OPEN

Community-based delivery of HIV treatment in Zambia: costs and outcomes

Brooke E. Nichols[a,b], Refiloe Cele[b,*], Lise Jamieson[b,*], Lawrence C. Long[a,b], Zumbe Siwale[c], Patrick Banda[d], Crispin Moyo[c] and Sydney Rosen[a,b]

Objective: The aim is to determine the total annual cost per patient treated and total cost per patient retained on antiretroviral therapy in Zambia in conventional care in facilities and across community-based differentiated service delivery (DSD) models.

Design: Economic evaluation was conducted using retrospective electronic record review.

Twenty healthcare facilities (13 with DSD models and 7 as comparison sites) in six of Zambia's 10 provinces were considered.

Methods: All individuals on antiretroviral therapy (ART) >18 years old at the study sites were enrolled in a DSD model or conventional care by site type, respectively, with at least 12 months of follow-up data. Accessing care through DSD models [community adherence groups (CAGs), urban adherence groups (UAGs), home ART delivery and care, and mobile ART services] or facility-based conventional care with 3-monthly visits. Total annual cost per patient treated and the annual cost per patient retained in care 12 months after model enrolment. Retention in care was defined as attending a clinic visit at 12 months ± 3 months.

Results: The DSD models assessed cost more per patient/year than conventional care. Costs ranged from an annual $116 to $199 for the DSD models, compared with $100 for conventional care. CAGs and UAGs increased retention by 2 and 14%, respectively. All DSD models cost more per patient retained at 12 months than conventional care. The CAG had the lowest cost/patient retained for DSD models ($140–157).

Conclusions: Although they achieve equal or improved retention in care, out-of-facility models of ART were more expensive than conventional care.

AIDS 2021, **35**:299–306

Keywords: community care, differentiated service delivery, economic evaluation, HIV, Zambia

Introduction

As of mid-2019, an estimated 1.2 million Zambians were HIV infected and, of these, 1.06 million were receiving antiretroviral therapy (ART), straining already overstretched existing healthcare infrastructure and human resources [1,2]. One way to reduce healthcare infrastructure strain and cope with the rising number of people on ART is to adopt differentiated service delivery (DSD) models for ART. Most existing DSD programs

[a]Department of Global Health, School of Public Health, Boston University, Boston, Massachusetts, USA, [b]Health Economics and Epidemiology Research Office, Department of Internal Medicine, School of Clinical Medicine, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa, [c]Right to Care Zambia, Lusaka, and [d]Ministry of Health, Lusaka, Zambia.

Correspondence to Brooke E. Nichols, 801 Massachusetts Ave., Crosstown Center 3rd Floor, Department of Global Health, Boston University School of Public Health, Boston, MA 02118, USA.

E-mail: brooken@bu.edu

* Refiloe Cele and Lise Jamieson contributed equally.

Received: 8 June 2020; revised: 21 August 2020; accepted: 2 September 2020.

DOI:10.1097/QAD.0000000000002737

ISSN 0269-9370

either (1) move services for stable ART patients away from overburdened healthcare facilities and into community-based locations, (2) reduce the number of provider interactions with patients, and/or (3) improve patient access by bringing the services closer to the patient and/or delivering ART refills in groups. DSD models often also shift tasks to lower cadres of providers, such as community health workers (CHW).

Early on in the national ART program, patients in Zambia were required to visit the health facility near-monthly for prescription refills [3]. Since 2016, multi-month dispensing—one form of differentiated care in which patients receive multiple months of antiretroviral medications at each visit—has been explicitly recommended in Zambia's national ART guidelines [4]. From 2016 until early 2019, 3-month ART dispensing was considered conventional care for patients regarded as stable on ART, with a 6-month dispensing policy not developed and implemented until mid-2019. In addition to the use of 3-month ART dispensing, Zambia introduced several other DSD models also aimed at reducing the healthcare system burden as numbers of ART patients continue to grow due to increased ART coverage [5].

Although a number of studies have evaluated the acceptability and clinical outcomes of DSD models [6,7], current evidence on the cost and cost-effectiveness of commonly implemented DSD models is limited [8]. There have been no published studies of patient-level healthcare resource utilization or provider costs for any DSD model of care in Zambia. Using primary data from patients enrolled in the five main models of care in use in Zambia in 2018, we estimated the annual cost of treatment, and cost per person retained in care, by model.

