

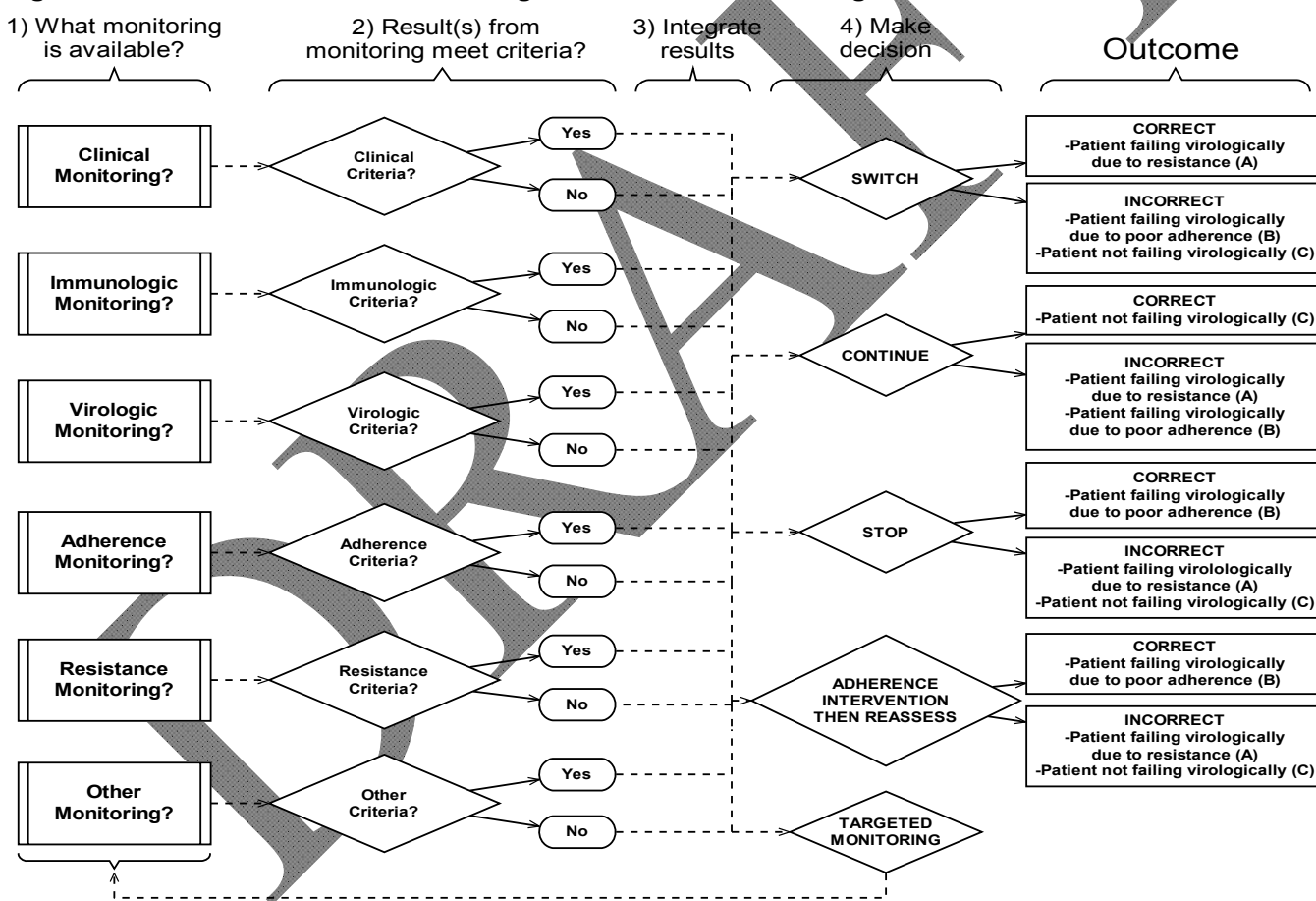
UCSF/WHO Systematic Review

Monitoring strategies for guiding when to switch first-line antiretroviral therapy regimens for treatment failure among adults and adolescents living with HIV in low-resource settings

BACKGROUND

One of the critical clinical decisions made in antiretroviral therapy (ART) is the decision on when to switch from one regimen to another for treatment failure. This is a complex decision which requires consideration of multiple factors including: (1) what type of monitoring (e.g. immunologic, virologic) is available to provide information to guide switching, (2) establishing criteria for treatment failure (e.g. viral load >10,000 copies/ml), (3) deciding how to integrate criteria (e.g. clinical and immunologic), and (4) making a decision and, if possible, follow-up and monitoring to determine outcomes (see Figure 1 and Appendix 1).

Figure 1: Decision tree for determining when to switch ART regimens.



The initial step in this model of deciding when to switch is determining what type of monitoring for guiding when to switch is available and appropriate. This review will seek to find and summarize evidence on monitoring strategies for guiding when to switch first-line regimens among adults and adolescents living with HIV in low-resource settings to assist with updating the 2006 WHO ART guidelines.¹

OBJECTIVE

To assess different monitoring strategies for guiding when to switch antiretroviral therapy (ART) regimens for first-line treatment failure among adults and adolescents living with HIV in low-resource settings.

METHODS

We followed standard Cochrane systematic review methodology and created standard GRADE evidence profiles.

Study inclusion criteria

- Study must evaluate a monitoring intervention/strategy that helps guide when to switch ART
- Randomised controlled trials
- Observational studies (cohort and case-control) which included comparators
- Systematic reviews and meta-analyses
- Cost-effectiveness/modelling studies were retained as articles of interest to help contextualize evidence with the intent to summarize these in narrative format only.

Study exclusion criteria

- Letter, editorial, non-systematic review, observational studies without comparators, case report, cross-sectional study design or descriptive studies
- Studies evaluating ART in patients failing more than 1 regimen
- Studies evaluating substituting rather than switching ART (as described in the WHO 2006 guidelines, substituting is for toxicities and usually involves single drug changes while switch is due to clinical, immunologic or virologic failure).

Population

- Adolescents and adults living with HIV failing or suspected of failing a first-line antiretroviral therapy regimen

Types of interventions and comparisons

This review looked for all studies with any potential comparisons of monitoring strategies (Table 1). However, comparisons of particular interest as originally defined in PICO (Population, Intervention, Comparator, Outcome format used to frame the question of interest) tables by organizers' consensus included clinical vs virologic, immunologic vs virologic, clinical + immunologic vs virologic, effect of different virologic threshold criteria for failure (e.g. virologic monitoring where criteria for failure was a VL < 10,000/ml vs. virologic criteria with VL ≥ 10,000 copies/ml), and role of adherence monitoring in comparisons.

Table 1: Potential comparisons of interest for guiding when to switch.

Switch guided by...	
Monitoring Strategy A vs.	Monitoring Strategy B
Clinical ± Immunologic ± Virologic ± Adherence ± Other Monitoring ^b	Clinical ± Immunologic ± Virologic ± Adherence ± Other Monitoring

^a While other potential monitoring strategies were postulated including other laboratory tests such as hemoglobin, targeted viral load monitoring, resistance monitoring, etc, this category was deliberately left undefined so that the review would be inclusive.

Outcomes of interest

Primary Outcomes (Critical)

1) Mortality

2) Morbidity-CDC AIDS-defining illness and/or WHO Stage IV

3) Virologic failure-pre-defined concentration of HIV-1 RNA, typically >400 or 500 copies/mL

- 4) **Immunologic response**-geometric mean or median increase in CD4 cell count from baseline
- 5) **Unnecessary switch**-participant switched while virologically suppressed
- 6) **Severe adverse events**-severe adverse events were classified according to grade 1 to 4 of the Adverse Event Toxicity Scale (Division of AIDS, NIH) where possible and reported as the proportion of participants that experienced grade 3 and 4 clinical and laboratory adverse events. Using this scale, grade 1 and 2 denote mild to moderate symptoms, grade 3 denotes serious symptoms and grade 4 denotes life-threatening events requiring significant clinical intervention.

Secondary Outcomes (Important)

- 1) **Development of antiretroviral resistance**-genotypic resistance only
- 2) **Missed virologic failure**-participant not identified while virologically failing
- 3) **Switch to second-line**

Search Methods

We used the HIV/AIDS Cochrane Collaborative Review Group search strategy. We formulated a comprehensive and exhaustive search strategy in an attempt to identify all relevant studies regardless of language or publication status (published, unpublished, in press or in progress). Full details of the Cochrane HIV/AIDS Review Group methods and the journals hand-searched are published in *The Cochrane Library* in the section on Collaborative Review Groups (<http://www.mrw.interscience.wiley.com/cochrane/clabout/articles/HIV/frame.html>). We combined the randomized controlled trial (RCT) strategy developed by The Cochrane Collaboration and detailed in the Cochrane Reviewers' Handbook in combination with terms specific to treatment failure and switching antiretroviral therapy regimens.²

Limits. The searches were performed without limits to language, setting or age. The searches were limited to human studies published from 1995 (designated start of combination ART era) to July 2009.

Electronic searches. We searched the following electronic databases:

Journal and trial databases

- MEDLINE
- EMBASE
- CENTRAL (Cochrane Central Register of Controlled Trials)

Conference databases

- NLM Gateway (for HIV/AIDS conference abstracts before 2005)
- We hand searched conference proceedings from the Conferences on Retroviruses and Opportunistic Infections, International AIDS Conferences and International AIDS Society Conferences on HIV Pathogenesis, Treatment, and Prevention from 2005 to 2009,

Clinical trials databases

- ClinicalTrials.gov (<http://clinicaltrials.gov/>)
- Current Controlled Trials (www.controlled-trials.com/)
- Pan-African Clinical Trials Registry (www.pactr.org)

Researchers and relevant organizations. We contacted individual researchers working in the field, such as the AIDS Clinical Trials Group, and policymakers based in inter-governmental organizations including the Joint United Nations Programme on HIV/AIDS (UNAIDS) and WHO to identify trials either completed or ongoing.

Reference lists. We checked the reference lists of all studies identified by the above methods and examined the bibliographies of any relevant systematic reviews, meta-analyses, or current guidelines we identified during the search process.

Search Terms

The following search string was used for MEDLINE:

Search #13 AND #17 AND #21 NOT (animals [mh] NOT human [mh]) Limits: Publication Date from 1995 to 2009
Search TREATMENT FAILURE
Search #14 OR #15 OR #16
Search (DRUG MONITORING) OR (THERAPEUTIC DRUG MONITORING)
Search VIRAL LOAD
Search CD4 LYMPHOCYTE COUNT
Search #11 AND #12
Search Antiretroviral Therapy, Highly Active[MeSH] OR Anti-Retroviral Agents[MeSH] OR Antiviral Agents[MeSH:NoExp] OR ((anti) AND (hiv[tw])) OR antiretroviral*[tw] OR ((anti) AND (retroviral*[tw])) OR HAART[tw] OR ((anti) AND (acquired immunodeficiency[tw])) OR ((anti) AND (acquired immunodeficiency[tw])) OR ((anti) AND (acquired immuno-deficiency[tw])) OR ((anti) AND (acquired immunodeficiency[tw])) OR ((anti) AND (acquired immun*) AND (deficiency[tw]))
Search HIV Infections[MeSH] OR HIV[MeSH] OR hiv[tw] OR hiv-1*[tw] OR hiv-2*[tw] OR hiv1[tw] OR hiv2[tw] OR hiv infect*[tw] OR human immunodeficiency virus[tw] OR human immunodeficiency virus[tw] OR human immuno-deficiency virus[tw] OR human immune-deficiency virus[tw] OR ((human immun*) AND (deficiency virus[tw])) OR acquired immunodeficiency syndrome[tw] OR acquired immunodeficiency syndrome[tw] OR acquired immuno-deficiency syndrome[tw] OR ((acquired immun*) AND (deficiency syndrome[tw])) OR "sexually transmitted diseases, viral"[MH]

We used a similar strategy with minor modifications for EMBASE (1995-2009), the Cochrane Controlled Trials Register, which contains mainly reference information to randomised controlled trials and controlled clinical trials in health care, and Gateway, a service of the U.S. National Library of Medicine, which contains conference abstracts. Keywords used for conference abstract searching and clinical trials database searching included the following: treatment failure, switch, and monitoring. There was overlap between the references retrieved in each database. All searches were conducted during July 2009.

