'removed' where all facilities covered are also covered by other laboratories: where there is a choice, the laboratory which is furthest from its maximum covered demand is removed. Facilities are re-allocated to nearest remaining laboratories and appropriate laboratory sizes can

thus be estimated. Finally, those facilities out of reach of any laboratory within travel time T are identified as being suitable for PoC machines.

The Relational-Algebraic Capacitated Location (RACL) algorithm
01. For each laboratory, $j \in J,$ produce a set, F_j, of all facilities, $i \in I$, with estimated time $t_{ij} \; <T$ to that laboratory.
02. At each laboratory, $j \in J$, the total possible demand is computed: $\sum_{i \in F_j} d_i$, where $d_i =$ demand at facility i [tests per day], $i \in I$.
03. Demand at each laboratory is reduced to a required maximum level, D_{max}, by removing facilities furthest away from the laboratory, thus producing a new set of facilities covered by each laboratory, F_j'.
04. A set, R, is computed of all laboratories where all of its facilities are also covered by other laboratories. These are termed 'removable' laboratories.
05 Initialise set J' to equal set J.
06. Do while $R \neq \emptyset$
07.   For each laboratory $r \in R$, select facility $f_r^* \in F_r'$ with $\max_{f_r \in F_r'}(d_{f_r})$, maximum demand covered by that laboratory.
08.   Select laboratory $r' \in R$ with $t_{f_r^*, r'} = \max_{r \in R}(t_{f_r^*, r})$, maximum time from the maximum demand covered.
09.   Remove laboratory r' from R and from J'.
10.   Re-compute R with the new, smaller set of laboratories J'.
11. End loop. Set J' represents the resulting set of laboratory sites for CD4 testing.
12. For $i \in I$, select $j^* \in J'$ s.t. $t_{i,j^*} = \min_{j \in J'}(t_{ij})$, i.e. j^* is the closest laboratory. Allocate demand at i to j^*.
13. Re-calculate total demand at each laboratory.
14. Extract the set of all uncovered facilities, $U = I - \bigcup_{j \in J'} F_j'$, i.e. facilities not

covered by any of the remaining laboratories within time *T*. Set *U*
represents the set of facilities where PoC machines are needed.

4.4 Software implementation and discussion of the system design

Modelling used Microsoft SQL Server Express 2008 R2, a free, basic level
version of Microsoft's SQL server.

The RACL algorithm was developed using Microsoft SQL Server in a C#
setting. The set covering approach employed by the RACL algorithm used
spatial commands available in the SQL server database, with modification for
the maximum daily throughput capacities of laboratory testing equipment.
Euclidean distances were calculated through SQL Server for this case study.

The version of SQL server used has full implementation of all spatial
operations offered on Microsoft's SQL server range, but is limited in size (less
than10Gb). These restrictions were found not to be a barrier to spatial
operations where the computational load is high per entry (i.e. it takes a lot of
effort to process a small amount of spatial data).

The development of the fast RACL algorithm to find optimal laboratory
locations was facilitated by the recent growth in relational databases which have
spatial query capabilities. Relational databases (as opposed to object- or
document-orientated databases) are written around relational algebra operations.
This facilitates fast execution of certain commands such as set selection,
intersection, union and difference. The new spatial commands provide similar
relational algebra commands, but to geographic and geometric data.

5. Results

We report firstly a comparison of the RACL algorithm to complete enumeration and provide a discussion of solution timings. Results for the NHLS problem then follow.

5.1 RACL Algorithm performance: findings on optimality of solution

The RACL algorithm was tested against complete enumeration on fictitious datasets with randomly generated locations and demand (see Appendix B), to compare optimality results (see Table 1). Runs were carried out with both 10 and 15 candidate locations. With 10 locations, the 600 runs were carried out in a total of 6 hours by full enumeration, while RACL took about 10 minutes, or 1 second per run. With 15 locations, just 64 runs took 15 hours with full enumeration, while RACL took about 1 minute.

