WHAT TO START

UCSF/CDC/WHO Systematic Review

Strategies for determining what initial antiretroviral therapy regimens to start among children and adults living with HIV in low-resource settings

BACKGROUND

Globally, there were 33 million people living with HIV in 2007, the majority of whom resided in sub-Saharan Africa (UNAIDS 2008). Since its introduction in 1996, highly active antiretroviral therapy (ART) has markedly reduced the morbidity and mortality of patients with HIV/AIDS (Palella 1998; Holtgrave 2005), and significant public and private resources have been devoted over the last 5 years to rapid scale-up efforts in low- and middle-income countries to provide access to first-line ART (UNAIDS 2008; Bendavid 2009). It is estimated that approximately 5 million (low estimates) to 10 million people (high estimates) from low- and middle-income countries will need ART by 2010 (UNAIDS/WHO 2006).

The World Health Organization (WHO), in collaboration with the Joint United Nations Programme on HIV/AIDS (UNAIDS), publishes guidelines on the use of antiretroviral therapy in adults, adolescents and children, which currently recommend standard initial treatment of two nucleoside reverse transcriptase inhibitors (NRTI) and a non-nucleoside reverse transcriptase inhibitor (NNRTI) (Gilks 2006; WHO 2006b). Guidelines were first published in 2002 (WHO, 2002) and updated in 2006 (WHO 2006b) and guidelines provided on scale up in 2003 (WHO, 2003). Additionally, ART guidelines have been published separately for pregnant and lactating women, infants and children (WHO 2004, WHO 2006a, WHO 2006b) and for general care and treatment of adolescents and adults with HIV infection (WHO 2006d, WHO 2008). A paediatric version of this latter guideline is in development. However, given the rapidly emerging scientific understanding of HIV treatment and care and the dynamic scale-up efforts in resource-limited settings, the WHO guidelines are updated routinely every few years. To the greatest extent possible, these guidelines are based on evidence of efficacy and effectiveness. Systematic reviews serve as the basis for compiling and assessing the evidence upon which these recommendations are updated. The current review represents a collaborative effort between the Cochrane Collaborative Review Group on HIV Infection and AIDS, the University of California, San Francisco (UCSF), the School of Public Health of the University of Minnesota, the U.S. Centers for Disease Control and Prevention (CDC), the South African Medical Research Council and WHO to address through systematic reviews, questions regarding the optimum first-line ART regimen in patients living with HIV in low- and middle-income countries.

OBJECTIVES

To assess the optimum first-line ART regimen for children >1 year old, adolescents and adults living with HIV in low-resource settings.

METHODS
We followed standard Cochrane systematic review methodology.

Criteria for considering studies for this review

Types of studies

Randomised controlled trials, other controlled trials, cohorts that included analysis of comparison of interest and case-control studies that included analysis for the comparison of interest

Types of participants

The study population is children and non-pregnant adolescents and adults diagnosed with HIV who are initiating a first-line antiretroviral therapy regimen

Relevant subpopulations are:
- Children 1-4 years old
- Children 5-13 years old
- Patients with concurrent hepatitis B virus infection
- Patients with concurrent hepatitis C virus infection
- Patients with concurrent tuberculosis

Types of interventions and comparisons

We examined three-drug antiretroviral regimens for the initial treatment of HIV infection. Regimens that used two or fewer drugs or four or more drugs were excluded as were regimens undertaken for pre- or post-exposure prophylaxis and in pregnant or lactating women, who are the subject of a separate review.

1-to-4-year-old children. See Table 1.

Children 5 years old and older, adolescents and adults. See Table 2.

Types of outcome measures

Primary (critical) outcomes
1. Mortality
2. Clinical response to ART
3. Severe adverse events
4. Virologic response to ART
5. Adherence, tolerance, retention

Secondary (important) outcomes
1. Immunologic response to ART
2. Prevention of sexual transmission
3. Development of ART drug resistance

Clinical response to ART. We assessed clinical response by the proportion of participants that progressed either to CDC-defined AIDS (that is stage III to stage IV disease) or who developed a second opportunistic infection or malignancy.
Table 1. Comparisons of interest for 1-to-4-year-old children.

<table>
<thead>
<tr>
<th>Intervention (current guideline or standard practice or “control” baseline) for 1st line</th>
<th>Comparator (&quot;new&quot; comparison intervention) for 1st line</th>
</tr>
</thead>
<tbody>
<tr>
<td>Double nucleoside reverse transcriptase inhibitor (NRTI) plus nevirapine (NVP)</td>
<td>Double NRTI plus efavirenz (EFZ)</td>
</tr>
<tr>
<td>Double NRTI backbone containing zidovudine (AZT)</td>
<td>Double NRTI backbone containing stavudine (d4T)</td>
</tr>
</tbody>
</table>

Table 2. Comparison of interest for children 5 years old and older, adolescents and adults.

<table>
<thead>
<tr>
<th>Intervention (current guideline or standard practice or “control” baseline) for 1st line therapy</th>
<th>Comparator (&quot;new&quot; comparison intervention) for 1st line therapy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Double NRTI plus NVP</td>
<td>Double NRTI plus EFZ</td>
</tr>
<tr>
<td>Double NRTI backbone containing d4T or AZT</td>
<td>Double NRTI backbone containing TDF*</td>
</tr>
<tr>
<td>Double NRTI backbone containing AZT</td>
<td>Double NRTI backbone containing d4T</td>
</tr>
<tr>
<td>Triple NRTI (AZT + lamivudine [3TC] + ABC)</td>
<td>Triple NRTI (AZT+3TC+TDF*)</td>
</tr>
</tbody>
</table>

*TDF is not approved for use in patients <18 years

**Adverse events.** Severe adverse events were classified according to grade 1 to 4 of the Adverse Event Toxicity Scale (Division of AIDS 2004) 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.

**Virologic response to ART.** Virologic response was reported as the proportion of participants that reached a pre-defined concentration of HIV-1 RNA, typically <400 copies/mL or <500 copies/mL, or who suppressed viral replication to non-detectable levels, typically <40 copies/mL. For purposes of meta-analysis we used the lowest reported value.

**Adherence, tolerance, retention.** We defined this variable to be the proportion of study participants that reached the end of the study on their initially assigned regimen. This category, therefore, includes participants whose regimens were altered because of toxicity, those lost to follow-up, those whose regimens were changed because of clinical or virologic failure and those who withdrew from the study for other reasons.

**Immunologic response to ART.** We defined immunologic response to ART as the mean change in the concentration of CD4 lymphocytes from baseline, as expressed in cells/µL. When studies presented median, instead of mean, we used the median values as reported.

**Search methods**

**Scope of search**

With the assistance of the HIV/AIDS Review Group Trials Search Co-ordinator, 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](http://www.mrw.interscience.wiley.com/cochrane/clabout/articles/HIV/frame.html)). We combined the randomised controlled trial (RCT) strategy developed by The Cochrane Collaboration and detailed in the Cochrane Reviewers' Handbook (*Higgins 08*) in combination with terms specific to initiation of antiretroviral therapy.

**Limits.** The searches were performed without limits to language or setting. The searches excluded studies conducted in pregnant or lactating women and infants <1 year of age. The searches were limited to human studies published from 1995 (start of the triple-drug combination antiretroviral therapy era) to the present.

**Electronic searches**

We searched the following electronic databases:

**Journal and trial databases**
- MEDLINE
- EMBASE
- CENTRAL (Cochrane Central Register of Controlled Trials)
- LILACS (Latin American and Caribbean health sciences literature)
- Cochrane HIV/AIDS Group Trials Register
- Web of Science

**Conference databases**
- Aegis
- AIDSearch: AIDSearch covers abstracts from a number of relevant international conferences including the International AIDS Conference, the International AIDS Society Conferences on HIV Pathogenesis, Treatment, and Prevention, the Conference on Retroviruses and Opportunistic Infections, the British HIV Association Conference and the International Congress on Drug Therapy in HIV infection.
- NLM Gateway (for HIV/AIDS conference abstracts before 2005)

We also 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.

**Ongoing trials**

We searched the following prospective trials registers:
- Current Controlled Trials ([www.controlled-trials.com/](http://www.controlled-trials.com/))
- Pan-African Clinical Trials Registry ([www.pactr.org](http://www.pactr.org))

**Other resources**

**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 examine the bibliographies of any systematic reviews, meta-analyses, or current guidelines we identify during the search process.

**Search terms**

Search terms were set broadly to identify not only clinical trials (see below) but also non-randomised trial designs and cohort and other observational studies. For all comparison we used the following search strings:


#2 Search Antiretroviral Therapy, Highly Active[MeSH] OR Anti-Retroviral Agents [MeSH] OR Antiviral Agents[MeSH:No Exp] 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 immune deficiency[tw])) OR ((anti) AND (acquired immuno-deficiency[tw])) OR ((anti) AND (acquired immun*[tw]) AND (deficiency[tw]))

#3 Search #1 and #2

#4 Specific drugs of interest were inserted here (see below)

#5 Search (#3 AND #4) NOT (animals[mh] NOT human[mh])

#6 Search (#3 AND #4) NOT (animals[mh] NOT human[mh]) Limits* Publication date: 1995-2009

We also added the following terms to identify controlled trials, which would have also been identified in the broader search:


#6 Search #3 AND #4 AND #5
#7 Search #3 and #4 and #5 Limits; Publication date: 1995-2009

We used a similar strategy with minor modifications for EMBASE (1996-2000) and EMBASE (2000-2009), the Cochrane Controlled Trials Register, which contains mainly reference information to randomised controlled trials and controlled clinical trials in health care, Lilacs, Aegis, Web of Science, AIDSearch and Gateway.

There was overlap between the references retrieved in each database. All searches were conducted on 9 July 2009.

To fully capture related conference abstracts we searched 1st through 5th Interantional AIDS Society (IAS) Conferences on HIV Pathogenesis, Treatment and Prevention (2001-2009); the 10th through 17th International AIDS Conferences (1994-2008); the 1st through 16th Conferences on Retroviruses and Opportunistic Infections (1994-2009); the United States National HIV Prevention Conference (1999, 2003, 2005); the 7th through 14th British HIV Association Conferences (2001-2008); and the 8th and 9th European AIDS Society Conference (2001, 2003), using these terms: 3TC, ABC, AZT, AZT, d4T, TDF, FTC, NRTI, NNRTI, nucleoside, nucleotide, protease, DLV, EFV, ETR, NVP, APV, ATV, DRV, IDV, LPV, RTV, NFV, TPV, T-20, MRC, Atripla, lamivudine, abacavir, zidovudine, stavudine, zalcitabine, didanosine, emtricitabine, epzicom, kivexa, trizivir, combivir, truvada, delavirdine, efavirenz, nevirapine, amnprenavir, fosamprenavir, atazanavir darunavir, indinavir, lopinavir, ritonavir, saquinavir, tipranavir, enfuvirtide, maraviroc, raltegravir, tenofovir, breast, mother, infant, baby, pregnant, pregnancy, perinatal, postnatal, feeding, breastfeeding, vertical, mtct, pmct, "when to start" OR timing OR ("early" AND "initia").

**AZT vs. d4T comparison.** For all searches we added the following string to identify trials that included d4T and AZT:

#4 Search (ZERIT OR STAVUDINE OR D4T) AND (ZIDOVUDINE OR RETROVIR OR AZT OR AZT)

MEDLINE (1996-2009) via PubMed yielded 916 records in total of which we selected 15 for full article review. EMBASE yielded 515 records, of which we selected 28 for full article review. The Cochrane Library Controlled Trials Register in July 2009 yielded 154 records in total, of which we selected 23 for full article review. NLM Gateway (1995-2009) using the PubMed strategy yielded 234 records, of which we selected 27 for the full review.

**EFV vs. NVP comparison.** For all searches we added the following string to identify trials that included EFV and NVP:

#4 Search (VIRAMUNE OR NEVIRAPINE OR NVP) AND (SUSTIVA OR STOCRIN OR EFAVIRENZE OR EFV)

MEDLINE (1996-2009) via PubMed yielded 571 records in total of which we selected 18 for full article review, EMBASE yielded 318 records, of which we selected 20 for full article review. The Cochrane Library Controlled Trials Register in July 2009 yielded 55 records in total, of which we selected 13 for full article review. NLM Gateway (1995-2009) using the PubMed strategy yielded 83 records, of which we selected seven for the full review.
**TDF vs. d4T or AZT comparison.** For all searches we added the following string to identify trials that included d4T and AZT:

#4 Search ((ZIDOVUDINE OR RETROVIR OR ZDV OR AZT) OR (ZERIT STAVUDINE OR D4T)) AND (VIREAD OR TENOFOVIR OR TDF)

MEDLINE (1996-2009) via PubMed yielded 277 records in total of which we selected 22 for full article review. EMBASE yielded 249 records, of which we selected 15 for full article review. The Cochrane Library Controlled Trials Register in July 2009 yielded 46 records in total, of which we selected 10 for full article review. NLM Gateway (1995-2009) using the PubMed strategy yielded 42 records, of which we selected 11 for the full review.

**TDF vs. ABC comparison.** For all searches we added the following string to identify trials that compared TDF and ABC:

#4 Search (VIREAD OR TENOFOVIR OR TDF) AND (ZIAGEN OR ABACAVIR OR ABC)

MEDLINE (1996-2009) via PubMed yielded 190 records in total of which we selected three for full article review. EMBASE yielded 185 records, of which we selected two for full article review. The Cochrane Library Controlled Trials Register in July 2009 yielded 18 records in total, of which we selected none for full article review. NLM Gateway (1995-2009) using the PubMed strategy. This yielded 268 records, of which we selected none for the full article review.

**Search outcomes**

GR and AS independently conducted the selection of potentially relevant studies by scanning the titles, abstracts, and descriptor terms of all downloaded material from the electronic searches. Irrelevant reports were discarded, and the full article was obtained for all potentially relevant or uncertain reports. GR and AS independently applied the inclusion criteria. NS 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 are ongoing.

**Search yield**

In total, from the 9,571 articles, abstracts, reports, research letter, meta-analyses and reviews identified in the search, we selected 874 for further examination. For the **d4T vs. AZT comparison**, we identified 99 abstracts, articles and published letters for full review. Thirty-three described 10 controlled trials, five were meta-analyses, three were other narrative reviews and 58 reported results from a variety of observational studies. For the **EFV vs. NVP comparison**, we identified 101 abstracts, articles and published letters for full review. Twenty-eight described nine controlled trials, four were meta-analyses, 63 reported results from a variety of observational studies and six were reviews. For the **d4T or AZT vs. TDF comparison**, we identified 58 abstracts, articles and published letters for more detailed scanning and 36 for full review. Eighteen described three controlled trials, two were meta-analyses, 10 reported results from a variety of observational studies and six were reviews. For the **ABC vs. TDF comparison**, we identified seven articles and abstracts that described three separate trials, one observational study and one meta-analysis.
Data extraction and coding

After initial search and article screening, two reviewers independently double-coded and entered information from each selected study onto standardised data extraction forms. 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
- **Interventions details**: Drug names, doses, duration, ancillary testing and monitoring, any other information on adherence or resistance
- **Outcome details**: mortality; response to ART [clinical (AIDS and non-AIDS events), virologic and immunologic]; severe adverse events; development of ART drug resistance; adherence/tolerability/retention; risk of sexual transmission of HIV.

Data analysis and presentation of findings

We used Review Manager 5 provided by the Cochrane Collaboration for statistical analysis and GradePro (GradePro 2008) software to produce Summary of Findings and Evidence Profile tables.

When interventions and study populations were sufficiently similar across different studies, we statistically pooled the outcomes and examined the differences between the two models using both fixed and random-effects models. Since there were no significant differences between the two models, final results are presented using random-effects models. We summarised dichotomous outcomes for effect in terms of risk ratio (RR), risk difference (RD) and number needed to treat (NNT) with their 95% confidence intervals. We summarised continuous outcomes with a weighted mean difference (WMD) and 95% confidence interval. We evaluated observational studies, non-randomised trials and randomised clinical trials separately.

We summarised the quality of evidence for each outcome for which data were available in GRADE Summary of Findings and GRADE Evidence Profile Tables (Guyatt 2008).

Subgroup analysis and investigation of heterogeneity

We examined heterogeneity among trials using the chi-square statistic with a significance level of 0.10 and the I-squared statistic. We interpreted an I-squared estimate greater than 50% as indicating moderate or high levels of heterogeneity and investigated its causes by sensitivity analysis. If heterogeneity persisted, we presented results separately and reported reasons for the observed heterogeneity.

If data were available, we performed sub-group analysis for studies in which participants were co-infected with hepatitis B, hepatitis C or tuberculosis. Heterogeneity was explored using further sub-group analyses by trial quality, setting (middle- or low- versus high-income country) or other sub-groups judged relevant.
Publication bias. We assessed the potential for publication bias using funnel plots. We minimised the potential for publication bias by our comprehensive search strategy that included evaluating published and unpublished literature.

Assessment of risk of bias for individual randomised studies. Application of GRADE (Guyatt 2008) and Cochrane Collaboration tools for risk of bias for each individual study will be applied and presented in summary tables. The GRADE and Cochrane approaches assess risk of bias in individual studies across six domains: sequence generation, allocation concealment, blinding, incomplete outcome data, selective outcome reporting and other potential biases (see Table 3).

Assessment of risk of bias for individual observational studies. Observational studies will be assessed for risk of bias using the above criteria in Table 3 and also the Newcastle-Ottawa Quality Assessment Scale (NOS) shown in Tables 4 and 5 (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.

