Journal of the Operational Research Society (2017) 68, 1068–1081

www.palgrave.com/journals

Open

Location of a hierarchy of HIV/AIDS test laboratories in an inbound hub network: case study in South Africa

Honora Smith[1]*, Daniel Cakebread[1], Maria Battarra[2], Ben Shelbourne[1], Naseem Cassim[3,4] and Lindi Coetzee[3,4]

[1]Mathematical Sciences, University of Southampton, Highfield, Southampton SO17 1BJ, UK; [2]School of Management, University of Bath, Bath BA2 7AY, UK; [3]National Health Laboratory Service, Johannesburg, South Africa; and [4]Department of Molecular Medicine and Haematology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

HIV/AIDS test laboratories in South Africa face growing demand for high quality, timely and efficient testing of blood samples in all regions of the country, however rural. The three main tests for HIV/AIDS, namely CD4, HIV Viral Load, and Infant PCR, are provided in a hierarchy of levels: CD4 is the most frequently needed test, with most laboratory coverage needed. HIV Viral Load is less frequently called for, and Infant PCR is the rarest test to be done, with correspondingly fewest laboratories needed. The National Health Laboratory Service (NHLS) of South Africa operates an inbound hub network for collection of blood samples and transfer to laboratories equipped to carry out the required tests: test results are transmitted electronically, so there is no outbound or return transport. This paper describes the development of modelling carried out over several years of collaboration with NHLS to advise decision-makers on an appropriate and efficient hub network. We present mixed integer programs to find efficient locations for both network hubs and locations for all levels of laboratory testing. Novel features include variable or range constraints on maximum travel times to test locations.

Journal of the Operational Research Society (2017) **68(9),** 1068–1081. doi:10.1057/s41274-017-0240-5;
published online 21 June 2017

Keywords: location; hierarchical health services; hub networks; HIV/AIDS testing; OR in developing countries

1. Introduction

The National Health Laboratory Service (NHLS) of South Africa is seeking efficiencies of both running costs and transportation costs for HIV/AIDS blood sample testing in the nationwide network of laboratories, in view of anticipated future increases in demand for services. HIV/AIDS tests have been given priority status in South Africa (NHLS, 2016; Cassim *et al*, 2017) in order to ensure that the whole population, wherever resident in the country, has adequate access to accurate and timely blood testing.

Blood samples collected in hospitals or clinics undergo one of three tests to diagnose and monitor the effects of treatment for HIV/AIDS: the most commonly performed test is CD4, while HIV Viral Load and Infant PCR are carried out successively less frequently (see Glossary for explanation of HIV/AIDS-related terms). The NHLS laboratories are *hierarchically* organised in a hub network. Once collected, the blood samples are transferred by courier to the closest local laboratory for testing. If the local laboratory does not have the equipment or the capacity to perform any of the tests needed, the blood samples are transferred to the allocated hub laboratory, as shown in Figure 1. Likewise, if the hub laboratory does not have the equipment to perform any tests needed, the blood samples are batched together and transferred to other hub laboratories in the network, termed 'tertiary laboratories'. CD4 tests take place as close as possible to source at the local or hub laboratories; HIV Viral Load and Infant PCR tests are carried out at tertiary laboratories.

The resulting hierarchical hub network allows for efficiencies of transfer in comparison with a fully connected network. However, this distributed network structure presents many operational challenges and great scope for optimisation in deciding the role of each laboratory in the hierarchy (local laboratory, hub laboratory, hub/tertiary laboratory) and in assigning the test equipment to each laboratory. Transportation costs have to be minimised, but distances should not be such as to compromise the quality of the blood samples.

The blood sample collection hub network differs from typical transportation hub location problems studied in the literature; in the latter, the emphasis is on efficient routing from origin to the final goods destination, whereas, for blood

*Correspondence: Honora Smith, Mathematical Sciences, University of Southampton, Highfield, Southampton SO17 1BJ, UK.
E-mail: honora.smith@soton.ac.uk

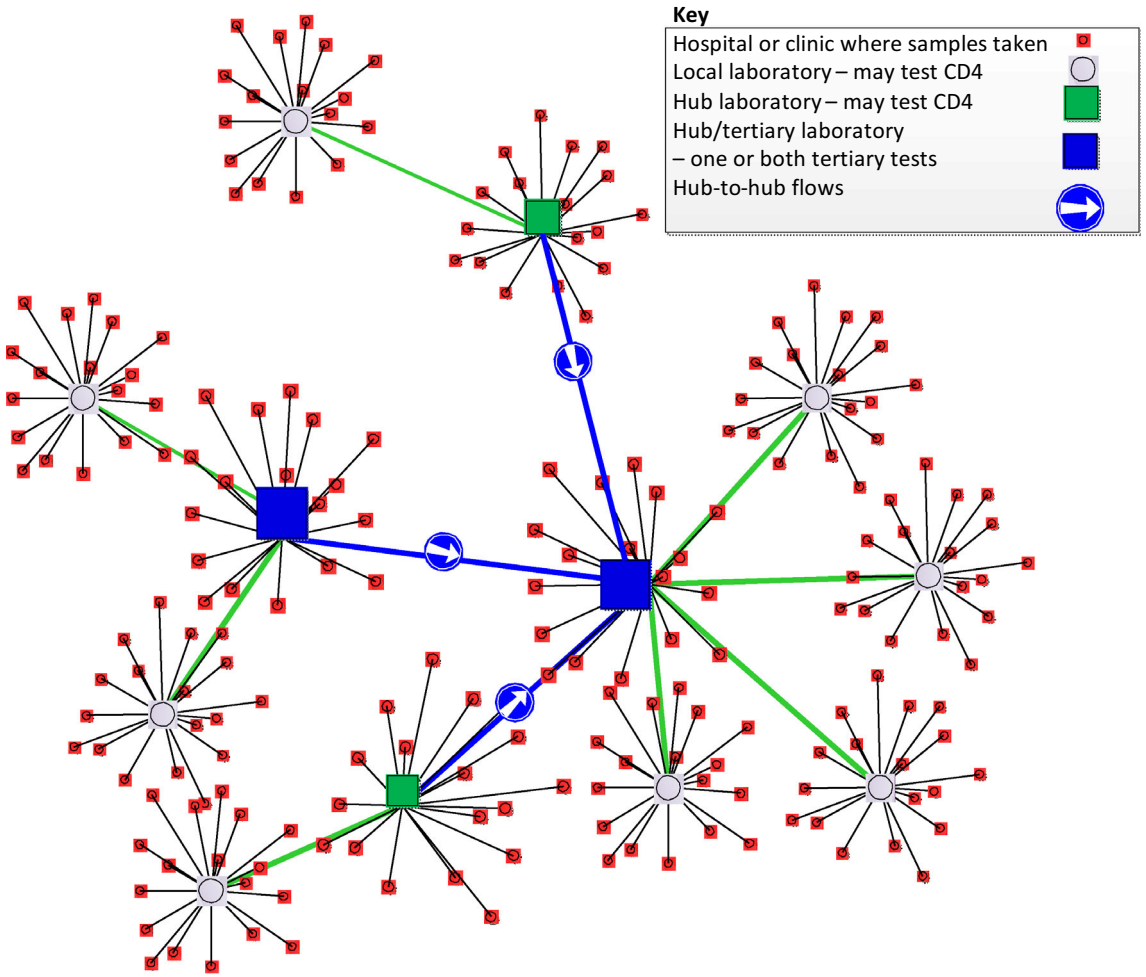

Figure 1 NHLS hub network: hospitals, local, hub and tertiary laboratories.

samples, the destination is a suitably close laboratory where samples are tested. From there, test results are transmitted by electronic means to the originating health facility and blood samples do not have to be routed back. We therefore term this an *inbound-only* hub network. Similarly, this description could apply to the transport of waste to landfill sites via transfer stations described by Eiselt (2007), who notes the similarity to hub location problems. It cannot be claimed that the blood sample network has the same hub-to-hub efficiencies of scale afforded by heavy goods transportation hub networks, but nevertheless, the need to refrigerate samples during onward transit makes a hub-to-hub network both efficient and one that protects the integrity of samples.

This case study highlights the complexities of operations of this network of laboratories, which gave rise to the path of analysis undertaken and modelling developed over several years of collaboration between the research team and NHLS. A single level of tests, for CD4 only, was first considered (Smith *et al*, 2017; Cassim *et al*, 2017); results were considered of sufficient interest to warrant a more complex hierarchical study considering all three HIV/AIDS tests simultaneously as well as the designation of laboratories within the hub network. In this paper, we propose mixed integer programming (MIP) models designed to find appropriate locations for the test equipment, as well as assigning to each component of the network the role of local laboratory, hub laboratory, or hub/tertiary laboratory. Special features of the models developed include variable or range constraints to suit the diverse terrains across which samples are delivered nationwide. Several scenarios (and successive MIP models) are considered and robust solutions explored to guide the NHLS in clearly identifying objectives, budget and operational constraints. Efficiencies are demonstrated in terms of numbers of laboratories, hubs operated, and distances travelled.