==================insert Table 1 here ====================

5.2 RACL algorithm performance: timings

Possible solutions for locations were generated within minutes. The whole system, which consists of custom-written C# code using SQL Server 2008 R2 Express, was running in a 32-bit Windows 7 VirtualPC which limited it to one processor of the base system, an Intel I7 K875 processor running at 3.31GHz. The disk rotational speed was 7,200rpm.

Table 2 shows the run times for the system on a VirtualPC on a Windows 7 Desktop PC. The unconstrained scenarios with 4- and 5-hour travel time limits required only one pass of data, since the existing CD4 laboratory locations were in reach of all health facilities with these time limits. The constrained scenarios with shorter travel time limits required two passes of data, since both existing

CD4 laboratories and the set of all other laboratory sites were needed (see Section 4.3.2).

In Table 2, the "number shared at each pass" represents the number of facilities covered by more than one laboratory site, an indicator of the total run time. The latter was high for scenarios with high redundancy of laboratories (the most unlikely scenario).

==================insert Table 2 here =====================

5.3 Results for NHLS scenarios

Table 3 gives results obtained from RACL runs for the following scenarios, which were meaningful to NHLS decision makers:

- Centralisation of services: no constraint on capacity at any laboratory and travel limits of 4 hours and 5 hours. (These scenarios with no limitations on test volumes would incur high staffing and other resource needs.)
- With a maximum capacity constraint of 3,600 tests per day, and travel limits of 1 hour, 2 hours, and 3 hours.)

===============Insert Table 3 here ========================

The results can be compared to the situation at the time of case study, of 60 CD4 laboratories. A trade-off can be observed between shorter drive times as travel limits decrease and larger numbers of laboratories and PoC devices needed. The capacity-constrained option with a 2-hour travel limit, illustrated in Figure 4, was considered suitable to offer a balance between equitable access and cost-effective and efficient service provision.

5.4 Visualisation of results

Geographic visualisation of solutions was achieved by means of Google®
Fusion Tables. An example is shown in Figure 4: for this scenario, all travel is
within two hours. Placemark icons show CD4 laboratories and small circles
mark locations for PoC devices. In comparison to Figure 1, the suggested
locations for CD4 laboratories and PoC devices clearly offer better coverage of
all regions.

===================insert Figure 4 here ====================

5.5 Validation

As part of the strategic plan, NHLS selected four health districts as a pilot of the
re-organisation of testing capabilities. The NHLS Team had developed a local
optimisation for these areas prior to this research. A useful comparison could
therefore be made of the NHLS and RACL approach, although the regional
pilots did not take into account the needs of neighbouring areas. Table 4 shows
the characteristics of the four pilot regions and Table 5 gives the comparison of
results. It was felt by the NHLS that these comparisons gave validation to the
RACL results and to the NHLS local optimisation process.

=================insert Table 4 here ======================

===============insert Table 5 here ==========================

6. Discussion and study limitations

This project's aims were to find the "ideal" number of CD4 laboratory sites, placed correctly for demand. This reflects the NHLS vision for its future: "a new consolidated service model that will respond tangibly to the affordability challenges facing the public health sector without compromising quality or access." (NHLS, 2013b, p9) To fulfil these objectives, a balance needed to be found between numbers of sites and travel distances incurred. Moreover, the covering potential of any site is complex to define, as sample quality may deteriorate during travel by varying extents with duration of travel, with road surfaces and vehicle type, and with external temperature or means of appropriate packaging and/or refrigeration during travel.

Supplementary calculations beyond the scope of our case study performed by NHLS, using the solutions obtained with these scenarios, showed that reduction of the number of laboratories could increase running costs above the savings achieved. This is mainly attributed to higher costs of transportation logistics incurred with transporting large numbers of samples into centralised facilities. Nevertheless, for NHLS, there would be personnel efficiency gains with CD4 testing at fewer laboratories. However, longer travel distances associated with smaller numbers of laboratories to process samples could also adversely affect sample quantity.