Data extraction and coding

After initial search and article screening, two reviewers (LC and JH) independently double-coded and entered onto a detailed and standardized data extraction form information from selected studies.

Extracted information included:

Study details: citation, start and end dates, location, study design and details

Participant details: study population eligibility (inclusion and exclusion) criteria, ages, population size, attrition rate, details of HIV diagnosis and disease and any clinical, immunologic or virologic staging or lab information, 1st line drug regimen details including drug name, dose and duration, 2nd line drug regimen details including drug name, dose and duration

Interventions details: Clinical, immunologic, virologic and/or other monitoring strategy for switching regimen, frequency of monitoring

Outcome details: mortality, clinical disease progression (AIDS and non-AIDS events), treatment response (CD4 recovery and viral load response), adherence, resistance, adverse events, frequency of switching, time to switch, if patient not switched when meeting criteria what was done (e.g. adherence intervention), unnecessary switches, missed failures.

Data analysis and presentation of findings

We used Review Manager 5 provided by Cochrane Collaboration for statistical analysis and GradePro software to produce Grade Evidence Profile tables.^{3,4}

Dichotomous outcomes for effect were summarized in terms of risk ratios (RR) and numbers needed to treat (NNT) with 95% confidence intervals. Time to event outcomes were summarized in terms of hazard ratios (HR) with 95% confidence intervals. Observational studies and randomized trials were evaluated separately.

Heterogeneity among trials was examined using the chi-square statistic with a significance level of 0.10 and the I-squared statistic. The I-squared estimate greater than 50% was interpreted as indicating moderate or high levels of heterogeneity and its causes were investigated. Where heterogeneity persisted, we presented results separately, if appropriate, and reported reasons for persistence. Where appropriate, we statistically pooled the outcomes and examined the differences between the two models using random-effects models. Summary statistics using meta-analytic methods were performed, if applicable, and presented in GRADE tables.

Subgroup analysis

Due to the small number of studies, no sub-group analysis was performed by trial quality, setting or other sub-groups.

Publication bias

Assessment of publication bias was performed through funnel plots; however the small number of studies did not allow for any definitive conclusions. Publication bias was minimized by a comprehensive search strategy that included evaluating published and unpublished literature.

Assessment of risk of bias for individual randomized studies

Application of GRADE and Cochrane Collaboration tools for risk of bias for each individual study was applied and presented in summary tables. The GRADE and Cochrane approaches assess risk of bias in individual studies across 6 domains: sequence generation, allocation concealment, blinding, incomplete outcome data, selective outcome reporting, and other potential biases (see Appendix 3).

Assessment of risk of bias for individual observational studies

Observational studies were assessed for risk of bias using the above criteria in Table 9 and also the Newcastle-Ottawa Quality Assessment Scale (NOS, see Appendix 4).⁵ The NOS is a validated scale from 0 to 9 that uses a 'star rating system' and assesses quality of cohort and case-control studies in 3 main areas: selection of study groups, comparability of study groups and ascertainment of exposure or outcome.

Assessment of quality of evidence across studies

The quality of evidence across a body of evidence was assessed with the GRADE approach (see Appendix 5), defining the quality of evidence for each outcome as, "the extent to which one can be confident that an estimate of effect or association is close to the quantity of specific interest".² The quality rating across studies has four levels: high, moderate, low or very low. Randomised trials are categorized as high quality but can be downgraded; similarly, observational studies can be upgraded. Factors that decrease the quality of evidence include limitations in study design, indirectness of evidence, inconsistency of results, imprecision of results, or high probability of publication bias. Factors that can increase the quality level of a body of evidence include a large magnitude of effect, if all plausible confounding would reduce a demonstrated effect, and if there is a dose-response gradient.

RESULTS

Search outcomes

From the search strategy, 2,361 titles were initially identified (see figure). LC performed an initial screen of these titles and abstracts, removing all titles which were clearly not fitting inclusion criteria or which met exclusion criteria, e.g. reviews, letters, clearly off topic studies. This initial screen resulted in 164 titles remaining. LC and EH then independently conducted the selection of potentially relevant studies by scanning the titles, abstracts, and descriptor terms of all downloaded material from the electronic searches for these 164 studies. Irrelevant reports were discarded, and the full article was obtained for all potentially relevant or uncertain reports. LC and EH independently applied the inclusion criteria. JH acted as arbiter where there was disagreement. Studies were reviewed for relevance, based on study design, types of participants, exposures and outcomes measures. Finally where resolution was not possible because further information was required, the study was allocated to the list of those awaiting assessment. Attempts to contact authors to provide further clarification of data were made.

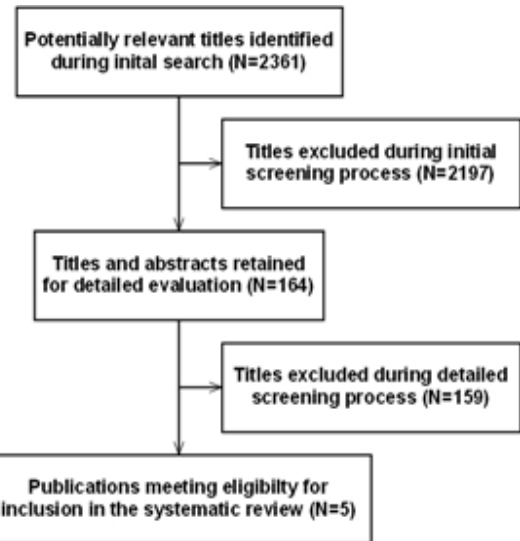


Figure: Flow diagram of citations during search process.

Search yield

In total, we identified 3 randomized trials and 2 observational studies with comparators for data extraction, coding, and potential meta-analysis. 3 ongoing trials were identified, and 8 cost-effectiveness/modeling studies were also identified.

DESCRIPTION OF THE STUDIES

Randomized Trials-Included

Home Based AIDS Care (H.B.A.C.) (2008, abstract only)⁶ is a three-arm randomized trial of the utility of clinical (weekly home visits) vs. clinical and immunologic (weekly home visits and CD4 cell counts every 12 weeks) vs. clinical, immunologic, and virologic monitoring (weekly home visits and CD4 cell counts and VLs every 12 weeks). It was conducted in Tororo, Uganda and randomised 1116 total patients enrolled in 2003 and 2004 who were followed a median of 3 years. The primary outcome was death or any new AIDS-defining illness.

Development of Antiretroviral Therapy in Africa (D.A.R.T.) (2009, abstract only)⁷ is a randomized trial of laboratory and clinical monitoring (chemistry, full blood count, and CD4 cell count every 12 weeks) versus clinical monitoring (full blood count and chemistry only returned if clinically indicated, CD4 never returned). It was conducted in Uganda and Zimbabwe and randomised 3316 total patients enrolled in 2003 to 2004 who were followed for a median of 4.9 years. The primary outcome was new WHO stage 4 event, death, or serious adverse event.

Randomized Trials-Excluded

Haubrich (2001)⁸ is a randomized study of frequent VL (at baseline and every 2 months) compared to infrequent VL monitoring (at baseline and twice yearly). It was conducted by the California Collaborative Treatment Group (CCTG) and consisted of 206 patients enrolled from 1996-1997. Patients were randomized (1.5:1) and the primary outcome was HIV RNA reduction at six months. We decided to

exclude this study because it was comparing effects of different frequencies of one type of monitoring rather than comparing one distinct monitoring strategy to another.

Observational studies with comparators-Included

Braitstein (2006)⁹ or **ART-LINC 2006** is a multicohort study of 30 ART programs in Africa, Asia, South America, Europe, and North America. This study was primarily intended to compare mortality of ART-naïve HIV-infected persons in the first year of ART between low-income (n=4810) and high-income countries (n=22,217). However, a substudy analysis compared mortality between low-income ART programs with and without routine monitoring of virologic response. Routine monitoring was not defined in the paper.

Egger (2009)¹⁰ or **ART-LINC 2009** is a multicohort study of 17 ART programs in low-resource settings comparing programs with (n=7) and without (n=10) routine VL monitoring (at least one measurement between 3 and 9 months after starting ART in at least 50% of patients). This is the same collaboration as the Braitstein study. This study analysed 20,113 patients over an unclear time period with outcomes including time to switching and CD4 cell counts at switching. Time to switch analyses were divided into three discrete time periods as the hazard was not proportional over time.

Cost-effectiveness/Modeling Studies-Summarized

Goldie (2006)¹² is a cost-effectiveness analysis of treatment strategies for a cohort of adults in Cote d'Ivoire. Several strategies were compared including: (a) no treatment; (b) cotrimoxazole prophylaxis alone; (c) cotrimoxazole and ART only; (d) cotrimoxazole, ART and CD4 testing to both guide ART starting and stopping criteria. The ICER for including CD4 testing (d) was \$1,180 (compared to GDP of \$708) and thus cost-effective.

Bishai (2007)¹³ is a cost-effectiveness analysis simulating cost and outcomes of 10,000 simulated HIV-infected patients followed every 6 months for 10 years in a low-resource settings. Five strategies were compared: (a) no ART; (b) with ART but without any laboratory markers of HIV other than positive serology; (c) ART plus total lymphocyte count; (d) ART plus CD4 cell counts; (e) ART plus CD4 cell count plus VLs. This study found that with second-line ART available, the incremental cost-effectiveness ratio (ICER) was \$8,636 for the CD4 only strategy (d) and \$13,670 for VL strategy (e), and concluded these strategies would save additional lives but at a high incremental cost.

Vijayaraghavan (2007)¹⁴ is a cost-effectiveness analysis of initiating and monitoring ART in South Africa. Several strategies were compared: (a) initiating ART at CD4 \leq 200 or Stage IV and monitor CD4 cell counts every 6 months; (b) initiate ART at \leq 350 or VL >100,000 copies/ml and (b1) monitoring CD4 cell count and VL every 6 months or (b2) monitoring CD4 cell count and VL every 3 months. The ICER for the 6 month CD4 and VL strategy (b1) was \$7,860 (compared to GDP per capita of \$4900 in 2005) and thus cost-effective.

Bendavid (2008)¹⁵ is a cost-effectiveness analysis of HIV monitoring strategies in southern Africa. Three main strategies were compared: (a) symptom-based monitoring; (b) CD4-based monitoring; (c) CD4 and VL-based monitoring. CD4-based monitoring (b) was found to result in a gain in life expectancy (6.5 months) at a discounted lifetime cost (\$464) if threshold for starting ART was 200 cells/ul and there was additional benefit if the threshold was raised to 350 cell/ul. CD4 and VL-based monitoring (c) had an ICER of \$5,414 which was not clearly cost-effective and noted to be highly dependent on test prices and rates of virologic failure.