Methods

Differentiated models of care for antiretroviral therapy service delivery in Zambia

There were five main models of care in use in Zambia in 2018: (1) conventional care, (2) community adherence groups (CAGs), (3) mobile ART, (4) urban adherence groups (UAGs), and (5) home ART delivery (Table 1). With the exception of mobile ART, all of these models enrolled only stable ART patients. In Zambia, stable patients are generally defined according to the 2016 consolidated WHO guidelines [9].

Site and sample selection

Study sites were selected from a census of public sector ART facilities ($n = 1475$) [10] in Zambia based on the presence or absence of DSD models during the study period 1 January 2015–31 December 2017. Other criteria for site selection included: (1) at least 12 months of potential follow-up from model entry date for a minimum of 10 patients; (2) availability of SmartCare data (the Zambian national electronic medical record for patients on ART) at the site level for the duration of the follow-up period; and (3) availability of data on model entry dates and patient identification numbers, to match the SmartCare dataset. We selected 13 sites that met these criteria and enrolled in the study a census of all adult ($\geq$18 years old) patients at these sites who entered a DSD model between January 2015 and December 2017.

Comparison cohort

We identified a cohort of patients receiving conventional care matched 1 : 1 with patients in the DSD cohort described above. For this comparison cohort, we first chose a convenience sample of seven sites across four provinces that were not actively implementing DSD models and had SmartCare data readily available for the study period. In two of these four provinces, there were also DSD model sites included from the DSD cohort. We then matched stable patients from these sites with the DSD cohort on the basis of sex, age, urban/rural location, year of ART initiation (with up to a 1-year difference). The number of years on ART was calculated at each visit date for the conventional care cohort and matched to DSD model patients' model start date (within 1 year), relative to when DSD patients started ART. For these conventional care patients, we defined an 'equivalent model entry date' as the visit date of the matched control selected. Due to the lack of viral load measures, patients were identified as 'stable' through reported clinical characteristics. In this analysis, stable was therefore defined as any patient in WHO stage 0–2, and received no treatment for opportunistic infections. From patients who met matching criteria at each site, a corresponding conventional care patient was randomly selected for our analysis. To make this group comparable to the DSD cohort, we also required the possibility of a 12-month follow-up period for each patient. This period started at the 'equivalent model start date' matched and continued for the following 12 months.

Patient outcomes

The primary outcome in this analysis was 12-month retention, defined as the patient having a facility visit between 9 and 15 months after DSD model entry or after the equivalent model entry date for conventional care patients. All patients were due for a facility-based clinical visit 12 months after model entry, and thus could reasonably be expected to have a clinical visit recorded approximately 12 months after model entry. Due to large quantities of absent laboratory data, viral load and $CD4^+$ cell count outcomes could not be used as primary outcomes in this analysis.

Resource utilization

For each patient, we calculated quantities of resources utilized for the first 12 months after model entry date. For

Table 1. Differentiated models of ART service delivery evaluated.

