Role of performance-based financing for HIV/AIDS control: a systematic review and meta-analysis

Introduction

HIV remains a leading cause of death and disability globally [1]. In 2013 there were an estimated 2.1 million new HIV infections, 1.5 million AIDS deaths, and 35 million people living with HIV [2]. Countries have made substantial progress in controlling the epidemic, with an estimated 12.9 million people receiving antiretroviral therapy (ART) at the end of 2013 [3]. Nonetheless, global resources to control the epidemic have not increased since 2011 [2]. Innovative health systems approaches may help countries achieve more with the same, or less, resources.

The WHO building blocks of health systems, including governance; financing; workforce; health commodities; strategic information; and service delivery, help ensure access to quality health services [4]. Health financing considers how domestic and external financing is mobilised, pooled, and strategically invested within health systems. Traditionally, health financing is allocated based on inputs. For example, governments may allocate resources based on the number of commodities needed, salary of health staff, and other logistical costs. In recent years, financing health services based on outputs (i.e. performance-based financing) has been implemented in a number of countries for potential increases in efficiency [5-10].

Previous systematic reviews have reviewed performance-based financing in a range of settings and programmes [11-13]. Since this time, evidence has emerged on performance-based financing for HIV services. We will systematically review this evidence to inform WHO guidelines.

Methods

Conduct

The systematic review will be conducted in accordance with PRISMA guidelines [14]. The PubMed and WHO Index Medicus databases will be systematically searched without language, publication, date, or any other limits (Appendix). Databases from the International AIDS Society, Conference on Retroviruses and Opportunistic Infections, and HIV/AIDS Implementers' Meeting will also be searched. The WHO International Clinical Trials Registry Platform, the Cochrane Central Register of Controlled Trials, the International Standard Randomised Controlled Trial Number Register, and ClinicalTrials.gov will be searched for future and on-going studies. Experts in the field will be contacted to identify unpublished research and on-going studies.

Selection criteria

Per recommendations from the PRISMA group, eligibility criteria will be based on key study characteristics: population, intervention, comparator, outcome, and design [14]. Specifically, studies will be included when (1) the study population was composed of people with HIV (infants, children, adolescents, or adults); (2) the intervention was performance-based financing (i.e. any program that rewards the health system for delivery of one or more outputs or outcomes by one or more incentives, financial or otherwise, upon verification that the agreed-upon result has actually been delivered [15]); (3) the comparator was no performance-based financing; (4) the outcomes were quality (including retention, viral suppression, adherence to national standard of care, patient satisfaction, patient centeredness, or harm), access (including testing uptake, testing coverage, linkage to care, treatment uptake, or treatment coverage), or cost of HIV services; and (5) the study design was a randomised trial or comparative contemporaneous or time series study. Importantly, performance-based financing focuses on health systems incentives rather than incentives to individuals or recipients of health care such as patients. Therefore, articles focusing exclusively on incentives for recipients of HIV services will not be included. Two of the team members will independently screen abstracts of all identified articles and then match the full texts of all articles selected during screening against the inclusion criteria. Articles meeting the inclusion criteria will be included in the reviews.

Data extraction

Two investigators will complete the data extraction using a standardised extraction spread-sheet comprising four tables. The first table will summarise the characteristics of study participants. The second table will summarise study methods. The third table will summarise the reported outcomes. The final table will focus on quality assessment.

Quality assessment

For the quality assessment, studies will be stratified based on study design (i.e. randomised controlled trial or observational study). Per recommendations from the Cochrane Collaboration, the Collaboration's 'Risk of bias' tool will be used to assess bias in randomised trials [16]. This tool rates studies based on four sources of bias: selection bias, performance and detection bias, attrition bias and reporting bias. Per recommendations from the Cochrane Collaboration [16], the *Newcastle-Ottawa Quality Assessment Scale* will be

used to assess bias in observational studies [17]. This scale rates studies based on eight criteria in three sources of bias: selection bias, confounding and measurement bias. The quality of evidence will be assessed using the *Grades of Recommendation, Assessment, Development, and Evaluation* (GRADE) system (Appendix, [18]).