Phillips (2008)¹⁶ is a computer simulation model of different monitoring strategies to decide when to switch for patients on ART in LRS. Several switching strategies were compared: (a) VL >500; (b) VL >10,000; (c) New WHO stage 3/4; (d) CD4 decline from peak; (e) current CD4 decline; (f) multiple WHO

stage 3 events/new WHO stage 4 even; (g) new WHO stage 4 event. No cost-effective benefit was seen for CD4-based strategies (d and e), and VL-based strategies had cost per life-year gained of \$1,500 (a) and \$4,011 (b) and were deemed by the others to be unlikely to be cost-effective in most low-resource settings.

Bendavid (2009)¹⁷ is a cost-effectiveness analysis of increasing the number of ART regimens in low-resource settings. Two main strategies were compared to the current WHO regimen and both were analyzed with CD4 cell count monitoring alone or with CD4 and VL monitoring available: (a) initial triple-NRTI regimen; (b) third regimen with second-generation PI. The strategy of using a third regimen with second-generation PI (b) with CD4 monitoring had an ICER of \$2,581 and with CD4 and VL monitoring \$6,519 which was deemed to be cost-effective in countries with GDP >\$2,000.

D.A.R.T. (2009, abstract only)¹⁸ is a cost-effectiveness analysis of the DART study. The ICER of the laboratory monitoring arm (CBC, chemistry, and CD4 cell count every 12 weeks) was \$1,693 which was not cost-effective (GDP per capita of \$400 in Uganda).

Ongoing studies

Lallemant (NCT00132682)¹⁹ is an ongoing randomized, multi-center study of adults in Thailand trying to determine if a decision to switch to a subsequent ART regimen based upon CD4 count (switch with confirmed 30% CD4 decline from baseline within 200 cells/ul from baseline) rather than based on VL (switch with confirmed VL>400 copies/ml) could ensure the same immunological and clinical outcomes while preserving future treatment options.

Laurent (NCT00301561)²⁰ is an ongoing randomized, multicentre trial of adults in 9 rural district hospitals in Cameroon evaluating outcomes of a clinical approach (VL and CD4 count every six months and clinical monitoring by physicians) versus a public health approach (clinical monitoring by nurses and physicians only) to the provision of ART.

Sagg (NCT00929604)²¹ is an ongoing randomized trial of adults in Lusaka, Zambia evaluating the use of targeted VL monitoring (VL performed if criteria for either immunologic or clinical failure met) versus routine VL monitoring (VL monitoring at 3, 6, 12, 18, 24, 30, and 36 months) to improve survival and decrease HIV disease progression in patients receiving ART.

PENPACT 1 (NCT00039741)²² is an ongoing randomized, open-label, multicountry study of ART and treatment-switching strategies (switching therapy at VL thresholds of 1,000 copies/ml versus 30,000 copies/ml) in ART-naïve children and adolescents >30 days and <18 years of age.

Study Characteristics and Methodological Quality-Randomized Trials

Two randomized studies (HBAC 2008 and DART 2009) were included in final analysis of study characteristics, methodological quality, and meta-analysis.

Table: Study Characteristics of H.B.A.C. 2008

Methods	Randomized Trial
Participants	Adults qualifying for ART
Interventions	Arm A: Clinical, Immunologic, and Virologic Monitoring (quarterly CD4 counts and viral loads) versus Arm B: Clinical and Immunologic Monitoring (quarterly CD4 counts) versus Arm C: Clinical Monitoring (weekly home visits)
Outcomes	Death; New AIDS-defining illness
Notes	Abstract Only, Highly pre-selected population within relatively well-resourced

	ART delivery program.
--	-----------------------

Table: Risk of bias table for H.B.A.C. 2008

Item	Judgment	Description
Adequate sequence generation?	Unclear	Not reported.
Allocation concealment?	Unclear	Not reported.
Blinding?	No	Blinding was not appropriate for this study.
Incomplete outcome data addressed?	Unclear	Lost-to follow-up analyses not extensively presented but absolute numbers were relatively small.
Free of selective reporting?	Yes	However, abstract only.
Free of other bias?	Yes	

Table: Study Characteristics of D.A.R.T. 2009

Methods	Randomized Trial
Participants	Adults qualifying for ART
Interventions	Clinical and Immunologic Monitoring (CD4 q12 weeks) versus Clinical Monitoring
Outcomes	Efficacy: Progression to a new WHO stage 4 HIV event or death; Safety: Any serious adverse event, which is not HIV related
Notes	Abstract Only, Highly pre-selected population within relatively well-resourced ART delivery program.

Table: Risk of bias table for D.A.R.T. 2009

Item	Judgment	Description
Adequate sequence generation?	Unclear	Not reported.
Allocation concealment?	Unclear	Not reported.
Blinding?	No	Blinding was not appropriate for this study.
Incomplete outcome data addressed?	Unclear	Lost-to follow-up analyses not extensively presented but absolute numbers were relatively small.
Free of selective reporting?	Yes	However, abstract only.
Free of other bias?	Yes	

Figure 1: Methodological quality summary for randomized trials.

	Adequate sequence generation?	Allocation concealment?	Blinding?	Incomplete outcome data addressed?	Free of selective reporting?	Free of other bias?
ARTLINC 2006						
ARTLINC 2009						
D.A.R.T. 2009	?	?	-	?	+	+
H.B.A.C. 2008	?	?	-	?	+	+

Study Characteristics and Methodological Quality-Observational Studies

Two observational studies (ART-LINC 2006 and 2009) were included in final analysis of observational study characteristics and methodological quality.

Table: Newcastle-Ottawa quality assessment scale for included observational studies.

Item	ART-LINC 2006	ART-LINC 2009
Representativeness of Cohort	1	1
Selection of Non-exposed Cohort	0	0
Ascertainment of Cohort	1	1
Outcome of Interest Not Present at Start of Study	1	1
Comparability of Cohorts 1	1	1
Comparability of Cohorts 2	0	0
Assessment of Outcome	1	1
Long Enough Follow-up	1	1
Adequacy of Follow-up	0	0
TOTAL	6	6

Comments: Selection of non-exposed cohort for both studies was not drawn from the same community as the exposed cohort. While adjusting did occur for age, sex, clinical stage, and CD4 cell count at baseline (most important factor), but not for other potential factors such as program characteristics. Follow-up was variable in both studies for the primary outcomes.

DRAFT

META-ANALYSES

Clinical (Experimental) versus Clinical + Immunologic Monitoring (Control)

Figure: Mortality

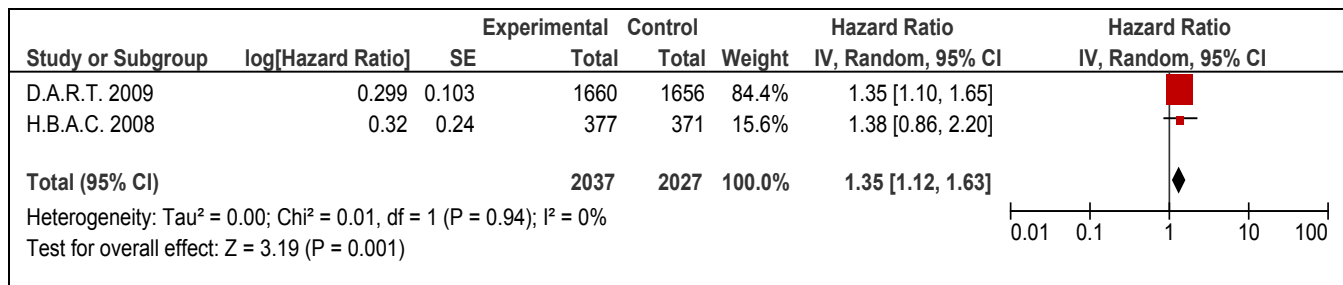


Figure: AIDS-defining illness or Mortality

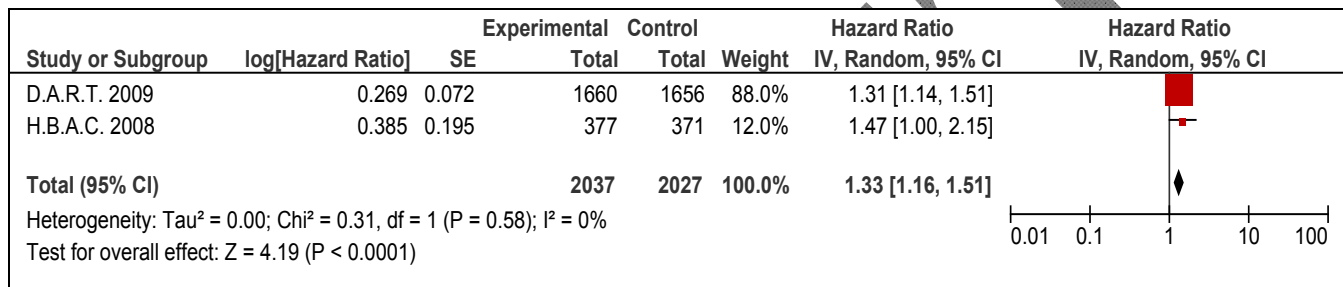


Figure: Serious Adverse Event

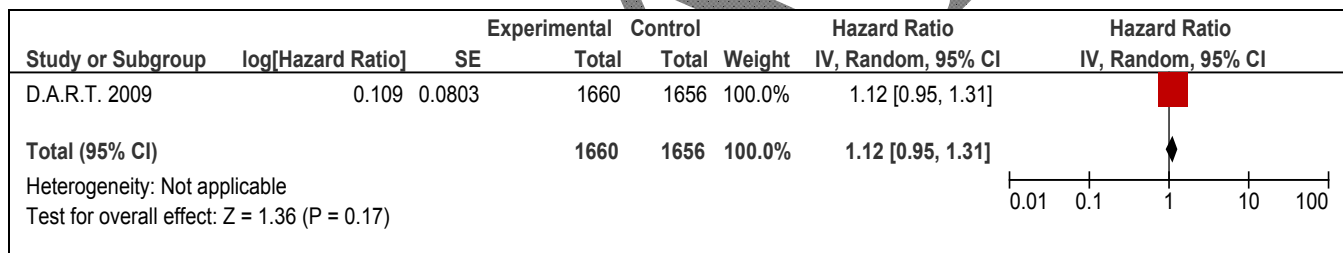


Figure: Unnecessary Switch

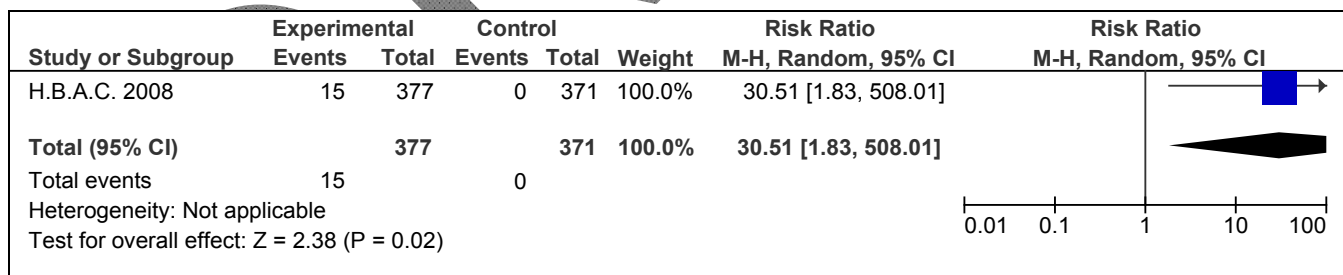
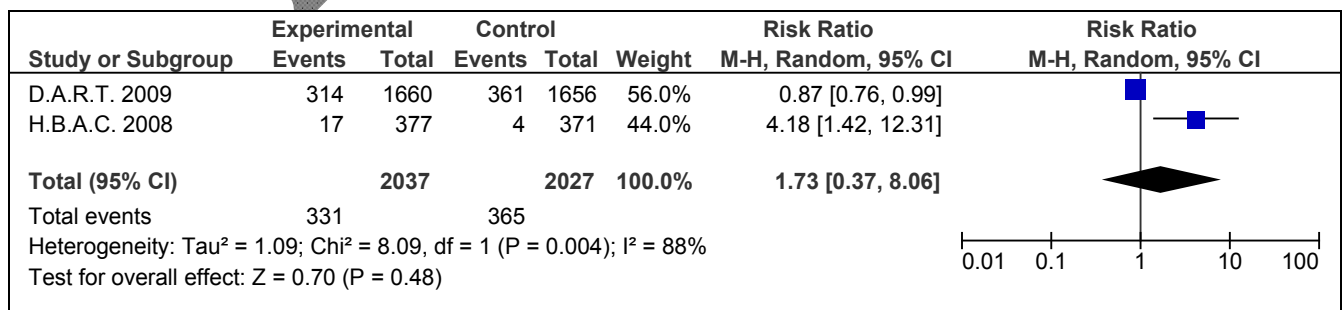