It should be noted that, although a reduced number of laboratories to be operating was desirable, it was not a concern to find the absolute minimum number possible. Some degree of surplus capacity in the network of laboratories was advisable, as demand is variable and there needs to be contingency planning for unexpected increases in test volumes or breakdown. It was thus appropriate to use an algorithm that was essentially constructive in nature, i.e. without an iterative improvement phase. Moreover, this methodology has provided NHLS with a demonstrably transparent method for choice of locations for testing: reduction in the number of laboratory sites is achieved by

'removing' those sites that are furthest from major demand, where there is redundancy of cover for facilities.

Preference was given in this case study to existing CD4 laboratory sites in achieving better coverage. Results were additionally compared with those from trials run with candidate sites including both existing CD4 sites and other laboratory sites. The RACL algorithm was able to use fewer laboratories with the extended set of candidate sites, but did not achieve such 'good' placements, i.e. close to main areas of demand. This could arise because existing CD4 laboratories are already well located regards high demand areas.

Runs of an unmodified LSCP mathematical programme were also made, achieving lower laboratory numbers than resulted from RACL. However, several unsuitable locations with respect to demand were obtained using this programme. Improved locations could have been obtained with use of further constraints added to the LSCP or an amended objective function. The use of the RACL algorithm ensured that sites were chosen individually with regard to proximity to health facilities with high demand.

A limitation to the research that should be highlighted is the area of travel time estimations. The RACL algorithm is capable of using any distance or travel time measure. However, in this research, full travel time estimates from website services were not available because of licence costs. Inaccuracies in the calculated travel times used in this research were known to be particularly important in the rural regions where roads are scarce, e.g. rural KwaZulu-Natal and districts in the mountainous northern part of the country and areas to the west, where there are very large distances between metropolitan districts.

A further distance-related limitation of this work is that courier routes were not taken into account: typically, in the NHLS, a courier collects samples from a number of health facilities, driving from facility to facility in a continuous loop sequence, before delivering them to a laboratory. Moreover, an ideal transport system may be a 'hub and spoke' type model, with routing from local to hub

laboratories. In this case study, it was felt that the approximation of keeping all direct travel times from health facility to laboratory below a certain maximum would both reduce courier costs significantly and reduce the concomitant risk to sample quality.

This study was also not able to consider the types of road surface, details of specific traffic patterns and high temperatures which can affect sample integrity: much more detailed information would have been needed to take these factors into account explicitly. However, expert opinion is in agreement that reduction in travel times to the limit chosen (of less than two hours) is effective in improving sample quality appreciably. Moreover, the use of cooling units including special transport containers to maintain temperatures around 20°C in courier vehicles would mitigate this risk.

General non-availability of data regarding demand from health facilities was another limitation to the research, as methods of estimation were shown to be inaccurate, particularly in the rural areas of low population.

This research has the potential to impact the lives of those making use of HIV-testing facilities, particularly in rural areas. Although sparsely-populated, rural areas may be home to some of the poorest population of the country, and visits to health facilities can mean time away from work as well as walking appreciable distances. Thus, especially in these hard-to reach areas, knowledge of local blood testing with a quick turn-around times (TATs) of results can effectively encourage people to seek testing. In such settings, quicker TAT for results could be achieved by use of PoCs, by reduced travel times from health facilities to laboratories, and by the use of Short Message Service (SMS) printers (WHO, 2013) which can relay results direct from laboratories to health facilities.

7. Conclusion

This case study contributes a real-world application of locational analysis to the literature. We have supplied a scenario-based methodology for approaching this "messy" problem of finding efficient locations, in terms of maximum travel distances.