This paper continues in Section 2 with a description of the NHLS hub network for the testing of HIV/AIDS blood samples. Section 3 provides a literature review of location analysis applied to hierarchical services and also to hierarchical hub networks. Section 4 presents two MIP models that are proposed to find locations for hierarchical test laboratories within a hub network. Section 5 provides computational results of the case study, with new model modifications for variable

or range constraints, and exploration of solutions close to optimal. Section 6 is the conclusion.

2. Background to the case study: HIV/AIDS blood testing in the NHLS network

The area of the Republic of South Africa is more than 1.2 million square kilometres, and terrain varies from mountainous to bush land (SA Government, 2015) and from desert to densely populated cities. In the NHLS system, blood samples are given by patients at 3490 health facilities, which can range in size from one-roomed clinics in remote rural areas to university hospitals in the major cities. The blood samples are taken to local laboratories, where some tests for HIV/AIDS may be performed. Otherwise, samples are further transported to a hub laboratory and thence to other hub laboratories for tertiary-level tests if needed.

There is more demand for HIV/AIDS tests in the eastern cities than in the remote regions of the west, but, for reasons of equity, it is a requirement that all health facilities, wherever they are located throughout the country, should be able to access laboratory service for HIV/AIDS and other tests. However, there are limitations to distances that can be covered by each laboratory, and sample quality can be affected by conditions of the roads, which vary from dirt tracks to highways, and also by the high temperatures sometimes encountered. A previous study of this problem (Smith *et al*, 2017; Cassim *et al*, 2017) recommended locations for point-of-care (POC) equipment for one HIV/AIDS test, namely CD4, that can be supplied where transportation is difficult, with instruments located directly in health facilities. CD4 testing is used to determine the need for antiretroviral treatment (ART) and is the most frequently used HIV/AIDS test, with over 3.8 million CD4 tests carried out by the NHLS in 2011/2012.

In this study, we consider simultaneously the locations for three tests used in HIV/AIDS treatment: CD4, and two less frequently needed tests, HIV Viral Load and Infant PCR. HIV Viral Load is used to determine treatment effectiveness, and Infant PCR is a test for infants and newborn babies thought to be at risk of having contracted HIV/AIDS. The distances that CD4, HIV Viral Load and Infant PCR samples can travel without deterioration vary with the stability of the sample type. CD4 tests are the most susceptible to damage during transit, due to vibrations from rough road surfaces and high temperatures if courier vehicles are not refrigerated (Smith *et al*, 2017). HIV Viral Load samples are less easily damaged than CD4, and Infant PCR samples the least, since the latter consist of dried blood or plasma. All health facilities must be within a 'suitable' distance of laboratories where all three tests can take place; however, a range of travel times is considered suitable by NHLS decision-makers for the different tests, rather than an exact maximum.

Clearly defined and quantified performance indicators for improvement were not readily available at the outset of this study. The investigation therefore proceeded by analysing the effects of varying the maximum travel time limits on both numbers of laboratory sites required and on total travel times, since both are of importance to decision-makers. It is generally desirable to minimise conveyance distances of samples from health facility to laboratory, to ensure sample quality is maintained and to improve turnaround time for test results to be returned to patients. However, reductions in travel times bring inefficiencies of increasing numbers of laboratories. It was therefore decided to provide decision-makers with alternative *scenarios*, to demonstrate trade-offs between these objectives with differing travel time targets for samples. More specifically, we use MIPs to define, for each set of travel time limits, the minimum number of laboratories needed to cover all demand, along with minimisation of the overall travel time.

Couriers transport blood samples from health facilities to the nearest local laboratory and thence to the network of hub laboratories, as shown in Figure 1. In this study, the allocations of health facilities to local laboratories are pre-assigned. All local laboratories are connected by courier to a single hub laboratory. Hub laboratories carry out sorting and rerouteing of samples, as well as the important function of quality management. If CD4 testing is not carried out at any local laboratory, then this test must be provided at the hub laboratory to which it is connected, since CD4 samples deteriorate quickly during transit. HIV Viral Load and Infant PCR samples are batched together at the hub laboratories for onward refrigerated transport where necessary to other hub laboratories ('tertiary' laboratories) equipped for these tests. The candidates sites for hub or hub/tertiary laboratories are all major hospital sites from which large volumes of tests originate. The hierarchy of tests carried out at the different laboratories is illustrated in Figure 2, and Table 1 summarises the tests that can be performed each type of laboratory. It should be noted that the HIV/AIDS tests are carried out at some, not all laboratories. If no tests are

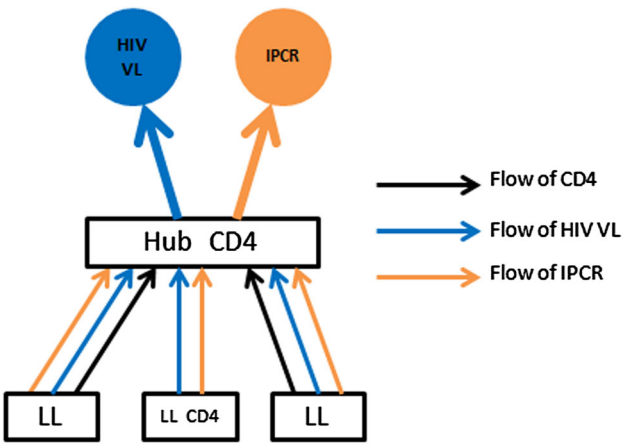

Figure 2 Hierarchy of HIV/AIDS tests carried out at local laboratories (LL), hub and hub/tertiary laboratories: CD4, HIV Viral Load (HIV VL) and Infant PCR (IPCR).

Table 1 Summary of HIV/AIDS tests that may be located at the different types of laboratory

Type of laboratory	*Tests that may be located*
Local laboratory	CD4
Hub/tertiary laboratory	CD4, HIV Viral Load, Infant PCR

carried out at any one local laboratory or hub laboratory, then functions of sorting and rerouteing only are performed there.

Providing appropriate laboratory test capacity is of importance to the NHLS, since there are limitations on numbers of tests that can be carried out at any one laboratory. Demand is highest for CD4 tests, whereas HIV Viral Load and Infant PCR tests are successively less frequently carried out. Moreover, some flexibility in the system is desirable, in terms of spare capacity, to allow for machine breakdowns and unexpected surges in demand.

To summarise, the objective of this project is to advise on appropriate locations for CD4, HIV Viral Load and Infant PCR testing, simultaneously with hub locations and allocations from local laboratories to hubs and hubs to hub/tertiary laboratories. The scenario approach chosen considers a range of maximum travel time limits from health facility to laboratory, which are used to gauge appropriate numbers of laboratories. More detailed modelling objectives are discussed in Section 3.

3. Literature review

The principles underlying two classical types of location modelling, *p*-median, and covering models, have relevance to the problem of design of the HIV/AIDS testing hub network. The *p*-median problem, due to Hakimi (1964, 1965), minimises total demand-weighted distance travelled with a given number of service centres; the assumption is made that all demand must be serviced. Covering problems consider demand covered within a certain travel distance or time. The set covering location problem (Toregas *et al*, 1971) finds minimum numbers of service centres needed to cover all demand, while the maximum covering problem (Church and ReVelle, 1974) finds the maximum demand that can be covered using a given number of service centres. In the NHLS problem, the travel times within which samples are brought in for testing are of importance, while the overall distance travelled on courier routes can benefit from cost reduction. Moreover, all demand must be covered: the principles of both the set covering and *p*-median problems are therefore both relevant here, although the requirements of hierarchical location of laboratories in this hub network result in different model formulations from the classical originals.

Location of hierarchical services has received much attention within locational literature, with reviews from Church and Eaton (1987), Eitan *et al* (1991), Galvão *et al* (2002), and Şahin and Süral (2007); most recently, Farahani *et al* (2014) analyse around 100 references. Schultz (1970), Narula (1984), Serra (1996) and Şahin and Süral (2007) provide classification of location problems where there is a hierarchical relationship between facilities. Classifying features include (1) the number of levels, (2) whether a particular type of services may be provided at several levels or at a unique level in the hierarchy, (3) whether flow is always to next higher (or lower) levels or to multiple levels, and (4) whether all demand assigned to one particular facility at one level is allocated together at the next level. With reference to such defining features, Eitan *et al* (1991) demonstrate use of a generalised hierarchical location–allocation model with flexible hierarchical relationships, referrals and variations in capacity constraints and costs. Marianov and Serra (2001) analyse hierarchical systems with congestion. Flow-based hierarchical formulations are provided by Şahin and Süral (2007) and Farahani *et al* (2014). Daskin (2013) classifies hierarchical systems in health care, schools, postal services and banks; generic median- and covering-based hierarchical facility location models are provided. Moreover, interacting hierarchical facilities are considered with referrals between levels and covering constraints between levels.