Figure: Switch to Second-Line



Clinical (Experimental) versus Clinical + Immunologic + Virologic Monitoring (Control)

Figure: Mortality

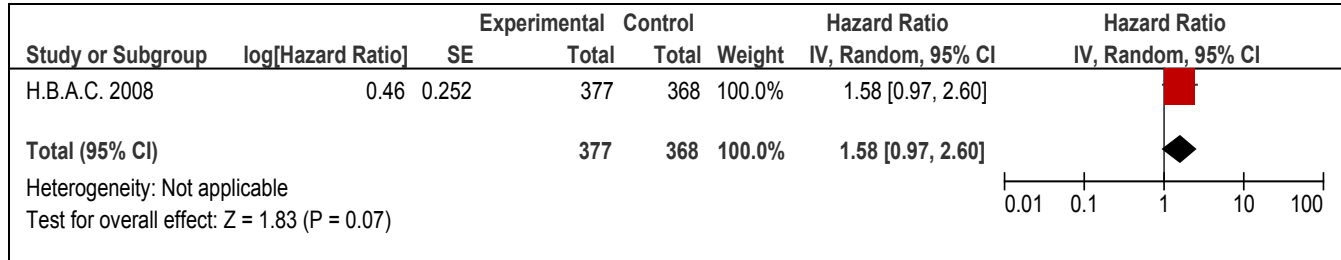


Figure: AIDS-defining illness or Mortality

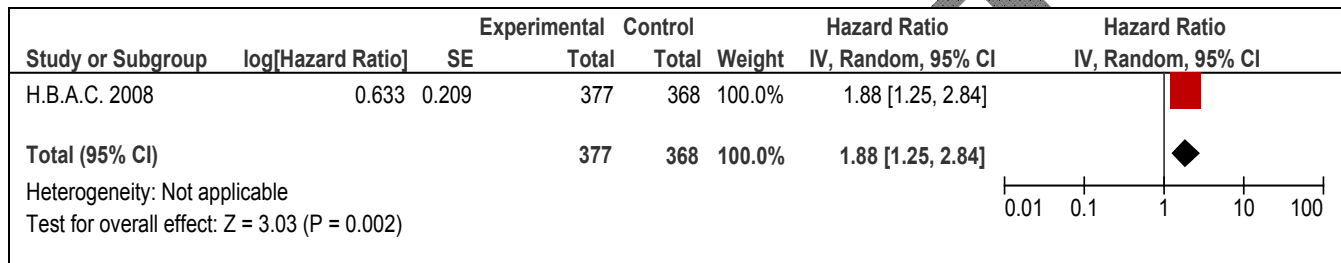


Figure: Unnecessary Switch

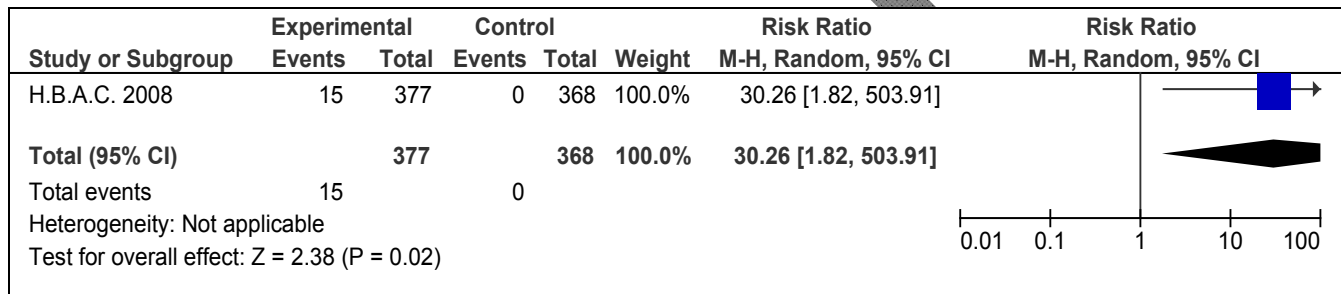


Figure: Virologic Treatment Failure

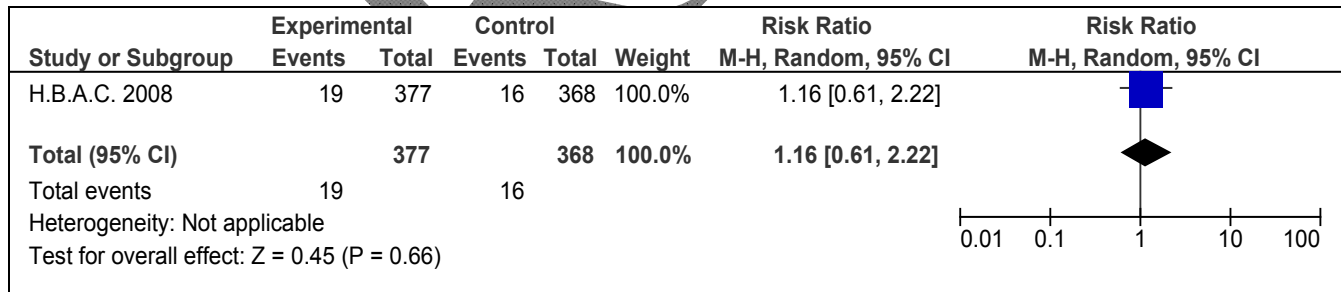
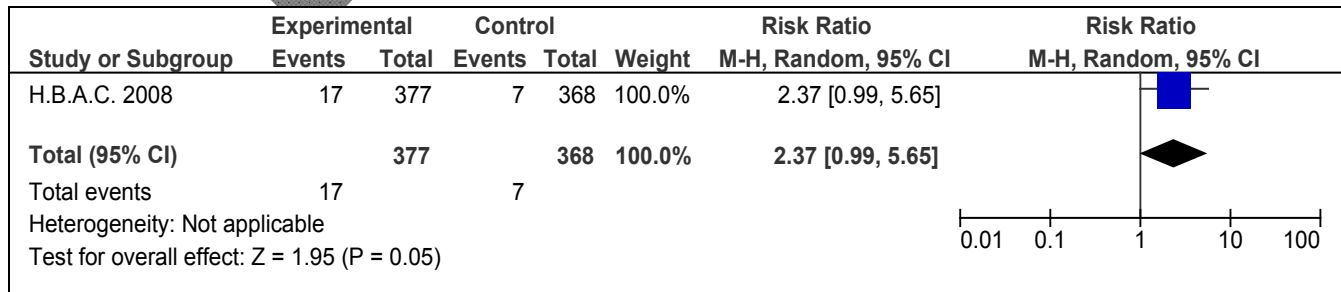


Figure: Switch to Second-line



Clinical + Immunologic (Experimental) versus Clinical + Immunologic + Virologic Monitoring (Control)