The RACL algorithm and associated methodology has provided decision makers with a methodology for finding appropriate sites for CD4 testing. However, the use of such a methodology must in practice be combined with local knowledge in creating a sustainable network of laboratories (Cassim et al., 2016). Moreover, the approach used proved efficient for a location problem with a large numbers of demand points and candidate service sites. The methodology employed made use of the software query capabilities now available. In tailoring the RACL algorithm to the NHLS problem, an efficient methodology has been made available for reducing numbers of laboratories whilst also reducing travel to laboratories: this method could be transferable to other applications of locational analysis. It is particularly to be recommended where decision makers need to demonstrate an objective and equitable rationale for choosing sites for services in the public sector.

Comparisons of results with the pilot study of four health districts suggestions were felt to give validation to the RACL results, despite the known limitations of this research in terms of estimated travel times, courier routes and estimated demand. Recommendations on CD4 laboratory site placements and closures, and sites for installation of PoC devices, have been made in detail using the results described here, with the scenarios described above (Glencross et al., 2014; Cassim et al., 2016).

Further locational analysis is continuing with an extension of the research to other tests needed for sufferers of HIV/AIDS including HIV Viral Load, Infant PCR, TB and other specialist laboratory tests. Rationalisation of the network of services available is being investigated, with the possibility of optimisation of courier routes for the collection of samples from health facilities.

Glossary of terms

HIV/AIDS: Human immunodeficiency virus infection and acquired immune deficiency syndrome (HIV/AIDS) is a spectrum of conditions caused by the human immunodeficiency virus (HIV). As the illness progresses, the immune system becomes compromised, causing susceptibility to other infections such as tuberculosis (TB).

ART: Antiretroviral therapy (ART) is a treatment for HIV/AIDS that suppresses the HIV and stops progression of the disease.

CD4: HIV causes AIDS by depleting the cluster of differentiation 4 (CD4) cells (also referred to as T-helper cells or T-cells). Laboratory tests are used to find CD4 levels in order to determine when to begin treatment during HIV infection.

HIV Viral Load: An HIV Viral Load test measures HIV nucleic acid (RNA), reported as number of HIV Virus "copies" per litre of blood. It is used to determine how well ART is controlling the virus.

Infant PCR: The infant polymerase chain reaction (PCR) is used for early infant HIV diagnosis where there has been a risk of mother-to-child transmission of HIV/AIDS.

References

Almiñana, M., Pastor, J.T., 1994. Two new heuristics for the Location Set Covering Problem. *TOP* 2, 2, 315-328.

Branas, C.C., ReVelle, C.S., 2001. An iterative switching heuristic to locate hospitals and helicopters. *Socio-Economic Planning Sciences* 35, 1, 11-30.

Bell, J.E., Griffis, S.E., Cunningham, W.A., Eberlan, J.A., 2011. Location optimization of strategic alert sites for homeland defense. *Omega* 39, 2, 151-158.

Burkey, M.L., Bhadury, J., Eiselt, H.A., 2012. A location-based comparison of health care services in four U.S. states with efficiency and equity. *Socio-Economic Planning Sciences* 46, 2, 157–163.

Cassim, N., Coetzee, L.M., Schnippel, K., Glencross, D.K., 2014. Estimating implementation and operational costs of an integrated tiered CD4 service including laboratory and point of care testing in a remote health district in South Africa. *PLOS ONE*, 9, 12, e115420. DOI: 10.1371/journal.pone.0115420

Cassim, N., Smith, H.K., Coetzee, L.M., Glencross, D.K., 2016. Programmatic implications of implementing the relational algebraic capacitated location (RACL) algorithm outcomes on the allocation of laboratory sites, test volumes, platform distribution and space requirements. Accepted for publication, *African Journal of Laboratory Medicine.*

Carling, K.; Han, M., Hakansson, J., 2012. Does Euclidean distance work well when the p-median model is applied in rural areas? *Annals of Operations Research* 201, 1, 83-97.

Chung, C.-H., 1986. Recent applications of the maximal covering location planning (M.C.L.P.) model. *Journal of the Operational Research Society* 37, 8, 735-746.