Hierarchical facility models have been applied to both public service and supply chain applications. Banerji and Fisher (1974) combine *p*-median and set covering modelling for integrated hierarchical planning of services in an area of rural India. Moore and ReVelle (1982) apply a maximal covering hierarchical location model to public health services in Honduras. Hodgson (1988) presents a hierarchical model for health facilities in developing countries. Hinojosa *et al* (2000) consider a manufacturing supply chain with two levels of distribution facilities for supplying customers. A mixed integer formulation is provided with heuristics based on lower bounds obtained by Lagrangian relaxation. Three-level hierarchical locations for maternity services are considered by Galvão *et al* (2002) in Rio de Janeiro, Brazil and by Baray and Cliquet (2013) for the whole of France. Teixeira and Antunes (2008) model the primary school system of Coimbra, Portugal, as a hierarchical facility location problem. Smith *et al* (2013) combine efficiency and equity objectives in both median- and covering-type models for location of hierarchical health services, illustrated with a case study in a rural region of north India.

Hub location problems have been widely researched in the field of locational analysis, with a focus on finding best locations for hub nodes, and allocation of local demand to hubs. A comprehensive review of network hub location problems is provided by Alumur and Kara (2008), emphasising that the classical types of location problems, such as *p*-median, *p*-centre and covering, have their equivalents in hub problems. In Alumur and Kara's review of more than 100 hub location papers, it is noted that *p*-median problems are more frequently encountered than covering ones, since the former address the common objective of reducing transportation costs. Only one multi-objective problem is noted (Costa *et al*, 2008) which is theoretically rather than application oriented. More recently, Campbell and O'Kelly (2012) give a historical perspective on the development of hub location problems. They point out that hub location is a network design problem,

with hubs and demand nodes constituting a two-level network design. Lowe and Sim (2013) attend to the class of hub problems with maximum distance constraints but without limitations on capacity. Farahani *et al* (2013) indicate the range of both exact solution techniques and heuristic or metaheuristic approaches employed in finding solutions to hub location problems, including continuous or plane locations. Examples of real-life case studies of hub location problems include landfill siting in Canada (Eiselt, 2007) and postal services in Turkey (Çetiner *et al*, 2010).

Hub networks may themselves be described as hierarchical, with a variety of designs. Goldman (1969) considers the scenario of different stages of processing carried out at different locations in a network such as that of a postal operation, in a paper later identified as a forerunner of hub network analysis (Campbell and O'Kelly, 2012); however, no examples are provided nor solution methods demonstrated. Several recent papers consider location problems in hierarchical hub networks, particularly concerning parcel deliveries. Smilowitz and Daganzo (2007) consider key costs of complex multimode, multiservice hierarchical networks such as package distribution systems with express and normal levels of service, such as Federal Express in the USA. Yaman (2008) finds hub locations for a star-star hub network with all connections made through a central hub, with the objective of minimising link installation costs; a case study is made of cargo deliveries in Turkey. Yu *et al* (2009) develop a cluster-based hierarchical model for the location of urban transit hubs, with a case study in an industrial park in China. A multi-hierarchical design is proposed, with three levels of hubs servicing different categories of routes. Lin (2010) studies integration of dual-level carrier services in Taiwan provided in a hierarchical hub-and-spoke network. Alumur *et al* (2012) analyse the postal delivery network of Turkey, which is arranged as a multimodal hierarchical hub network: a star-shaped nationwide airline network with hubs that are individually accessed by local truck hub networks. This network design represents an example of networks within networks. Telecommunications also provide examples of hierarchical hub networks, with multilevel concentrators and routers through which connections are made (Chardaire *et al*, 1999; Ignacio *et al*, 2008).

4. Model formulations

Two models for hierarchical services in a hub network (HH—hierarchical hub) are proposed. The first model, HH-mN, finds the minimum total number of laboratories (mN—minimum total number), summing over the three types of tests, needed to ensure that the travel times are within the defined limits. A second model follows: HH-mT, which minimises the total travel time (mT—minimum total time) given particular numbers of laboratories for the three test types. Use of these two models facilitates a two-stage decision process: firstly on numbers of laboratories to utilise and secondly on where to site them.

It should be noted that this represents a simplification of the system in that there may be a waiting time required for batches to form for courier transport, particularly at smaller laboratories.

The data needed for these models are: numbers of samples (demand) for each type of testing brought to each local laboratory; maximum travel times from health facilities into each local laboratory; and travel times from local laboratories to candidate hub laboratory locations and travel times between candidate hub locations. We summarise the technical problem details that apply to both models:

- The hub network is assumed to have a homogeneous design throughout the country. Local laboratories are considered to be the demand points for this model, since samples from hospitals are transported to the nearest local laboratory. Each local laboratory is connected to one hub laboratory; no flows are allowed from local laboratories to other hubs. This is termed *single allocation* by Alumur and Kara (2008). A hub may itself be a local laboratory for nearby health facilities.
- If a local laboratory does not provide CD4 testing, this must be provided at the hub laboratory to which it is connected. Each hub is connected to up to two other hub laboratories that carry out tertiary testing for HIV Viral Load and Infant PCR tests; such hub/tertiary laboratories may carry out both HIV Viral Load and Infant PCR tests or just one of these tests. A set of candidate sites is given for both hub and hub/tertiary laboratories. Hub/tertiary laboratories may receive samples for processing from several other hubs.
- There are constraints on the capacity of the hub laboratories. For this study, it is assumed that a maximum of one piece of equipment for one test can be placed at a given location. However, there are no constraints put on CD4 demand to be satisfied at local laboratories, since demand must be satisfied at these laboratories if necessary to maintain sample quality. Moreover, since hub locations can also be local laboratories, it is possible for a hub to receive CD4 samples both as a local laboratory and as a hub laboratory, with the possibility of requiring extra equipment.
- There are also constraints on the time from when the blood sample is collected to when it is tested. Only inbound travel times to laboratories are considered, since test results are communicated electronically to health facilities.
- Since the candidate sites for HIV Viral Load are the same as for Infant PCR, the same travel time targets can be used for these two tests, although the smaller demand for Infant PCR tests will require fewer laboratories for these tests.

This section continues with details of the different models proposed, in particular the different objective functions.

4.1. HH-mN formulation: minimisation of total equipment numbers given travel time targets

The objective of HH-mN is to minimise the total numbers of pieces of test equipment, summed over the different types

of test given by set B = {1 (CD4), 2 (HIV Viral Load), 3 (Infant PCR)}. This set-covering-type objective is a proxy for the operating costs of laboratory equipment; it assumes that facility costs at all sites are identical. Several tests can be carried out at one laboratory site in the case of hub/tertiary laboratories. Optimal locations for both hub and tertiary laboratories are sought from a set of candidate locations, J. Set I comprises the local laboratory sites, and $J \subset I$, since hub/tertiary laboratories also act as local laboratories. Demand from local laboratory i, $i \in I$, for test type b, $b \in B$, is denoted by d_i^b. The maximum travel time to local laboratory i, $i \in I$, from health facilities for which local laboratory i is the nearest, is given by m_i. The travel time from local laboratory i to hub laboratory j, $i \in I, j \in J$, is given by t_{ij}. Since $J \subset I$, t_{jk} gives travel times between hubs for $j,k \in J$. Constraints are imposed through test equipment capacity, D^b, and maximum permitted total travel time, T^b, from local laboratory to test site, for test type b, $b \in B$.

Decision variables for locations of tests and connections to hubs and tertiary laboratories are as follows. It should be noted that flow-based rather than assignment-based decision variables are used, for manageable matrix sizes, since samples from any one local laboratory are taken to more than one tertiary laboratory.

$$x_i^L = \begin{cases} 1 & \text{if local laboratory } i \text{ does CD4 testing,} \\ 0 & \text{otherwise,} \end{cases} \quad i \in I.$$

$$x_j^H = \begin{cases} 1 & \text{if hub laboratory } j \text{ does CD4 testing,} \\ 0 & \text{otherwise,} \end{cases} \quad j \in J.$$

$$x_k^{T,b} = \begin{cases} 1 & \text{if tertiary laboratory } k \text{ does type } b \text{ testing,} \\ 0 & \text{otherwise,} \end{cases}$$
$$k \in J, b \in \{2,3\}.$$

$$y_{ij}^H = \begin{cases} 1 & \text{if laboratory } i \text{ is connected to hub } j, \\ 0 & \text{otherwise,} \end{cases} \quad i \in I, j \in J.$$

$$y_{jk}^{T,b} = \begin{cases} 1 & \text{if hub } j \text{ is connected to tertiary } k \text{ for type } b \text{ tests,} \\ 0 & \text{otherwise,} \end{cases}$$
$$j,k \in J, b \in \{2,3\}.$$

Auxiliary variables are thus:

$u_{ij}^{H,b}$ = flow of type b demand from laboratory i to hub j, $i \in I,\ j \in J,\ b \in B$.