Assessment of quality of evidence across studies

We assessed the quality of evidence across a body of evidence with the GRADE approach (see Table 6), 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 controlled trials are categorised as high quality but can be downgraded; similarly, other types of controlled trials and observational studies are categorised as low quality but 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>
<thead>
<tr>
<th>Domain</th>
<th>Description</th>
<th>Review authors’ judgment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sequence generation</strong></td>
<td>Describe the method used to generate the allocation sequence in sufficient detail to allow an assessment of whether it should produce comparable groups.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Was the allocation sequence adequately generated?</td>
<td></td>
</tr>
<tr>
<td><strong>Allocation concealment</strong></td>
<td>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.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Was allocation adequately concealed?</td>
<td></td>
</tr>
<tr>
<td><strong>Blinding of participants, personnel and outcome assessors</strong></td>
<td>Assessments should be made for each main outcome (or class of outcomes). 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.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Was knowledge of the allocated intervention adequately prevented during the study?</td>
<td></td>
</tr>
<tr>
<td><strong>Incomplete outcome data</strong></td>
<td>Assessments should be made for each main outcome (or class of outcomes). 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 randomised participants), reasons for attrition/exclusions where reported, and any re-inclusions in analyses performed by the review authors.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Were incomplete outcome data adequately addressed?</td>
<td></td>
</tr>
<tr>
<td><strong>Selective outcome reporting</strong></td>
<td>State how the possibility of selective outcome reporting was examined by the review authors, and what was found.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Are reports of the study free of suggestion of selective outcome reporting?</td>
<td></td>
</tr>
<tr>
<td><strong>Other sources of bias</strong></td>
<td>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.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Was the study apparently free of other problems that could put it at a high risk of bias?</td>
<td></td>
</tr>
</tbody>
</table>
Table 4. 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
Table 5. Newcastle-Ottawa quality assessment scale for case-control studies.

Note: A study can be awarded a maximum of one star for each numbered item within the Selection and Exposure categories. A maximum of two stars can be given for comparability.

### Selection

1) Is the case definition adequate?
   - a) yes, with independent validation ★
   - b) yes, eg record linkage or based on self reports
   - c) no description

2) Representativeness of the cases
   - a) consecutive or obviously representative series of cases ★
   - b) potential for selection biases or not stated

3) Selection of Controls
   - a) community controls ★
   - b) hospital controls
   - c) no description

4) Definition of Controls
   - a) no history of disease (endpoint) ★
   - b) no description of source

### Comparability

1) Comparability of cases and controls 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.)

### Exposure

1) Ascertainment of exposure
   - a) secure record (eg surgical records) ★
   - b) structured interview where blind to case/control status ★
   - c) interview not blinded to case/control status
   - d) written self report or medical record only
   - e) no description

2) Same method of ascertainment for cases and controls
   - a) yes ★
   - b) no

3) Non-Response rate
   - a) same rate for both groups ★
   - b) non respondents described
   - c) rate different and no designation
Table 6. GRADE approach to assessing the quality of evidence across studies.

<table>
<thead>
<tr>
<th>Quality of Evidence (summary score)</th>
<th>Study Design</th>
<th>Downgrading Factors</th>
<th>Upgrading Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High (4)</strong> = Further research is very unlikely to change our confidence in the estimate of effect.</td>
<td>Randomised trials or valid accuracy studies for diagnostic tests begin with a score of High (4)</td>
<td>Study limitations: -1 Serious -2 Very serious</td>
<td>Large effect +1 Large +2 Very Large</td>
</tr>
<tr>
<td><strong>Moderate (3)</strong> = Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.</td>
<td></td>
<td>Consistency: -1 Serious -2 Very serious</td>
<td>Plausible confounding would change the effect +1</td>
</tr>
<tr>
<td><strong>Low (2)</strong> = 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.</td>
<td>Observational studies or indirect accuracy studies for diagnostic tests begin with a score of Low (2).</td>
<td>Directness: -1 Serious -2 Very serious</td>
<td>Dose-response gradient +1 if present</td>
</tr>
<tr>
<td><strong>Very low (1)</strong> = Any estimate of effect is very uncertain.</td>
<td></td>
<td>Precision: -1 Serious -2 Very serious</td>
<td></td>
</tr>
</tbody>
</table>

**Note:** We specifically considered whether evidence directly addressed low- and middle-income country 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 lack of directness.
DESCRIPTION OF THE STUDIES

Table 7 contains a list of all randomised controlled studies included in this review.

Table 7. Completed and ongoing randomised trials.

<table>
<thead>
<tr>
<th>d4T vs. AZT</th>
<th>EFV vs. NVP</th>
<th>TDF vs. d4T or AZT</th>
<th>TDF vs. ABC</th>
</tr>
</thead>
</table>

*Ongoing studies

d4T vs. AZT

We identified 101 abstracts, articles and published letters for full review. All but one met our inclusion criteria. Five of the reviewed studies were meta-analyses (Carr 2009; Enanoria 2004; Hill 2009; Moyle 2004; Siegfried 2006), and three were narrative reviews (Gulick 1998; Havlir 1998; Wood 2006). Thirty-three articles and abstracts described nine controlled trials, and 58 reported results from a variety of observational studies.

Randomised controlled trials

Two basic NRTI-backbones were compared in the nine controlled trials we identified, AZT+3TC vs. d4T+3TC and AZT+3TC vs. d4T+ddl (Table 8). We present these two sets of comparisons separately below. We excluded data from one study because it focused solely on subcutaneous fat thickness as its sole outcome (Shlay 2008).

AZT+3TC vs. d4T+3TC. For the AZT+3TC vs. d4T+3TC comparison, we identified 16 articles abstracts from five randomised controlled trials, including OzCombo 1 (Amin 2003; Carr 1999a; Carr 1999b; Carr 2000; Hudson 1998), ESS4002 (Kumar 2003; Kumar 2004; Kumar 2006), Li (2008), START I (Gulick 1998; Pavia 2002; Squires 1998; Squires 1999; Squires 2000a; Squires 2000b) and the Tshepo Study (Bussmann 2009; Wester 2007).

OzCombo 1 was a randomised, open-label, three-arm study that examined three NRTI backbones (AZT+3TC, d4T+3TC and d4T+ddl) in combination with indinavir (IDV) among ART-naïve adults. It was conducted in Australia and primarily funded by the Australian National Council and AIDS and Related Diseases. Patients were followed for 52 weeks, and the primary
Table 8. Characteristics of included studies, AZT vs. d4T.

<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Participants</th>
<th>Interventions</th>
<th>Outcomes</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carr 2000</td>
<td>Randomised, open label, multicentre</td>
<td>106 ART-naïve adults in Australia</td>
<td>1. AZT+3TC+IDV</td>
<td>Undetectable viral load at 52 weeks, drug toxicity requiring regimen change</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2. d4T+3TC+IDV</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3. d4T+ddI+IDV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eron 2000</td>
<td>Randomised, open label, multicentre</td>
<td>205 ART-naïve adults in USA and Puerto Rico</td>
<td>1. AZT+3TC+IDV</td>
<td>HIV-1 RNA &lt;50 copies/mL at 48 weeks</td>
<td>No SD for CD4 reported</td>
</tr>
<tr>
<td>START II</td>
<td></td>
<td></td>
<td>2. d4T+ddI+IDV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>French 2002</td>
<td>Randomised, open label, multicentre</td>
<td>65 ART-naïve adults in Australia</td>
<td>1. AZT+3TC+NVP</td>
<td>Undetectable viral load, drug toxicity requiring regimen change at 52 weeks</td>
<td>Used real-time viral load as opposed to stored samples</td>
</tr>
<tr>
<td>OzCombo 2</td>
<td></td>
<td></td>
<td>2. d4T+3TC+NVP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arm 1 vs. 2</td>
<td></td>
<td></td>
<td>3. d4T+ddI+NVP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>and Arm 1 vs. 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gathe 2002</td>
<td>Randomised, open label, multicentre</td>
<td>511 ART-naïve adults with ≥200 CD4 cells/µL in North America, Europe, South American, Russia, South Africa, Australia</td>
<td>1. AZT+3TC+NVP</td>
<td>HIV-1 RNA &lt;400 copies/mL at 48 weeks at 48 weeks</td>
<td>No SD for CD4 reported</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2. d4T+ddI (enteric coated)+NVP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Geijo Martinez 2006</td>
<td>Randomised, open label, multicentre</td>
<td>86 ART-naïve adults in Spain</td>
<td>1. AZT+3TC+IDV</td>
<td>Decline in HIV-1 RNA at 48 weeks</td>
<td>No SD for CD4 reported</td>
</tr>
<tr>
<td>GEMCEI</td>
<td></td>
<td></td>
<td>2. AZT+3TC+RTV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arms 1 +2 vs. 3</td>
<td></td>
<td></td>
<td>3. d4T+ddI+IDV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ 4</td>
<td></td>
<td></td>
<td>4. d4T+ddI+RTV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kumar 2006</td>
<td>Phase IV, randomised open-label parallel group</td>
<td>261 ART-naïve adults in Dominican Republic, Guatemala, Panama, Puerto Rico, USA</td>
<td>1. AZT+3TC+ABC</td>
<td>Change in LDL cholesterol from baseline, virologic failure at 96 weeks</td>
<td>Used HIV-1 RNA &lt;50 copies/mL for outcome; SAEs not reported; no SD for CD4 reported</td>
</tr>
<tr>
<td>ESS40002</td>
<td></td>
<td></td>
<td>2. AZT+3TC+NFV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arm 2 vs. 3</td>
<td></td>
<td></td>
<td>3. d4T+3TC+NFV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Li 2008</td>
<td>Randomised, open-label, multicentre</td>
<td>198 ART-naïve adults in China</td>
<td>1. AZT+ddI+NVP</td>
<td>HIV-1 RNA &lt;50 copies/mL at 52 weeks</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2. 3TC+d4T+NVP</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3. 3TC+AZT+NVP</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
squares 2000

START I

Randomised, open-label, multicentre

204 ART-naive adults with >200 CD4 cells/μL in USA and Puerto Rico

1. AZT+3TC+IDV

2. d4T+3TC+IDV

HIV-1 RNA <500 copies/mL at 24 and 48 weeks

Only clinical SAEs reported; no SD for CD4 reported

SAE, severe adverse events; SD, standard deviation

endpoints were time to full viral suppression (HIV-1 RNA <50 copies/mL) and the proportion of patients with drug toxicities sufficiently severe to warrant dose adjustment or change. The primary results are found in Carr 2000. Carr 1999a, Carr 1999b and Hudson 1998 are conference presentations for various stages of data, and Amin 2003 is a follow-up study of patients enrolled in OzCombo 1 and OzCombo 2.

ESS4002 was an industry-sponsored, phase IV, parallel group study of random-dose ABC-3TC-AZT (Trizivir®), random-dose AZT+3TC (Combivir®)+nelfinavir (NFV) and d4T+3TC+NFV in ART-naive patients in the Dominican Republic, Guatemala, Panama, Puerto Rico and the United States of America (Kumar 2006). Its primary focus was on lipid metabolism with a primary outcome of change from baseline for low-density lipoprotein cholesterol. The study also collected efficacy endpoints, including HIV-1 RNA <50 and <400 copies/mL at 96 weeks.

Li (2008) and colleagues compared three different NRTI backbones, AZT+ddl+NVP, AZT+3TC+NVP and d4T+3TC+NVP, in a randomised controlled trial in China that enrolled 198 patients with between 100 and 350 CD4 cells/μL and HIV-1 RNA >500 copies/mL. Its primary endpoints were immunologic, virologic and clinical response.

START I was a large industry-sponsored, randomised, open-label trial conducted in Puerto Rico and the United States of America that examined AZT+3TC+IDV and d4T+3TC+IDV in ART-naive adults (Squires 2000b). Its primary endpoints were suppression of viral replication to HIV-1 RNA <50 and <500 copies/mL at 24 and 48 weeks.

The Tshepo Study is an on-going study in Botswana that is employing a 3 x 2 x 2 factorial design to study three different NRTI backbones (AZT+3TC, d4T+3TC and AZT+ddl), two NNRTIs (EFV and NVP) and two different adherence strategies in ART-naive adults infected with subtype C virus (Wester 2007). In April 2006, the Data Safety Monitoring Board recommended discontinuation of the AZT+ddl arm for reasons of inferiority (virological failure and acquired resistance). The one published analysis of this study (Bussman 2009) compares the AZT+ddl arm to the AZT+3TC and the d4T+AZT arms but not the AZT+3TC arm to the d4T+3TC arm, the comparison of interest for this review. Thus, until additional data are published, this trial does is non-contributory.

AZT+3TC vs. d4T+ddl. For the AZT+3TC vs. d4T+ddl comparison, we identified 20 articles abstracts from six randomised controlled trials, including ACTG 384 and its substudies A5005s and A 5152s (Dubé 2004; Dubé 2005; Dubé 2007a; Dubé 2007b; Gandhi 2006;
Mulligan 2006; Robbins 2003; Shafer 2003; Smeaton 2001; Torriani 2008), the GEMCEI study (Geijo Martinez 2006), OzCombo 2 (Amin 2003; French 2002; Hudson 1998), START II (Eron 1998a; Eron 1998b; Eron 2000; Squires 2000a) and another study of enteric-coated ddl conducted by Gathe (2002).

ACTG 384 and substudies A5005s and A5152s were a group of studies conducted in Italy and the United States of America and sponsored by the Instituto Superiore di Sanità and the National Institutes of Health (Smeaton 2001). The main study was a multicentre, partially blinded, factorial trial of NRTI therapies as sequential therapy (that is, with predetermined second-line regimens for use following virologic failure) with either NVP or EFV. We examined the pre-virologic failure arms, comparing AZT+3TC+EFV with d4T+ddl+EFV and AZT+3TC+NVP with d4T+ddl+NVP. The endpoints were virologic failure, defined as a decrease in HIV-1 RNA <10-fold by week 8 on therapy or an increase of >10-fold above a nadir of <200 copies/mL after week 24, and toxicity-related failure. Outcomes were measured through 144 weeks of follow up. The main results paper is Robbins 2003; Gandhi 2006 reported on immunologic recovery, and Shafer 2003 reported on efficacy of pre-planned second-line therapy. Other publications and presentations include studies on specific toxicities, including dyslipidaemias and fat redistribution (Dubé 2004; Dubé 2007a; Dubé 2005; Dubé 2007b; Mulligan 2006) and endothelial function (Torriani 2008).

Gathe (2002) and colleagues studies the use of once-daily enteric-coated ddl in a multicentre trial conducted in North America, South America, Europe, Russia, South Africa and Australia and sponsored by Bristol-Myers Squibb. The basic comparison involved AZT+3TC+NFV and d4T+enteric coated ddl+NFV. The main outcome variable was HIV-1 RNA <400 copies/mL and <50 copies/mL at 48 weeks.

Geijo Martinez (2006) and colleagues conducted a four-arm study that compared two NRTI backbones, AZT+3TC and d4T+ddI, and two protease inhibitors, IDV and ritonvair (RTV), in combination in ART-naive patients. The study was conducted in Spain and funded by local government. The primary endpoints were decline in HIV-1 RNA compared to baseline, and the follow-up period was 48 weeks.

OzCombo 2 was a randomised, open-label, three-arm study that examined three NRTI backbones (AZT+3TC, d4T+3TC and d4T+ddI) in combination with NVP among ART-naive adults. It was conducted in Australia and primarily funded by the Australian National Council and AIDS and Related Diseases. Patients were followed for 52 weeks, and the primary endpoints were time to full viral suppression (HIV-1 RNA <50 copies/mL) and the proportion of patients with drug toxicities sufficiently severe to warrant dose adjustment or change. The primary results are found in French 2002. Hudson 1998 is conference presentation of initial data from OzCombo 1 and 2, and Amin 2003 is a follow-up study of patients enrolled in OzCombo 1 and OzCombo 2.

START II was a multicentre, randomised, open-label study that compared AZT+3TC+IDV with d4T+ddl+IDV in ART-naive patients in Puerto Rico and the United States of America. The study was sponsored by Bristol-Myers-Squibb. The primary endpoints were HIV-1 RNA ≤50 copies/mL and <500 copies/mL and changes in CD4 cell counts at 48 weeks. The primary results are found in Eron 2000; two conference abstracts presented preliminary results (Eron 1998a; Eron 1998b), and Squires (2000b) examined differences in response by gender from both START studies.

Observational studies


Ongoing studies

We identified one study that was not yet completed that is comparing AZT and d4T-containing regimens, although not as a primary comparison. The Optimizing Pediatric HIV-1 Treatments in Infants with Prophylactic Exposure to Nevirapine study is a randomised, open-label, parallel assignment, efficacy study being conducted in Nairobi, Kenya (NCT00427297). The study has five first-line ART treatment arms, including AZT+3TC+NVP, d4T+3TC+NVP, AZT+3TC+ABC and d4T+3TC+ABC, and the subjects will be 6-18 month olds who were exposed to NVP prophylaxis. Its primary endpoints will be mortality, CD4 percentage and viral suppression in three NVP-containing arms compared to two triple NRTI arms over 24 months. A major predictor of response will be baseline NVP resistance. The anticipated enrolment is 100 children, and the study completion date is estimated to be December 2010.
EFV vs. NVP

We reviewed 101 studies, which all met our inclusion criteria, including 21 reports of randomised controlled trials, four systematic reviews (Neuwelt 2003; Sheran 2005; Siegfried 2006; Torre 2001), six narrative reviews (Arribas 2001; Moyle 2000; Moyle 2003; Skowron 2001; Stern 2004; van Leth 2006d), seven non-randomised trials and 63 observational studies.

Randomised controlled trials

Twenty-one abstracts and articles reported results from six randomised controlled trials, including the 2NN trial (Ananworanich 2005; Kappelhoff 2005a; Kappelhoff 2005b; van Leth 2004a; van Leth 2004b; van Leth 2004c; van Leth 2004d; van Leth 2005a; van Leth 2005b; van Leth 2006a; van Leth 2006b; van Leth 2006c; Wit 2007), the Tshepo trial (Ayala Gaytan 2004), a Mexican trial (Bussman 2009), the Spanish Efavirenz vs. Nevirapine Comparison (SENC) trial (Núñez 2002), a Senegalese study (Sow 2006) and the FIRST or CPCRA 058 study (MacArthur 2001; van den Berg-Wolf 2006; van den Berg-Wolf 2008) (Table 9).