Figure: Mortality

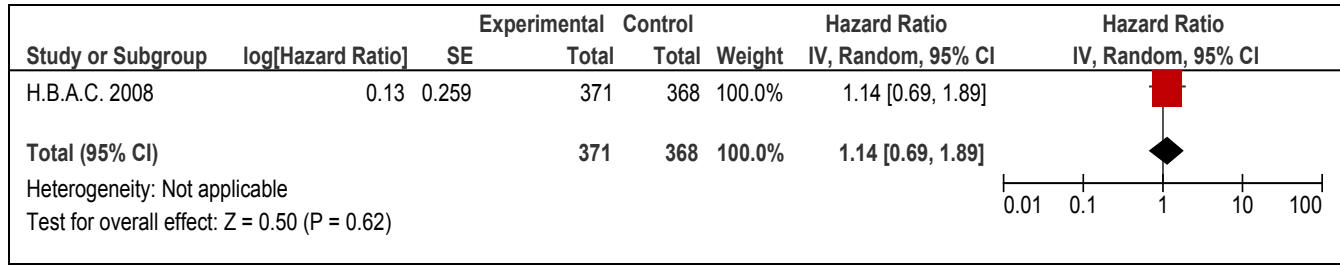


Figure: AIDS-defining illness or Mortality

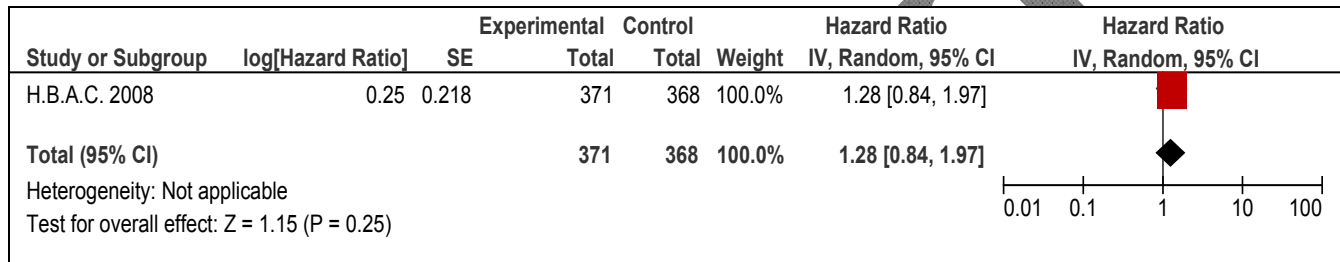


Figure: Unnecessary Switch

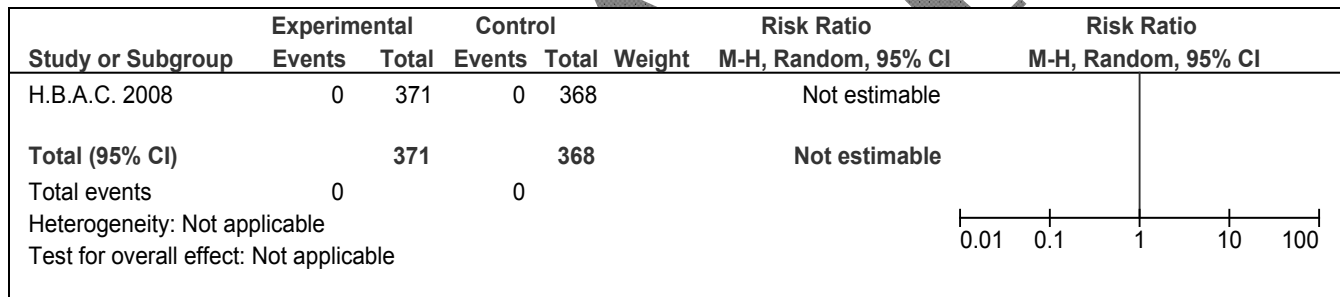


Figure: Virologic Treatment Failure

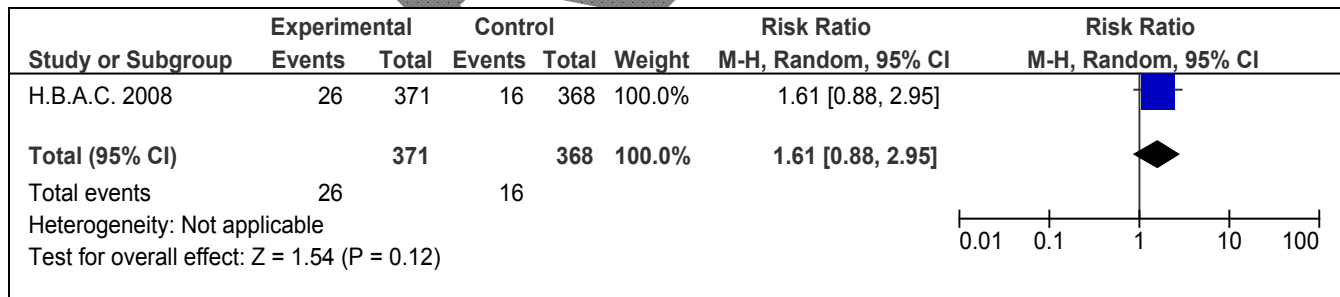
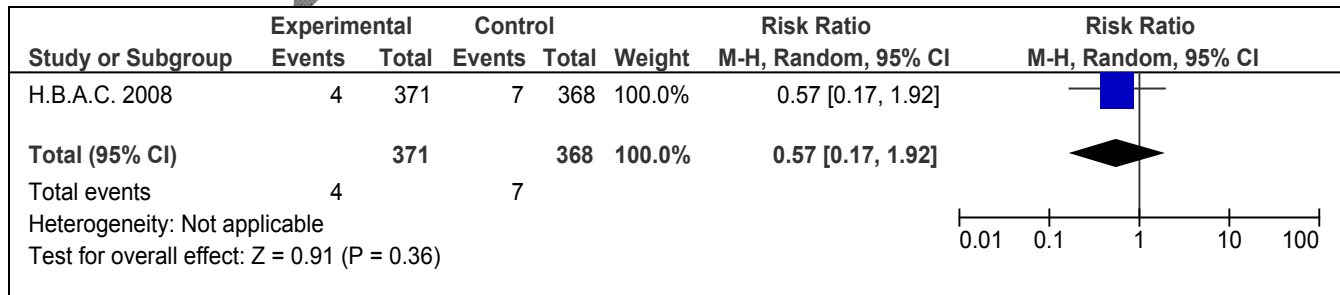


Figure: Switch to Second-line



Clinical + Immunologic (Control) versus Clinical + Immunologic + Virologic Monitoring (Experimental)

Figure: Mortality

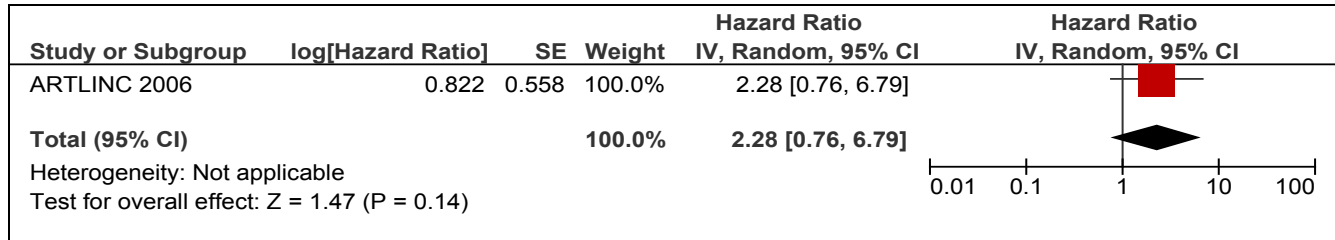


Figure: Rate of Switching

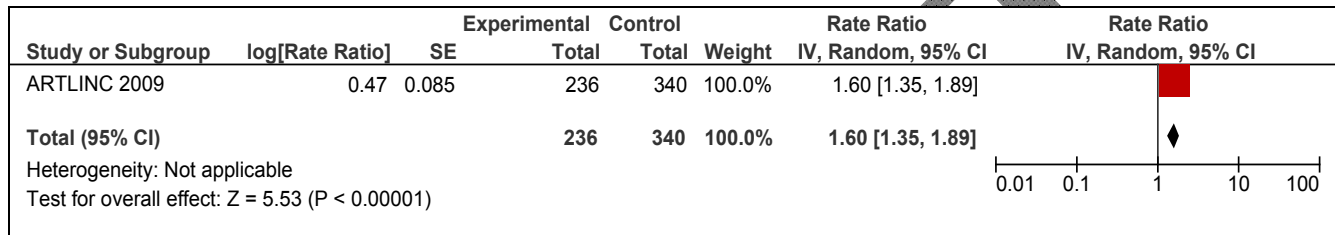


Figure: Time to Switch (7-18 months)

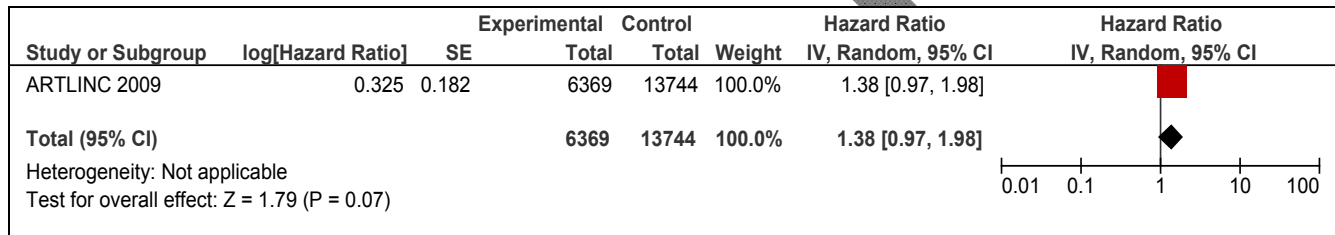


Figure: Time to Switch (19-30 months)

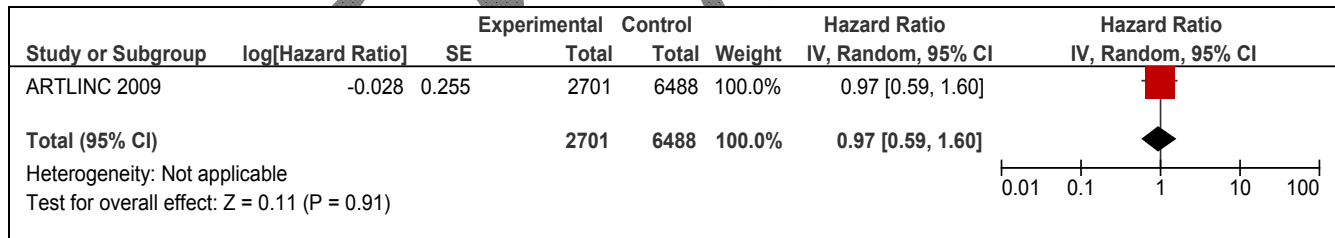


Figure: Time to Switch (31-42 months)

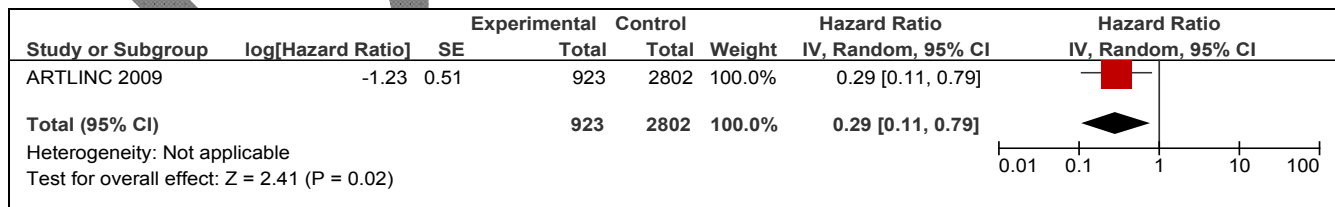
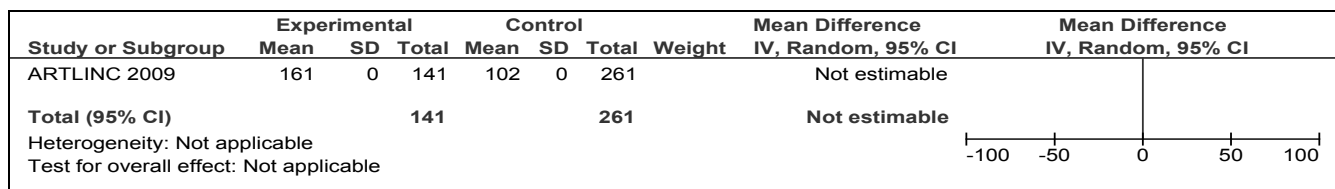


Figure: CD4 Count at Time of Switch



GRADE Tables

Author(s): Larry William Chang, Jamal Harris

Date: 2009-08-12

Question: Should Clinical Monitoring vs Immunologic and Clinical Monitoring be used in guiding when to switch first-line antiretroviral therapy in adults in low-resource settings?

Settings: Low-resource settings

Bibliography: H.B.A.C. 2008; D.A.R.T. 2009

Quality assessment							Summary of findings					Importance
							No of patients		Effect		Quality	
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Clinical Monitoring	Immunologic and Clinical Monitoring	Relative (95% CI)	Absolute		
Mortality (follow-up median 3-5 years)												
2	randomised trials	serious ¹	no serious inconsistency	no serious indirectness ²	serious ³	none ⁴	?/2037 ⁵	?/2027 ⁵	HR 1.35 (1.12 to 1.63)	-	⊕⊕OO LOW	CRITICAL
AIDS-defining illness - not reported												
0	-	-	-	-	-	-	-	-	-	-		CRITICAL
AIDS-defining illness or Mortality (follow-up median 3-5 years)												
2	randomised trials	serious ¹	no serious inconsistency	no serious indirectness ²	no serious imprecision	none ⁴	547/2037 (26.9%) ⁶	414/2027 (20.4%) ⁶	HR 1.33 (1.16 to 1.51)	58 more per 1000 (from 29 more to 88 more)	⊕⊕⊕O MODERATE	CRITICAL
Serious Adverse Event (follow-up median 5 years)												
1 ⁷	randomised trials	serious ¹	no serious inconsistency	no serious indirectness ²	serious ³	none ⁴	?/1660 ⁸	?/1656 ⁸	HR 1.12 (0.94 to 1.31)	-	⊕⊕OO LOW	CRITICAL
Unnecessary Switch (Switch to Second-line with Undetectable Viral Load) (follow-up median 3 years)												
1 ⁷	randomised trials	serious ¹	no serious inconsistency	no serious indirectness ²	serious ³	none ⁴	15/377 (4%)	0/371 (0%)	RR 30.5 (1.83 to 508)	-	⊕⊕OO LOW	CRITICAL
Switch to Second-Line (follow-up median 3-5 years)												
2	randomised trials	serious ¹	serious ⁹	no serious indirectness ²	no serious imprecision	none ⁴	331/2037 (16.2%)	365/2027 (18%)	RR 1.73 (0.37 to 8.06)	13 more per 100 (from 11 fewer to 127 more)	⊕⊕OO LOW	

¹ Unclear sequence generation and allocation concealment and blinding was not possible for both studies; lost-to follow-up analyses not extensively presented for either trial but absolute numbers were relatively small.