Model	Design	Location of services and visit schedule	Other characteristics	Number of facility visits[a]	Number of DSD interactions[a]
Conventional care	Conventional service delivery, without differentiation.	Clinical visits, ART dispensation at health facility 3-monthly (stable patients), monthly (nonstable patients).		4	0
Community adherence groups (CAGs)	Group of ±6 people, based on residential proximity or patient preference, meet monthly at a designated place in the community. Members collect medication at clinical appointments for other CAG members, in a rotating fashion.	Monthly CAG meetings at community meeting places; twice yearly clinical appointments at the health facility and medication collection for other CAG members.	Patients considered 'stable' at model entry. Early implementation defined stability clinically (absence of clinical findings), viral load is used to determine stability where available.	2	12
Mobile ART services (mobile ART)	A mobile ART team comprised of medical professionals from a district hospital conduct biweekly visits to select rural health centers (RHCs) to provide ART services.	Patients visit the mobile ART services at a rural health center every 2 months.	Not limited to stable patients. Accepts advanced disease patients and newly initiated patients. Intended to serve individuals living far from a health facility that offered standard ART services.	0	6
Urban adherence groups (UAGs)	Group of 20–30 people. Patients receive group adherence counseling by a lay healthcare worker (HCW), followed by prepacked ART dispensation. are dispensed ART.	2–3 Monthly UAG meetings at health facility, generally outside normal clinic hours; twice yearly clinical appointments at health facility.	Patients considered 'stable' at model entry. Early implementation defined stability clinically (through an absence of clinical findings), viral load is used to determine stability where available.	2	4
Community HIV epidemic control (CHEC) model (home ART delivery)	Trained community health workers (CHWs) linked to facilities conduct home visits to deliver ART, conduct health screening, monitor adherence and refer patients as required. All community services are captured on a tablet-based SmartCare linked Community HTC or Community ART module.	Visits occur at the home monthly for the first three months and quarterly thereafter; there are once to twice-yearly clinical appointments at the health facility.	Patients considered 'stable' at model entry. Early implementation defined stability clinically (through an absence of clinical findings), viral load is used to determine stability where available.	1	6

DSD, differentiated service delivery.
[a]As per guidelines, not necessarily 'as implemented'.

facility visits, patient-level resource utilization data were sourced from the SmartCare database (e.g. visits by type, medications dispensed, laboratory tests). Average number of facility visits over the 12-month period are reported as pharmacy or nonpharmacy visits (i.e. clinical visit). These visit types were calculated separately as pharmacy visits occurred occasionally without a clinical visit. If ART was recorded as being dispensed in either the clinical or pharmacy record at a clinical visit, we included the cost of pharmacy dispensing.

Data on DSD interactions were not systematically collected for DSD models during the study period. We therefore inferred the number of DSD interactions based on the design of the DSD models and existing facility-based data. To do this, we modeled two scenarios: (1) full DSD use, in which we assumed that all DSD interactions called for by model guidelines actually occurred; (2) proportional DSD use, in which we assumed that DSD interactions occurred in proportion to the number of clinic visits adhered to in the 12-month period (See Table, Supplemental Digital Content 1, http://links.lww.com/QAD/B893, for a description of both scenarios). The methods to analyze each scenario represent different ways to consider and impute missing DSD-visit, related ART dispensing, and patient-level resource utilization. The two scenarios define what we think likely to be the plausible highest and lowest resource usage for the DSD models. These two scenarios apply to the CAG, UAG, and home ART delivery DSD models only. Resource utilization was fully documented in SmartCare for mobile ART and for conventional care.

Cost estimates

Costs were estimated from a provider perspective using bottom-up micro-costing techniques. Provider perspective was chosen given available data for this study, and can aid in decision making from the perspective of the payer. Unit costs for antiretroviral drugs and other medications recorded in pharmacy data were sourced from the Medical Stores Limited (2016) catalog (prices unchanged from 2016 to 2018). Costs of laboratory tests ($CD4^+$ cell count, viral load, creatinine, hemoglobin, full blood count, alanine transaminase, Sputum, rapid plasma regain syphilis testing) were derived from the CIDRZ Kalingalinga laboratory pricelist (2018 USD). All laboratory tests, ART, and non-ART medications reported to be used by a patient during the follow-up period of interest were included in the total annual cost per patient (see Text and Tables, Supplemental Digital Content 2–4, http://links.lww.com/QAD/B893, providing additional costing details).

To determine a unit cost per visit to a healthcare facility or per DSD interaction, we collected costs on all resources utilized to provide services, including for staff (management, administration, support, and direct service delivery), equipment, staff training, vehicle maintenance in the case of mobile ART, managerial oversight, and other operational needs (e.g. printing, airtime, travel/transport). Unit costs were calculated using the Healthcare Cost Outcomes Model (HCOM) [11]. The HCOM synthesizes the information on individual resource-use per visit (including staff time and consumables) and allocates fixed and shared costs (including equipment, shared staff, infrastructure, and overhead costs) across the health facility, and assigns the correct proportion of costs to ART patients based on the proportion of ART patient visits to non-ART patient visits.