The 2NN study (main results are in van Leth 2004c) was a randomised, four-arm, open label, multicentre trial in which 1 216 antiretroviral naïve patients were assigned to receive d4T and 3TC plus one of the following: NVP 400 mg daily, NVP 200 mg twice daily, EFV 600 mg daily or NVP 400 mg + EFV 600 mg daily. The primary endpoint was the proportion of patients with treatment failure, defined as <10-fold decline in plasma HIV-1 RNA in the first 12 weeks or two consecutive measurements of more than 50 copies/mL from week 24 onwards, disease progression (new Centers for Disease Control and Prevention grade C event or death) or change of allocated treatment. Analyses were by intention to treat. Other published papers presented results on incidence and risk factors for rash (Ananworanich 2005), efficacy by initial CD4 stratum or HIV-1 RNA plasma viral load (van Leth 2004a; van Leth 2005a), quality of life (van Leth 2004b), rates of early plasma HIV-1 RNA decline (van Leth 2005b), lipid profile changes (van Leth 2004d; van Leth 2006a), relationship between plasma concentration of EFV and NVP and adverse events (Kappelhoff 2005a) and pharmacokinetics of EFV and NVP (Kappelhoff 2005b; van Leth 2006b). The 2NN Study continued to follow patients in a cohort study until 144 weeks (NCT00127972), which was completed in April 2007. No data have as yet been published from this continuation study.

The Tshepo study (Bussman 2009) is a randomised, open-label study of 650 antiretroviral-naïve patients conducted in Botswana. It used a 3 x 2 x 2 factorial design to compare the efficacy and tolerability of three different NRTI backbones (AZT+3TC, AZT+didanosine [ddI] and d4T+3TC), two different NNRTIs (EFV and NVP) and two different adherence strategies. Patients were stratified by CD4 count (≤200 and 201-350 cell/µL), and the primary endpoints were virologic failure and antiretroviral resistance. Following DSMB review in April 2006, all AZT+ddl-containing arms were discontinued due to inferiority.

Ayala Gaytan (2004) and colleagues conducted a randomised, open-label, pilot trial of AZT+3TC+EFV compared to AZT+3TC+NVP among 58 HIV-infected antiretroviral-naive adults with CD4 counts <350 cells/µL and detectable plasma HIV-1 RNA >55,000 copies/mL and/or AIDS defining conditions at a single hospital in Mexico City. Endpoints were the proportion of patients achieving plasma HIV-1 RNA <400 copies/mL at 24 and 48 weeks, changes of CD4 counts, emergence of AIDS-defining conditions or death and development of antiviral-related toxicities. Analysis was by intention to treat.
2NR study (Manosuthi 2009) examined the efficacy of EFV- (600 mg daily) and NVP- (400 mg daily) based antiretroviral therapy among 142 HIV-infected Thai patients with concurrent tuberculosis being treated with rifampicin. Rifampicin is a commonly used first-line antimycobacterial drug in many countries and has been previously shown to increase the metabolism of NVP through a cytochrome P450 mechanism. Endpoints of the study were virologic failure at 48 weeks and CD4 counts. Subgroup analysis also examined patients who achieved subtherapeutic plasma levels of EFV and NVP. Note that this study is asking a different question than the other trials, which are examining EFV and NVP in the absence of rifampicin.

The Spanish Efavirenz vs. Nevirapine Comparison (SENC) trial (Núñez 2002) was the first published head-to-head comparison of EFV and NVP in 67 treatment-naïve patients; the NRTI backbone used in the study was ddI plus d4T. Inclusion criteria were >100 CD4 cells/µL and plasma HIV-1 RNA <100 000 copies/mL. Endpoints were the proportion of patients reaching plasma HIV-1 RNA <50 copies/mL and/or developing NNRTI-related toxicities leading to drug discontinuation at 48 weeks.

Sow (2006) and colleagues conducted a small 70-patient trial in Senegal, which has been published only in abstract form. They compared AZT+3TC+NVP and AZT+3TC+EFZ with 18-month endpoints of efficacy, safety, adherence and quality of life.

The FIRST Study (CPCRA 058) is a complex long-running study that has at its core a comparison of EFV and NVP (MacArthur 2001; van den Berg-Wolf 2006; van den Berg-Wolf 2008). There are two groups, one participating in a randomised controlled trial (N=228) and the other participating in a cohort study (N=423). In the randomised trial there were three possible assignments: a protease inhibitor plus an NRTI (not further analysed in this review), a protease inhibitor plus an NRTI plus an NNRTI (3-class therapy, not further analysed in this review) and two NRTIs plus an NNRTI. In each NNRTI-containing arm, EFV or NVP was assigned randomly. For the randomised substudy, the primary endpoint was plasma HIV-1 RNA >50 copies/mL after eight months of therapy or death. Genotypic resistance testing was done at the time of virological failure, defined as plasma HIV-1 RNA >1,000 copies/mL at or after 4 months.

Non-randomised controlled studies

We identified an additional seven articles from three controlled trials from China (Han 2005), Germany (Bruck 2008; Hartmann 2005a; Hartmann 2005b) and Italy (Manfredi 2004; Manfredi 2005; Manfredi 2006).

Han (2005) and colleagues reported on a community-controlled study undertaken in three provinces in China. Each province received a different regimen, so this public health intervention in essence is a community-controlled trial. Assignment of which province received treatment was presumably non-random, although it was not explicitly stated. Endpoints were viral suppression and CD4 recovery along with extensive surveillance for NRTI and NNRTI resistance mutations.

The NEEF Cohort (Hartmann 2005b) was a German multicentre, non-randomised, prospective, two-arm comparison of 662 antiretroviral-naïve patients initiating an EFV- or a NVP-containing regimen. Endpoints were plasma HIV-1 RNA <50 copies/mL and CD4 count at 48 weeks. A second paper (Hartmann 2005a) reported on the incidence of rashes in these two groups. An earlier report from the same group (Bruck 2008) describes the incidence of
Table 9. Characteristics of included studies, EFV vs. NVP.

<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Participants</th>
<th>Interventions</th>
<th>Outcomes</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ayala Gaytan 2004</td>
<td>Randomised, open label</td>
<td>58 ART-naïve adults from single center in Monterrey, Mexico</td>
<td>1. AZT+3TC+NVP 2. AZT+3TC+EFV</td>
<td>Undetectable viral load at 48 weeks</td>
<td>Used &lt;400 copies/mL as undetectable, no SD for CD4 reported</td>
</tr>
<tr>
<td>Manosuthi 2009 2NR</td>
<td>Randomised, open label</td>
<td>142 ART-naïve adults with ≥2 months of TB therapy from Thailand</td>
<td>1. d4T+3TC+NVP 2. d4T+3TC+EFV</td>
<td>HIV-1 RNA &lt;50 copies/mL at 48 weeks</td>
<td>No SD for CD4 reported, primary outcome C₁₂</td>
</tr>
<tr>
<td>Núñez 2002 SENC</td>
<td>Randomised, open label</td>
<td>67 ART-naïve adults from Spain</td>
<td>1. d4T+ddI+NVP 2. d4T+ddI+EFV</td>
<td>HIV-1 RNA &lt;50 copies/mL at 48 weeks; grade 3/4 adverse events</td>
<td>Only gave total CD4 at 48 weeks without SD</td>
</tr>
<tr>
<td>Sow 2006</td>
<td>Randomised controlled trial</td>
<td>70 ART-naïve patients from Senegal</td>
<td>1. AZT+3TC+NVP 2. AZT+3TC+EFV</td>
<td>Efficacy, safety, compliance, quality of life, adverse events</td>
<td>Reported mean decrease in viral load, no SD for CD4 reported</td>
</tr>
<tr>
<td>van den Berg-Wolf 2008</td>
<td>Randomised multicentre</td>
<td>228 ART-naïve adolescents aged ≥13 years old from USA and Puerto Rico (NNRTI 058 substudy)</td>
<td>1. Two NRTIs+ NVP 2. NRTI+one PI+NVP 3. Two NRTIs + EFV 4. One NRTI + one PI + EFV</td>
<td>Virologic failure, clinical progression or death, grade 4 toxicity, acquired drug resistance, discontinuation due to toxicity at 8 months</td>
<td>8 months follow up, no SD for CD4 reported</td>
</tr>
<tr>
<td>van Leth 2004c 2NN</td>
<td>Randomised controlled multicentre</td>
<td>1 216 ART-naïve adults from Argentina, Australia, Belgium, Brazil, Canada, France, Germany, Greece, Ireland, Italy, Poland, Portugal, South Africa, Switzerland, Thailand, UK, USA</td>
<td>1. d4T+3TC+NVP 2. d4T+3TC+NVP bid 3. d4T+3TC+EFV 4. d4T+3TC+NVP+EFV</td>
<td>Virologic failure at 48 weeks, clinical progression or death, discontinuation</td>
<td>No SD for CD4 reported; CD4 change reported as median rather than mean</td>
</tr>
</tbody>
</table>

qd, daily; bid, twice daily; SD, standard deviation
hepatotoxicity in the first 296 study enrollees.

Manfredi (2004) and colleagues conducted an open-label, prospective, 18-month evaluation of the safety and tolerability of EFV and NVP. One hundred fifty-eight patients were antiretroviral-naive; other study groups included patients switching from protease inhibitors and patients entering a salvage regimen. Endpoints were virologic and immunologic. Additional publications from this group include reports on hepatotoxicity in the overall study (Manfredi 2005) and a specific analysis of NVP hepatotoxicity in women with >250 CD4 cells/µL at baseline from an expanded clinical series (N=742) (Manfredi 2006).

Observational studies

Ongoing studies

Three additional studies are currently ongoing in Cameroon, India, Mozambique and Senegal. Protocols remain open, as well, for 2NN, 2NR and CPCRA 058.

Comparison of Nevirapine and Efavirenz for the Treatment of HIV-TB co-Infected Patients (ANRS 12146 CARINEMO, (NCT00495326), conducted by the Agence nationale de recherches sur le SIDA et les hépatites virales (ANRS) is a two-arm randomised controlled, non-inferiority trial comparing EFV and NVP among 570 antiretroviral-naïve patients coinfected with tuberculosis and being treated with rifampicin in Mozambique. It has virological and clinical endpoints and has an estimated completion date of March 2011.

A phase III clinical trial (NCT00332306) being conducted in India is similarly evaluating the safety and efficacy of EFV and NVP with a ddI+d4T backbone in a randomised controlled study of patients who have completed two months of tuberculosis therapy. They are seeking to recruit 180 patients, and the anticipated completion date is December 2011.

ANRS 12115 DAYANA (NCT00573001) is a four-arm Phase III randomised, open-label evaluation of four new simplified antiretroviral treatment regimens being conducted in Senegal and Cameroon. The intervention arms are random dose combination TDF-FTC-EFV (Atripla®), random dose combination TDF-FTC (Truvada®) + NVP, random dose combination TDF-FTC+AZT and TDF+LPV/r. The first two arms will allow a head-to-head comparison of EFV and NVP, albeit with a different NRTI backbone than those employed in the other studies reviewed here. The primary outcome measure will be virologic, and patients will be followed for up to 96 weeks; it will also assess adverse events, immunological improvement, plasma concentrations of drugs and quality of life. Estimated enrolment is 120 patients, and the study is scheduled to be completed in July 2010.

TDF vs. d4T or AZT

We identified 36 abstracts, articles and published letters for full review. There were three meta-analyses (Anonymous 2005; Carr 2009; Parienti 2009) and five narrative reviews (Clotet 2008; Frampton 2005; Frampton 2006; Makinson 2008; Masho 2007), 18 publications described three controlled trials and 10 reported results from a variety of observational studies.

Randomised controlled trials

Two basic NRTI-backbones were compared in the three randomised controlled trials we identified: AZT+3TC vs. TDF+3TC and AZT+3TC vs. TDF+FTC. We present these two sets of comparisons separately below (Table 10).

AZT+3TC vs. TDF+3TC. For the AZT+3TC vs. TDF+3TC comparison, we identified 14 articles abstracts from two randomised controlled trials, including Study 903 (Cassetti 2007; Gallant 2002; Gallant 2003; Gallant 2004a; Gallant 2004b; Gallant 2008; Izzedine 2005;...
Margot 2006; McGowan 2001; Pozniak 2003; Staszewski 2002; Staszewski 2003a; Staszewski 2003b) and the DAUFIN study (Rey 2009).

Study 903. Study 903 was a large industry-sponsored randomised controlled multicentre non-inferiority trial conducted in Europe, Latin America and North America that compared d4T+3TC+EFV with TDF+3TC+EFV. It was blinded for the d4T and TDF but not the other components. Its primary efficacy end point was the proportion of patients with HIV-1 RNA levels of <400 copies/mL at week 48. Secondary efficacy end points were the proportion of patients with HIV RNA levels of less than 50 copies/mL and change in CD4 cell count from baseline at weeks 48, 96, and 144. The main 144-week results were published in Gallant 2004a, but results from the 336-week single-arm extension in Latin American sites were also published in Cassetti 2007. Aside from a variety of interim reports in conference abstracts (Gallant 2002; Gallant 2003; Gallant 2004b; Pozniak 2003; Staszewski 2002; Staszewski 2003a; Staszewski 2003b), the study has also reported on a variety of longer-term safety and efficacy issues, including renal toxicity (Gallant 2008; Izzedine 2005), antiretroviral resistance (Margot 2006) and osteopoenia and osteoporosis (McGowan 2001).

DAUFIN Study. The DAUFIN Study was a small Spanish study that compared twice-daily AZT+3TC+NVP with once-daily TDF+3TC+NVP (Rey 2009). Initially, the investigators planned to enroll 250 patients, but the trial was stopped after an unplanned interim analysis with the first 71 patients. Eight (22%) of 36 patients in the once-daily TDF+3TC+NVP arm had not responded to therapy and an additional two (one in each arm) had rebounded after reaching viral suppression. Reasons for failure remain unclear and were discussed extensively in the review article by Clotet (2008).

AZT+3TC vs. TDF+FTC. For the AZT+3TC vs. TDF+FTC comparison, we identified four articles and abstracts from one randomised controlled trial, Study 934 (Arribas 2008; Gallant 2006; Gallant 2008; Pozniak 2006).

Study 934. Study 934 was also a large industry-sponsored multicentre, randomised controlled, open-label trial conducted in Europe and the United States of America that compared twice-daily random-dose AZT+3TC+EFV and once-daily random dose TDF+FTC+EFV. Results from the 144-week analyses were published by Arribas (2008), from the 96-week analysis by Pozniak (2006) and from the 48-week analysis by Gallant (2006). Primary study endpoints were HIV RNA load <400 copies/mL at 48 weeks, and secondary endpoints were HIV-1 RNA load <400 copies/mL at 96 and 144 weeks, HIV RNA load <50 copies/mL at 48 weeks, CD4 levels and safety. Gallant (2008) also published a study of three-year renal toxicity that combined data from Studies 903 and 934.

Observational studies. We also identified an additional 10 observational studies (Fernandez Lison 2005; Havlir 2005; Jones 2005; Lewis 2004; Lodwick 2008; Manfredi 2006; Mocroft 2006; Sanchez-de-la-Rosa 2008; Willig 2008; Wolbers 2007). Five studies dealt primarily with efficacy (Havlir 2005; Lodwick 2008; Mocroft 2006; Willig 2008; Wolbers 2007), two were cost-effectiveness analyses (Fernandez Lison 2005; Sanchez-de-la-Rosa 2008) and three described the association between therapy and adverse events, such as dyslipidaemia and hypocholesterolaemia (Jones 2005; Manfredi 2006) and renal dysfunction (Lewis 2004). Of these 10 studies, one study met the criteria for inclusion in the GRADE table and provided data that were able to be coded (Mocroft 2006); no restrictions were placed based on location for inclusion.
### Table 10. Characteristics of included studies, TDF vs. d4T or AZT.

<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Participants</th>
<th>Interventions</th>
<th>Outcomes</th>
<th>Notes</th>
</tr>
</thead>
</table>
| Arribas 2008  | Randomised controlled non-inferiority multicentre trial | 517 ART-naïve adults in France, Germany, Italy, Spain, UK, USA | 1. AZT+3TC+EFV  
2. TDF+FTC+EFV | HIV-1 RNA <50 copies/mL at 48 weeks; <400 copies/mL at 48, 96 and 144 weeks; CD4 count; safety | SAEs were reported as grade 2-4; no SD for CD4 reported |
| Gallant 2004a | Randomised controlled multicentre trial | 602 ART-naïve adults in Argentina, Brazil, Dominican Republic, France, Germany, Italy, Spain, Switzerland, UK, USA | 1. d4T+3TC+EFV  
2. TDF+3TC+EFV | HIV-1 RNA <50 copies/mL at 48 weeks; <400 copies/mL at 48, 96 and 144 weeks; CD4 count; safety | No SD for CD4 reported |
| Rey 2009 DAUFIN | Randomised, open label, multicentre, non-inferiority trial | 71 ART-naïve adults in France | 1. AZT+3TC+NVP  
2. TDF+3TC+NVP | HIV-1 RNA <400 copies/mL at 96 weeks | No SD for CD4 reported |

SAE, severe adverse events; SD, standard deviation

**Ongoing studies**

We identified one additional ongoing study, a phase IV, prospective, randomised, open-label evaluation of the efficacy of once-daily PI/NRTI-containing therapy combinations for the initial treatment of HIV-1-infected individuals from resource-limited settings (the PEARLS Trial, NCT00084136). This study, sponsored by ACTG (A5175), is a three-arm trial that is comparing AZT+3TC+EFV, FTC+3TC+NVP and TDF+FTC+EFV in poor communities in Africa, Asia, Haiti, South America and the United States of America; it also has a pre-planned regimen for patients who fail therapy. It is planned to have an estimated enrolment of 1574 participants in the three arms and is estimated to be completed in December 2009.

**TDF vs. ABC**

We reviewed seven studies, of which one was a meta-analysis of fat dysregulation *(Hill 2009)*; three reported on two randomised controlled trials *(Pappa 2008; Sax 2008; Smith 2009)*, two reported on one non-randomised trial and one was an observational study reporing on a cohort.

**Randomised controlled trials**
We reviewed two randomised controlled trials, which met our inclusion criteria (Sax 2008; Smith 2009) (Table 11). We also identified one non-randomised trial and two ongoing studies. All the studies examined the use of fixed-dose combinations of abacavir 600 mg and lamivudine 300 mg (Epzicom® or Kivexa®) and TDF 300 mg and emtricitabine (FTC) 200 mg (Truvada®) as the nucleoside/nucleotide backbone. In one group of studies (Sax 2008; Manfredi 2008a, Manfredi 2008b, IMCJ-H19-466) they were combined with EFV or a ritonavir-boosted (r) protease inhibitor; in the other group (Smith 2009; ASSERT) they were examined in combination with ritonavir-boosted lopinavir (LPV)/r in one study and EFV in the other.