$u_{jk}^{T,b}$ = flow of type b demand from hub j to tertiary k, $j,k \in J,\ b \in \{2,3\}$.

w_j^H = an upper bound to all travel times from local laboratories to hub j, $j \in J$.

Since $J \subset I$, it should be noted that if $y_{jj}^H = 1, \quad j \in J$, this means that a hub is placed at j, since demand from j is allocated to j. Similarly, $y_{kk}^{T,b} = 1, \quad k \in J$, means that a tertiary for type b tests, $b \in \{2,3\}$, is placed at k.

The formulation is as follows:

Minimise

$$\sum_{i \in I} x_i^L + \sum_{j \in J} x_j^H + \sum_{b \in \{2,3\}} \sum_{k \in J} x_k^{T,b}, \tag{1}$$

subject to:

$$\sum_{j \in J} y_{ij}^H = 1, \quad \forall\, i \in I, \tag{2}$$

$$\sum_{k \in J} y_{jk}^{T,b} = y_{jj}^H, \quad \forall\, j \in J,\ b \in \{2,3\}, \tag{3}$$

$$y_{ij}^H \le y_{jj}^H, \quad \forall\, i \in I, j \in J, \tag{4}$$

$$y_{jk}^{T,b} \le y_{kk}^{T,b}, \quad \forall\, j,k \in J,\ b \in \{2,3\}, \tag{5}$$

$$x_j^H \le y_{jj}^H, \quad \forall\, j \in J, \tag{6}$$

$$x_k^{T,b} = y_{kk}^{T,b}, \quad \forall\, k \in J,\ b \in \{2,3\}, \tag{7}$$

$$u_{ij}^{H,b} \le M y_{ij}^H, \quad \forall\, i \in I,\ j \in J,\ b \in B, \tag{8}$$

$$u_{jk}^{T,b} \le M y_{jk}^{T,b}, \quad \forall\, j,k \in J,\ b \in \{2,3\}, \tag{9}$$

$$d_i^1 x_i^L + \sum_{j \in J} u_{ij}^{H,1} = d_i^1, \quad \forall\, i \in I, \tag{10}$$

$$\sum_{j \in J} u_{ij}^{H,b} = d_i^b, \quad \forall\, i \in I,\ b \in \{2,3\}, \tag{11}$$

$$\sum_{k \in J} u_{jk}^{T,b} = \sum_{i \in I} u_{ij}^{H,b}, \quad \forall\, j \in J,\ b \in \{2,3\}, \tag{12}$$

$$\sum_{j \in J} \sum_{k \in J} u_{jk}^{T,b} = \sum_{i \in I} d_i^b, \quad \forall\, b \in \{2,3\}, \tag{13}$$

$$\sum_{i \in I} u_{ij}^{H,1} \le D^1 x_j^H, \qquad \forall\, j \in J, \tag{14}$$

$$\sum_{j \in J} u_{jk}^{T,b} \le D^b x_k^{T,b}, \quad \forall\, k \in J,\ b \in \{2,3\}, \tag{15}$$

$$-M x_i^L + m_i + \sum_{j \in J} t_{ij} y_{ij}^H \le T^1, \quad \forall\, i \in I, \tag{16}$$

$$w_j^H \ge y_{ij}^H (m_i + t_{ij}), \quad \forall\, i \in I,\ j \in J, \tag{17}$$

$$w_j^H + \sum_{k \in J} t_{jk} y_{jk}^{T,b} \le T^b, \quad \forall\, j \in J,\ b \in \{2,3\}, \tag{18}$$

$$x_i^L, x_j^H, x_k^{T,b}, y_{ij}^H, y_{jk}^{T,b} \in \{0,1\}, \quad \forall\, i \in I,\ j,k \in J,\ b \in \{2,3\}, \tag{19}$$

$$u_{ij}^{H,b}, w_j^H \geq 0, \quad \forall\, i \in I, j \in J,\ b \in B, \tag{20}$$

$$u_{jk}^{T,b} \geq 0, \quad \forall\, j,k \in J,\ b \in \{2,3\}. \tag{21}$$

The objective function (1) minimises the total numbers of pieces of all test equipment in all laboratories.

Constraints (2)–(5) make the connections between local laboratories, hubs and tertiary laboratories. Constraint (2) ensures that each local laboratory is connected to exactly one hub. Constraint (3) connects each hub laboratory to exactly one tertiary laboratory of each type; non-hub laboratories are not connected to tertiaries. Constraint (4) mandates that a local laboratory can only be connected to a located hub. These constraints also ensure that demand originating at a hub will be sent to that hub. Similarly, constraint (5) ensures that hubs can only be connected to open tertiary laboratories, for each test.

Constraints (6)–(13) are concerned with the placing of equipment in laboratories for carrying out the various tests, and the flows of samples between laboratories. Constraint (6) ensures that equipment is placed only in opened hubs. Constraint (7) mandates that a tertiary laboratory for a particular test is opened only when equipment for that test is placed there; these constraints do, in fact, remove the necessity for $x_k^{T,b}$, but this is retained for clarity. Constraints (8) and (9) mandate that flows of samples from local laboratory i to hub j and from hub j to tertiary k correspond to the connections made; M is a large positive number. Constraint (10) causes the flow of CD4 samples between local laboratory i and hub j to equal demand at i, unless CD4 testing is done at i. All CD4 testing of samples from i is done either at local laboratory i or at the connected hub, but not in both locations. Constraint (11) mandates that the flow of HIV Viral Load and Infant PCR samples from local laboratory i to the connected hub equals the demand from i. Constraint (12) is for flow balancing: all tertiary demand that reaches any hub must be sent on to a tertiary. Constraint (13) ensures that all HIV Viral Load and Infant PCR demand is allocated to tertiary laboratories.

Constraints (14) and (15) ensure that equipment is placed at hubs and tertiaries for all tests for which there is demand and that capacities at hubs and tertiaries are not exceeded. The assumption is made that only one piece of equipment for any one test type can be placed in any one laboratory. Constraints (16)–(18) keep travel times for all types of sample within the appropriate threshold. Constraint (16) mandates that CD4 equipment is placed at a local laboratory if travel from any health facilities to the assigned hub is longer than the CD4 travel limit. Through constraint (17), auxiliary variable w_j^H is an upper bound on total travel times from health facilities into hub j, $j \in J$. Constraint (18) ensures that the total time from health facilities to tertiary laboratories is within limits.

4.2. HH-mT formulation: minimisation of total travel time

HH-mT minimises total travel time, given preset numbers of laboratories carrying out the different tests. This model can thus be run with numbers of laboratories resulting from HH-mN runs, or for any chosen set of numbers of laboratories. An additional decision variable is introduced: $z_{jk},\ j,k \in J$, which equals 1 if $y_{jk}^{T,2}$ or $y_{jk}^{T,3} = 1$ (or both), and 0 otherwise. This models the combined transportation of both HIV Viral Load and Infant PCR samples where both tests are carried out at one tertiary laboratory.

The objective function minimises the total travel time from local laboratories to hubs and from hubs to hub/tertiary laboratories. The constraints of HH-mN are retained, with added specification of numbers of laboratories for the individual test types. Numbers of laboratories carrying out each type of test are set using parameters p_1, p_2 and p_3 for CD4, HIV Viral Load and Infant PCR tests, respectively. Such laboratory numbers may be those found from HH-mN or may be specified to represent other scenarios.

Minimise

$$\sum_{i \in I, j \in J} t_{ij}\, y_{ij}^H + \sum_{j,k \in J} t_{jk} z_{jk}, \tag{22}$$

subject to constraints (2)–(21) and as follows:

$$\sum_{i \in I} x_i^L + \sum_{j \in J} x_j^H = p_1, \tag{23}$$

$$\sum_{k \in J} x_k^{T,2} = p_2, \tag{24}$$

$$\sum_{k \in J} x_k^{T,3} = p_3, \tag{25}$$

$$z_{jk} \geq y_{jk}^{T,2}, \qquad \forall\, j,k \in J, \tag{26}$$

$$z_{jk} \geq y_{jk}^{T,3}, \qquad \forall\, j,k \in J, \tag{27}$$

$$z_{jk} \in \{0,1\}, \qquad \forall\, j,k \in J. \tag{28}$$

Constraints (23)–(25) specify numbers of laboratories carrying out testing for CD4, HIV Viral Load and Infant PCR tests, respectively. Constraints (26) and (27) register the link between hubs j and k, $j,k \in J$, if tertiary demand of type 2 or 3 is sent from j to k.