Cost data were collected from implementing partners' expense records and from in-depth interviews with each implementing partner. Equipment costs were annualized at a discount rate of 5% over their expected useful life years [12]. Training costs were annualized according to the frequency of retraining staff or hiring new staff. Costs were collected in Zambian Kwacha (ZMW) and reflect 2018 market prices. For reporting purposes, costs were then converted to US dollars (USD) at an annual (January 2017–December 2018) average exchange rate of ZMW 10.00 per USD [13]. Costs are reported in 2018 USD.

Cost per outcome analysis

Unit costs were multiplied by resourced used at the patient level to determine the total monthly and annual cost of that patient accessing care. Means and standard deviations were then used to describe the costs of all patients accessing a respective model of care. The production cost of a successful outcome (one person retained in care) was calculated as the total costs of a model of care divided by the total number of patients retained at 12 months. Cost-outcomes analyses were conducted in SAS 9.4 (SAS Institute Inc., Cary, North Carolina, USA) through the combination of the unit costs developed in the HCOM and the resource-use data from the SmartCare database. We did not conduct a formal cost-effectiveness analysis because the underlying patient populations accessing each model differed, preventing the models from being regarded as practical substitutes for one another.

Sensitivity analysis

As the frequency of DSD interactions for CAGs, UAGs, and home ART delivery, and the quantities of antiretrovirals dispensed during these interactions were unknown, we undertook a probabilistic sensitivity analysis for these variables, using 10 000 Monte Carlo individual-level simulations. For each DSD model, we sampled for the number of DSD visits using triangular and uniform distributions (see Table, Supplemental Digital Content 5, http://links.lww.com/QAD/B893, with distribution parameters). While the triangular distribution allows us to make an assumption about the shape of the distribution, the uniform distribution is more conservative and allows for more variation in our sampling. The respective estimates sampled from the triangular and uniform distributions were then scaled by multiplying by *[(maximum − minimum) + minimum]* for both DSD visits and days of antiretrovirals dispensed. For DSD visits this minimum value is 0, whereas the maximum value corresponds to the design of the specific DSD model. For ART dispensed, the minimum is zero, whereas the maximum is equal to the balance of ART days that should be covered within the 12-month period, after accounting for the amount of ART dispensed at the facility. We assumed that 12-month retention (as per primary analysis) in each of the DSD models remained unchanged regardless of the number of DSD interactions assumed to have been made.

Ethical considerations

The study was approved by the Human Research Ethics Committee (Medical) of the University of the Witwatersrand, Johannesburg (M170967); the Boston University Institutional Review Board (H-36929); and ERES Converge (Zambia) (2017-Oct-001).

Results

Study population characteristics

We enrolled a total of 2506 patients from 20 healthcare facilities (13 with DSD models and seven as comparison sites) in six Zambian provinces (Table 2). Most patients (67–74%) were female, with a median age at ART initiation of 35–37 years. Patients enrolled in DSD models were on ART for an average of 4–6 years at the time of model entry.

Table 2. Demographic characteristics and 12-month retention in care by differentiated model of care.

Characteristic	Model of care				
	CAG	Mobile ART	UAG	Home ART delivery	Conventional care
Number analyzed	754	216	193	169	1174
Site type					
Urban	686 (91%)	0 (0%)	193 (100%)	52 (31%)	806 (69%)
Rural	68 (9%)	216 (100%)	0 (0%)	117 (69%)	368 (31%)
Sex, % female (*n*)	527 (70%)	139 (67%)	138 (72%)	125 (74%)	829 (71%)
Age at ART initiation (years)					
Median (IQR)	35 (30–41)	36 (27–45)	35 (30–41)	37 (31–45)	35 (30–42)
18–24	52 (7%)	30 (14%)	8 (4%)	10 (6%)	95 (8%)
≥25 years	694 (93%)	177 (86%)	185 (96%)	159 (94%)	1079 (92%)
Age at DSD model start (years)[a]					
Median (IQR)	41 (36–48)	36 (27–45)	41 (36–48)	42 (35–47)	40 (34–47)
18–24 years	19 (3%)	30 (14%)	1 (1%)	7 (4%)	47 (4%)
≥25 years	727 (97%)	177 (86%)	192 (99%)	162 (96%)	1127 (96%)
Years from ART initiation to model start date[a] (median (IQR))	6 (3–9)	0 (0–0)	6 (3–9)	4 (1–5)	4 (2–7)
Outcomes					
Retained at 12 months, *n* (%)	627 (83%)	148 (69%)	183 (95%)	134 (79%)	948 (81%)[b]

CAG, community adherence group; UAG, urban adherence group.
[a]For SOC stable this is at the equivalent model entry date.
[b]SOC stable outcomes calculated 12 months after at the equivalent model entry date, with a 90-day window period.