The HEAT Study (Smith 2009) was a randomised controlled trial of once-daily ABC-3TC 600 mg/300 mg plus LPV/r 800 mg/200 mg compared to once-daily TDF-FTC 300 mg/200 mg plus LPV/r 800 mg/200 mg. The study enrolled 688 antiretroviral-naïve patients. The study had a double-blind, placebo-matched, multicentre, non-inferiority design and primary endpoints were the proportion of patients with HIV-1 RNA below 50 copies/mL at week 48 and the proportion experiencing adverse events over 96 weeks.

AIDS Clinical Trials Group (ACTG) 5202 (Sax 2008) is an ongoing phase IIIIB, randomised four-arm, double-blind comparison of ABC-3TC and TDF-FTC in combination with either open-label EFV or ATV/r. Primary endpoints are time to virologic failure (HIV-1 RNA >1000 copies/mL at 16-24 weeks or >200 copies/mL at 24 weeks) and time to first Grade 3/4 adverse events. The study enrolled 1,858 patients. Due to shorter time to virologic failure and shorter time to grade 3/4 adverse events in the subgroup that presented with high viral loads (≥100,000 copies/mL) in one arm, the study was unblinded by the Data and Safety Monitoring Board for this stratum (797 patients, 43%), and it was found that both were associated with ABC-3TC (time to virologic failure, hazard ratio [HR]=2.33, 95% confidence interval [CI] 1.46-3.72 and time to Grade 3/4 adverse events HR =1.87, 95% CI 1.43-2.43). The study is continuing for patients who entered with <100,000 copies/mL. Primary endpoints are time to virologic failure, time to development of a Grade 3/4 abnormality and time to treatment discontinuation. Pappa (2008) reviewed six prior trials of ABC+3TC-containing regimens and found that efficacy and safety endpoints were similar in both HIV-1 RNA strata of <100 000 copies/mL and ≥100 000 copies/mL.

Non-randomised controlled trials

Manfredi (2008a, 2008b) was a non-randomised 12-month open-label prospective study that compared random-dose combination ABC-3TC plus EFV or boosted protease inhibitors to random-dose combination TDF-FTC plus EFV or boosted protease inhibitors in 118 patients in a single center in Bologna, Italy. Of these 118 patients, 53 (45%) were antiretroviral-naïve at the start of treatment. Eleven (21%) of these received ABC-3TC (in 10 of the 11 the third drug was a boosted protease inhibitor and in one EFV), and 42 (79%) received TDF-FTC (in 16 of the 42 the third drug was a boosted protease inhibitor and in 26 EFV). Primary endpoints were virologic failure and virologic suppression (the proportion of patients with HIV-1 RNA <50 copies/mL at any time in the trial).

Observational studies

One observational study and one meta-analysis were identified for this comparison. Anastos (2007) examined treatment patterns and risk of dyslipidaemia in the Women’s Interagency Cohort Studies in the United States of America. Because it was conducted in a high-income country, it was not included in the GRADE table.

Ongoing studies
Table 11. Characteristics of included studies, TDF vs. ABC.

<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Participants</th>
<th>Interventions</th>
<th>Outcomes</th>
<th>Notes</th>
</tr>
</thead>
</table>
| Sax 2009 | Phase IIIB, randomised, double-blind 4-arm study | 1 858 ART-naïve participants from USA and Puerto Rico (analysis is of 797 participants with baseline HIV-1 RNA >100 000 copies/mL) | 1. ABC+3TC+ (EFV or ATZ/r)  
2. TDF+FTC+ (EFV or ATZ/r) | Virologic failure, Grade 3/4 adverse event, drug discontinuation | SAEs were not specifically reported, but subjects receiving ABC/3TC had shorter time to grade 3/4 AEs (HR=1.87, 95% CI 1.43-2.43, p<0.0001), predominantly general body aches and lipid increases. Suspected drug hypersensitivity was reported in 7% of each NRTI group. |
| Smith 2009 | Randomised, double-blind, placebo-matched multicentre, non-inferiority study | 694 ART-naïve participants from USA and Puerto Rico | 1. ABC+3TC+ LPV/r  
2. TDF+FTC+ LPV/r | Virologic failure, Grade 3/4 adverse events | For outcome of immunologic response, median values were reported but coded as mean values. SD values calculated using the IQR, averaging the values on each side of median. |

Two studies are currently ongoing, one in Europe and one in Japan. The European study has completed enrolment whereas the Japanese study is currently recruiting subjects.

**ASSERT (NCT00549198).** This study, sponsored by GlaxoSmithKline, is a Phase IV study of random-dose combination ABC-3TC plus EFV and TDF-FTC plus EFV currently being conducted in Europe. Its primary endpoint at 24, 48 and 96 weeks is renal function (glomerular filtration rate as measured by the modified diet in renal disease equation), and its secondary endpoints are osteopoenia, lipid metabolism abnormalities, adverse events necessitating discontinuation of therapy, other markers of renal function and viral resistance. Estimated date of study completion is October 2009.

**IMCJ-H19-466 (NCT00544128)** is sponsored by the International Medical Center of Japan and is comparing ABC-3TC-ATV/r with TDF-FTC-ATV/r. It is a Phase IV randomised, open-label, multicenter, non-inferiority study with safety and efficacy endpoints at 48 weeks and is being conducted at a single center in Tokyo. Overall follow up is planned for 144 weeks. Estimated date for completion of primary endpoints is March 2011, and overall completion is estimated for March 2013.

**METHODOLOGICAL QUALITY**

Below we review quality of randomised trials. Information on quality scoring non-randomised or observational studies is available on request.
**d4T vs. AZT**

Risk of bias for the studies included in the ABC vs. TDF analysis is shown in Table 12.

**Generation of allocation sequence**


**Allocation concealment**

Only *Robbins (2003)* reported that they had concealed allocation.

**Blinding**

All the studies except *Robbins (2003)* were open label. *Robbins (2003)* blinded the EFV and NFV components of the study but not the AZT vs. d4T.

**Incomplete outcome data**

Mortality was not reported by *Carr (2000)* or *Geijo Martinez (2006)*. *Kumar (2006)* did not reported clinical progression, and acquired drug resistance was reported only by *Kumar (2006)*.

**Free of selective reporting**

There was no evidence of selective reporting.

**Free of other bias**

*Kumar (2006)* and *Squires 2000* were funded by GlaxoSmithKline and Bristol-Myers Squibb, respectively. All other studies were funded from governmental sources.

**EFV vs. NVP**

Risk of bias for the studies included in the ABC vs. TDF analysis is shown in Table 13.

**Generation of allocation sequence**

Only *van Leth (2004c)* reported that there had been random generation of the allocation sequence. *Ayala Gaytán (2004), Manosuthi (2009), Núñez (2002)* and *Sow (2006)* did not report how they generated the allocation sequence.

**Allocation concealment**

Only *van Leth (2004)* reported that allocation was concealed.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Carr 2000 OzCombo 1</td>
<td>Yes, allocated in equal proportions</td>
<td>Not reported</td>
<td>Open-label</td>
<td>Mortality, acquired drug resistance not reported</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Eron 2000 START II</td>
<td>Yes, randomised in 1:1 ratio using permuted blocks</td>
<td>Not reported; centers used blocks but concealment not reported</td>
<td>Open-label</td>
<td>Clinical progression, drug resistance not reported</td>
<td>Yes, ITT analysis used</td>
<td>Industry funded</td>
</tr>
<tr>
<td>French 2002 OzCombo 2</td>
<td>Yes, allocated in equal proportions</td>
<td>Not reported</td>
<td>Open-label</td>
<td>Drug resistance not reported</td>
<td>Yes, ITT analysis used</td>
<td>Yes</td>
</tr>
<tr>
<td>Gathe 2002</td>
<td>Yes, allocated in equal proportions</td>
<td>Not reported</td>
<td>Open-label</td>
<td>Mortality, clinical progression, drug resistance not reported</td>
<td>Yes, ITT analysis used</td>
<td>Industry funded</td>
</tr>
<tr>
<td>Geijo Martínez 2006 GEMCEI</td>
<td>Random number table with block design</td>
<td>Not reported</td>
<td>Open-label</td>
<td>Mortality, drug resistance not reported</td>
<td>Yes, ITT analysis used</td>
<td>Yes</td>
</tr>
<tr>
<td>Kumar 2006 ESS40002</td>
<td>Yes, 1:1 allocation</td>
<td>Not reported</td>
<td>Open-label</td>
<td>Clinical progression not reported</td>
<td>Yes, ITT analysis used</td>
<td>Industry funded</td>
</tr>
<tr>
<td>Li 2008</td>
<td>Subjects were randomly allocated to treatment arm</td>
<td>Not reported</td>
<td>Open-label</td>
<td>Mortality, clinical progression, drug resistance not reported</td>
<td>Yes, ITT analyses used</td>
<td>Yes</td>
</tr>
<tr>
<td>Robbins 2003 ACTG 384</td>
<td>Yes, muted block computerised algorithm</td>
<td>Not reported</td>
<td>Open-label for NRTIs; EFV and NFV blinded</td>
<td>Drug resistance not reported</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Squires</td>
<td>Yes, 1:1</td>
<td>Not reported</td>
<td>Drug</td>
<td>Yes, ITT analysis used</td>
<td>Yes, ITT</td>
<td>Industry funded</td>
</tr>
</tbody>
</table>
2000 START I using permuted blocks label resistance not reported analysis used funded

ITT, intention-to-treat

Blinding

Manosuthi (2009), Núñez (2002) and van Leth (2004c) were reported as open-label studies. Ayala Gaytan (2004) was open-label. Neither of the other two studies reported if they were blinded or not.

Incomplete outcome data


Free of selective reporting

There was no evidence of selective reporting.

Free of other bias

Manosuthi (2009), Núñez (2002), Sow (2006) and van den Berg (2009) were all funded from governmental sources. The 2NN study (van Leth 2004c) was funded by Boehringer-Ingelheim. The source of funding for Ayala Gaytan (2004) was not reported.

TDF vs. d4T or AZT

Risk of bias for the studies included in the ABC vs. TDF analysis is shown in Table 14.

Generation of allocation sequence

Both Arribas (2008) and Gallant (2004) had random sequence generation. Rey (2009) did not report how the allocation sequence was generated.

Allocation concealment

Both Arribas (2008) and Gallant (2004) described how allocation was concealed. Rey (2009) did not report if allocation was concealed.

Blinding

Only Gallant (2004) reported that they had blinded the TDF and d4T components of the study. Neither of the other studies reported using placebos.

Incomplete outcome data

Neither Arribas (2008) nor Rey (2009) reported mortality, and none of the three studies
### Table 13. Risk of bias, EFV vs. NVP.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Ayala Gaytan 2004</td>
<td>Yes, proper allocation was used</td>
<td>No</td>
<td>Open-label</td>
<td>Drug resistance not reported</td>
<td>Yes, ITT analysis used</td>
<td>Unclear, funding source not reported</td>
</tr>
<tr>
<td>Manosuthi 2009 2NR</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Open-label</td>
<td>Clinical response, drug resistance not reported</td>
<td>Yes, ITT analysis used</td>
<td>Yes</td>
</tr>
<tr>
<td>Núñez 2002 SENC</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Open-label</td>
<td>Mortality, clinical progression, drug resistance not reported</td>
<td>Yes, ITT analysis used</td>
<td>Yes</td>
</tr>
<tr>
<td>Sow 2006</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Mortality, clinical progression, drug resistance not reported</td>
<td>Unclear how virologic response is reported</td>
<td>Yes</td>
</tr>
<tr>
<td>van den Berg-Wolf 2008</td>
<td>Yes, 1:1:1 allocation</td>
<td>Yes participants called hotline for assignment</td>
<td>Not reported</td>
<td>No, all outcomes reported</td>
<td>Yes, ITT analysis used</td>
<td>Yes</td>
</tr>
<tr>
<td>van Leth 2004c 2NNn</td>
<td>Yes, sequence generation and implementation were fully separated</td>
<td>Treatment allocation was done at the central study coordination center</td>
<td>Open-label</td>
<td>Drug resistance not reported</td>
<td>Yes, ITT analysis used</td>
<td>Industry funded</td>
</tr>
</tbody>
</table>

CPCRA, Terry Beirn Community Programs for Clinical Research on AIDS; ITT, intention-to-treat reported clinical progression. All reported drug resistance data.

**TDF vs. ABC**

Risk of bias for the studies included in the ABC vs. TDF analysis is shown in Table 15.

**Generation of allocation sequence**
### Table 14. Risk of bias, d4T or AZT vs. TDF.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Arribas 2008, Study 934</td>
<td>Yes</td>
<td>Yes</td>
<td>Open-label</td>
<td>Mortality, clinical progression not reported</td>
<td>Yes, ITT analysis used</td>
<td>Industry funded</td>
</tr>
<tr>
<td>Gallant 2004, Study 903</td>
<td>Yes, allocated in 1:1 ratio</td>
<td>Yes</td>
<td>Yes, double blinded</td>
<td>Clinical progression not reported</td>
<td>Yes, ITT analysis used</td>
<td>Industry funded</td>
</tr>
<tr>
<td>Rey 2009, DAUFIN</td>
<td>Yes, allocated in 1:1 ratio</td>
<td>Not reported</td>
<td>Open-label</td>
<td>Mortality, clinical response not reported</td>
<td>Yes, ITT analysis used</td>
<td>Yes</td>
</tr>
</tbody>
</table>

ITT, intention-to-treat

Neither Sax (2008) nor Smith (2009) reported how they generated the allocation sequence.

**Allocation concealment**

Neither Sax (2008) nor Smith (2009) reported if they concealed allocation.

**Blinding**

Both Sax (2008) and Smith (2009) blinded the ABC and TDF components of the study regimen.

**Incomplete outcome data**

Neither of the studies reported mortality or drug resistance, and only Smith (2009) reported clinical progression.

**Free of selective reporting**

Sax (2009) reported primary outcome variables only for the subset of patients with >100 000 CD4 cells/µL at baseline (the subset analysed by the Data Safety Monitoring Board). There was
Table 15. Risk of bias, TDF vs. ABC.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sax 2008 ACTG 5202</td>
<td>Unclear, not reported</td>
<td>Unclear, not reported</td>
<td>Yes, assignment to NNRTI combinations was blinded</td>
<td>Mortality, clinical progression, retention, immunologic response, drug resistance not reported</td>
<td>Yes</td>
<td>NIH funded</td>
</tr>
<tr>
<td>Smith 2009 HEAT</td>
<td>Yes, 1:1 randomisation</td>
<td>Yes</td>
<td>Double-blinded study</td>
<td>Mortality not reported</td>
<td>Yes, ITT analysis used</td>
<td>Industry funded</td>
</tr>
</tbody>
</table>

ACTG, AIDS Clinical Trials Group; ITT, intention-to-treat analysis.

no other evidence of selective reporting.

Free of other bias

Sax 2008 was funded by the U.S. National Institutes of Health (through ACTG); Smith 2009 was funded by GlaxoSmithKline.

Free of selective reporting

There was no evidence of selective reporting although Rey (2009) was terminated early because of unexpectedly high virologic failure rate.

Free of other bias

All three studies were industry-funded (Arribas 2009 and Gallant 2004 by Gilead and Rey 2009 by Boehringer-Ingelheim).

RESULTS

AZT vs. d4T

Randomised controlled trials

We identified eight completed randomised controlled trials and two ongoing trials that compared AZT with d4T with a third NRTI, an NNRTI or a protease inhibitor:
- AZT+3TC+ABC vs. d4T+3TC+ABC (Optimizing Pediatric HIV-Treatments in Infants with Prophylactic Exposure to Nevirapine)
- AZT+3TC+EFV vs. d4T+3TC+EFV (Tshepo)
- AZT+3TC+NVP vs. d4T+3TC+NVP (OzCombo 2 arms 1 and 2, Tshepo, Optimizing Pediatric HIV-Treatments in Infants with Prophylactic Exposure to Nevirapine)
- AZT+3TC+IDV vs. d4T+3TC+IDV (OzCombo 1 arms 1 and 2, START I)
- AZT+3TC+NFV vs. d4T+3TC+NFV (ESS4002 arms 2 and 3, ACTG 384 arms 3 and 4)
- AZT+3TC+EFV vs. d4T+ddI+EFV (ACTG 384 arms 1 and 2)
- AZT+3TC+NVP vs. d4T+ddI+NVP (OzCombo 2 arms 1 and 3)
- AZT+3TC+IDV vs. d4T+ddI+IDV (OzCombo 1 arms 1 and 3, GEMCEI arms 1 and 3, START II)
- AZT+3TC+NFV vs. d4T+ddI+NFV (Gathe 2002)
- AZT+3TC+RTV vs. d4T+ddI+RTV (GEMCEI arms 2 and 4)

ACTG 384 and its substudies A5005s and A5152s randomised 620 patients into four arms: AZT+3TC+EFV, d4T+ddI+EFV, AZT+3TC+NFV and d4T+ddI+NFV, allowing for direct comparison between AZT+3TC and d4T+ddI with both an NNRTI and a protease inhibitor (Robbins 2003). Six patients died, and 20 progressed clinically in the trial, although it was not reported in which arm these events occurred. Among patients receiving EFV, 155 (87.1%) of 181 in the AZT+3TC arm reached viral suppression at 144 weeks compared to 106 (88.4%) of 119 in the d4T+ddI arm. Ninety-four (60.6%) in the AZT+3TC arm remained on their original regimen at 144 weeks compared to 60 (38.7%) in the d4T+ddI arm. Among patients receiving NFV, 92 (59.4%) of 155 maintained viral suppression at 144 weeks in the AZT+3TC arm compared to 97 (62.5%) of 155 in the d4T+ddI arm. Sixty-four (41.3%) in the AZT+3TC arm remained on their original regimen compared to 49 (31.6%) of 155 in the d4T+ddI arm. No significant differences were reported in immunologic recovery across the four arms. In examining time to a grade 3 or 4 adverse event, the initial use of AZT+3TC delayed occurrence of the first serious toxic effect and the occurrence of the first symptom or diagnosis of peripheral neuropathy compared to d4T+ddI (p<0.001 for both).