² Patient populations pre-selected and within relatively well-resourced ART delivery programs; however, as setting(s) was low-resource, no downgrading occurred.

³ Total number of events is small.

⁴ Abstract(s) only, no peer-reviewed print publication(s) of these data are available; however, as a significant amount of data was available from abstracts/conference presentations no downgrading occurred.

⁵ Number with event not reported in either study. DART mortality in clinical arm 2.94/100 P-Y, in immunologic + clinical arm 2.18/100 P-Y.

⁶ In DART in clinical arm 6.94 events/100 P-Y, in immunologic + clinical arm, 5.24 events/100 P-Y. In HBAC in clinical arm 7.57 events/100 P-Y, in immunologic + clinical arm 5.97 events/100 P-Y.

⁷ DART study only.

⁸ Number with event not reported.

⁹ Number of events and point estimate varied widely between the two studies.

DRAFT: When to switch review, last edit 10/8/2009

Author(s): Larry William Chang, Jamal Harris

Date: 2009-08-12

Question: Should Clinical Monitoring vs Virologic, Immunologic, and Clinical Monitoring be used for guiding when to switch first-line antiretroviral therapy in adults in low-resource settings?

Settings: Low-resource settings

Bibliography: H.B.A.C. 2008

Quality assessment							Summary of findings					Importance
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	No of patients		Effect		Quality	
							Clinical Monitoring	Virologic, Immunologic, and Clinical Monitoring	Relative (95% CI)	Absolute		
Mortality (follow-up median 3 years)												
1	randomised trials	serious ¹	no serious inconsistency	no serious indirectness ²	serious ³	none ⁴	7/377 ⁵	7/368 ⁵	HR 1.58 (0.97 to 2.6)	-	⊕⊕○○ LOW	CRITICAL
AIDS-defining illness - not reported												
0	-	-	-	-	-	-	-	-	-	-		CRITICAL
AIDS-defining illness or Mortality (follow-up median 3 years)												
1	randomised trials	serious ¹	no serious inconsistency	no serious indirectness ²	serious ³	none ⁴	72/377 (19.1%) ⁶	47/368 (12.8%) ⁶	HR 1.88 (1.25 to 2.84)	99 more per 1000 (from 29 more to 194 more)	⊕⊕○○ LOW	CRITICAL
Unnecessary Switch (Switch to Second-line with Undetectable Viral Load) (follow-up median 3 years)												
1	randomised trials	serious ¹	no serious inconsistency	no serious indirectness ²	serious ³	none ⁴	15/377 (4%)	0/368 (0%)	RR 30.3 (1.82 to 504)	-	⊕⊕○○ LOW	CRITICAL
Virologic Treatment Failure (follow-up median 3 years)												
1	randomised trials	serious ¹	no serious inconsistency	no serious indirectness ²	serious ³	none ⁴	19/377 (5%)	16/368 (4.3%)	RR 1.16 (0.6 to 2.19)	7 more per 1000 (from 17 fewer to 52 more)	⊕⊕○○ LOW	IMPORTANT
Switch to Second-line (follow-up median 3 years)												
1	randomised trials	serious ¹	no serious inconsistency	no serious indirectness ²	serious ³	none ⁴	17/377 (4.5%)	7/368 (1.9%)	RR 2.37 (0.99 to 5.65)	26 more per 1000 (from 0 fewer to 88 more)	⊕⊕○○ LOW	

¹ Unclear sequence generation and allocation concealment, lost-to follow-up analyses not extensively presented but absolute numbers were relatively small, and blinding was not possible.

² Patient populations pre-selected and within relatively well-resourced ART delivery programs; however, as this study was in a low-resource setting it was not downgraded.

³ Total number of events was small.

⁴ Abstract(s) only, no peer-reviewed print publication(s) of these data are available; however, as a significant amount of data was available from abstracts/conference presentations no downgrading occurred.

⁵ Number with event not reported.

⁶ In clinical arm 7.57 events/100 P-Y, in virologic + immunologic + clinical arm 4.80 events/100 P-Y.

Author(s): Larry William Chang, Jamal Harris

Date: 2009-08-12

Question: Should Clinical and Immunologic Monitoring vs Virologic, Immunologic, and Clinical Monitoring be used in guiding when to switch first-line antiretroviral therapy in adults in low-resource settings?

Settings: Low-resource settings.

Bibliography: H.B.A.C. 2008

Quality assessment							Summary of findings					Importance
							No of patients		Effect		Quality	
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Clinical and Immunologic Monitoring	Virologic, Immunologic, and Clinical Monitoring	Relative (95% CI)	Absolute		
Mortality (follow-up median 3 years)												
1	randomised trials	serious ¹	no serious inconsistency	no serious indirectness ²	serious ³	none ⁴	?/371 ⁵	?/368 ⁵	HR 1.14 (0.7 to 1.9)	-	⊕⊕⊕ LOW	CRITICAL
AIDS-defining illness - not reported												
0	-	-	-	-	-	-	-	-	-	-		CRITICAL
AIDS-defining illness or Mortality (follow-up median 3 years)												
1	randomised trials	serious ¹	no serious inconsistency	no serious indirectness ²	serious ³	none ⁴	58/371 (15.6%) ⁶	47/368 (12.8%)	HR 1.28 (0.84 to 1.97)	33 more per 1000 (from 19 fewer to 108 more)	⊕⊕⊕ LOW	CRITICAL
Unnecessary Switch (Switch to Second-line with Undetectable Viral Load) (follow-up median 3 years)												
1	randomised trials	serious ¹	no serious inconsistency	no serious indirectness ²	very serious ⁷	none ⁴	0/371 (0%)	0/368 (0%)	Not estimable	-	⊕⊕⊕ VERY LOW	CRITICAL
Virologic Treatment Failure (follow-up median 3 years)												
1	randomised trials	serious ¹	no serious inconsistency	no serious indirectness ²	serious ³	none ⁴	26/371 (7%)	16/368 (4.3%)	RR 1.61 (0.88 to 2.95)	27 more per 1000 (from 5 fewer to 85 more)	⊕⊕⊕ LOW	IMPORTANT
Switch to Second-Line (follow-up median 3 years)												
1	randomised trials	serious ¹	no serious inconsistency	no serious indirectness ²	serious ³	none ⁴	4/371 (1.1%)	7/368 (1.9%)	RR 0.57 (0.17 to 1.92)	8 fewer per 1000 (from 16 fewer to 18 more)	⊕⊕⊕ LOW	

¹ Unclear sequence generation and allocation concealment, lost-to-follow-up analyses not extensively presented but absolute numbers were relatively small, and blinding was not possible.

² Patient populations pre-selected and within relatively well-resourced ART delivery programs; however, as this study was in a low-resource setting it was not downgraded.

³ Total number of events is small.

⁴ Abstract(s) only, no peer-reviewed print publication(s) of these data are available; however, as a significant amount of data was available from abstracts/conference presentations no downgrading occurred.

⁵ Number with event not reported.

⁶ In clinical + immunologic arm 5.97 events/100 P-Y, in virologic + immunologic + clinical arm 4.80 events/100 P-Y.

⁷ Total number of events is very small.

Author(s): Larry William Chang, Jamal Harris

Date: 2009-09-14

Question: Should Virologic, Immunologic, and Clinical Monitoring vs Immunologic and Clinical Monitoring be used in guiding when to switch first-line antiretroviral therapy in adults in low-resource settings?

Settings: Low-resource settings.

Bibliography: ARTLINC 2006, 2008

Quality assessment							Summary of findings					Importance
							No of patients		Effect		Quality	
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Virologic, Immunologic, and Clinical Monitoring	Immunologic and Clinical Monitoring	Relative (95% CI)	Absolute		
Mortality (follow-up 12 months)												
1	observational studies	serious ¹	no serious inconsistency	no serious indirectness	no serious imprecision	none	See comment. ²	See comment. ²	HR 2.28 (0.76 to 6.79)	-	⊕○○○ VERY LOW	CRITICAL
Rate of switching												
1	observational studies	serious ³	no serious inconsistency	no serious indirectness	no serious imprecision	none	236/6369 (3.7%)	340/13744 (2.5%)	RR 1.60 (1.35 to 1.89) ⁴	15 more per 1000 (from 9 more to 22 more)	⊕○○○ VERY LOW	
Time to switch (7-18 months)												
1	observational studies	serious ³	no serious inconsistency	no serious indirectness	no serious imprecision	none	?/6369 ⁵	?/13744 ⁵	HR 1.38 (0.97 to 1.98)	-	⊕○○○ VERY LOW	
Time to switch (19-30 months)												
1	observational studies	serious ³	no serious inconsistency	no serious indirectness	no serious imprecision	none	?/2701 ⁵	?/6488 ⁵	HR 0.97 (0.58 to 1.6)	-	⊕○○○ VERY LOW	
Time to switch (31-42 months)												
1	observational studies	serious ³	no serious inconsistency	no serious indirectness	no serious imprecision	none	?/923 ⁵	?/2802 ⁵	HR 0.29 (0.11 to 0.79)	-	⊕○○○ VERY LOW	
CD4 cell count at time of switch												
1	observational studies	serious ³	no serious inconsistency	no serious indirectness	no serious imprecision	none	141 patients	261 patients	See comment. ⁶	-	⊕○○○ VERY LOW	

¹ This outcome was a subgroup analysis, selection of non-exposed cohorts were not drawn from same communities as the exposed cohorts.

² Neither number with event or at risk reported.

³ Selection of non-exposed cohorts were not drawn from the same communities as the exposed cohorts; incomplete follow-up data on many participants.

⁴ Programs with virologic monitoring rate of switching was 3.2/100 P-Y (95% CI 2.2-2.6) versus 2.0/100 P-Y (95% CI 1.9-2.3) in those without (p<0.0001); RR here is a rate ratio.

⁵ Number with event not reported.

⁶ Programs with virologic monitoring CD4 cell count at time of switching was 161 cells/ul compared to 102 cells/ul in those without (p=0.001).

SUMMARY

Clinical vs. Immunologic and Clinical Monitoring. Based upon two randomized trials, clinical monitoring alone results in increased mortality (low quality evidence), increased AIDS-defining illnesses and mortality as a composite endpoint (moderate), no difference in serious adverse events (low), and increased numbers of unnecessary switches (low) compared to immunological and clinical monitoring.

Clinical vs. Virologic, Immunologic, and Clinical Monitoring. Based upon a single randomized trial, clinical monitoring alone results in increased mortality (low), increased AIDS-defining illnesses and mortality as a composite endpoint (low), increased unnecessary switches (low), and no difference in virologic treatment failures (low) compared to virologic, immunologic, and clinical monitoring.