Outcomes

For stable conventional care patients, the 12-month retention rate was 81%. Within the DSD models, UAGs had the highest 12-month retention rate (95%), followed by CAGs (83%), home ART delivery (79%), and mobile ART (69%). The poor retention rate in the mobile ART model can be explained by the fact that this model enrolled patients at ART initiation. Since loss to follow-up is generally highest during the first few months on ART, high attrition from the mobile ART model was expected.

Cost per clinic visit and differentiated service delivery interaction

The average cost for a standard facility clinical follow-up visit was $3.65 (Fig. 1). The costs per CAG and UAG interaction were less, at $0.99 and $2.35, respectively. The low cost of the CAG interaction was due to the fact that CAGs are largely client-run, with minimal professional oversight. UAG interactions were more expensive as a clinic staff member was present at each UAG meeting and dispensed ART. The costs for a home ART delivery and a mobile ART visit were substantially higher than for a standard facility visit, at $12.21 and $9.38, respectively. The high cost of home ART delivery visit was driven largely by personnel and oversight. The mobile ART visit cost was driven largely by vehicle maintenance and personnel.

Total cost and cost per outcome

The mean cost per patient treated in conventional care during the 12-month follow-up period was $100. Under

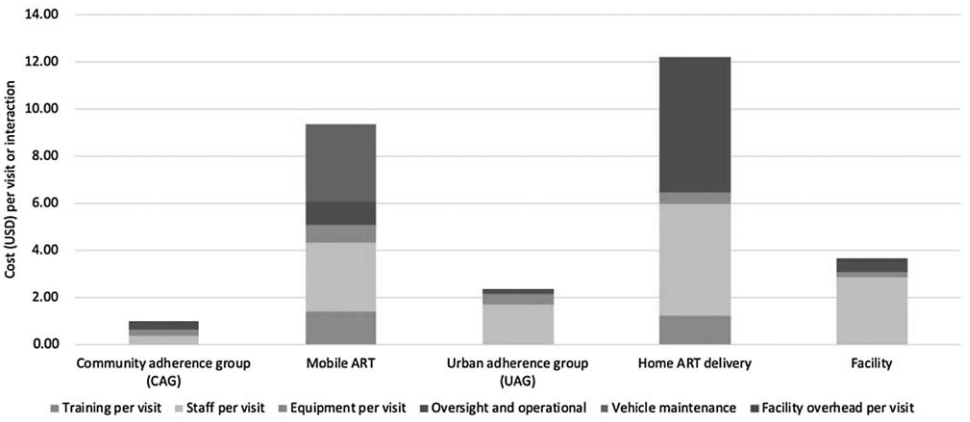

Fig. 1. Unit cost of clinic visits and differentiated service delivery (DSD) model interactions by cost categories.

Table 3. Total costs (2018 USD) of visits and tests conducted at facilities for 12 months, categorized by differentiated model of care.