Statistical analyses

If studies are similar enough to combine after stratification, meta-analyses will be performed and statistical heterogeneity will be assessed. Depending on the number of studies available, subanalyses may be stratified by outcome and/or performance-based financing approach. Random-effects models will be used for all analyses [19]. An I-squared statistic will be used to measure heterogeneity in the magnitude of pooled estimates [20]. If there is moderate to significant heterogeneity in estimates, potential causes may be explored using sensitivity analyses. If there are at least ten studies meeting eligibility criteria for meta-analyses, a funnel plot with the effect measures on the *x*-axis and standard error of the log for the effect measures on the *y*-axis will be created to assess publication bias and the Egger and Begg tests will be used to test the funnel plot's asymmetry. STATA version 13.0 will be used for all analyses.

REFERENCES

1. Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet. 2015; **385**(9963): 117-71.
2. Joint United Nations Programme on HIV and AIDS. The Gap Report. 2014 [cited 10 April 2015]; Available from:
3. World Health Organization. Global update on the health sector response to HIV, 2014. 2014 [cited 4 June 2015]; Available from: http://www.who.int/iris/bitstream/10665/128494/1/9789241507585_eng.pdf
4. World Health Organization. Monitoring the building blocks of health systems. 2010 [cited 4 June 2015]; Available from: http://www.who.int/healthinfo/systems/WHO_MBHSS_2010_full_web.pdf
5. Eichler R, Auxila P, Pollock J. Output-based health care : paying for performance in Haiti. Public Policy for the Private Sector. 2001; **236**: 1-4.
6. Hecht R, Batson A, Brenzel L. Making health care accountable. Finance and Development. 2004; **41**(1): 16-9.
7. Soeters R, Griffiths F. Improving government health services through contract management: a case from Cambodia. Health Policy Plan. 2003; **18**(1): 74-83.
8. Low-Beer D, Afkhami H, Komatsu R, Banati P, Sempala M, Katz I, et al. Making performance-based funding work for health. PLoS Med. 2007; **4**(8): e219.
9. Basinga P, Gertler PJ, Binagwaho A, Soucat AL, Sturdy J, Vermeersch CM. Effect on maternal and child health services in Rwanda of payment to primary health-care providers for performance: an impact evaluation. Lancet. 2011; **377**(9775): 1421-8.
10. World Health Organization. Performance-based grants for reproductive health in the Philippines. 2011 [cited 4 June 2015]; Available from: http://whqlibdoc.who.int/hq/2011/WHO_RHR_11.04_eng.pdf
11. Eldridge C, Palmer N. Performance-based payment: some reflections on the discourse, evidence and unanswered questions. Health Policy Plan. 2009; **24**(3): 160-6.
12. Oxman AD, Fretheim A. Can paying for results help to achieve the Millennium Development Goals? Overview of the effectiveness of results-based financing. J Evid Based Med. 2009; **2**(2): 70-83.
13. Houle SK, McAlister FA, Jackevicius CA, Chuck AW, Tsuyuki RT. Does performance-based remuneration for individual health care practitioners affect patient care?: a systematic review. Ann Intern Med. 2012; **157**(12): 889-99.
14. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med. 2009; **6**(7): e1000100.
15. World Health Organization. Technical update on HIV incidence assays for surveillance and monitoring purposes. 2015 [cited 11 April 2015]; Available from: http://www.unaids.org/sites/default/files/media_asset/HIVincidenceassayssurveillancemonitoring_en.pdf
16. The Cochrane Collaboration. Cochrane Handbook for Systematic Reviews of Interventions. [cited 19 April 2011]; Available from: http://www.cochrane-handbook.org/
17. Wells G, Shea B, O'Connell D, Peterson J, Welch V, Losos M, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. [cited 23 February 2011]; Available from: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp

18. Schünemann H, Brozek J, Oxman A, editors. GRADE Handbook: for grading the quality of evidence and the strength of recommendations. 2009 [cited 29 March 2012]; Available from: www.who.int/hiv/topics/mtct/grade_handbook.pdf
19. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986; **7**(3): 177-88.
20. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002; **21**(11): 1539-58.