Chung, C.-H., Schilling, D.A., Carbone, R., 1983. The capacitated maximal covering problem: a heuristic solution. *Modeling and Simulation* 14, 1383-1388.

Church, R.L., ReVelle, C., 1974. The maximal covering location problem. *Papers of the Regional Science Association* 32, 1, 101-118.

Curtin, K.M., Hayslett-McCall, K., Qiu, F., 2010. Determining optimal police patrol areas with maximal covering and backup covering location models. *Networks & Spatial Economics* 10, 1, 125-145.

Erkut, E., Ingolfsson, A., Sim, T., Erdoğan, G., 2009. Computational comparison of five maximal covering models for locating ambulances. *Geographical Analysis* 41, 1, 43–65.

Finkel, R., Bentley, J.L., 1974. Quad Trees: A data structure for retrieval on composite keys. *Acta Informatica* 4, 1, 1–9.

Glencross, D.K., Coetzee, L.M., Faal, M., Masango, M., Stevens, W.S., Venter, W.D.F., Osih, R., 2012. Performance evaluation of the Pima™ point-of-care CD4 analyser using capillary blood sampling in field tests in South Africa. *Journal of the International AIDS Society* 15, 1, 3, doi:10.1186/1758-2652-15-3.

Glencross, D.K., Coetzee, L.M., Cassim, N., 2014. An integrated tiered service delivery model (ITSDM) based on local CD4 testing demands can improve turn-around times and save costs whilst ensuring accessible and scalable CD4 services across a national programme. *PLOS ONE* 9, 12, e114727. DOI: 10.1371/journal.pone.0114727

Hakimi, S.L., 1964. Optimum locations of switching centres and the absolute centres and medians of a graph. *Operations Research* 12, 3, 450-459.

Hakimi, S.L., 1965. Optimum distribution of switching centers in a communication network and some related graph theoretic problems. *Operations Research* 13, 3, 462–475.

Hamers, R.L., Sawyer, A.W., Tuohy, M., Stevens, W.S., Rinke de Wit, T.F., Hill, A.M., and for the ART-A Consortium, 2012. Cost-effectiveness of laboratory monitoring for management of HIV treatment in sub-Saharan Africa: a model-based analysis. *Aids* 26, 13, 1663-1672.

Higgins, A.J., Laredo, L.A., 2006. Improving harvesting and transport planning within a sugar value chain. *Journal of the Operational Research Society* 57, 4, 367-376.

Li, X., Zhao, Z., Zhu, X., Wyatt, T., 2011. Covering models and optimization techniques for emergency response facility location and planning: a review. *Mathematical Methods of Operations Research* 74, 3, 281-310.

Marianov, V., Serra, D., 2011. Median problems in networks. In: Eiselt, H.A., Marianov, V. (Eds), *Foundations of Locations Analysis*. Springer, New York, pp. 39-59.

Nahid, P., Kim, P.S., Evans, C.A., Alland, D., Barer, M., Diefenbach, J., *et al*. (2012). Clinical research and development of tuberculosis diagnostics: moving from silos to synergy. *Journal of Infectious Diseases* 205, S2, S159-S168.

NDOH (2010). HIV and AIDS. National Department of Health, Republic of South Africa. http://www.doh.gov.za/list.php?type=HIV and AIDS&year=2010, accessed 30 May 2010.

NDOH (2013). National Department of Health, Republic of South Africa: The South African Antiretroviral Guidelines, 2013. http://www.kznhealth.gov.za/medicine/2013_art_guidelines.pdf, accessed 29 June 2014.

NHLS (2013a). National Health Laboratory Service Priority Programmes. http://www.nhls.ac.za/?page=priority_areas&id=11, accessed 30 May 2013.

NHLS (2013b). National Health Laboratory Service (NHLS) Annual Report 2012-13. http://www.nhls.ac.za/assets/files/NHLS_AR2012-13.pdf, accessed 4 August 2014.