	Model of care				
	CAG ($N=754$)	Mobile ART ($N=216$)	UAG ($N=193$)	Home ART delivery ($N=169$)	Conventional care ($N=1174$)
12-month retention, % (n)	627 (83%)	148 (69%)	183 (95%)	134 (79%)	948 (81%)
Scenario 1, full DSD use					
Mean annual costs (USD) (SD)					
Facility visits (clinical)	5.63 (5.15)	–	4.87 (3.54)	1.10 (2.55)	4.99 (3.27)
Pharmacy pick-ups	4.00 (2.49)	–	6.29 (1.70)	2.60 (1.55)	4.32 (1.91)
DSD interactions	11.93 (0.00)	45.71 (22.33)	11.26 (1.80)	51.44 (15.88)	–
Laboratory testing	6.92 (10.77)	–	23.24 (16.57)	4.56 (8.50)	4.61 (7.68)
Non-antiretroviral drugs	0.10 (0.64)	3.45 (32.03)	0.18 (0.53)	0.18 (0.50)	0.13 (0.53)
Antiretroviral drugs	101.21 (48.38)	73.30 (43.64)	114.66 (51.36)	126.63 (19.78)	87.96 (61.35)
Total cost (USD)	98 368	26 452	30 953	31 519	117 510
Mean (SD) cost per patient enrolled	130 (51.9)	122 (70.1)	160 (57.0)	186 (23.9)	100 (61.6)
Average cost per person-month	12	13	14	16	9
Production cost (total cost per each person retained)	157	179	169	235	124
Scenario 2, proportional DSD use					
Mean annual costs (USD) (SD)					
Facility visits (clinical)	5.63 (5.15)		4.87 (3.54)	1.10 (2.55)	4.99 (3.27)
Pharmacy pick-ups	4.00 (2.49)		6.29 (1.70)	2.60 (1.55)	4.32 (1.91)
DSD interactions	9.92 (0.00)	45.71 (22.33)	10.68 (1.71)	40.78 (12.59)	
Laboratory testing	6.92 (10.77)		23.24 (16.57)	4.56 (8.50)	4.61 (7.68)
Non-antiretroviral drugs	0.10 (0.64)	3.45 (32.03)	0.18 (0.53)	0.18 (0.50)	0.13 (0.53)
Antiretroviral drugs	89.01 (60.66)	73.30 (43.64)	101.87 (49.70)	87.96 (61.35)	86.04 (58.90)
Total cost (USD)	87 655	26 452	28 373	23 183	117 510
Mean (SD) cost per patient enrolled	116.25 (67.83)	122.46 (70.10)	147.01 (57.15)	137.18 (57.02)	100.09 (61.59)
Average cost per person-month	10	13	13	12	9
Production cost (total cost per each person retained)	140	179	155	173	124

ART, antiretroviral therapy; CAG, community adherence group; UAG, urban adherence group.

Scenario 1, full DSD use, the mean cost per patient treated was \$130, \$122, \$160, and \$186 for the CAG, mobile ART, UAG, and home ART delivery models, respectively (Table 3). Although mobile ART appeared to be the least costly of the DSD models per patient treated, this is likely an underestimate of the true cost of the mobile ART model, as all laboratory test data for this model were missing. The average cost per person per month spent in a DSD model ranged from \$12 to \$16, compared with \$9 in conventional care. The mean production cost per patient retained in care 12 months after model entry was \$124, \$157, \$180, \$169, and \$235 for conventional care, CAGs, mobile ART, UAGs, and home ART delivery, respectively (Table 3).

Under Scenario 2, proportional DSD use, the mean cost per patient treated was \$116, \$122, \$147, and \$137 for the CAG, mobile ART, UAG, and home ART delivery models respectively (Table 3); the mean production cost per patient retained in care 12 months after model entry was \$140, \$122, \$155, and \$173 for CAGs, mobile ART, UAGs, and home ART delivery. Despite the smaller number of DSD interactions reflected in Scenario 2, costs for all DSD models were higher than those for conventional care (\$100/patient treated; \$124 production cost). Total costs stratified by those retained for a full 12 months and for <12 months are reported for both scenarios in Supplemental Digital Content 6–7, http://links.lww.com/QAD/B893.

Sensitivity analysis

In the event that our two primary scenarios incorrectly estimated the mean number of DSD interactions that occurred, we conducted a probabilistic sensitivity analysis at the individual level. All 10 000 Monte Carlo simulations resulted in a cost per person retained in CAGs, UAGs, and home ART delivery (triangular distribution: \$154, \$204, and \$179, respectively; uniform distribution: \$148, \$186, and \$171, respectively) to be greater than that for stable conventional care patients (\$124), with none of the simulations resulting in a cost less than the cost of conventional care. There was no overlap between the DSD models (See Table, Supplemental Digital Content 8, http://links.lww.com/QAD/B893).

When sampling from this wide range of patient-level DSD interactions, we found it highly unlikely that our conclusions would change with more information on DSD interactions.