ESS4002 was an industry-sponsored study that randomised 261 patients into three arms, including AZT+3TC+ABC, AZT+3TC+NFV and d4T+3TC+NFV (Kumar 2006). This analysis is restricted to the latter two arms. After 96 weeks there had been no deaths; clinical progression and grade 3/4 adverse events were not reported. Thirty-four (37.4%) of 91 in the AZT+3TC+NFV arm had achieved viral suppression, as measured by HIV-1 RNA <50 copies/mL, compared to 27 (32.5%) of 83 in the d4T+ddI+NFV arm. CD4 cell counts increased by an average of 270 cells/µL in the AZT+3TC+NFV arm and 289 in the d4T+3TC+NFV arm, and there were 10 acquired mutations in the AZT arm than in the d4T arm. The principal comparison of interest for the authors was comparing risk of dyslipidaemia and adherence in the AZT+3TC+ABC (Trizivir®) arm compared to the other two, for both which it was superior.

Gathe (2002) and colleagues studied the equivalence of once-daily enteric-coated ddI in a trial that compared AZT+3TC+NFV and d4T+enteric coated ddI+NFV in patients with ≥200 CD4 cells/µL. They randomised 511 ART-naive patients into the two arms and followed for 48 weeks. The proportion with HIV-1 RNA <400 copies/mL was 54% in both arms and the proportion with HIV-1 RNA <50 copies/mL was 32% in both arms. There was a mean increase in CD4 cells of 189 cells/µL in the AZT+3TC arm compared to 157 cells/µL in the d4T+ddI arm (p=NS). Grade 3/4 adverse events occurred in 8% of the AZT+3TC arm and 13% of the d4T+ddI arm.

The GEMCEI study randomised 98 patients into four arms: AZT+3TC+IDV, d4T+ddI+IDV, AZT+3TC+RTV and d4T+ddI+RTV (Geijo Martinez 2006). Deaths were not reported, but four patients progressed clinically in the two AZT+3TC arms compared to three in the two d4T+ddI arms. Viral load was non-detectable (HIV-1 RNA <400 copies/mL) in 15 (34.1%) of the 44
patients in the two AZT+3TC arms compared to 17 (40.4%) of 42 in the two d4T+ddl arms at 48 weeks. Fifteen (34.1%) of 44 remained on their assigned regimens at the completion of the trial in the two AZT+3TC arms compared to 17 (40.4%) of 42 in the two d4T+ddl arms. Sixteen (36.4%) discontinued therapy because of adverse events in the two AZT+3TC arms compared to 12 (28.6%) in the d4T+3TC arms.

Li (2008) randomised 198 patients with 100-300 CD4 cells/µL into three arms, AZT+ddl, AZT+3TC and d4T+3TC, all in combination with NVP. At 52 weeks, the proportion of patients with HIV-1 RNA <50 copies/mL was 68% in those receiving AZT+3TC and 69% in those receiving d4T+3TC. One patient on AZT+3TC and none on d4T+3TC acquired antiretroviral resistance. Hepatotoxicity was common (20%) and was associated with HCV infection and CD4 counts >250 cells/µL.

OzCombo 1 randomised 106 patients into three arms, including AZT+3TC, d4T+3TC and d4T+ddl, all with IDV (Carr 2000). At 52 weeks, one patient in each arm had progressed clinically. The proportion of patients with HIV-1 RNA <50 copies/mL were 66% in the AZT+3TC arm, 59% in the d4T+3TC arm and 48% in the d4T+ddl arm (p=0.34). Fifty-three percent in the AZT+3TC arm completed the study on their original regimen compared to 60% in the d4T+3TC arm and 35% in the d4T+ddl arm (p=0.18). CD4 cell counts increased most in the d4T+3TC arm (+237 cells/µL) compared to the AZT+3TC arm (+175 cells/µL) and the d4T+ddl arm (+176 cells/µL). No data were presented on mortality or acquired resistance.

OzCombo 2 randomised 70 patients into three arms: AZT+3TC+NVP, d4T+3TC+NVP and d4T+ddl+NVP (French 2002). There were no deaths and no cases of clinical progression in any of the arms. Eleven (55%) of 20 patients in the AZT+3TC arm had complete viral suppression at 52 weeks as measured by HIV-1 RNA <50 copies/mL compared to 13 (59%) of 22 in the d4T+3TC arm and 16 (70%) of 23 in the d4T+ddl arm (p=NS). Eighty percent of patients in the AZT+3TC arm completed 52 weeks of assigned therapy compared to 73% in the d4T+3TC arm and 70% in the d4T+ddl arm. Mean increases in CD4 counts were 172 cells/µL, 201 cells/µL and 190 cells/µL, respectively. Four (20%) patients in the AZT+3TC arm had severe adverse events, compared to 8 (36%) in the d4T+3TC arm and 5 (23%) in the d4T+ddl arm. Five (11%) of the 45 patients in the d4T-containing arms experienced peripheral neuropathy compared to none in the AZT+3TC arm.

START I was an industry-sponsored study that randomised 204 patients into two arms, AZT+3TC+IDV and d4T+3TC+IDV. There was a single death in the d4T arm and a single clinical progression in the AZT arm. At 48 weeks, 49 (47.6%) of 103 patients in the AZT arm had successfully suppressed viral replication as measured by HIV-1 RNA <50 copies/mL compared to 56/101 (55.4%) in the d4T arm. Sixty-four (63.4%) patients in the AZT arm completed the trial on their originally assigned regimen compared to 68 (66.0%) in d4T arm. The median CD4 cell count increase from baseline was 198 in the AZT arm and 227 in the d4T arm. Twenty-two (%) of 102 patients that received study drugs in the AZT arm had grade 3/4 clinical toxicity and 25 (%) had grade 3/4 laboratory toxicity compared to 30 (30.0%) and 35 (35.0%) of 100 in the d4T arm, respectively. There were no data presented on drug resistance.

START II randomised 205 patients with CD4 cell counts ≥200 cells/µL to AZT+3TC+IDV and d4T+ddl+IDV. Two patients died in the d4T+ddl arm; clinical progression was not reported. Forty-two (41%) of 103 patients randomised to the AZT+3TC arm had HIV-1 RNA <500 copies/mL between 40 and 48 weeks compared to 54 (53%) of 102 patients in the d4T+ddl arm. Thirty-five percent of patients in the AZT+3TC arm had HIV-1 RNA <50 copies/mL at 48 weeks.
compared to 41% in the d4T/ddI group (p>0.2). There was a significantly greater increase in CD4 cell counts from baseline in the d4T/ddI arm (+214 cells/µL) than in the AZT+3TC arm (+142 cells/µL, p=0.026). Fifty-eight (56%) of 103 in the AZT+3TC arm completed the trial on their assigned treatment compared to 64 (63%) in the d4T/ddI arm. Eight grade 3/4 adverse events occurred in each arm, including one death from pancreatitis in the d4T/ddI arm.

**Table 16. Pooled effect estimates, AZT vs. d4T.**

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Studies</th>
<th>Participants</th>
<th>Statistical method</th>
<th>Effect estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mortality</td>
<td>5</td>
<td>1 136</td>
<td>MH risk ratio, random effects</td>
<td>0.74 (0.18-2.93)</td>
</tr>
<tr>
<td>Clinical response</td>
<td>5</td>
<td>564</td>
<td>MH risk ratio, random effects</td>
<td>1.84 (0.49-6.90)</td>
</tr>
<tr>
<td>Severe adverse events</td>
<td>6</td>
<td>1 075</td>
<td>MH risk ratio, random effects</td>
<td>0.79 (0.62-1.03)</td>
</tr>
<tr>
<td>Virologic response</td>
<td>7</td>
<td>1 249</td>
<td>MH risk ratio, random effects</td>
<td>0.96 (0.86-1.07)</td>
</tr>
<tr>
<td>Adherence/tolerability/retention</td>
<td>9</td>
<td>1 869</td>
<td>MH risk ratio, random effects</td>
<td>1.07 (0.93-1.21)</td>
</tr>
<tr>
<td>Immunologic response</td>
<td>7</td>
<td>1 249</td>
<td>WMD, random effects</td>
<td>-16.18 (-53.16, +20.79)</td>
</tr>
<tr>
<td>Antiretroviral drug resistance</td>
<td>1</td>
<td>174</td>
<td>Single study, no meta-analysis</td>
<td>1.52 (0.58-4.00)</td>
</tr>
<tr>
<td>Sexual transmission of HIV</td>
<td>0</td>
<td>0</td>
<td>Not estimatable</td>
<td></td>
</tr>
</tbody>
</table>

MH, Mantel-Haenszel; WMD, weighted mean difference; RR>1.0 favours d4T

*Note that Carr 2000, French 2002 and Robbins 2003 were counted as two separate studies for purposes of meta-analysis.

Meta-analysis

Nine trials contributed to the meta-analysis (ACTG 384, OzCombo 1 and 2, ESS4002, START I and II, GEMCEI, Gathe 2002, Li 2008) (Table 16, Figure 1). There were no differences between AZT- and d4T-containing regimens with respect to death (3/573 [0.5%] in the AZT arm vs. 5/563 [0.9%] in the d4T arm, RR<sub>MHRE</sub> 0.74, 95% CI 0.18-2.93 [5 studies]), clinical progression (6/282 [2.1%] vs. 3/282 [1.1%], RR<sub>MHRE</sub> 1.84, 95% CI 0.49-6.90 [4 studies]), severe adverse events (77/538 [14.3%] vs. 101/537 [18.8%], RR<sub>MHRE</sub> 0.79, 95% CI 0.62-1.03 [6 studies]), virologic response (309/629 [49.1%] vs. 320/620 [51.6%], RR<sub>MHRE</sub> 0.96, 95% CI 0.86-1.07 [7 studies]), adherence/tolerability/retention (547/939 [58.3%] vs. 505/930 [54.3%], RR<sub>MHRE</sub> 1.06, 95% CI 0.93-1.21 [9 studies]) and immunologic response (N=629 in AZT arm vs. 620 in d4T arm, WMD<sub>RE</sub> -16.18, 95% CI -53.16, +20.79 [7 studies]). Only a single study (Kumar 2006) contributed data on acquired drug resistance, and the investigators found no relationship (RR 1.52, 95% CI 0.58-4.00).

GRADE

Based on the nine randomised controlled trials included in this review, the quality of evidence for all five critical outcomes and both important outcomes is low (Figures 2-3, Tables 17-18). This is due largely to design limitation (six of the eight studies were open-label studies), only three of the studies being conducted at least partly in developing country settings, indirect
comparisons (e.g., comparing AZT+3TC to d4T+ddl, especially with protease inhibitors, which will not be widely available in low- and middle-income countries), small sample sizes and the

**Figure 1. Forest plot by outcome, d4T vs. AZT.**

**Figure 1a. Mortality**

QuickTime™ and a decompressor are needed to see this picture.

**Figure 1b. Clinical response.**

QuickTime™ and a decompressor are needed to see this picture.

**Figure 1c. Severe adverse events.**

QuickTime™ and a decompressor are needed to see this picture.

**Figure 1d. Virologic response.**
industry funding in the four of the larger studies (Eron 2000; Gathe 2002; Kumar 2006; Squires 2000), which accounted for almost half of all patients randomised.
Observational studies

We identified six studies that met criteria for inclusion in the observational study GRADE tables (George 2009; Laurent 2008; Mocroft 2006; Nijorge 2009; Pazare 2008, Toure 2008). However, only five studies provided data that could be coded; authors for the remaining study have been contacted for additional follow-up. There was no evidence of a differential effect among regimens containing AZT and those containing d4T with the exception of the outcome of virologic response, which was less robust in the AZT arms (RR 0.67, 95% CI 0.48-0.92). Given design limitations and other issues, the quality of the observational literature was rated as low or very low.
Figure 2. Assessment of risk of bias assessment, AZT vs. d4T.

Figure 3. GRADE risk of bias assessment, AZT vs. d4T.
Table 17. GRADE table (randomised controlled trials), AZT vs. d4T.

<table>
<thead>
<tr>
<th>Quality assessment</th>
<th>Summary of findings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No of patients</td>
</tr>
<tr>
<td></td>
<td>AZT</td>
</tr>
<tr>
<td></td>
<td>Risk</td>
</tr>
<tr>
<td></td>
<td>No of patients</td>
</tr>
<tr>
<td></td>
<td>3/593 (0.5%)</td>
</tr>
</tbody>
</table>

Mortality (follow-up 3 studies at 48 weeks, 1 study at 52 weeks, 1 study at 96 weeks)

<table>
<thead>
<tr>
<th>No of studies</th>
<th>Design</th>
<th>Limitations</th>
<th>Inconsistency</th>
<th>Indirectness</th>
<th>Imprecision</th>
<th>Other considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>serious</td>
<td>reporting bias</td>
</tr>
<tr>
<td>7</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
<tr>
<td>9</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
<tr>
<td>10</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
<tr>
<td>12</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
<tr>
<td>10</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious im precision</td>
<td>reporting bias</td>
</tr>
</tbody>
</table>

Immunologic response (follow-up median 48 weeks; Better indicated by higher values)

<table>
<thead>
<tr>
<th>No of studies</th>
<th>Design</th>
<th>Limitations</th>
<th>Inconsistency</th>
<th>Indirectness</th>
<th>Imprecision</th>
<th>Other considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
<tr>
<td>8</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
<tr>
<td>9</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
<tr>
<td>10</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
<tr>
<td>12</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
</tbody>
</table>

Adherence/ tolerability/retention (follow-up 4 studies at 48 weeks, 2 studies at 52 weeks, 1 study at 96 weeks, 1 study at 144 weeks)

<table>
<thead>
<tr>
<th>No of studies</th>
<th>Design</th>
<th>Limitations</th>
<th>Inconsistency</th>
<th>Indirectness</th>
<th>Imprecision</th>
<th>Other considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
<tr>
<td>8</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
<tr>
<td>9</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
<tr>
<td>10</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
<tr>
<td>12</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
</tbody>
</table>

Virologic response (follow-up 4 studies at 48 weeks, 2 studies at 52 weeks, 1 study at 96 weeks)

<table>
<thead>
<tr>
<th>No of studies</th>
<th>Design</th>
<th>Limitations</th>
<th>Inconsistency</th>
<th>Indirectness</th>
<th>Imprecision</th>
<th>Other considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
<tr>
<td>8</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
<tr>
<td>9</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
<tr>
<td>10</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
<tr>
<td>12</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
</tbody>
</table>

Severe adverse events (follow-up 4 studies at 48 weeks, 2 studies at 52 weeks)

<table>
<thead>
<tr>
<th>No of studies</th>
<th>Design</th>
<th>Limitations</th>
<th>Inconsistency</th>
<th>Indirectness</th>
<th>Imprecision</th>
<th>Other considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
<tr>
<td>8</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
<tr>
<td>9</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
<tr>
<td>10</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
<tr>
<td>12</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
</tbody>
</table>

Clinical progression (follow-up 3 studies at 48 weeks, 2 studies at 52 weeks)

<table>
<thead>
<tr>
<th>No of studies</th>
<th>Design</th>
<th>Limitations</th>
<th>Inconsistency</th>
<th>Indirectness</th>
<th>Imprecision</th>
<th>Other considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
<tr>
<td>8</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
<tr>
<td>9</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
<tr>
<td>10</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
<tr>
<td>12</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
</tbody>
</table>

Virologic response (follow-up 4 studies at 48 weeks, 2 studies at 52 weeks, 1 study at 96 weeks)

<table>
<thead>
<tr>
<th>No of studies</th>
<th>Design</th>
<th>Limitations</th>
<th>Inconsistency</th>
<th>Indirectness</th>
<th>Imprecision</th>
<th>Other considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
<tr>
<td>8</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
<tr>
<td>9</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
<tr>
<td>10</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
<tr>
<td>12</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
</tbody>
</table>

Severe adverse events (follow-up 4 studies at 48 weeks, 2 studies at 52 weeks)

<table>
<thead>
<tr>
<th>No of studies</th>
<th>Design</th>
<th>Limitations</th>
<th>Inconsistency</th>
<th>Indirectness</th>
<th>Imprecision</th>
<th>Other considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
<tr>
<td>8</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
<tr>
<td>9</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
<tr>
<td>10</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
<tr>
<td>12</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
</tbody>
</table>

Mortality (follow-up 3 studies at 48 weeks, 1 study at 52 weeks, 1 study at 96 weeks)

<table>
<thead>
<tr>
<th>No of studies</th>
<th>Design</th>
<th>Limitations</th>
<th>Inconsistency</th>
<th>Indirectness</th>
<th>Imprecision</th>
<th>Other considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>serious</td>
<td>reporting bias</td>
</tr>
</tbody>
</table>

Clinical progression (follow-up 3 studies at 48 weeks, 2 studies at 52 weeks)

<table>
<thead>
<tr>
<th>No of studies</th>
<th>Design</th>
<th>Limitations</th>
<th>Inconsistency</th>
<th>Indirectness</th>
<th>Imprecision</th>
<th>Other considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
</tbody>
</table>

Severe adverse events (follow-up 4 studies at 48 weeks, 2 studies at 52 weeks)

<table>
<thead>
<tr>
<th>No of studies</th>
<th>Design</th>
<th>Limitations</th>
<th>Inconsistency</th>
<th>Indirectness</th>
<th>Imprecision</th>
<th>Other considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
</tbody>
</table>

Virologic response (follow-up 4 studies at 48 weeks, 2 studies at 52 weeks, 1 study at 96 weeks)

<table>
<thead>
<tr>
<th>No of studies</th>
<th>Design</th>
<th>Limitations</th>
<th>Inconsistency</th>
<th>Indirectness</th>
<th>Imprecision</th>
<th>Other considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
</tbody>
</table>

Adherence/ tolerability/retention (follow-up 4 studies at 48 weeks, 2 studies at 52 weeks, 1 study at 96 weeks, 1 study at 144 weeks)

<table>
<thead>
<tr>
<th>No of studies</th>
<th>Design</th>
<th>Limitations</th>
<th>Inconsistency</th>
<th>Indirectness</th>
<th>Imprecision</th>
<th>Other considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
</tbody>
</table>

Immunologic response (follow-up median 48 weeks; Better indicated by higher values)

<table>
<thead>
<tr>
<th>No of studies</th>
<th>Design</th>
<th>Limitations</th>
<th>Inconsistency</th>
<th>Indirectness</th>
<th>Imprecision</th>
<th>Other considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>no serious imprecision</td>
<td>reporting bias</td>
</tr>
</tbody>
</table>

DRAFT: What to Start
Drug resistance (follow-up 96 weeks)

<table>
<thead>
<tr>
<th>Group</th>
<th>Serious</th>
<th>No Serious Inconsistency</th>
<th>Serious</th>
<th>Serious</th>
<th>None</th>
</tr>
</thead>
<tbody>
<tr>
<td>Randomised trials</td>
<td>10/91</td>
<td>(11%)</td>
<td>6/83</td>
<td>7.2%</td>
<td></td>
</tr>
<tr>
<td>RR</td>
<td>1.52 (0.58 to 4)</td>
<td></td>
<td>38 more per 1000 (from 30 fewer to 217 more)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sexual transmission of HIV - not reported

<table>
<thead>
<tr>
<th>Group</th>
<th>No</th>
<th>None</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-</td>
<td>0/0 (0%)</td>
<td>0/0 (0%)</td>
</tr>
</tbody>
</table>

1 Separate comparison arms from 3 RCT studies (Carr, French, Robbins) contributed more than once for a number of outcomes. There were 9 studies in total.
2 6 out of 9 RCT studies were open-label studies and some studies had large rates of loss to follow-up, but studies were not downgraded based on these facts.
3 5 out of 9 RCT studies looked at indirect comparisons of drug regimens.
4 Number of events <300 and/or confidence intervals include potential harm and benefit.
5 7 out of 9 RCT studies were industry funded, although some were funded simultaneously by competitors.
6 This outcome was coded as the proportion that finished the trial out of those who were initially assigned the treatment.
7 Only 1 RCT study (Kumar) reported on drug resistance, suggesting selective reporting.