5. Case study results

This section provides results of running HH-mN and HH-mT and highlights modifications made to these models suggested by the results.

For all model runs, optimisation was implemented using the FICO^{TM} solver Xpress-IVE 7.8 64 bit. Runs were carried out

on a computer with Intel®Core™ i5-3340M CPU at 2.70GHz, with 4GB RAM. Suitable optimality of solutions, for the purposes of this case study, was obtained without the need for further solution methods. Maps were prepared using Google® Fusion Tables.

A total of 266 local laboratories are considered to be the demand points for the models, and a set of 39 candidate sites is given for both hub and hub/tertiary laboratories. The current number of laboratories processing CD4 samples is 58, while 15 laboratories offer HIV Viral Load tests and 8 laboratories carry out the tests for Infant PCR. Scenarios are of interest that use similar numbers of laboratories to those in the current NHLS network, or fewer laboratories. If expected demand for tests alone is considered, the absolute minimum numbers of laboratories needed for CD4, Viral Load and Infant PCR would be 17, 7 and 2 laboratories, respectively, but without consideration for travel times from remote places.

5.1. HH-mN experiments: variations on maximum travel times

The experiments carried out using model HH-mN were designed to make comparisons with the current situation in terms of numbers of laboratories and demonstrate areas for possible improvement in locations.

Table 2 shows results of experiments with model HH-mN. The first column gives the travel time targets for (1) CD4, (2) HIV Viral Load, and (3) Infant PCR. Minimum numbers of laboratories were found (see columns 2 and 3), indicating operating efficiencies that can be gained for different scenarios of maximum travel time. Geographical considerations preclude lowering the travel time to tertiary laboratories below 11 h, because of travel from remote regions. As expected, shorter travel times were shown to require more laboratories. We highlight that with a 3-h target for CD4 testing and targets to tertiary laboratories for HIV Viral Load and Infant PCR of 11 h, a total of 70 laboratories is required: 60 CD4, 7 HIV Viral Load and 3 Infant PCR. This is a comparable number of laboratories to the current number (58) for CD4 and less than the current numbers (15 and 8) for HIV Viral Load and Infant PCR, respectively. It was decided to take this configuration of 70 laboratories forward for further modelling with HH-mT, since maximum 3-h travel for CD4 samples is a desirable target; moreover, optimality was reached in model execution. The optimality gaps otherwise achieved (all less than 5%) were considered sufficiently small at this stage in the study. It is clear that better solutions (i.e. lower optimality gaps) could be obtained more quickly with the more constrained model runs, as is expected with smaller search spaces.

Geomapping of the resulting hub network locations, however, showed a confusing picture of delivery routes with crossovers. This is because travel was not necessarily to the closest facility. Closest assignment constraints could have been used to address this problem, but this is also achieved using minimisation of total travel times, which is the objective in the HH-mT experiments that follow.

5.2. HH-mT experiments: minimising total travelling time given numbers of laboratories

Runs of model HH-mT are summarised in Table 3, which includes numbers of laboratories given, travel time limits set and resulting minimised total travel times. Firstly, taking results from HH-mN (see Table 2), a run was undertaken of model HH-mT to find optimal locations for 70 laboratories (60 CD4, 7 HIV Viral Load, 3 Infant PCR) with minimised total travel time, using travel time bounds of (3 − 11 − 11 h). Figure 3 shows the resulting geographical locations. Small yellow circles show local CD4 laboratories, while the small red circles are local laboratories sending CD4 samples to hub laboratories. Hub laboratories not performing tertiary tests are shown by green placemark icons. Hubs with the HIV Viral Load tertiary test only are shown by blue icons with the number 1; hubs performing both tertiary tests are coloured blue with the number 2. Thin red lines join local laboratories to hubs; it is notable that the local laboratories more distant from hubs provide CD4 tests. Thicker lines join hubs to hubs for tertiary tests: these are coloured green (pale) for hub to HIV Viral Load, blue for hub to Infant PCR and blue/green for hub to both tests. It can be observed in this figure that the resulting network does not contain any route crossovers, because of distance minimisation.

Secondly, for comparison purposes, an HH-mT run was carried out with actual NHLS numbers of laboratories, with 81 laboratories (58 CD4, 15 HIV Viral Load, 8 Infant PCR), to

Table 2 Results of runs of HH-mN to determine minimum total laboratory numbers needed to achieve travel time targets for (1) CD4, (2) HIV Viral Load, (3) Infant PCR

Max travel times (h) (1) − (2) − (3)	*Min numbers of labs (1) + (2) + (3)*	*Min total number*	*Optimality gap (%)*	*Elapsed time (s)*
4 − 15 − 15	35 + 7 + 3	45	4.26	8621.0
3 − 15 − 15	57 + 7 + 2	66	2.12	7862.3
2 − 15 − 15	112 + 7 + 2	121	0	3025.3
4 − 11 − 11	36 + 7 + 3	46	2.99	480.6
3 − 11 − 11	60 + 7 + 3	70	0	195.1
2 − 11 − 11	116 + 7 + 3	126	0	72.0

Table 3 Results of runs of HH-mT to determine minimum total travel time, given numbers of laboratories and travel time limits

Numbers of labs (1) + (2) + (3)	*Travel time limits (1) − (2) − (3) (h)*	*Total travel time (h)*	*Optimality gap (%)*	*Elapsed time*
70 = 60 + 7 + 3	3 − 11 − 11	445.6	2.06	14.3 h
81 = 58 + 15 + 8	3 − 11 − 11	Infeasible		
81 = 58 + 15 + 8	4 − 11 − 11	324.154	0	115.4 s

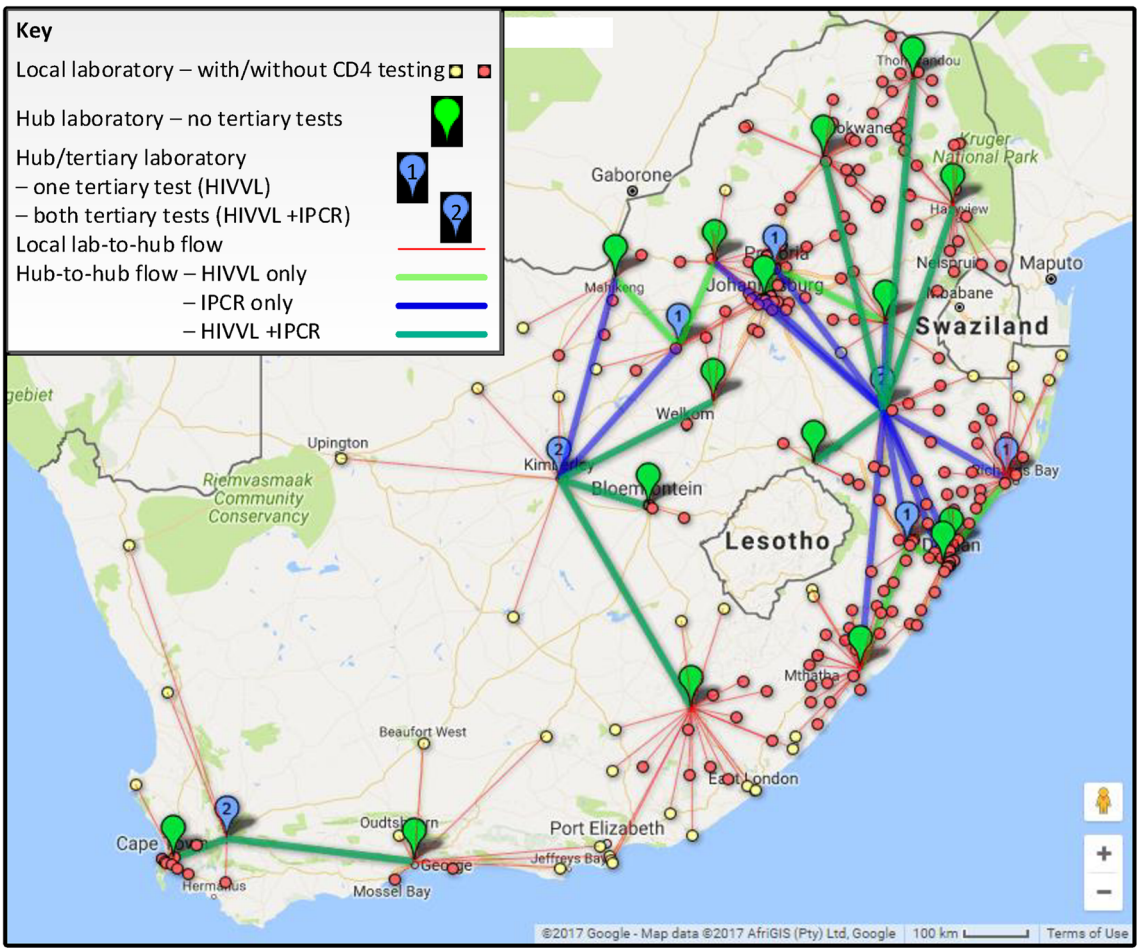

Figure 3 HH-mT results, minimised travel time with 60 CD4, 7 HIV Viral Load, 3 Infant PCR laboratories and travel time limits 3 − 11 − 11 h.

find locations that minimise total travel time. With travel bounds of (3 − 11 − 11) h, there was no feasible solution, due to insufficient CD4 laboratories. With bounds of (4 − 11 − 11) h, an optimal solution of 324 h was achieved: a total daily travel time saving of over 120 h results from the 11 extra laboratories (compared to the 70-laboratory run), but a large number of hubs are used (in 34 of the available 39 sites). Figure 4 shows the hierarchical hub network resulting from this solution. It may be observed that the hubs for tertiary tests are spread across the country, with more concentration around densely populated areas such as Johannesburg in comparison with Figure 3; there is consequently much less travel between hubs.