Discussion

Although it is widely expected that DSD implementation will result in cost savings for providers, all models evaluated in this study were more expensive than conventional care per patient retained in care. Conventional clinic-based care with three-month dispensing was

the low-cost option for public sector facilities in Zambia, when considered solely from the provider's perspective. Although these models of care are more costly from the provider perspective, it is possible that even those that do not improve patient outcomes are preferable for patients and/or reduce patient-level costs. The full range of DSD model outcomes, and not solely provider costs, should be considered when planning the wide-scale implementation of any DSD model.

While perhaps unwelcome, the finding that community-based DSD models in Zambia cost more than conventional care is not surprising. Each of the DSD models considered here was designed to invest more resources in patient care through additional interactions with the health system, rather than fewer, making conventional care the least resource-intensive model. Strategies for reducing operating costs of these models, or reducing patient touch points, could be explored as one opportunity to reduce the cost per patient enrolled in these models. Since costs for medications and laboratory tests do not vary by model, it is unlikely that DSD models, as currently designed, will be cost-saving.

These are among the first empirical estimates of the costs of managing ART through DSD models in sub-Saharan Africa. One prior study conducted in Malawi found that DSD models cost less than conventional care, but this analysis was an estimate of resource utilization based on guidelines, not on resource utilization from patient records [14]. Another study of adherence clubs in South Africa (similar to UAGs in Zambia) also found that adherence clubs are cost-saving compared with conventional care. In this instance, the clubs dispensed medications to a group of 25–30 people at one time, and ART was provided by a lower cadre of staff at adherence club meetings compared with standard individual appointments [15]. Broadly, the cadre of staff implementing a model of care and integration of the model of care within the required clinical visits drives whether or not these models are likely to be cost-saving. The only study we identified that utilized an empirical costing approach to assess the full annual cost of a streamlined model of care in Uganda and Kenya did not report a comparison group, limiting comparability to our study [16].

This study had several limitations. First, our analysis does not include all models of care that are currently being implemented in Zambia as of 2019. By mid-2019, three additional models were introduced in Zambia: 6-month ART dispensing, fast-track refills, and chronic centralized medicine dispensing and delivery. These new models were not evaluated as part of this analysis as they had not yet accrued sufficient follow-up time to observe outcomes or estimate costs. Less resource-intensive facility-based models of care, such as fast-track ART dispensing or 6-month ART scripting and dispensing, may be better positioned to reduce overall costs. Future work should aim to empirically compare the costs and outcomes of those less resource-intensive models to the models described in this manuscript.

Second, while each DSD model called for a specific number of clinic visits and/or interactions, the study team's experience at facilities offering DSD models in Zambia suggests that fidelity to guidelines is inconsistent. Instead, facilities and patients interpret the guidelines based on their own resources and preferences. The actual costs estimated here thus do not reflect consistent, compliant implementation of the DSD models, but rather what is happening on the ground. This is not a limitation per se, but it should be kept in mind in interpreting our results.

Third, some data were missing nonuniformly across models. In particular, all laboratory values were missing from those receiving care in the mobile ART model of care. Including these costs would only further increase the cost of this model of care, and mobile ART would remain one of the two most costly models of care. The absence of laboratory results for most participants prevented using viral suppression or $CD4^+$ cell gain as outcomes.

Fourth, estimating the true level of effort of provider time supporting DSD models can be difficult when multiple tasks are being completed concurrently. These models of care are still likely to cost more than conventional care given the overall increase in interactions with the health system regardless of routine under/overestimated staff costs due to imprecise person-time quantification of. Fifth, this study is limited by the missingness of DSD interaction resource utilization that may influence the cost per person and cost per person retained point estimates. Though the point estimate may have been affected by this missingness, the results of our comprehensive sensitivity analysis demonstrate that this missingness does not affect our primary conclusions. Finally, the majority of individuals that entered into these DSD models had initiated ART several years prior to model entry. It is possible that patient outcomes in particular—retention at 12 months post entry into a DSD model—may differ in a cohort of individuals who started ART more recently.