Bibliography:
### Table 18. GRADE table (observational studies), AZT vs. d4T.

<table>
<thead>
<tr>
<th>Quality assessment</th>
<th>Summary of findings</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No of patients</td>
<td>Effect</td>
</tr>
<tr>
<td></td>
<td>AZT</td>
<td>d4T</td>
</tr>
<tr>
<td><strong>Mortality (observational)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 observational studies</td>
<td>8/85 (9.4%)</td>
<td>11/84 (13.1%)</td>
</tr>
<tr>
<td><strong>Severe adverse events (observational)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 observational studies</td>
<td>14/415 (3.4%)</td>
<td>383/1941 (19.7%)</td>
</tr>
<tr>
<td><strong>Virologic response (observational)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 observational studies</td>
<td>33/85 (38.8%)</td>
<td>49/84 (58.3%)</td>
</tr>
<tr>
<td><strong>Adherence/tolerability/retention (observational)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 observational studies</td>
<td>112/137 (81.8%)</td>
<td>108/135 (80%)</td>
</tr>
<tr>
<td><strong>Immunologic response (observational) (Better indicated by higher values)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 observational studies</td>
<td>13123</td>
<td>7423</td>
</tr>
<tr>
<td><strong>Drug resistance (observational)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 observational studies</td>
<td>4/85 (4.7%)</td>
<td>7/84 (8.3%)</td>
</tr>
</tbody>
</table>
1 Number of events <300 and/or confidence intervals include potential harm and benefit.
2 1 out of 5 observational studies were industry funded, although some were funded simultaneously by competitors.
3 1 observational study (Laurent) reported on drug resistance, suggesting selective reporting.

EFV vs. NVP

Randomised controlled trials

We identified six randomised controlled trials that compared EFV and NVP that made five basic comparisons:

- d4T+3TC+NVP vs. d4T+3TC+EFV (relevant arms in the 2NN and the Tshepo Studies)
- d4T+ddI+NVP vs. d4T+ddI+EFV (SENC)
- AZT+3TC+NVP vs. AZT+3TC+EFV (Ayala Gaytan, Sow and relevant arms in the Tshepo Study)
- AZT+ddI+NVP vs. AZT+ddI+EFV (relevant arms in Tshepo Study)
- Two NRTIs (not specified) + NVP vs. two NRTIs (not specified) + EFV (FIRST study)

The 2NR Study was really a study of whether rifampicin therapy for concurrent tuberculosis would effect EFV- or NVP-containing ART regimens more.

2NN study. Treatment failure occurred in 96 (44%) of 220 patients assigned to d4T+3TC+NVP once daily, 169 (44%) of 387 assigned to d4T+3TC+NVP twice daily, 151 (37.8%) of 400 assigned to d4T+3TC+EFV, and 111 (53.1%) of 209 assigned to d4T+3TC+NVP+EFV. The difference between NVP twice daily and EFV was 5.9% (95% CI -0.9 to 12.8). There were no significant differences among the study groups in the proportions with plasma HIV-1 RNA concentrations below 50 copies/mL at week 48 or the increases in CD4 cells. NVP plus EFV was associated with the highest frequency of clinical adverse events, and NVP once daily with significantly more hepatobiliary laboratory toxicities than EFV. Of 25 observed deaths, two were attributed to NVP toxicity.

Tshepo Study. Bussman (2009) reported only comparison between the AZT+ddI arms and the combined AZT+3TC and d4T+3TC arms to isolate the effect of ddI vs. 3TC and nothing on the comparison between EFV and NVP. Thus, at this point, the Tshepo study is non-contributory on the question of EFV versus NVP.

Ayala Gaytan (2004). At 24 weeks 18 (64%) of 28 patients in the NVP arm and 24 (81%) of 30 patients in the EFV arm achieved undetectable plasma viral load (<400 HIV-1 RNA copies/mL), which declined to 15 (54%) of 28 in the NVP arm and 20 (68%) of 30 in the EFV arm by 48 weeks. In each arm, patients had similar increases in CD4 cell counts, and there were no AIDS defining conditions or deaths in either group. Adverse events led to discontinuation in two patients each in the NVP and EFV arms.

2NR Study (Manosuthi 2009). At baseline the mean body weight of patients was 53 kg, the mean CD4 cell count was 65 cells/µL and the median plasma HIV-1 RNA level was 5.8 log10 copies/mL. At 12 hours post dose, the mean concentration (C12) (± standard deviation) of the 121 patients in the EFV group was 3.54 ±3.78 mg/L and of the 121 in the NVP group 5.6 ±2.65 mg/L. At week 12, 3.1% of patients in the EFV group and 21.3% in the NVP group had C12 values that were less than the recommended minimum concentrations (OR, 8.40; 95% CI 1.81-39.00). At 48 weeks 73.2% of patients in the EFV group and 71.8% of those in the NVP group had plasma HIV-1 RNA levels <50 copies/mL, with respective mean CD4 cell counts of 274 and 252 cells/µL (p=NS). In multivariate analysis, patients with low C12 values and those with a body weight <55 kg were 3.6 and 2.4 times more likely, respectively, to develop all-cause treatment failure (p<.05).
SENC Study (Núñex 2002). At 48 weeks 23 (74%) of 31 patients in the EFV group and 23 (64%) of patients in the NVP achieved plasma HIV-1 RNA viral loads of <50 HIV RNA copies/mL. Adverse events led to NNRTI discontinuation in four patients in the EFV arm and three in the NVP arm. There were no statistically significant differences between groups for any primary endpoint.

FIRST Study (CPCRA 058). One hundred eleven patients randomised to the EFV arm and 117 randomised to the NVP arm were followed for a median of 5.0 years. The incidence of virological failure was 41.2 per 100 person-years in the EFV arm and 42.8 per 100 person-years in the NVP arm. The proportion of patients with plasma HIV-1 RNA <50 copies/mL was similar throughout follow-up (p = .24), as were average increases in CD4 cells (p = .30).

Meta-analysis

Six studies contributed to the randomised controlled trial meta-analysis (Ayala Gaytan 2004; Manosuthi 2009; Núñez 2002; Sow 2008; van den Berg-Wolf 2008; van Leth 2004) (Table 19, Figure 4). There were no differences between EFV and NVP-containing regimens with respect to mortality (29/582 [5%] for EFV arm vs. 33/575 [5.7%] for NVP arm, RR_MHRE 0.89, 95% CI 050-1.57 [3 studies]), clinical progression (44/541 [8.2%] in EFV arm vs. 34/532 [6.4%] in NVP arm, RR_MHRE 1.31, 95% CI 0.78-2.20 [3 studies], serious adverse events (115/643 [17.9%] in the EFV arm vs. 152/639 [23.8%] in the NVP arm, RR_MHRE 0.55, 95% CI 0.33-1.31 [5 studies]), virological response 508/643 [79%] in EFV arm vs. 500/639 in NVP arm, RR_MHRE 0.99, 95% CI 0.95-1.28 [5 studies]) or retention (335/673 [49.4%] in EFV arm vs. 304/674 [45.1%] in NVP arm, RR_MHRE 1.11, 95% CI 0.95-1.28 [6 studies]), or immunologic response (mean increase in CD4 cell count (N=678 in EFV arm vs. 674 in NVP arm, WMD_RE -1.85, 95% CI -23.16, 19.46 [6 studies]). A single study (van den Berg-Wolf 2008) contributed data on viral resistance and found that acquired resistance was lower in the EFV-containing arm (32/111 [28.8%] in EFV arm vs. 49/117 [41.9%] in NVP arm, RR_MHRE 0.69, 95% CI 0.48-0.99).

GRADE

Based on the six randomised controlled trials included in this review, the quality of evidence for two of the five critical outcomes (virologic response and adherence/tolerability/retention) and one of the three important outcomes (immunologic response) was high and for the other three critical outcomes (mortality, clinical response, severe adverse events) was moderate; the evidence for drug resistance is low (Figures 5-6, Tables 20-21). Four of the six studies were open-label (design limitations), the underpowering of individual studies (imprecision) and only a single study (Manosuthi 2008) was conducted in a developing country and then in a population of people being treated for tuberculosis. Additionally, the largest study, the 2NN study, which contributed the majority of patients to the analysis was industry funded, although, given the specificity of the endpoints (mortality, clinical progression, virological response, immunologic response), this is unlikely to have led to serious bias.

Non-randomised controlled studies

Han (2005). In Liaoning Province treatment consisted of EFV plus IDV, in Jilin Province d4T+ddI+EFV and in Henan Province d4T+ddI+NVP. At 6 months the authors reported that there were similar proportions of patients with non-detectable plasma HIV-1 RNA viral loads and comparable CD4 counts from Jilin and Henan provinces as well as similar proportions of patients with NNRTI mutation K103N but emergence of Y181C in Henan.
Table 19. Pooled effect estimates, EFV vs. NVP.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Studies</th>
<th>Participants</th>
<th>Statistical method</th>
<th>Effect estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mortality</td>
<td>3</td>
<td>1 157</td>
<td>MH risk ratio, random effects</td>
<td>0.90 (0.56-1.44)</td>
</tr>
<tr>
<td>Clinical response</td>
<td>3</td>
<td>1 073</td>
<td>MH risk ratio, random effects</td>
<td>1.33 (0.89-1.99)</td>
</tr>
<tr>
<td>Severe adverse events</td>
<td>5</td>
<td>1 282</td>
<td>MH risk ratio, random effects</td>
<td>0.83 (0.53-1.31)</td>
</tr>
<tr>
<td>Virologic response</td>
<td>5</td>
<td>1 282</td>
<td>MH risk ratio, random effects</td>
<td>0.99 (0.91-1.09)</td>
</tr>
<tr>
<td>Adherence/tolerability/retention</td>
<td>6</td>
<td>1 352</td>
<td>MH risk ratio, random effects</td>
<td>1.11 (0.95-1.28)</td>
</tr>
<tr>
<td>Immunologic response</td>
<td>6</td>
<td>1 352</td>
<td>WMD, random effects</td>
<td>-1.85 (-23.16, 19.46)</td>
</tr>
<tr>
<td>Antiretroviral drug resistance</td>
<td>1</td>
<td>228</td>
<td>MH risk ratio, random effects</td>
<td>0.69 (0.48-0.99)</td>
</tr>
<tr>
<td>Sexual transmission of HIV</td>
<td>0</td>
<td>0</td>
<td></td>
<td>Not estimatable</td>
</tr>
</tbody>
</table>

MH, Mantel-Haenzsel; WMD, weighted mean difference; RR>1.0 favours EFV

**NEEF Cohort (Hartmann 2005b).** The difference in success between the two arms was 4.5% (-11.5%, 19.0%, p=0.578) and results were similar for pre-treated and ART-naive subjects. Non-significant results appeared for all secondary analyses. Of the treated patients, 4.5% (n=30) developed rashes (nevirapine: 2.4% and efavirenz: 6.4%). In four patients treatment was not interrupted. Three patients were re-exposed to the initial drug without any side effects. In a smaller initial study, involving 151 patients in an EFV arm and 145 in the NVP arm, hepatitis C virus and hepatitis B virus were detected in 10.1% and 4.1% of patients, respectively. The overall rate of severe hepatotoxicity (grade 3 to 4 elevations in aspartate aminotransferase and/or alanine aminotransferase) was 2(1.3%) of 151 in patients receiving EFV and 3 (2.1%) of 145 in patients prescribed NVP. Mild-to-moderate hepatotoxicity (grade 2 elevation) was observed in 60% (EFV) and 3.4% (NVP) of patients. Incidence of mild-to-moderate and severe hepatotoxicity did not differ significantly between the study groups. Three (2.1%) of 14 patients with grade 2 elevations of liver enzymes and 4 (80%) of 5 patients with grade 3 to 4 elevations were symptomatic. The only risk factor for the development of mild-to-moderate hepatotoxicity was hepatitis C coinfection.

**Manfredi (2004).** Antiretroviral-naive patients experienced greater efavirenz activity at 3-12 months (maximum plasma HIV-1 RNA decline = -2.4 log_{10} copies/mL), associated with a significantly higher rate of complete viral suppression. A limited virologic and immunologic advantage of efavirenz was observed at the first 12-month assessment of antiretroviral-naive patients. Short- and long-term toxicity and withdrawal rates from the two arms were similar. In the separate reports of hepatotoxicity (Manfredi 2005, Manfredi 2006) found that elevations in transaminases were associated with NVP administration but were no worse among women or among those more severely immunosuppressed.

**Observational studies**
Twenty-six studies contributed to the observational study meta-analysis (Annan 2009; Aranazabal 2005; Aurpibul 2007; Bannisters 2008; Berenguer 2008; Boulle 2007; Boulle 2008). Figure 4. Forest plot by outcome, EFV vs. NVP.

**Figure 4a. Mortality**

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>EFV Events</th>
<th>NVP Events</th>
<th>Total Events</th>
<th>Total Weight</th>
<th>M-H, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manosuthi 2008</td>
<td>2</td>
<td>71</td>
<td>73</td>
<td>12.2%</td>
<td>0.33 (0.07, 1.80)</td>
</tr>
<tr>
<td>von den Berg-Wolf 2008</td>
<td>20</td>
<td>111</td>
<td>131</td>
<td>19.6%</td>
<td>1.17 (0.65, 2.09)</td>
</tr>
<tr>
<td>von Leth 2004</td>
<td>7</td>
<td>400</td>
<td>407</td>
<td>28.1%</td>
<td>0.74 (0.28, 2.00)</td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td><strong>502</strong></td>
<td><strong>575</strong></td>
<td><strong>1077</strong></td>
<td><strong>100.0%</strong></td>
<td><strong>0.89 (0.50, 1.57)</strong></td>
</tr>
</tbody>
</table>

Total events: 293
Heterogeneity: Tau² = 0.06, Ch² = 2.46, df = 2 (P = 0.29), P = 19%
Test for overall effect: Z = 0.42 (P = 0.68)

**Figure 4b. Clinical response.**

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>EFV Events</th>
<th>NVP Events</th>
<th>Total Events</th>
<th>Total Weight</th>
<th>M-H, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaytan 2004</td>
<td>0</td>
<td>30</td>
<td>30</td>
<td>0</td>
<td>Not estimable</td>
</tr>
<tr>
<td>van den Berg-Wolf 2008</td>
<td>34</td>
<td>111</td>
<td>145</td>
<td>39.7%</td>
<td>1.56 (0.98, 2.47)</td>
</tr>
<tr>
<td>von Leth 2004</td>
<td>10</td>
<td>400</td>
<td>410</td>
<td>30.3%</td>
<td>0.88 (0.38, 2.05)</td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td><strong>541</strong></td>
<td><strong>532</strong></td>
<td><strong>1073</strong></td>
<td><strong>100.0%</strong></td>
<td><strong>1.31 (0.78, 2.20)</strong></td>
</tr>
</tbody>
</table>

Total events: 443
Heterogeneity: Tau² = 0.04, Ch² = 1.37, df = 1 (P = 0.26), P = 27%
Test for overall effect: Z = 1.02 (P = 0.31)

**Figure 4c. Severe adverse events.**

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>EFV Events</th>
<th>NVP Events</th>
<th>Total Events</th>
<th>Total Weight</th>
<th>M-H, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaytan 2004</td>
<td>4</td>
<td>30</td>
<td>34</td>
<td>9.1%</td>
<td>0.93 (0.28, 3.38)</td>
</tr>
<tr>
<td>Manosuthi 2006</td>
<td>5</td>
<td>71</td>
<td>76</td>
<td>12.4%</td>
<td>0.50 (0.18, 1.39)</td>
</tr>
<tr>
<td>Nunez 2002</td>
<td>19</td>
<td>31</td>
<td>48</td>
<td>22.0%</td>
<td>1.84 (1.07, 3.16)</td>
</tr>
<tr>
<td>von den Berg-Wolf 2008</td>
<td>24</td>
<td>111</td>
<td>135</td>
<td>26.1%</td>
<td>0.56 (0.38, 0.90)</td>
</tr>
<tr>
<td>von Leth 2004</td>
<td>63</td>
<td>400</td>
<td>463</td>
<td>28.6%</td>
<td>0.73 (0.55, 0.99)</td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td><strong>643</strong></td>
<td><strong>639</strong></td>
<td><strong>1282</strong></td>
<td><strong>100.0%</strong></td>
<td><strong>0.83 (0.53, 1.31)</strong></td>
</tr>
</tbody>
</table>

Total events: 1152
Heterogeneity: Tau² = 0.16, Ch² = 12.42, df = 4 (P = 0.01), P = 68%
Test for overall effect: Z = 0.79 (P = 0.43)

**Figure 4d. Virologic response.**

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>EFV Events</th>
<th>NVP Events</th>
<th>Total Events</th>
<th>Total Weight</th>
<th>M-H, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaytan 2004</td>
<td>13</td>
<td>30</td>
<td>43</td>
<td>2.4%</td>
<td>0.93 (0.53, 1.65)</td>
</tr>
<tr>
<td>Manosuthi 2006</td>
<td>51</td>
<td>71</td>
<td>122</td>
<td>15.0%</td>
<td>0.88 (0.93, 1.20)</td>
</tr>
<tr>
<td>Nunez 2002</td>
<td>23</td>
<td>31</td>
<td>54</td>
<td>6.9%</td>
<td>1.16 (0.84, 1.60)</td>
</tr>
<tr>
<td>von den Berg-Wolf 2008</td>
<td>82</td>
<td>111</td>
<td>193</td>
<td>26.0%</td>
<td>0.88 (0.77, 1.01)</td>
</tr>
<tr>
<td>von Leth 2004</td>
<td>339</td>
<td>400</td>
<td>739</td>
<td>49.7%</td>
<td>1.04 (0.90, 1.21)</td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td><strong>643</strong></td>
<td><strong>639</strong></td>
<td><strong>1282</strong></td>
<td><strong>100.0%</strong></td>
<td><strong>0.99 (0.91, 1.09)</strong></td>
</tr>
</tbody>
</table>

Total events: 593
Heterogeneity: Tau² = 0.00, Ch² = 5.76, df = 4 (P = 0.22), P = 31%
Test for overall effect: Z = 0.12 (P = 0.91)
However, only 24 studies provided data that could be coded; authors for the remaining two studies have been contacted for additional follow-up but have not yet responded. Among these 24 studies, there were no differences between EFV and NVP-containing regimens with respect to any critical or important outcome although there was a trend to fewer serious adverse events with EFV (RRMHE 0.70, 95% CI 0.49-1.01).