Of particular interest are the distributions of times taken to transport samples, particularly CD4, the least robust sample: see Figure 5 for travel times from health facilities to local laboratories and Figure 6 for times from local laboratories to hubs. The latter times result from the 70-laboratory HH-mT run described in this section. Recall that constraint (16) locates CD4 testing at the local laboratories if travel times from health facilities to local laboratories plus times from local laboratories to hubs of longer than the permitted CD4 travel time, in this case 3 h. It is evident that most facility-to-local laboratory travel times are less than 2.0 h, while most local laboratory-to-hub times are less than 3.0 h. This occurs because, in South Africa, most health facilities are situated in populous areas,

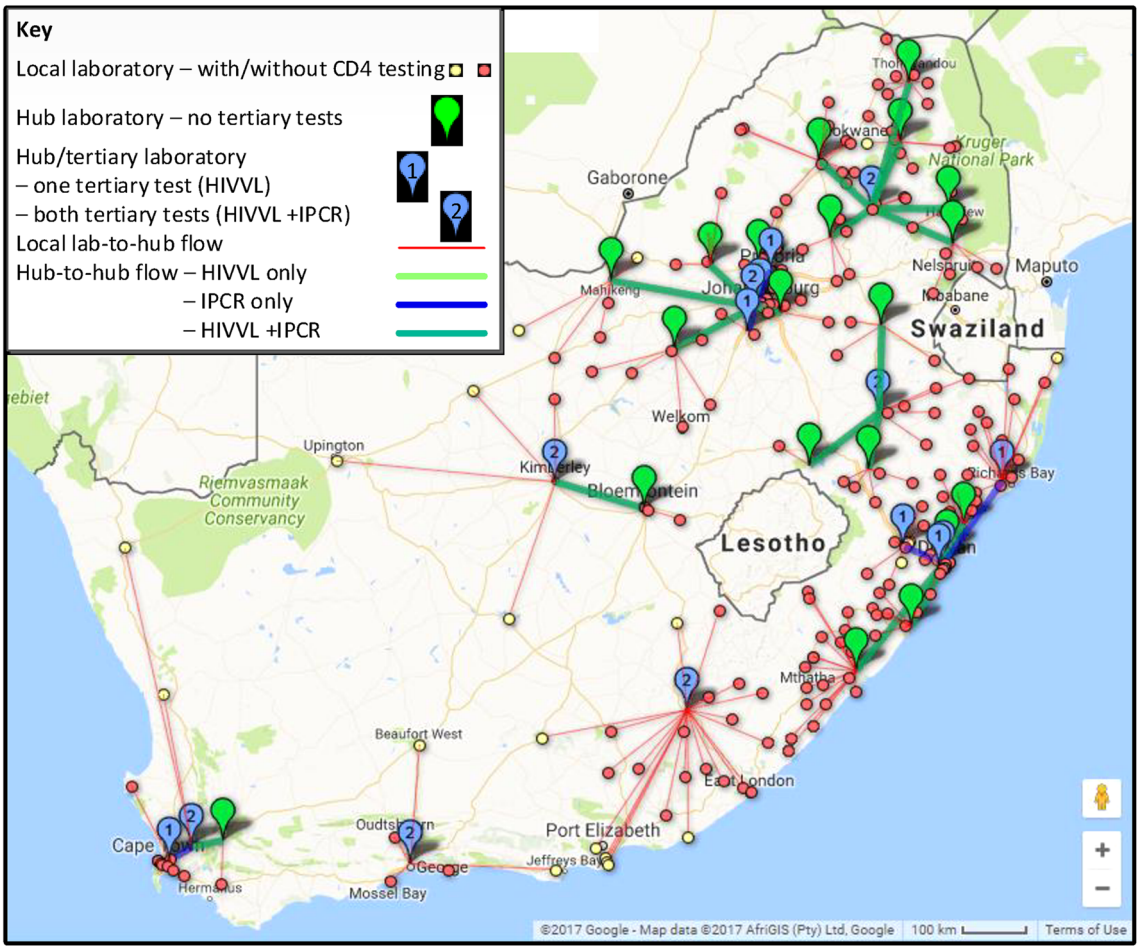

Figure 4 HH-mT results, minimised distances with 58 CD4, 15 HIV VL, 8 IPCR as currently, with travel time limits of 4 − 11 − 11 h, respectively, 34 hubs.

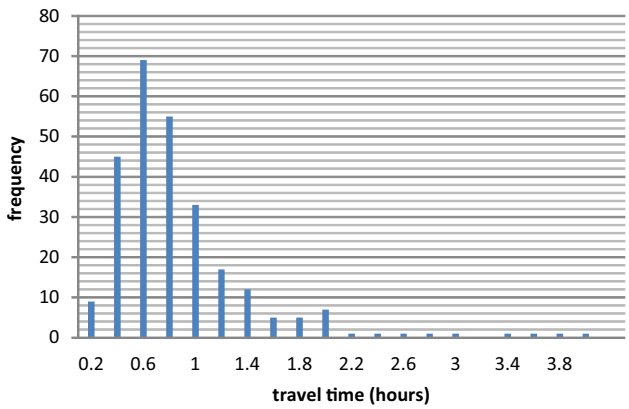

Figure 5 Histogram of travel times from health facilities to local laboratories.

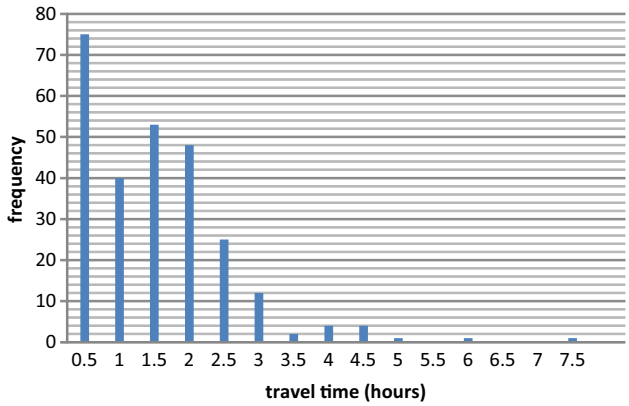

Figure 6 Histogram of travel times from local laboratories to hubs, model HH-mT, 60 + 7 + 3 laboratories, 3 − 11 − 11 h travel limits.

close to laboratories; few health facilities are in the rural areas, remote from laboratories.

The results of running HH-mN and HH-mT (presented in Section 5) highlight the differences of operation of the NHLS network in the rural west, compared with the denser populated areas of the eastern coast and inland. In the latter regions, there is high demand for tests and numerous laboratory sites; in the west there is less demand and there are greater distances to be

travelled to laboratory sites. This means that a maximum travel constraint that cannot be achieved in the west, with a given number of laboratories, may be achievable in the eastern coast and inland regions. We therefore propose further models with variable or range constraints for times to CD4 testing, which we describe in Section 5.3.

5.3. HH-mNvT and HH-mTvT, with variable or range time constraints

The results of running HH-mN and HH-mT (presented in Section 5) suggest modifications on these models; two additional models are therefore proposed, HH-mNvT and HH-mTvT (see Section 5), which employ variable or range constraints on travel times (vT—variable time constraints), enabling appropriate service to be provided throughout the country within an acceptable range of travel times.

Models HH-mNvT and HH-mTvT are modified from HH-mN and HH-mT with constraint (16a) varying the times to CD4 testing, in place of constraint (16):

$$-Mx_i^L + m_i + \sum_{j \in J} t_{ij} y_{ij}^H \leq T_i^1, \qquad \forall\ i \in I. \tag{16a}$$

These models allow for a range of acceptable time limits to be specified, according to the situation of particular local laboratories. So, in the examples that follow, while the majority of the 266 local laboratories may provide CD4 testing within a 3-h travel limit, a small number may offer testing within the relaxed time of 4 h, providing a more desirable solution than constraining all travel within 4 h.