Pizzolato, N.D., 1994. A heuristic for large-size *p*-median location problems with application to school location. *Annals of Operations Research* 50, 1, 473-485.

Rehle, T., Shisana, O., Glencross, D.K., Colvin, M., 2005. *HIV-Positive Educators in South African Public Schools: Predictions for prophylaxis and antiretroviral therapy*. HSRC Publishers, South Africa.

Rehle, T.M., Hallett, T.B., Shisana, O., Pillay-van Wyk, V., Zuma, K., Carrara, H., Jooste, S., 2010. A decline in new HIV infections in South Africa: estimating HIV incidence from three national HIV surveys in 2002, 2005 and 2008. *PLOS ONE* 5, 6, e11094. DOI:10.1371/journal.pone.0011094.

ReVelle, C.S., Swain, R.W., 1970. Central facilities location. *Geographical Analysis* 2, 1, 30-42.

Shariat-Mohaymany, A., Babaei, M., Moadi, S., Amiripour, S.M., 2012. Linear upper-bound unavailability set covering models for locating ambulances: Application to Tehran rural roads. *European Journal of Operational Research* 221, 1, 263-272.

Shuklaa, N., Wickramasuriya, R., Millera, A., Pereza, P., 2015. An approach to plan and evaluate the location of radiotherapy services and its application in the New South Wales, Australia. *Computer Methods and Programs in Biomedicin* 122, 2, 245-256.

Smith, H.K., Laporte, G., Harper, P.R., 2009. Locational analysis: highlights of growth to maturity. *Journal of the Operational Research Society* 60, S1, S140-S148.

Smith, H.K., Harper, P.R., Potts, C.N., 2013. Bicriteria efficiency/equity hierarchical location models for public service application. *Journal of the Operational Research Society* 64, 4, 500–512.

Toregas, C., Swain, R., ReVelle, C. Bergman, L., 1971. The location of emergency service facilities. *Operations Research* 19, 6, 1363–1373.

WHO (2013) World Health Organisation. SMS printers aid early infant diagnosis of HIV/AIDS in Nigeria: CHAI's SMART. http://apps.who.int/iris/bitstream/10665/92801/1/WHO_RHR_13.19_eng.pdf?ua=1, accessed 29 June 2014.

Ye, L., Ye, C., Chuang, Y.-F., 2011. Location set covering for waste resource recycling centers in Taiwan. *Resources, Conservation and Recycling* 55, 11, 979-985.

Appendix A. The QuadTree clustering method

The QuadTree clustering method (Finkel and Bentley, 1974) is used as follows:

1. The geography is split into rectangles which contain a maximum number of health facilities that can be processed efficiently: the resulting rectangles are called leaf nodes. By experimentation it was found that 250 facilities per leaf node provided reasonable performance.

2. Clustering of a leaf node:

   a. Having reached a leaf node, SQL server spatial commands are used to find all facilities within cluster distance C km of each other in that leaf node.

   b. Facilities are then clustered by searching for chains of facilities that are within distance C, e.g. if A is close to B and B is close to C, then the cluster consists of A, B and C.

3. Clustering leaves at branch level then takes place. When recursing back up the QuadTree, clusters in one leaf node are combined that touch a cluster (or clusters) in another leaf node. The leaf nodes are referred to as Northwest (NW), Southwest (SW), Northeast (NE) and Southeast (SE). The procedure is as follows:

   a. First, any touching clusters in NW and SW are combined to form a new group called West.

   b. Similarly, any touching clusters in NE and SE are combined forming a new group called East.

   c. West and East are then combined and handed back as a group of clusters, as if as a leaf node.

d. Next, proceeding back to the next higher level: if this is not the top then steps 3a) to 3c) are executed again.

Experimentation was carried out with different levels of clustering: it was found by trial and error that a distance of C = 4km was the best compromise to assist the subsequently-used RACL algorithm, without excessive demand per cluster. The cluster distance of 4km reduced the number of demand points from 3,266 facilities to 1,880 clustered facilities. Clustering mainly affected groups of facilities in the densely populated areas of Durban and Johannesburg.