Immunologic and Clinical vs. Virologic, Immunologic, and Clinical Monitoring. Based upon a single randomized trial, immunologic and clinical monitoring results in no difference in mortality (low), no difference in AIDS-defining illnesses and mortality as a composite endpoint (low), no difference in unnecessary switches (very low), and no difference in virologic treatment failures (low) compared to virologic, immunologic, and clinical monitoring. Observational studies appear to demonstrate that programs with virologic, immunologic, and clinical monitoring switch therapy more frequently (very low), and earlier (very low), and at higher CD4 counts (very low) compared to programs with only immunologic and clinical monitoring.

DRAFT

REFERENCES

1. World Health Organization. Antiretroviral therapy for HIV infection in adults and adolescents: towards universal access. Recommendations for a public health approach, 2006, 2006.
2. Higgins JPT, Green S, eds. Cochrane Handbook for Systematic Reviews of Interventions. Chichester: John Wiley & Sons Ltd, 2008.
3. Review Manager (RevMan) [Computer program]. Version 5.0. [program]. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2008.
4. GRADEPro Version 3.2 for Windows [Computer program]. [program]: GRADE Working Group, 2008.
5. Wells GA, Shea B, O'Connell D, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Available from: http://www.ohri.ca/programs/clinical_epidemiology/oxford.htm Accessed 22 June 2009.
6. HBAC, Coutinho A, Mermin J, et al. Utility of routine viral load, CD4 cell count, and clinical monitoring among HIV-infected adults in Uganda: a randomized trial. Conference on Retroviruses and Opportunistic Infections 2008, Montreal, Canada.
7. DART, Mugenyi P, Walker S, et al. Impact of routine laboratory monitoring over 5 years after antiretroviral therapy (ART) initiation on clinical disease progression of HIV-infected African adults: the DART trial final results. 5th IAS Conference on HIV Pathogenesis, Treatment and Prevention 2009, Cape Town, South Africa.
8. Haubrich RH, Currier JS, Forthal DN, et al. A randomized study of the utility of human immunodeficiency virus RNA measurement for the management of antiretroviral therapy. *Clin Infect Dis* 2001;**33**(7):1060-8.
9. Braitstein P, Brinkhof MW, Dabis F, et al. Mortality of HIV-1-infected patients in the first year of antiretroviral therapy: comparison between low-income and high-income countries. *Lancet* 2006;**367**(9513):817-24.
10. Egger M, Keiser O. Switching to second-line antiretroviral therapy in resource-limited settings: comparison of programmes with and without viral load monitoring. *Aids* 2009.
11. Orrell C, Harling G, Lawn SD, et al. Conservation of first-line antiretroviral treatment regimen where therapeutic options are limited. *Antivir Ther* 2007;**12**(1):83-8.
12. Goldie SJYYLEWMCAXWRPHHEKAHCKJ. Cost-effectiveness of HIV treatment in resource-poor settings - The case of Cote d'Ivoire. *New England Journal of Medicine* 2006;**355**(11):1141-1153.
13. Bishai D, Colchero A, Durack DT. The cost effectiveness of antiretroviral treatment strategies in resource-limited settings. *Aids* 2007;**21**(10):1333-40.
14. Vijayaraghavan A, Efrusy MB, Mazonson PD, Ebrahim O, Sanne IM, Santas CC. Cost-effectiveness of alternative strategies for initiating and monitoring highly active antiretroviral therapy in the developing world. *J Acquir Immune Defic Syndr* 2007;**46**(1):91-100.
15. Bendavid E, Young SD, Katzenstein DA, Bayoumi AM, Sanders GD, Owens DK. Cost-effectiveness of HIV monitoring strategies in resource-limited settings: a southern African analysis. *Arch Intern Med* 2008;**168**(17):1910-8.
16. Phillips AN, Pillay D, Miners AH, Bennett DE, Gilks CF, Lundgren JD. Outcomes from monitoring of patients on antiretroviral therapy in resource-limited settings with viral load, CD4 cell count, or clinical observation alone: a computer simulation model. *Lancet* 2008;**371**(9622):1443-51.
17. Bendavid E, Wood R, Katzenstein DA, Bayoumi AM, Owens DK. Expanding Antiretroviral Options in Resource-Limited Settings-A Cost-Effectiveness Analysis. *J Acquir Immune Defic Syndr* 2009.
18. DART, Gilks C. Cost-effectiveness analysis of routine laboratory or clinically driven strategies for monitoring anti-retroviral therapy in Uganda and Zimbabwe. 5th IAS Conference on HIV Pathogenesis, Treatment and Prevention 2009, Cape Town, South Africa.
19. Lallemand M. Monitoring highly active antiretroviral therapy in HIV-infected patients in Thailand. NCT00162682.
20. Laurent C. Antiretroviral treatment simplified follow-up management assessment (ANRS 12110 STRATALL). NCT00301561.
21. Sagg M. HIV viral load monitoring in resource-poor settings. NCT00929604.
22. PENPACT1. Anti-HIV Drug Regimens and Treatment-Switching Guidelines in HIV Infected Children. NCT00039741.

Appendix 1: Current WHO Guidelines-Summarized and Condensed

Current Adult WHO Guidelines

Key points on 2006 treatment failure guidelines:

- Clinical events occurring during first six months are excluded from definition of failure.
- New opportunistic infections should be differentiated from immune reconstitution syndrome (though the exclusion of events during first six months should largely account for this).
- Unless adherence has been assessed and optimized, clinical failure can not be concluded.
- In general, do not switch if the CD4 cell count is above 200 cells/mm³ (regardless of other criteria).
- In general, do not switch if the viral load is undetectable (regardless of other criteria).
- In general, do not switch therapy during first six months (regardless of other criteria).

Table 1: Definitions for Treatment Failure

Clinical failure	New or recurrent WHO stage 4 condition ^{b,c}
CD4 cell failure^d	<ul style="list-style-type: none"> • Fall of CD4 count to pre-therapy baseline (or below); or • 50% fall from the on-treatment peak value (if known); or • Persistent CD4 levels below 100 cells/mm³^e
Virological failure	Plasma viral load above 10 000 copies/ml ^f

b Certain WHO stage 3 conditions (e.g. pulmonary TB, severe bacterial infections), may be an indication of treatment failure & require consideration of switch.

c Some WHO stage 4 conditions (lymph node TB, uncomplicated TB pleural disease, oesophageal candidiasis, recurrent bacterial pneumonia) may not be indicators of treatment failure and thus do not require consideration of switch, response to appropriate therapy should be used to evaluate the need to switch therapy.

d Without concomitant infection to cause transient CD4 cell decrease.

e Some experts consider that patients with persistent CD4 cell counts below 50/mm³ after 12 months on ART may be more appropriate.

f Optimal viral load value has not been defined. However, >10 000 have been associated with subsequent clinical progression and CD4 cell count decline.

Table 2. Clinical staging events to guide decision-making on switching for adults

New or recurrent event on ART^a	Recommendations	Additional management options
Asymptomatic (T1)	Do not switch regimen	<ul style="list-style-type: none"> • Maintain scheduled follow-up visits, including CD4 monitoring (if available) • Continue to offer adherence support
Stage 2 event (T2)	Do not switch regimen	<ul style="list-style-type: none"> • Treat and manage staging event • Assess and offer adherence support • Check if on treatment for at least six months • Assess continuation or reintroduction of OI prophylaxis • Schedule earlier visit for clinical review and consider CD4 (if available)^c
Stage 3 event (T3)	Consider switching regimen	<ul style="list-style-type: none"> • Treat and manage staging event and monitor response • Assess and offer adherence support • Check if on treatment for at least six months • Check CD4 cell count (if available)^c • Assess continuation or reintroduction of OI prophylaxis • Institute more frequent follow-up
Stage 4 event (T4)	Switch regimen	<ul style="list-style-type: none"> • Treat and manage staging event and monitor response • Check if on treatment for at least six months • Assess continuation or reintroduction of OI prophylaxis • Check CD4 cell count (if available)^c • Assess and offer adherence support

a Refers to clinical stages while on ART for at least six months (termed T1, T2, T3, T4).

c Treat and manage the staging event before measuring CD4 cell count.

Table 3: Integrating clinical status, CD4 cell count, and viral load to guide switching.

Failure Criteria	WHO Stage 1	WHO Stage 2	WHO Stage 3	WHO Stage 4
CD4 failure (Viral load testing not available)	Do not switch regimen. Follow patient for development of clinical signs or symptoms. Repeat CD4 cell count in three months.	Do not switch regimen. Follow patient for evidence of further clinical progression. Repeat CD4 cell count in three months.	Consider switching to second-line regimen.	Recommend switching to second-line regimen.
CD4 failure and viral load failure	Consider switching to second-line regimen.	Consider switching to second-line regimen.	Recommend switching to second-line.	Recommend switching to second-line.

Key points on 2006 monitoring guidelines

- Clinical monitoring 2, 4, 8, 12, 24 weeks after ART initiation and every 6 months thereafter (A-IV)
- CD4 cell count every six months (A-IV)
- Viral load is not recommended for monitoring patients (B-IV)
- Viral load may be considered in connection with diagnosing treatment failure earlier and in more complex cases, such as those with discordant clinical and immunologic responses (B-IV)

Table 4: WHO 2006 lab monitoring guidelines.

Test	Pre-ART	At initiation of first-line or second-line ART	Every six months	As required
HIV Diagnostic	✓			
Haemoglobin ^a		✓		✓
WBC and Differential ^b		✓		✓
CD4 Cell Count ^c	✓	✓	✓	
Pregnancy Testing ^d		✓		✓
Full Chemistry ^{e,f}				✓
Viral Load ^g				✓

a Haemoglobin monitoring for patients on AZT is recommended at baseline and at weeks 4, 8 and 12 after initiation of AZT.

b Monitoring at week 4, 8 and 12 after initiation of ART is optional.

c Patients who are not yet eligible for ART should be monitored with measurement of CD4 every six months. For patients who develop WHO stage 2 events, or whose CD4 measurements approach threshold values, the frequency of CD4 measurement can be increased. Patients on ART should have CD4 measurement every six months if stable. More frequent CD4 monitoring may be necessary for deciding when to start or switch ART.