APPENDIX

Search strategy for MEDLINE

(HIV Infections[MeSH] OR HIV[MeSH] OR hiv[tiab] OR hiv-1*[tiab] OR hiv-2*[tiab] OR hiv1[tiab] OR hiv2[tiab] OR hiv infect*[tiab] OR "human immunodeficiency virus"[tiab] OR "human immunedeficiency virus"[tiab] OR "human immuno-deficiency virus"[tiab] OR "human immune-deficiency virus"[tiab] OR ((human immun*[tiab]) AND (deficiency virus[tiab])) OR AIDS[tiab] OR "acquired immunodeficiency syndrome"[tiab] OR "acquired immunedeficiency syndrome"[tiab] OR "acquired immuno-deficiency syndrome"[tiab] OR "acquired immune-deficiency syndrome"[tiab] OR ((acquired immun*[tiab]) AND (deficiency syndrome[tiab])) OR "sexually transmitted diseases, Viral"[MeSH:NoExp]) AND (performance[tiab] OR results[tiab] OR output[tiab] OR delivery[tiab] OR conditional[tiab] OR contract*[tiab]) AND (financ*[tiab] OR subsid*[tiab] OR remunerat*[tiab] OR pay*[tiab] OR incentive[tiab] OR cash[tiab])

Search strategy for other databases

(hiv OR hiv-1* OR hiv-2* OR hiv1 OR hiv2 OR hiv infect* OR "human immunodeficiency virus" OR "human immunedeficiency virus" OR "human immuno-deficiency virus" OR "human immune-deficiency virus" OR ((human immun*) AND (deficiency virus)) OR AIDS OR "acquired immunodeficiency syndrome" OR "acquired immunedeficiency syndrome" OR "acquired immuno-deficiency syndrome" OR "acquired immune-deficiency syndrome" OR ((acquired immun*) AND (deficiency syndrome)) AND (performance OR results OR output OR delivery OR conditional OR contract*) AND (financ* OR subsid* OR remunerat* OR pay* OR incentive OR cash)

GRADE

For systematic reviews, the GRADE approach defines the quality of a body of evidence as the extent to which one can be confident that an estimate of effect or association is close to the quantity of specific interest. The quality rating across studies has four levels: high, moderate, low, or very low. High quality indicates that further research is very unlikely to change our confidence in the estimate of effect. Moderate quality indicates that further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate. Low quality indicates that further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate. Very low quality indicates that any estimate of effect is very uncertain. By default randomised trials are categorised as high quality and can be downgraded while cohort studies are categorised as low quality and can be upgraded or downgraded. The GRADE Profiler software will be used for performing the GRADE assessment (GRADEprofiler version 3.2.2).

There are five factors that can decrease the quality of a body of evidence. The first factor is major limitations in study design or execution that are likely to result in a biased assessment of the effect estimate. This factor will be gauged by assessing the risk of bias across studies. When the proportion of information from studies at high risk of bias is sufficient to affect the interpretation of results, the risk of bias across studies will be ‘high’. When most data included in the GRADE review is insufficient to affect the interpretation of results, the risk of bias across studies will be ‘low’. When most information included in the GRADE review is from studies at an unclear risk of bias, the risk of bias across studies will be ‘unclear’. ‘Low’ risk of bias will indicate ‘no limitation’, an ‘unclear’ risk of bias will indicate ‘no limitation’ or ‘serious limitation’, and a ‘high’ risk of bias will indicate ‘serious limitation’ or ‘very serious limitation.’ The second factor that can decrease the quality of a body of evidence is indirectness of evidence. Indirectness of evidence refers to bodies of literature that do not correspond to the population, intervention, comparator, and outcome specified in the inclusion criteria. The third factor that can decrease the quality of a body of evidence is inconsistency of study results. This would primarily be when studies yield widely different estimates of effect in terms of heterogeneity or variability in results. The fourth factor that can decrease the quality of a body of evidence is imprecision of results, i.e. when there are few participants, few events, and wide confidence intervals. The fifth and final factor that can decrease the quality of a body of evidence is high probability of publication bias. This would be when investigators fail to publish studies or outcomes on the basis of their results.

There are three factors that can increase the quality level of a body of evidence. The first factor is a large magnitude of effect. In the absence of plausible confounders, a large effect (i.e. RR > 2 or RR < 0.5) increases the quality one level while a very large effect (i.e. RR > 5 or RR < 0.2) increases the quality two levels. The second factor is plausible confounding that reduces the effect demonstrated in the included studies. The third factor is the presence of a dose-response gradient.