Appendix B. Random generation of test locations

Fictitious datasets used in testing the RACL algorithm (see Section 5.1) were generated as follows. The locations of candidate sites for laboratories were generated within a first, inner, circle, of predetermined radius. Health facility locations were generated within an outer circle, of larger radius. In each circle, latitudes and longitudes of locations were generated separately using a pseudo-random number generator (.net class Random) with seeds taken from the system clock. Locations were first generated within a square: it was then calculated whether a location had been generated within the required circle. If outside the circle, that location was discarded and more locations were generated until the required numbers of locations was obtained.
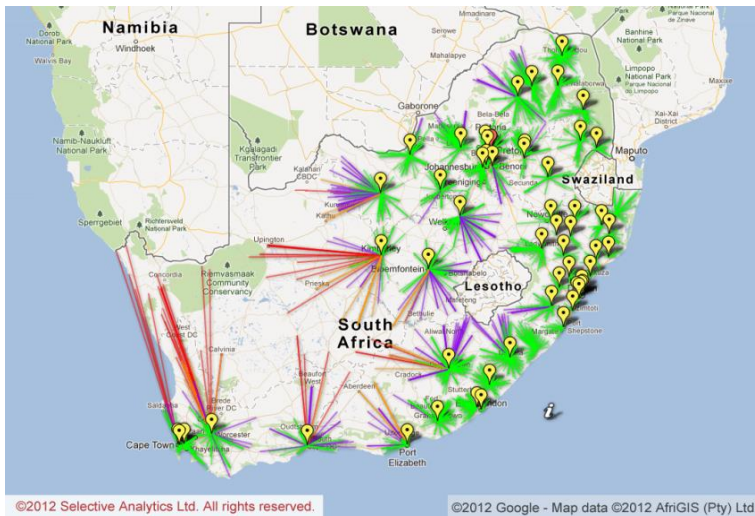
Figure 1: NHLS laboratories throughout South Africa undertaking CD4 testing in 2011, showing links from health facilities. Key: yellow placemark icon CD4 laboratory, straight lines links from health facilities.
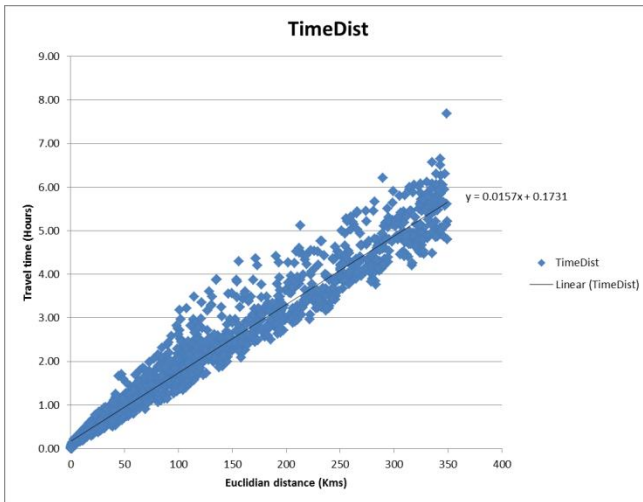
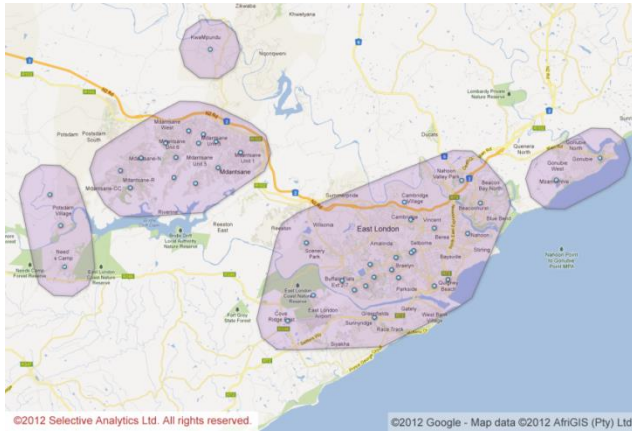**Figure 2**: Drive time against Euclidean distances on selected routes