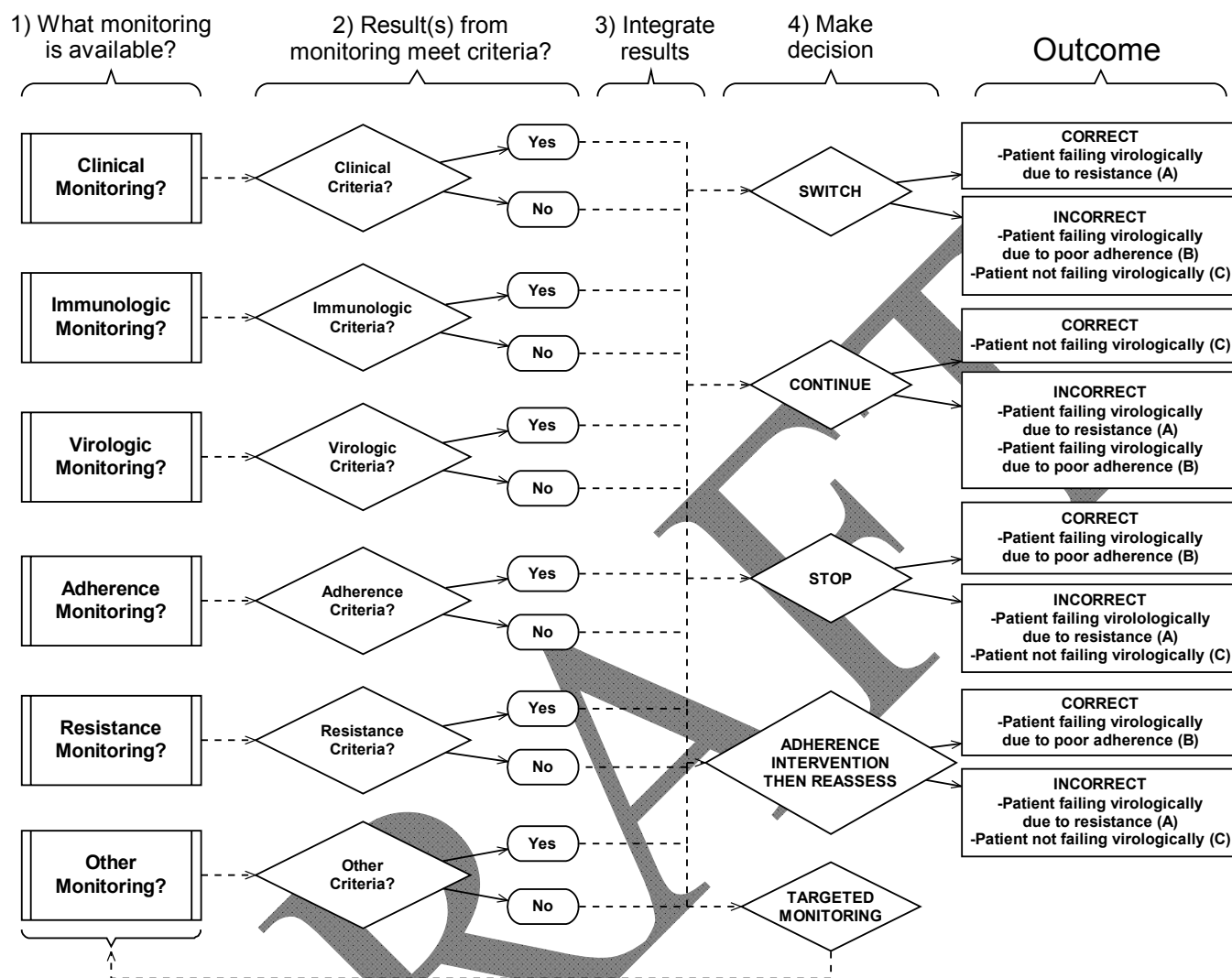
d Pregnancy testing for women initiating a first-line regimen containing EFV, and if pregnancy is suspected in women who are receiving an EFV-based regimen.

e The predictive value of routine liver enzyme monitoring is considered very low by some experts. WHO recommends liver enzyme monitoring in response to symptoms. However, regular monitoring during the first three months of treatment and symptom-directed measurement of liver enzymes thereafter has been considered by some experts for certain patients using nevirapine-based regimens, in particular for women with CD4 cell counts above 250 cells/mm³ and for those coinfecting with hepatitis B or hepatitis C virus or with other hepatic disease.

f Regular monitoring (every six months), if available, of full chemistry tests, particularly lipid levels, ALT and renal function, should be considered for patients receiving second-line drugs.

g Viral load measurement is not recommended for decision-making on the initiation or regular monitoring of ART in resource-limited settings. It is recommended primarily for the definitive diagnosis of HIV infection in HIV exposed children aged under 18 months and may be considered in connection with diagnosing treatment failure earlier or to assess discordant clinical and CD4 findings in patients in whom it is suspected that ART has failed.

Appendix 2: Model for deciding when to switch an ART regimen.



One of the critical clinical decisions made in antiretroviral therapy (ART) is the decision on when to switch from one regimen to another for treatment failure. This is a complex decision which requires consideration of multiple factors including: (1) what type of monitoring (e.g. immunologic, virologic) is available to provide information to guide switching, (2) establishing criteria for treatment failure (e.g. viral load >10,000 copies/ml), (3) deciding how to integrate criteria (e.g. clinical and immunologic), and (4) making a decision and, if possible, determining it's correctness and consequences.

The initial step in this model of deciding when to switch is determining what type of monitoring for guiding when to switch is available and appropriate. In the next step, treatment failure is typically defined according to set criteria. Treatment failure can be described generally as a suboptimal response to antiretroviral therapy and is typically measured in three different ways: (i) clinically, as evidenced by disease progression; (ii) immunologically, as evidenced by trends in CD4 counts over time; and, (iii) virologically, as evidenced by measurement of HIV RNA levels. The 2006 WHO Antiretroviral Therapy for HIV Infection in Adults and Adolescents guidelines recognized that definitions for treatment failure were not standardized and that there was an urgent need for agreement in definitions and in antiretroviral monitoring strategies to determine when to switch. However, given the need for normative standards, WHO outlined in the 2006 guidelines a set of definitions for treatment failure based upon available evidence at that time and suggested strategies for deciding when to switch (See Appendix 1).

Appendix 3: The Cochrane Collaboration's tool for assessing risk of bias.

Assessment of risk of bias for individual randomized studies

Application of GRADE (Guyatt 2008) and Cochrane Collaboration tools for risk of bias for each individual study was applied and presented in summary tables. The GRADE and Cochrane approaches assess risk of bias in individual studies across 6 domains: sequence generation, allocation concealment, blinding, incomplete outcome data, selective outcome reporting, and other potential biases (see Appendix).

Table: The Cochrane Collaboration's tool for assessing risk of bias.

Domain	Description	Review authors' judgement
Sequence generation.	Describe the method used to generate the allocation sequence in sufficient detail to allow an assessment of whether it should produce comparable groups.	Was the allocation sequence adequately generated?
Allocation concealment.	Describe the method used to conceal the allocation sequence in sufficient detail to determine whether intervention allocations could have been foreseen in advance of, or during, enrolment.	Was allocation adequately concealed?
Blinding of participants, personnel and outcome assessors <i>Assessments should be made for each main outcome (or class of outcomes).</i>	Describe all measures used, if any, to blind study participants and personnel from knowledge of which intervention a participant received. Provide any information relating to whether the intended blinding was effective.	Was knowledge of the allocated intervention adequately prevented during the study?
Incomplete outcome data <i>Assessments should be made for each main outcome (or class of outcomes).</i>	Describe the completeness of outcome data for each main outcome, including attrition and exclusions from the analysis. State whether attrition and exclusions were reported, the numbers in each intervention group (compared with total randomized participants), reasons for attrition/exclusions where reported, and any re-inclusions in analyses performed by the review authors.	Were incomplete outcome data adequately addressed?
Selective outcome reporting.	State how the possibility of selective outcome reporting was examined by the review authors, and what was found.	Are reports of the study free of suggestion of selective outcome reporting?
Other sources of bias.	State any important concerns about bias not addressed in the other domains in the tool. If particular questions/entries were pre-specified in the review's protocol, responses should be provided for each question/entry.	Was the study apparently free of other problems that could put it at a high risk of bias?

Appendix 4: Newcastle-Ottawa quality assessment scale for cohort studies.

Assessment of risk of bias for individual observational studies

Observational studies were assessed for risk of bias using the above criteria in Table 9 and also the Newcastle-Ottawa Quality Assessment Scale (NOS) (Wells 2009). The NOS is a validated scale from 0 to 9 that uses a 'star rating system' and assesses quality of cohort and case-control studies in 3 main areas: selection of study groups, comparability of study groups and ascertainment of exposure or outcome.

Table: Newcastle-Ottawa quality assessment scale for cohort studies.

Note: A study can be awarded a maximum of one star (★) for each numbered item within the Selection and Outcome categories. A maximum of two stars can be given for Comparability.

Selection

- 1) Representativeness of the exposed cohort
 - a) truly representative of the average _____ (describe) in the community ★
 - b) somewhat representative of the average _____ in the community ★
 - c) selected group of users, eg nurses, volunteers
 - d) no description of the derivation of the cohort
- 2) Selection of the non exposed cohort
 - a) drawn from the same community as the exposed cohort ★
 - b) drawn from a different source
 - c) no description of the derivation of the non exposed cohort
- 3) Ascertainment of exposure
 - a) secure record (eg surgical records) ★
 - b) structured interview ★
 - c) written self report
 - d) no description
- 4) Demonstration that outcome of interest was not present at start of study
 - a) yes ★
 - b) no

Comparability

- 1) Comparability of cohorts on the basis of the design or analysis
 - a) study controls for _____ (select the most important factor) ★
 - b) study controls for any additional factor ★ (This criteria could be modified to indicate specific control for a second important factor.)

Outcome

- 1) Assessment of outcome
 - a) independent blind assessment ★
 - b) record linkage ★
 - c) self report
 - d) no description
- 2) Was follow-up long enough for outcomes to occur
 - a) yes (select an adequate follow up period for outcome of interest) ★
 - b) no
- 3) Adequacy of follow up of cohorts
 - a) complete follow up - all subjects accounted for ★
 - b) subjects lost to follow up unlikely to introduce bias - small number lost - > ____ % (select an adequate %) follow up, or description provided of those lost)
 - c) follow up rate < ____ % (select an adequate %) and no description of those lost
 - d) no statement

Appendix 5: GRADE approach to assessing the quality of evidence across studies.

Assessment of quality of evidence across studies

The quality of evidence across a body of evidence was assessed with the GRADE approach, defining the quality of evidence for each outcome as, “the extent to which one can be confident that an estimate of effect or association is close to the quantity of specific interest” (Higgins 2008). The quality rating across studies has four levels: high, moderate, low or very low. Randomised trials are categorized as high quality but can be downgraded; similarly, observational studies can be upgraded. Factors that decrease the quality of evidence include limitations in design, indirectness of evidence, unexplained heterogeneity or inconsistency of results, imprecision of results, or high probability of publication bias. Factors that can increase the quality level of a body of evidence include a large magnitude of effect, if all plausible confounding would reduce a demonstrated effect and if there is a dose-response gradient.

Table: GRADE approach to assessing the quality of evidence across studies.

Quality of Evidence (summary score)	Study Design	Downgrading Factors	Upgrading Factors
High = Further research is very unlikely to change our confidence in the estimate of effect.	Randomized trials or valid accuracy studies for diagnostic tests begin with a score of High	Study limitations: -1 Serious -2 Very serious	Large effect +1 Large +2 Very Large
Moderate = Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.		Inconsistency: -1 Serious -2 Very serious	All plausible confounding +1 Would reduce a demonstrated effect +1 Would suggest a spurious effect when results show no effect
Low = 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.	Observational studies or indirect accuracy studies for diagnostic tests begin with a score of Low .	Indirectness: -1 Serious -2 Very serious	
Very low = Any estimate of effect is very uncertain.		Imprecision: -1 Serious -2 Very serious	Dose-response gradient +1 if Present
		Publication Bias -1 Likely -2 Very likely	

Note: We will specifically consider whether evidence directly addresses low-resource settings in assessing quality of evidence. If the question being addressed only has evidence from high-resource settings, the quality of evidence will be downgraded by -1 for **Indirectness**.

Prepared by Larry William Chang, MD, MPH¹, Jamal Harris, MD², and Eliza Humphreys, MD, MPH³

¹ Johns Hopkins School of Medicine, Baltimore, MD USA;

² Center for AIDS Prevention Studies, University of California, San Francisco, CA USA;

³ Cochrane Collaborative Group on HIV/AIDS, University of California, San Francisco, CA USA