In conclusion, we find that out-of-facility models of ART delivery that do not reduce the number of healthcare system interactions made by patients should be expected to be more expensive than conventional, facility-based care. To account for the full health system impact of DSD models, future studies should focus on the simultaneous evaluation of multiple outcomes of DSD implementation; consequences for the full population of ART patients, and not just those eligible for DSD models; and on the newer facility-based models of care currently implemented in Zambia, such as fast-track ART refills and 6-month ART scripting and dispensing.

Acknowledgements

This work was funded by the United States Agency for International Development through the following cooperative agreement: AID-OAA-A-15-00070.

Conflicts of interest

The authors declare no conflicts of interest.

References

1. Ministry of Health Zambia. *District Health Information System 2 (DHIS2)*. Lusaka, Zambia: Ministry of Health Zambia. 2019.
2. UNAIDS. Country: Zambia. 2020. https://www.unaids.org/en/regionscountries/countries/zambia. [Accessed 9 August 2020]
3. Scott CA, Iyer HS, McCoy K, Moyo C, Long L, Larson BA, *et al*. **Retention in care, resource utilization, and costs for adults receiving antiretroviral therapy in Zambia: a retrospective cohort study.** *BMC Public Health* 2014; **14**:29611 pages.
4. Republic of Zambia Ministry of Health. Zambia consolidated guidelines for treatment & prevention of HIV infection. Lusaka, Zambia: Ministry of Health; 2016.
5. Ministry of Health. Zambia: Differentiated service delivery framework. Lusaka, Zambia: Ministry of Health; 2018.
6. Kuchukhidze S, Long LC, Pascoe S, Huber AN, Nichols BE, Fox MP, *et al*. *Differentiated models of service delivery in sub-Saharan Africa: A review of the gray literature as of June 2019*. Boston: Boston Univerity and HERO; 2019.
7. Long L, Kuchukhidze S, Pascoe S, Nichols B, Cele R, Govathson C, *et al*. *Differentiated service delivery models for antiretroviral treatment of HIV in sub-Saharan Africa: A rapid systematic review*. Boston: Boston University and HERO; 2020.
8. Kuchukhidze S, Long LC, Rosen S, Pascoe S, Huber A, Fox M, *et al*. *Provider costs associated with differentiated models of service delivery for HIV treatment in sub-Saharan Africa*. Boston: Boston University and HERO; 2019.
9. WHO. *Consolidated guidelines for the use antiretroviral drugs for treating and preventing HIV infection: recommendations for a public health approach*. 2nd ed. Geneva: WHO; 2016.
10. Nichols BE, Girdwood SJ, Shibemba A, Sikota S, Gill CJ, Mwananyanda L, *et al*. **Cost and impact of dried blood spot versus plasma separation card for scale-up of viral load testing in resource limited settings.** *Clin Infect Dis* 2020; **70**:1014–1020.
11. Long L, Rosen S, Girdwood SJ. The Healthcare Costs and Outcomes Model (HCOM). 2017. http://www.heroza.org/research-tools/the-healthcare-cost-and-outcomes-model-hcom/ [Accessed 5 March 2020].
12. South African Revenue Service, editor. Binding general ruling (income tax) 7 (Issue 3): Wear-and-tear or depreciation allowance. South Africa: South African Revenue Service; 2020.
13. Oanda. Oanda Currency Converter. 2020. www.oanda.com [Accessed 20 July 2020].
14. Prust ML, Banda CK, Nyirenda R, Chimbwandira F, Kalua T, Jahn A, *et al*. **Multimonth prescriptions, fast-track refills, and community ART groups: results from a process evaluation in Malawi on using differentiated models of care to achieve national HIV treatment goals.** *J Int AIDS Soc* 2017; **20 (Suppl 4)**:41–50.
15. Bango F, Ashmore J, Wilkinson L, van Cutsem G, Cleary S. **Adherence clubs for long-term provision of antiretroviral therapy: cost-effectiveness and access analysis from Khayelitsha, South Africa.** *Trop Med Int Health* 2016; **21**:1115–1123.
16. Shade SB, Osmand T, Luo A, Aine R, Assurah E, Mwebaza B, *et al*. **Costs of streamlined HIV care delivery in rural Ugandan and Kenyan clinics in the SEARCH Studys.** *AIDS* 2018; **32**:2179–2188.