The quality of the evidence from the observational studies was rated as low, primarily because of design issues.
Figure 5. Assessment of risk of bias assessment, NVP vs. EFV.

Figure 6. GRADE risk of bias assessment, NVP vs. EFV.
Table 20. GRADE table (randomised controlled trials only), NVP vs. EFV.

<table>
<thead>
<tr>
<th>Quality assessment</th>
<th>Summary of findings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No of patients</td>
</tr>
<tr>
<td>No of studies</td>
<td>Design</td>
</tr>
<tr>
<td>Mortality (follow-up 2 studies at 48 weeks, 1 study at 65 months)</td>
<td>3 randomised trials</td>
</tr>
<tr>
<td>Clinical progression (follow-up 2 studies at 48 weeks, 1 study at 65 months)</td>
<td>3 randomised trials</td>
</tr>
<tr>
<td>Serious adverse events (follow-up 4 studies at 48 weeks, 1 study at 65 months)</td>
<td>4 randomised trials</td>
</tr>
<tr>
<td>Virologic response (follow-up 4 studies at 48 weeks, 1 study at 65 months)</td>
<td>5 randomised trials</td>
</tr>
<tr>
<td>Adherence/tolerability/retention (follow-up 4 studies at 48 weeks, 1 study at 65 months, 1 study did not report follow up period)</td>
<td>6 randomised trials</td>
</tr>
</tbody>
</table>
DRAFT: What to Start

**Immunologic response (follow-up 4 studies at 48 weeks, 1 study at 65 months, 1 study at 6 months; better indicated by higher values)**

<table>
<thead>
<tr>
<th>Randomised trials</th>
<th>no serious limitations</th>
<th>no serious inconsistency</th>
<th>no serious indirectness</th>
<th>serious</th>
<th>none</th>
<th>MD 3.95 higher (11.58 lower to 19.48 higher)</th>
<th>⊕⊕⊕O MODERATE IMPORTANT</th>
</tr>
</thead>
</table>

**Drug resistance (follow-up median 65 months)**

<table>
<thead>
<tr>
<th>Randomised trials</th>
<th>no serious inconsistency</th>
<th>no serious indirectness</th>
<th>serious</th>
<th>none</th>
<th>32/111 (28.8%)</th>
<th>49/117 (41.9%)</th>
<th>RR 0.69 (0.48 to 0.99)</th>
<th>⊕⊕ LOW IMPORTANT</th>
</tr>
</thead>
</table>

**Sexual transmission of HIV - not reported**

| - | - | - | - | none | 0/0 (0%) | 0/0 (0%) | - | - |

1. 4 out of 6 RCT studies were open-label; 1 of the remaining studies did not provide sufficient information on blinding (Sow) and the other was blinded (van den Berg-Wolf) but studies were not downgraded based on these facts.
2. 1 RCT study (van den Berg) looked at multiple indirect comparisons. Also, only 1 out of 6 RCT studies was only conducted in a developed country setting (Manosuthi).
3. Number of events <300 and/or confidence intervals include potential harm and benefit.
4. 1 out of 6 RCT studies were industry funded (van Leth), while 1 out of 6 studies had a funding source that was unclear.
5. This outcome was coded as the proportion that finished the trial out of those who were initially assigned the treatment.
6. None of the included studies provided standard deviations for the mean outcome so the same estimated SD value was used for all studies.
7. Only 1 RCT study reported on drug resistance (van den Berg-Wolf), suggesting selective reporting.

**Bibliography:**

1. Ayala Gaytan JJ, de la Garza ERZ, Garcia-Macho-Chavez SBV. Nevirapine or efavirenz in combination with two nucleoside analogues in HIV infected antiretroviral naïve patients. Med Intern Mex 2004; 20:24
Table 21. GRADE tables (observational studies), NVP vs. EFV.

<table>
<thead>
<tr>
<th>Quality assessment</th>
<th>Summary of findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>No of patients</td>
<td>Effect</td>
</tr>
<tr>
<td></td>
<td>No of studies</td>
</tr>
<tr>
<td>Mortality (observational)</td>
<td>5</td>
</tr>
<tr>
<td>Serious adverse events (observational)</td>
<td>14</td>
</tr>
<tr>
<td>Virological response (observational)</td>
<td>11</td>
</tr>
<tr>
<td>Adherence/tolerability/retention (observational)</td>
<td>5</td>
</tr>
<tr>
<td>Immunologic response (observational) (Better indicated by lower values)</td>
<td>4</td>
</tr>
</tbody>
</table>

* None of the included studies provided standard deviations for the mean outcome so the same estimated SD value was used for all studies.

**TDF vs. d4T or AZT**

**Randomised controlled trials**

We identified three completed and one ongoing randomised controlled trials that compared TDF with either AZT or d4T in four different comparisons:

- d4T+3TC+EFV vs. TDF+3TC+EFV (Study 903)
- AZT+3TC+EFV vs. TDF+3TC+EFV (PEARLS Arms A and C); this is an ongoing study
- AZT+3TC+NVP vs. TDF+3TC+NVP (DAUFIN)
- AZT+3TC+EFV vs. TDF+FTC+EFV (Study 934)

**Study 903.** Gallant (2004a), Pozniak 2003 and Cassetti 2007 found that 253 (84%) of 301 participants in the d4T+3TC+EFV arm had <400 HIV-1 RNA copies/mL at 48 weeks compared to 239 (80%) of 299 in the TDF+3TC+EFV arm (p=NS), that at 96 weeks 214 (71%) of 301 and 226 (76%) of 299, respectively, had reached <400 HIV-1 RNA copies/mL (p=0.004) and that at 144 weeks 193 (67%) of 301 and 214 (71%) of 301, respectively, had <400 HIV-1 RNA copies/mL (p=0.004). Using the more stringent endpoint of <50 HIV-1 RNA copies/mL, at 48 weeks 240 (80%) of 301 in the d4T arm achieved viral suppression compared to 228 (76%) of 299 in the TDF arm (weighted difference 95% CI -9.8 to 3.0 with a pre-defined lower confidence bound of -10). At 96 weeks, these proportions were 204 (68%) of 299 and 226 (76%) of 299, respectively (p=NS), and at 144 weeks, they were 188 (63%) of 301 and 202 (68%) of 299, respectively (p=0.08). By week 48 100 (33%) of 301 in the d4T arm had discontinued the study regimen compared to 82 (27%) of 299 in the TDF arm. Grade 3/4 clinical adverse events were similar in both arms (25%-27%), as were Grade 3/4 laboratory abnormalities (42% and 36%). Compared to patients in the TDF arm, patients in the d4T arm had greater increases in fasting triglycerides, total cholesterol and low-density lipoprotein cholesterol and smaller increases in high-density lipoprotein cholesterol. Additionally, mitochondrial toxicities (peripheral neuropathy, lipodystrophy and lactic acidosis) occurred more commonly in the d4T arm (83/301 [28%] vs. 17/299 [6%], p<0.001). Renal function was comparable in the two groups.

**DAUFIN Study.** The much smaller DAUFIN study used AZT instead of d4T and NVP instead of EFV but otherwise provided a direct comparison of AZT and TDF: AZT+3TC+NVP vs. TDF+3TC+NVP. This study was originally planned for 250 patients fro 96 weeks but was halted, as noted above, after 71 patients were enrolled. Eight patients, all in the TDF+3TC+NVP arm, had early virologic failure, defined as a <2 log_{10} decrease in HIV-1 RNA by week 12 or rebound from a non-detectable HIV-1 RNA specimen, and two, one in each group, had late virologic failure, defined as two detectable HIV-1 RNA specimens one month apart after reaching non-detectability. By 36 weeks, 15 (42%) of 35 patients in the AZT arm had discontinued their original regimen compared to 14 (3%) of 36 in the TDF arm. Acquired resistance was substantial in the TDF arm with multiple NNRTI resistance mutations and K65R conferring NRTI cross-resistance. There were no differences in NVP trough concentrations between the two arms or between patients who failed and fully suppressed viral replication. Patients who failed had higher baseline viral loads and lower CD4 counts, but the reasons for failure remain unclear. The authors cite two single-armed studies (Towner 2004; Lapadula 2008) as having experienced similarly high failure rates with regimens containing both NVP and TDF.

**Study 934.** Study 934 examined random-dose combination backbones of AZT/3TC and TDF/FTC plus EFV. The study enrolled 517 patients into the two arms, and results were published for 48 (Gallant 2006a), 96 (Pozniak 2006) and 144 weeks (Arribas 2008). At 48
weeks 177 (73%) of 243 patients without baseline EFV resistance in the AZT/3TC arm reached
had suppressed viral replication to <400 copies/mL compared to 206 (84%) of 244 patients in

Table 22. Pooled effect estimates, TDF vs. d4T or AZT.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Studies</th>
<th>Participants</th>
<th>Statistical method</th>
<th>Effect estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mortality</td>
<td>1</td>
<td>602</td>
<td>MH risk ratio, random effects</td>
<td>1.18 (0.37-3.84)</td>
</tr>
<tr>
<td>Severe adverse events</td>
<td>3</td>
<td>1182</td>
<td>MH risk ratio, random effects</td>
<td>1.03 (0.80-1.32)</td>
</tr>
<tr>
<td>Virologic response</td>
<td>3</td>
<td>1129</td>
<td>MH risk ratio, random effects</td>
<td>0.99 (0.76-1.29)</td>
</tr>
<tr>
<td>Adherence/tolerability/retention</td>
<td>3</td>
<td>1133</td>
<td>MH risk ratio, random effects</td>
<td>1.14 (1.03-1.26)</td>
</tr>
<tr>
<td>Immunologic response</td>
<td>2</td>
<td>1055</td>
<td>WMD, random effects</td>
<td>+10.39 (-49.47, +70.09)</td>
</tr>
<tr>
<td>Antiretroviral drug resistance</td>
<td>2</td>
<td>673</td>
<td>MH risk ratio, random effects</td>
<td>6.12 (1.43-26.15)</td>
</tr>
<tr>
<td>Sexual transmission of HIV</td>
<td>0</td>
<td>1</td>
<td>Not estimatable</td>
<td></td>
</tr>
</tbody>
</table>

MH, Mantel-Haenzsel; WMD, weighted mean difference; RR>1.0 favours TDF

the TDF/FTC arm (p=0.002); at 96 weeks these results were 143 (62%) of 231 and 173 (75%) of 232, respectively (p=0.004), and at 144 weeks they were 133 (56%) of 229 and 161 (71%) of 227, respectively (p=0.004). Using a more stringent measure of viral suppression of <50 copies/mL, at 144 weeks there was a trend toward greater suppression in the TDF/FTC arm (130/231 [56%] in the AZT/3TC arm vs. 146/227 [64%] in the TDF/FTC arm, p=0.08). The proportions with Grade 2-4 clinical and laboratory adverse events were similar between the two arms. The AZT/3TC arm was associated with greater reductions in limb fat and with greater increases from baseline in fasting total cholesterol, low- and high-density lipoprotein cholesterol and mean fasting triglyceride levels. Overall at 144 weeks 41% of patients in the AZT/3TC arm had discontinued the original regimen compared to 29% in the TDF/FTC arm. Additionally, there was a trend to greater CD4 cell recovery in the TDF/FTC arm (+271 cells/µL vs. +312, p=0.09), and a greater proportion in the AZT/3TC arm developed the M184V/I mutation. Laessig (2006) pointed out 51 patients discontinued the study before 48 weeks and suggested that a press release describing an interim analysis at 24 weeks may have led the patients to do so, since this was an open-label study. She concluded, however, that while this source of bias may have led to overestimation of the results, in 10 simulation studies, they did not change the basic conclusion of the study. In responding, Gallant (2006b) pointed out that 45 of the discontinuations had occurred prior to the press release, and thus, even if the press release led to patients discontinuing the AZT/3TC arm, it would not have changed the basic 48-week conclusion.

Meta-analysis

All three randomised controlled trials contributed to the meta-analysis (Table 22, Figure 7). There were no differences between TDF-containing regimens and those using AZT or d4T as the comparator in terms of mortality (6/303 [2%] in the TDF arm vs. 5/299 [1.7%] in the AZT or d4T arm, RR_{MHRE} 1.18, 95% CI 0.37-3.84 [1 study]), severe adverse events (241/592 [40.7%] vs. 236/590 [40%], RR_{MHRE} 1.03, 95% CI 0.8-1.32 [3 studies], virologic response (334/566 [59%] vs. 56
341/563 [61%], RR_{MHRE} 0.99, 95% CI 0.76-1.29 [3 studies]) or immunologic response (N=530 for TDF arm, N=528 for d4T or AZT arm, WMD = +10.31, 95% CI -49.47, +70.09).
Figure 7. Forest plot by outcome, d4T or AZT vs. TDF.

Figure 7a. Mortality

Figure 7b. Severe adverse events.

Figure 7c. Virologic response.

Figure 7d. Adherence/tolerability/retention.

Figure 7e. Immunologic response.
However, the proportion able to remain on TDF-containing regimens was higher (395/566 [69.8%]) than in the d4T or AZT arm (347/567 [61.2%], RR\textsubscript{MHRE} 1.14, 95% CI 1.03-1.26 [3 studies]), but the risk of developing antiretroviral resistance was also substantially higher in the TDF arm (18/335 [5.4%]) than in the d4T or AZT-containing arms (2/338 [0.6%], RR\textsubscript{MHRE} 6.12, 95% CI 1.43-26.15 [2 studies]). None of the three studies reported on clinical progression.

GRADE

The quality of evidence for four critical outcomes (mortality, severe adverse events, virologic response and adherence/tolerability/retention) was moderate, as it was for the one of the important outcomes (resistance); the evidence for immunologic response was high (Figures 8-9, Table 22). Issues with the trials included using open-label drugs in two studies (Arribas 2008, Rey 2009), one large study (Arribas 2008) compared two drugs (AZT+3TC vs. TDF+FTC), only one study (Gallant 2004) had participants from low- and middle-income countries, Rey 2009 had small sample sizes and low precision and all three studies were industry sponsored.

Observational studies

None of the efficacy cohorts or the toxicity studies was from a low or middle-income country and were thus not graded because of indirectness. In general, however, they showed that TDF-containing backbones were associated with a higher proportion of non-detectable viraemia (Havlir 2005), a lower rate of change due to toxicity (Lodwick 2008) and overall greater durability (Willig 2008) but slower rates of CD4-cell increase (Mocroft 2006; Wolbers 2007). One cost-effectiveness analyses compared random-dose AZT/3TC plus EFV to random-dose TDF/FTC plus EFV and found that the cost of random-dose TDF/FTC was €9 734 less expensive at 48 weeks and would have 13% more viral suppression at 24 months (Sanchez-de-la-Rosa 2008). The other compared AZT with TDF and found that AZT-based regimens would have less cost in the short term but the cost of TDF-based regimens would be €129 less per year after three years (Fernandez Lison 2005).
Figure 8. Assessment of risk of bias assessment, d4T or AZT vs. TDF.

Figure 9. GRADE risk of bias assessment, d4T or AZT vs. TDF.

Legend:
- Green: Yes (low risk of bias)
- Yellow: Unclear
- Red: No (high risk of bias)
Table 22. GRADE table, d4T or AZT vs. TDF.