Table 4 presents experiments with model HH-mNvT, with maximum time to CD4 testing in the range 3–4 h. Laboratories chosen for the 4-h limit were those with longer than a given time to hub in runs of HH-mN (see the first column of Table 4). Results are compared with HH-mN results where all local laboratories have a 3-h time constraint. The reductions in numbers of CD4 laboratories needed as time limits are relaxed is evident.

Table 5 shows results of runs of HH-mTvT, finding minimum total distances for two configurations, with variable time limits as determined in HH-mNvT runs described previously. A comparison with Table 3 for the current numbers of laboratories ($81 = 58 + 15 + 8$) shows that an increase in total travel time from 324 h to 370 h is needed for this improvement in service from all CD4 laboratories at 4 h to a variable range of 3/4 h. This configuration is shown in Figure 7: noticeably, there are fewer hubs and correspondingly fewer hub-to-hub routes than in Figure 4 with the same number of laboratories.

In this case study, several of the models have been solved to optimality for the required parameters in very short time scales using commercial software. An exception to this has been with the time minimisation models; a future study could investigate different methods of improving solution times for this problem instance (Klotz and Newman 2013). However, for the purposes of this case study, the degree of optimality is sufficient since 'good' rather than 'ideal' solutions are of interest, as we discuss in Section 5.4.

5.4. Comparison of solutions for robust decision-making

Optimal and near-optimal solutions have been found for locating hubs, CD4 test laboratories and hub/tertiary laboratories for HIV Viral Load and Infant PCR, under several scenarios for travel time limits. However, decision-makers in NHLS are interested in which locations are robust choices for

Table 4 Results of runs of HH-mNvT to determine minimum total numbers of laboratories, given variable range of travel time limits

Lab-to-hub time (h)	*# Labs at 4-h relaxed time*	*Optimal solution Tot = (1) + (2) + (3)*	*Time to optimality (s)*
2	52	50 = 40 + 7 + 3	67
2.5	28	61 = 51 + 7 + 3	153.2
2.7	21	66 = 56 + 7 + 3	213.9
2.9	18	68 = 58 + 7 + 3	175.4
3	16	69 = 59 + 7 + 3	77.2
3.2	15	69 = 59 + 7 + 3	76
3.5	14	69 = 59 + 7 + 3	76.5
3.7	8	69 = 59 + 7 + 3	198.0
n/a	0	70 = 60 + 7 + 3	195.5

Table 5 Results of runs of HH-mTvT to determine minimum total travel time, given numbers of laboratories and variable range of travel time limits

# Labs (1) + (2) + (3)	*Travel time limits (1) – (2) – (3) (h)*	*# Labs relaxed time*	*Total travel time (h)*	*Optimality gap (%)*	*Elapsed time*
68 = 58 + 7 + 3	Range 3/4 – 11 – 11	18	434.983	1.93	11 h
81 = 58 + 15 + 8	Range 3/4 – 11 – 11	18	369.798	0	1076.8 s

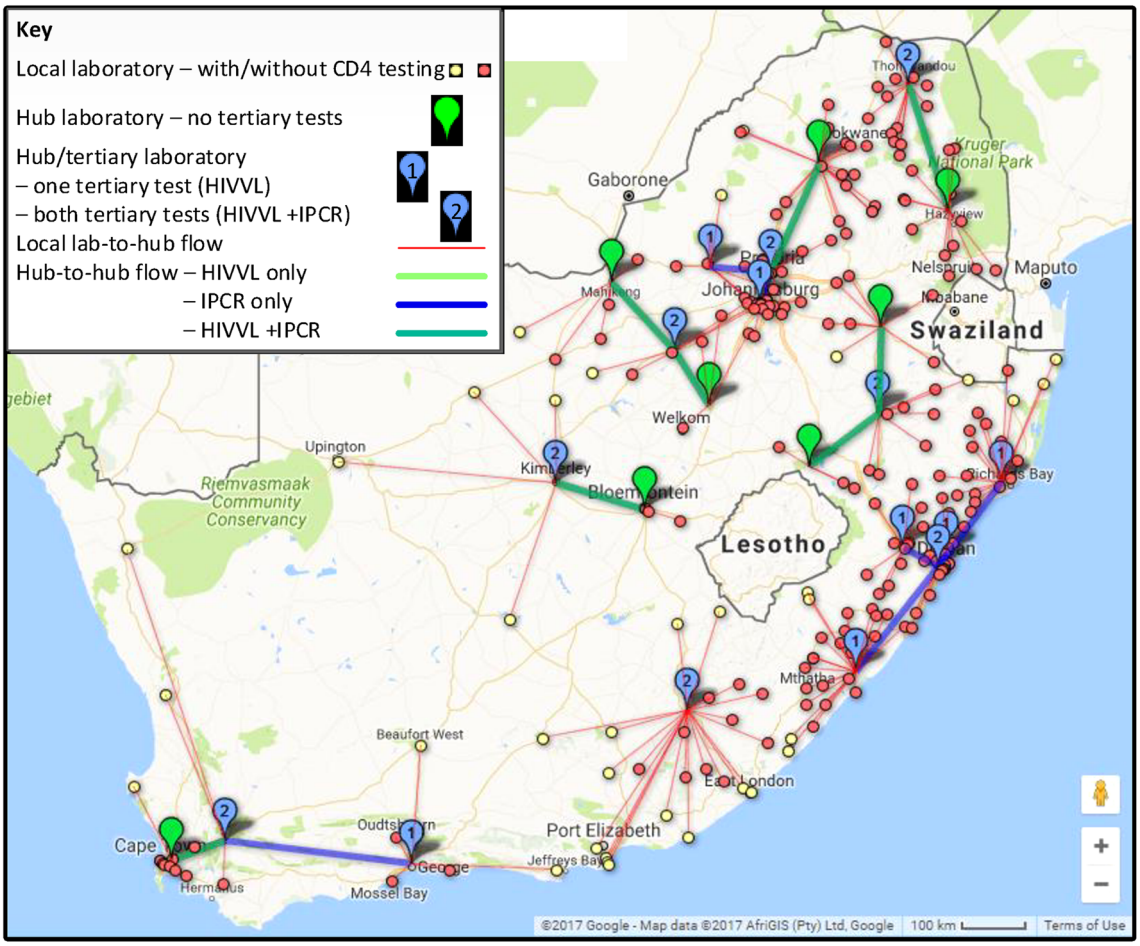

Figure 7 HH-mNvT results, current numbers of laboratories, variable range 3/4-h CD4 travel limits, 23 hubs.

hubs and laboratories under several scenarios. We therefore analyse and compare a set of representative solutions, to note patterns that may be emerging for locations, both for hubs and for tertiary laboratories. In particular, we compare:

1. Current numbers of laboratories (58 CD4, 15 HIV Viral Load and 8 Infant PCR), (a) fixed 4-h CD4 travel time limit and (b) variable range of 3/4 h, optimal solutions of 324.2 and 369.8 h, respectively;
2. Reduced number of laboratories (60 CD4, 7 HIV Viral Load and 3 Infant PCR), fixed 3-h travel time limits, solutions at 446.3 h (gap 2.84%) and 445.6 h (gap 2.06%);
3. Further reduced number of laboratories (58 CD4, 7 HIV Viral Load and 3 Infant PCR), variable 3/4-h travel limits, solutions at 444.6 h (gap 2.239%) and 443.5 h (gap 1.926%).

For brevity's sake, results of comparisons between solutions are not included in detail. However, the following may be noted, providing strategic recommendations to NHLS:

- In all solutions, Infant PCR laboratories are co-located with HIV Viral Load laboratories;
- The subset of hub and tertiary locations appearing in solutions for reduced numbers of laboratories (items 2 and 3) are included in optimal locations for current numbers (item 1): this subset of locations is therefore suggested as efficient hub/tertiary sites;
- Optimal locations for current numbers of laboratories (item 1) use 34 hubs (out of the total of 39) for fixed 4-h limits with shorter total travel time, but 23 for variable range of 3/4 h for fixed 4-h limits, again suggesting that the latter 23 sites are efficient hub/tertiary locations;
- The time improvement in solutions with reduced numbers of laboratories (item 2) is gained with 2 less hubs and one different HIV Viral Load test location;
- The time improvement with further reduced numbers of laboratories (item 3) is gained with a different HIV Viral Load test location.

So, in summary, the hub and tertiary locations for reduced numbers of tertiary laboratories (see Figure 3) provide robust locations, i.e. these locations are suited to several scenarios with current numbers. An increase in the number of hubs

operating may result in reduction in total travel time, but this is not always the case.