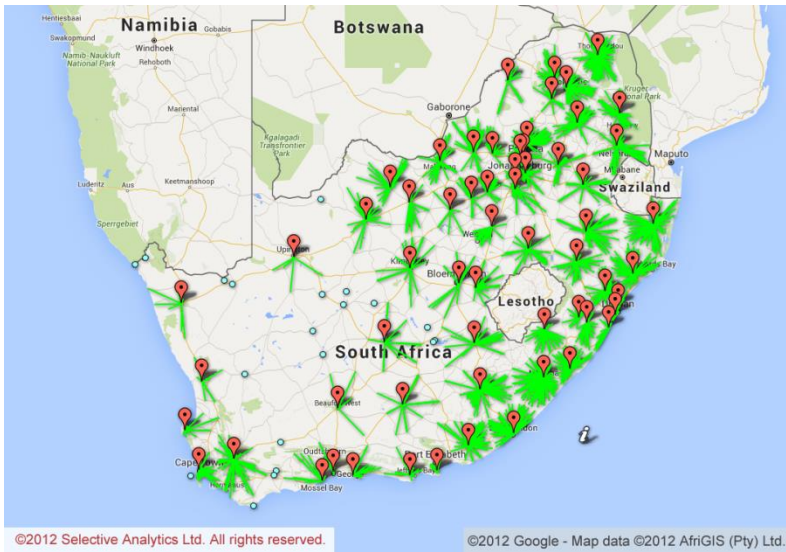**Figure 3**: Clustering of health facilities in the East London area of South Africa.

Figure 4: Results for capacity-constrained scenario, 2-hour travel limit. Key: red placemark icon CD4 laboratory, small circles PoC placement, green lines links from clustered health facilities.

Table 1: Test optimality comparisons of the RACL algorithm with complete enumeration

#locations	#runs	% optimal	% sub-optimal by 1 laboratory	% sub-optimal by 2 laboratories
10	600	97.67	2.17%	0.17%
15	64	92.19	7.8%	0%

Table 2: RACL algorithm run times by scenario

Run characteristics			Timings		
Scenario	Travel limit	#passes	Run time (RACL) hr:min:sec	#shared (pass 1)	#shared (pass 2)
Unconstrained	5hr	1	00:14:21	7,067,672	
Unconstrained	4hr	1	00:07:05	3,830,260	
Constrained	3hr	2	00:00:29	20,316	14,458
Constrained	2hr	2	00:00:23	11,750	5,738
Constrained	1hr	2	00:00:28	1,296	5,132

Table 3: Results obtained for various scenarios

Capacity constraint	Travel limit	Numbers of laboratories needed	Numbers of PoC devices needed
None	5 hours	10	0
None	4 hours	15	0
Yes	3 hours	41	2
Yes	2 hours	61	20
Yes	1 hour	127	190

Table 4: Pilot district characteristics

District number	Province	Population.	Estimated % HIV prevalence
1	Northern Cape	182,000	13%
2	KwaZulu-Natal	852,827	36%
3	Gauteng	862,679	32%
4	Free State	752,906	32%

Table 5: Comparison of pilot district and RACL suggestions

District Number	Pilot suggestions	RACL suggestions
1	1. Improve logistics to (place A), or 2. Place new CD4 Lab in (place B)	Place new CD4 Lab in (place B)
2	Area well serviced. One lab, or labs, could be closed	Close labs in two places
3	Re-open lab (at place C)	Re-open lab (at place C). Close two other labs.
4	Good service. No change.	3 hour travel: no change. 2 hour travel: open lab (at place D)