6. Conclusion

The operation of the complex, large-scale NHLS network of laboratories has given rise to several questions: "What are 'appropriate' numbers of laboratories?" and "What are 'efficient' (rather than 'the optimal') locations for hubs and laboratories?" Our approach has therefore been to present different scenarios for acceptable travel time targets for the different HIV/AIDS tests, which differ in both capacities needed and toleration of road travel. We have used principles of classical locational analysis to develop hub location models that reflect the hierarchical nature of both the tests and the hub network structure. Moreover, model versions with ranges of acceptable travel time constraints have been used to make access times appropriate across differing conditions of rural and city travel. An equitable time to test within acceptable ranges for samples has the potential to improve blood sample testing for HIV/AIDS sufferers in the poorest rural regions of the country.

This case study has provided NHLS laboratory management with an objective methodology for validating planning decisions concerning the hub network. Model runs have demonstrated efficiencies that can be gained in terms of numbers of laboratories and hubs, and total travel times. Results have suggested several hub locations that are robust under several different scenarios for laboratory numbers; alternative designs of hub networks are of interest to the decision-makers. It is demonstrated that laboratories for the less frequently needed tertiary-level test, Infant PCR, should be co-located with HIV Viral Load laboratories, although the latter needs additional laboratory provision. The presentation and visualisation of results, in terms of numbers and locations for laboratories and hubs, has helped decision-makers to better identify the challenges and opportunities offered by alternative network designs.

The unusual approach of defining variable range travel limits was investigated for CD4 tests, the most frequently needed tests, with high volumes of tests carried out annually by NHLS. For future work, this approach could also be applied for the less frequently used tertiary tests. Long-distance travel from a few rural hospitals affects the nationwide times within which tests can be brought to tertiary laboratories, in the same manner as for CD4 tests; the tertiary test samples are, however, less subject to damage during travel.

In view of uncertainty in the system caused by machine breakdowns and unexpected surges in demand, future studies could include exploration of approaches such as stochastic programming or robust optimisation.

Acknowledgements—The authors are grateful to the anonymous referees, whose constructive comments have led to considerable improvements to this paper. This research was partly funded by the UK Engineering and Physical Sciences Research Council: Science and Innovation Grant EP/F033982/1. Xpress-IVE was made available under academic licence to the University of Southampton. Our grateful thanks go to Jon Smith of Selective Analytics Ltd, for project support, particularly with geomapping.

References

Alumur SA, Yaman H, and Kara BY (2012). Hierarchical multimodal hub location problem with time-definite deliveries. *Transportation Research Part E, Logistics and Transportation Review* **48**(6):1107–1120.

Alumur S and Kara BY (2008). Network hub location problems: the state of the art. *European Journal of Operational Research* **180**(1):1–21.

Banerji SH and Fisher HB (1974). Hierarchical location analysis for integrated area planning in rural India. *Papers of the Regional Science Association* **33**(1):177–194.

Baray J and Cliquet G (2013). Optimizing locations through a maximum covering/p-median hierarchical model: maternity hospitals in France. *Journal of Business Research* **66**(1):127–132.

Campbell JF and O'Kelly ME (2012). Twenty-five years of hub location research. *Transportation Science* **46**(2):153–169.

Cassim N, Smith H, Coetzee LM and Glencross DK (2017). 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. *African Journal of Laboratory Medicine* **6**(1). doi:10.4102/ajlm.v6i1.545.

Çetiner S, Sepil C and Süral H (2010). Hubbing and routing in postal delivery systems. *Annals of Operations Research* **181**(1):109–124.

Chardaire P, Lutton J-L and Sutter A (1999). Upper and lower bounds for the two-level simple plant location problem. *Annals of Operations Research* **86**(0):117–140.

Church RL and Eaton DJ (1987). Hierarchical location analysis using covering objectives. In: Ghosh A and Rushton C (Eds.) *Spatial Analysis and Location-Allocation Models*. Van Nostrand Reinhold: New York, pp 163–185.

Church RL and ReVelle CS (1974). The maximal covering location problem. *Papers of the Regional Science Association* **32**(1):101–118.

Costa MDG, Captivo ME and Clímaco J (2008). Capacitated single allocation hub location problem—a bi-criteria approach. *Computers & Operations Research* **35**(11):3671–3695.

Daskin MS (2013). *Network and Discrete Location: Models, Algorithms, and Applications*, 2nd ed. Wiley: New Jersey.

Eiselt HA (2007). Locating landfills—optimization vs. reality. *European Journal of Operational Research* **179**(3):1040–1049.

Eitan Y, Narula SC and Tien JM (1991). A generalized approach to modeling the hierarchical location-allocation problem. *IEEE Transactions on Systems, Man and Cybernetics* **21**(1):39–46.

Farahani RZ, Hekmatfar M, Arabani AB and Nikbakhsh E (2013). Hub location problems: a review of models, classification, solution techniques, and applications. *Computers and Industrial Engineering* **64**(4):1096–1109.

Farahani RZ, Hekmatfar M, Fahimnia B and Kazemzadeh N (2014). Survey: hierarchical facility location problem: models, classifications, techniques, and applications. *Computers and Industrial Engineering* **68**(February):104–117.

Galvão RD, Espejo L and Boffey B (2002). A hierarchical model for the location of perinatal facilities in the municipality of Rio de Janeiro. *European Journal of Operational Research* **138**(3):495–517.

Goldman AJ (1969). Optimal location for centers in a network. *Transportation Science* **3**(4):352–360.

Hakimi SL (1964). Optimum locations of switching centers and the absolute centers and medians of a graph. *Operations Research* **12**(3):450–459.
Hakimi SL (1965). Optimum distribution of switching centers in a communication network and some related graph theoretic problems. *Operations Research* **13**(3):462–475.
Hinojosa Y, Puerto J and Fernandez FR (2000). A multiperiod two-echelon multicommodity capacitated plant location problem. *European Journal of Operational Research* **123**(2):271–291.
Hodgson MJ (1988). An hierarchical location-allocation model for primary health care delivery in a developing area. *Social Science and Medicine* **26**(1):153–161.
Ignacio AAV, Filho VJMF and Galvão RD (2008). Lower and upper bounds for a two-level hierarchical location problem in computer networks. *Computers & Operations Research* **35**(6):1982–1998.
Klotz E, Newman, AM (2013). Practical guidelines for solving difficult mixed integer linear programs. *Surveys in Operations Research and Management Science* **18**(1–2):18–32.
Lin CC (2010). The integrated secondary route network design model in the hierarchical hub-and-spoke network for dual express services. *International Journal of Production Economics* **123**(1):20–30.
Lowe TJ and Sim T (2013). The hub covering flow problem. *Journal of the Operational Research Society* **64**(7):973–981.
Marianov V and Serra D (2001). Hierarchical location–allocation models for congested systems. *European Journal of Operational Research* **135**(1):195–208.
Moore GC and ReVelle CS (1982). The hierarchical service location problem. *Management Science* **28**(7):775–780.
Narula SC (1984). Hierarchical location-allocation problems: a classification scheme. *European Journal of Operational Research* **15**(1):93–99.
NHLS (2016). National Health Laboratory Service Priority Programmes. http://www.nhls.ac.za/?page=priority_areas&id=11. Accessed 23 Aug 2016.
SA Government (2015). *South African Government*. About SA: Geography and Climate. http://www.gov.za/about-SA/geography-and-climate. Accessed 23 Oct 2015.
Şahin G and Süral H (2007). A review of hierarchical facility location models. *Computers & Operations Research* **34**(8):2310–2331.
Schultz GP (1970). The logic of health care facility planning. *Socio-Economic Planning Sciences* **4**(3):383–393.
Serra D (1996). The coherent covering problem. *Papers in Regional Science* **75**(1):79–101.
Smilowitz KR and Daganzo CF (2007). Continuum approximation techniques for the design of integrated package distribution systems. *Networks* **50**(3):183–196.
Smith HK, Harper PR and Potts CN (2013). Bicriteria efficiency/equity hierarchical location models for public service application. *Journal of the Operational Research Society* **64**(4):500–512.
Smith HK, Smith JP, Glencross DK, Cassim N, Coetzee LM, Carmona S and Stevens W (2017). Siting of HIV/AIDS diagnostic equipment in South Africa: a case study in locational analysis. *International Transactions in Operational Research* advance online publication 8 February. doi:10.1111/itor.12366.
Teixeira JC and Antunes AP (2008). A hierarchical location model for public facility planning. *European Journal of Operational Research* **185**(1):92–104.
Toregas C, Swain R, ReVelle C and Bergman L (1971). The location of emergency service facilities. *Operations Research* **19**(6):1363–1373.
Yaman H (2008). Star p-hub median problem with modular arc capacities. *Computers & Operations Research* **35**(9):3009–3019.
Yu J, Liu Y, Chang G-L, Ma W and Yang X (2009). Cluster-based hierarchical model for urban transit hub location planning: formulation, solution, and case study. *Transportation Research Record* **2112**(0):8–16.

Glossary

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 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.

Received 25 August 2016;
accepted 18 April 2017