<table>
<thead>
<tr>
<th>No of studies</th>
<th>Design</th>
<th>Limitations</th>
<th>Inconsistency</th>
<th>Indirectness</th>
<th>Imprecision</th>
<th>Other considerations</th>
<th>Summary of findings</th>
<th>Quality</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No of patients</td>
<td>Effect</td>
<td>Absolute</td>
</tr>
<tr>
<td>Mortality (follow-up mean 144 weeks)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TDF (d4T or AZT)</td>
<td>Relative (95% CI)</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>randomised trials</td>
<td>serious(^1)(^2)</td>
<td>no serious inconsistency</td>
<td>no serious indirectness</td>
<td>serious(^4)</td>
<td>none</td>
<td>6/306 (2%)</td>
<td>5/299 (1.7%)</td>
<td>RR 1.18 (0.37 to 3.84)</td>
</tr>
<tr>
<td>Clinical progression - not reported</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0/0 (0%)</td>
<td>0/0 (0%)</td>
<td></td>
</tr>
<tr>
<td>Severe adverse events (follow-up 1 study at 36 weeks, 2 studies at 144 weeks)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>256/551 (46.3%)</td>
<td>247/595 (41.8%)</td>
<td>OR 1.04 (0.81 to 1.34)</td>
</tr>
<tr>
<td>Virologic response (follow-up 1 study at 36 weeks, 2 studies at 144 weeks)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>384/595 (64.8%)</td>
<td>384/593 (64.8%)</td>
<td>RR 1 (0.76 to 1.3)</td>
</tr>
<tr>
<td>Adherence/tolerability/retention (follow-up 1 study at 36 weeks, 2 studies at 144 weeks)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>445/591 (75.3%)</td>
<td>400/597 (67%)</td>
<td>RR 1.13 (1.05 to 1.21)</td>
</tr>
<tr>
<td>Immunological response (follow-up 144 weeks: better indicated by higher values)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>559</td>
<td>558</td>
<td></td>
</tr>
<tr>
<td>Drug resistance (follow-up 1 study at 36 weeks, 1 study at 144 weeks)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>18/335 (5.4%)</td>
<td>2/338 (0.6%)</td>
<td>RR 6.12 (1.43 to 30 more per 1000)</td>
</tr>
</tbody>
</table>

DRAFT: What to Start
26.15) (from 3 more to 149 more)

Sexual transmission of HIV - not reported

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>0/0 (0%)</th>
<th>0/0 (0%)</th>
</tr>
</thead>
</table>

1 Only 1 out of 3 RCT studies reported on mortality (Gallant), suggesting selective reporting.
2 2 out of 3 RCT studies were open-label (Gallant and Rey), but studies were not downgraded based on these facts.
3 1 study out of 3 RCT studies was an indirect comparison of TDF/FTC/EFV vs. ZDV/3TC/EFV (Gallant) and 2 studies out of 3 (Gallant, Rey) were conducted only in developed country settings, but studies were not downgraded based on these facts.
4 Number of events <300 and/or confidence intervals include potential harm and benefit.
5 Assessment of adherence/retention/tolerability or assessment of adverse events may be subject to bias in an open-label study, so downgraded for this outcome.
6 All 3 RCT studies were industry funded; not downgraded for this, however, as study drug did not show benefit so less concern for reporting bias.
7 This outcome was coded as the proportion that finished the trial out of those who were initially assigned the treatment.
8 None of the included studies provided standard deviations for the mean outcome so the same estimated SD value was used for all studies.

Bibliography:
**TDF vs. ABC**

Randomised controlled trials

We identified two randomised controlled trials, both of which compared ABC+3TC to TDF+FTC with either EFV or a RTV-boosted protease inhibitor. There were no trials identified that directly compared ABC to TDF using the same NRTI and NNRTI as the third and fourth drugs.

**HEAT Study (Smith 2009).** This trial compared random-dose combination ABC-3TC with TDF-FTC both in combination with LPV/r. There were no reported clinical outcomes, except for adverse events, and the primary endpoint was virological failure. Protocol-defined virologic failure occurred in 14% of patients in both groups. At week 96, 243 (71%) of 343 in the ABC-3TC group compared to 231 (67%) of 345 in the TDF-FTC group had HIV-1 RNA <50 copies/mL using intent to treat analysis (p=0.913). In addition, efficacy of both regimens was similar in patients with baseline HIV-1 RNA ≥100,000 copies/mL or CD4 cell counts below 50 cells/µl. At 96 weeks, median CD4 recovery was +250 cells/µL for ABC-3TC and +247 cells/µL for TDF-FTC. Premature study discontinuation due to adverse events occurred in 103 (30%) of 343 in the ABC+3TC arm compared to 97 (28%) of 345 in the TDF+FTC arm by 48 weeks.

**ACTG 5202 (Sax 2008).** This four-armed trial compared random-dose combination ABC-3TC with TDF-FTC in combination with either EFV or ATV/r. As noted above, patients’ data for patients presenting with HIV-1 RNA viral loads ≥100,000 copies/mL have been removed from the study, based on clear inferiority of ABC-3TC compared to TDF-FTC. Specifically, of the 797 patients enrolled with HIV-1 RNA viral loads ≥100,000 copies/mL, 85% were men, 26% Black, 25% Hispanic; mean baseline RNA=5.1 log c/mL, CD4=181/mm³. Median follow-up was 60 weeks. Virological failure occurred in 57 patients in the ABC-3TC arm and 26 in the TDF-FTC arm regardless of EFV or ATV/r assignment. Time to VF was significantly shorter in the ABC-3TC than TDF-FTC arm (HR=2.33, 95% CI 1.46-3.72). In a secondary cross-sectional analysis (prior VF and regimen changes included), the proportion with HIV-1 RNA <50 copies/mL at week 48 was 75% (95% CI 69%-80%) for ABC/3TC and 80% (95% CI 74%-85%) for TDF/FTC (p=0.20).

No additional data have as yet been reported for the subgroup with HIV-1 RNA <100,000 copies/µL at study entry.

In the subgroup presenting with ≥100,000 copies/mL, patients receiving ABC-3TC had shorter time to grade 3/4 adverse events (HR=1.87, 95% CI 1.43-2.43), which were described as predominantly general body aches and lipid increases. Suspected drug hypersensitivity was reported in 7% of each group.

**Meta-analysis**

The two randomised controlled trials contributed to the meta-analysis (Table 23, Figure 10). Only one study contributed data to clinical progression, severe adverse events, adherence/tolerability/retention and immunologic response, and none of these analyses showed a difference between ABC and TDF. Neither study had mortality or antiretroviral resistance outcomes, but both had virologic response outcomes. For virologic response There were no differences between ABC- and TDF-containing regimens with respect to 559 (75%) of 744 in the TDF arm maintained viral suppression throughout the trial compared to 533 (72%) of 741 in the ABC arm (RRMHRE 1.03, 95% CI 0.95-1.11).
Table 23. Pooled effect estimates, TDF vs. ABC.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Studies</th>
<th>Participants</th>
<th>Statistical method</th>
<th>Effect estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mortality</td>
<td>0</td>
<td>0</td>
<td>Not estimatable</td>
<td></td>
</tr>
<tr>
<td>Clinical response</td>
<td>1</td>
<td>688</td>
<td>Single study, no meta-analysis</td>
<td>2.98 (0.12-72.6)</td>
</tr>
<tr>
<td>Severe adverse events</td>
<td>1</td>
<td>688</td>
<td>Single study, no meta-analysis</td>
<td>0.94 (0.74-1.18)</td>
</tr>
<tr>
<td>Virologic response</td>
<td>2</td>
<td>1485</td>
<td>MH risk ratio, random effects</td>
<td>1.03 (0.95-1.11)</td>
</tr>
<tr>
<td>Adherence/tolerability/retention</td>
<td>1</td>
<td>688</td>
<td>Single study, no meta-analysis</td>
<td>0.94 (0.84-1.05)</td>
</tr>
<tr>
<td>Immunologic response</td>
<td>1</td>
<td>688</td>
<td>Single study, no meta-analysis</td>
<td>3.0 (-12.69, +18.69)</td>
</tr>
<tr>
<td>Antiretroviral drug resistance</td>
<td>0</td>
<td>0</td>
<td>Not estimatable</td>
<td></td>
</tr>
<tr>
<td>Sexual transmission of HIV</td>
<td>0</td>
<td>0</td>
<td>Not estimatable</td>
<td></td>
</tr>
</tbody>
</table>

MH, Mantel-Haenszel; WMD, weighted mean difference; RR>1.0 favours ABC

**GRADE**

Based on the three studies included in this review, one of which was a non-randomised trial, the evidence for using TDF as opposed to ABC is somewhat compelling. The quality of the evidence for three critical outcomes (severe adverse events and virologic response) was rated as moderate and for one (clinical events) as low; for the important outcome of immunologic response, the evidence for was rated as moderate (Figures 11-12, Table 24). The strength of this literature lies in the randomization and blinding of the HEAT study and ACTG 5202, and the weakness lies in indirectness; all three studies compared ABC+3TC to TDF+FTC rather than ABC to TDF directly, three different third drugs were used (EFV, non-specified PI, ATV/r and LPV/r) making direct comparisons difficult and none was conducted in a middle- or low-income country. Additionally, the non-randomised design and small sample size in Manfredi 2008 were problematic, making this study largely non-contributory. Other issues to note are that the large HEAT study, which contributed almost 40% of the patients to this analysis, was industry funded and there was only partial reporting of results from ACTG 5202 (for patients with HIV-1 RNA >100 000 copies/mL at baseline).

**Non-randomised controlled trials**

**Manfredi (2008b)**: This two-arm non-randomised comparison study examined the use of random dose ABC-3TC and TDF-FTC in both treatment-experienced and treatment-naïve patients. The third drug was either EFV or a RTV-boosted PI. Among the 53 treatment-naïve patients, 10 (91%) of 11 on ABC-3TC and 38 (90%) of 42 on TDF-FTC had <50 HIV-1 RNA copies/mL at 52 weeks (p=NS). The change in CD4 counts was similar in both arms, +109 in the ABC-3TC arm and +117 in the TDF-FTC arm (p=NS). Two (18%) of 11 patients in the ABC-3TC arm discontinued therapy because of early ABC hypersensitivity, but there were no other changes reported. However, it should be noted that by design no patient who changed the NRTI backbone after one month on therapy was analysed in this study.

**Observational studies**
Anastos (2007) found that ABC-containing regimens (among others) were associated with higher triglycerides and that TDF-containing regimens were associated with lower triglycerides in the Women’s Interagency HIV Study. This was similar to the findings in a meta-analysis of six trials conducted by Hill (2009), who found that ABC-3TC-containing regimens, among others, were more likely to be associated with hypercholesterolaemia, hypertriglyceridaemia and elevations of both low-density and high-density lipoprotein cholesterol compared to patients...
taking TDF-FTC-containing regimens. Neither study had data that met inclusion criteria and so was not included in GRADE tables.

Figure 11. Assessment of risk of bias assessment, ABC vs. TDF.

Figure 12. GRADE risk of bias assessment, ABC vs. TDF.
Table 24. GRADE table, ABC vs. TDF.

<table>
<thead>
<tr>
<th>No of studies</th>
<th>Design</th>
<th>Limitations</th>
<th>Inconsistency</th>
<th>Indirectness</th>
<th>Imprecision</th>
<th>Other considerations</th>
<th>TDF</th>
<th>ABC</th>
<th>Relative (95% CI)</th>
<th>Absolute</th>
<th>Quality</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>none</td>
<td>0/0 (0%)</td>
<td>0/0 (0%)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Clinical progression (follow-up mean 96 weeks)</td>
<td>1 randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious¹</td>
<td>serious²</td>
<td>none³</td>
<td>1/345 (0.3%)</td>
<td>0/343 (0%)</td>
<td>RR 2.93 (0.12 to 72.96)</td>
<td>0 more per 1000 (from 0 fewer to 0 more)</td>
<td>⊕⊕⊕Ο</td>
<td>CRITICAL</td>
</tr>
<tr>
<td>Severe adverse events (follow-up mean 96 weeks)</td>
<td>1 randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency</td>
<td>serious¹</td>
<td>no serious imprecision</td>
<td>none³</td>
<td>977/345 (28.1%)</td>
<td>103/343 (30%)</td>
<td>RR 0.94 (0.74 to 1.18)</td>
<td>18 fewer per 1000 (from 78 fewer to 54 more)</td>
<td>⊕⊕Ο</td>
<td>CRITICAL</td>
</tr>
<tr>
<td>Virologic response (follow-up 1 study at 48, 1 study at 96 weeks)</td>
<td>2 randomised trials</td>
<td>no serious limitations</td>
<td>no serious inconsistency⁴</td>
<td>serious¹</td>
<td>no serious imprecision</td>
<td>none³</td>
<td>550/744 (73.9%)</td>
<td>533/741 (71.9%)</td>
<td>RR 1.03 (0.95 to 1.11)</td>
<td>22 more per 1000 (from 36 fewer to 79 more)</td>
<td>⊕⊕Ο</td>
<td>CRITICAL</td>
</tr>
<tr>
<td>Adherence/tolerability/retention (follow-up mean 96 weeks)</td>
<td>1 randomised trials</td>
<td>no serious limitations⁵</td>
<td>no serious inconsistency</td>
<td>serious¹</td>
<td>no serious imprecision</td>
<td>none³</td>
<td>221/345 (64.1%)</td>
<td>234/343 (68.2%)</td>
<td>RR 0.94 (0.84 to 1.05)</td>
<td>41 fewer per 1000 (from 109 fewer to 34 more)</td>
<td>⊕⊕Ο</td>
<td>CRITICAL</td>
</tr>
<tr>
<td>Immunologic response (follow-up mean 96 weeks; better indicated by higher values)</td>
<td>1 randomised trials</td>
<td>no serious limitations⁶</td>
<td>no serious inconsistency</td>
<td>serious¹</td>
<td>serious²</td>
<td>none³</td>
<td>345</td>
<td>343</td>
<td>-</td>
<td>MD 3 higher (12.69 lower to 18.69 higher)</td>
<td>⊕⊕Ο</td>
<td>IMPORTANT</td>
</tr>
<tr>
<td>Drug resistance - not reported</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>none</td>
<td>0/0 (0%)</td>
<td>0/0 (0%)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sexual transmission of HIV - not reported</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>none</td>
<td>0/0 (0%)</td>
<td>0/0 (0%)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

DRAFT: What to Start

67 of 106
Both studies looked at the indirect basic comparison of TDF+FTC vs. ABC+3TC. One study was conducted only in developed country settings (Smith); the final study did not report a location for the study.

Number of events <300 and/or confidence intervals include potential harm and benefit.

One study was industry funded (Smith) while the source of funding for the other (Sax) was unclear; studies were not downgraded based on these facts.

Treatment failure in high PVL group (viral load >=100,000 copies/mL) inconsistent with findings from a meta-analysis (Pappa 2008) of patients starting ABC-3TC regimens in which patients with HIV-1 RNA levels of <100,000 and >=100,000 had similar experiences and that between 87% and 95% did not experience virological failure.

This outcome was coded as the proportion that finished the trial out of those who were initially assigned the treatment.

**Bibliography:**

ART to prevent the sexual transmission of HIV

Several cohort studies have examined this outcome either directly or indirectly (Barroso 2000; Bunnell 2008; Cu-Uvin 2000; Garcia 1999; Gupta 1997; Kaplan 1996; Musicco 1994; Operaksalski 1997; Pedraza 1999; Porco 2004; Quinn 2000; Ragni 1998; Vernazza 1997a; Vernazza 1997b; Vernazza 2000), and a large trial is ongoing (HIV Prevention Trials Network [HPTN 052]). At this point in time there is no direct RCT evidence that HIV will prevent sexual transmission of HIV, although there is clear evidence that it can prevent perinatal transmission. Observational studies have clearly shown that lower viral loads are associated with lower risks of transmission and have suggested that ART is associated with decreased risk of transmission.

Studies in paediatric populations

No randomised controlled trials enrolled paediatric patients. There is a large prospective study of NVP-containing and NVP-sparing regimens for babies exposed perinatally to NVP in Nairobi, Kenya, and a large cohort of children being followed in Thailand by the Netherlands-Thai Collaboration (NCT00476606). Six additional observational studies (Aurpibul 2007; Kamya 2007; Lapphra 2008; McConnell 2009; Wamalwa 2007; Zhang 2009) compared the use of EFV to NVP with mixed results. Kamya (2007) found a clear virologic response advantage to EFV compared to NVP (OR 2.46, 95% CI 1.23-4.90) whereas Lapphra (2008) in Thailand did not.

Studies in populations co-infected with Mycobacterium tuberculosis

Only one study (Manosuthi 2009) dealt explicitly with patients with tuberculosis. The 2NR study randomised 142 patients with HIV and M. tuberculosis co-infection who had received 4-16 weeks of rifampicin therapy to receive either d4T+3TC+NVP or d4T+3TC+EFV. At 48 weeks, virologic, immunologic and adverse events outcomes were similar. Six other observational studies (Boulle 2008; Manosuthi 2008; Sathie 2008; Shipton 2009; Sungkanaparph 2006; Varma 2009) compared outcomes in patients using EFV and NVP with mixed results. Boulle (2008) in South Africa found worse virologic outcomes among patients with tuberculosis taking NVP than those without tuberculosis whereas there was no difference for EFV, but Manosuthi (2008) and Varma (2009) in Thailand and Shipton (2009) in Botswana did not.

Studies in populations co-infected with hepatitis B or hepatitis C

No randomised controlled trial death specifically with patients with hepatitis B or hepatitis C co-infections. Five observational studies (Aranzabal 2005; Berenguer 2008; Martin-Carbonero 2003; Mocroft 2005; Sulkowski 2002) did examine differences in outcomes with exposures to EFV and NVP with mixed results. Aranzabal (2005) and Martin-Carbonero (2003) in Spain and Sulkowski (2002) in the the United States of America found that in patients with HCV co-infection NVP was more likely to cause progression of hepatic disease, whereas Berenguer (2008) found that NNRTI use in general and NVP use in particular was associated with reversal of fibrosis.

DISCUSSION

AZT vs. d4T
Ten studies have compared or are comparing AZT and d4T, albeit in 10 different combinations. The multiplicity of combinations, including using 3TC as the second NRTI in some studies and ddI in others and using ABC, EFV, NVP, IDV, NFV and RTV as the third drug make this a complex literature. However, laying aside the confounding that may accompany making comparisons across these various regimens, there is a reasonably robust literature to suggest that the regimens are roughly equivalent with perhaps AZT being slightly better tolerated (RRMRE 1.17 in the meta-analysis for maintenance of regimen to completion of study). Moreover, d4T has a host of well-recognised side effects including mitochondrial toxicity and dyslipidaemia, while AZT is associated with anaemia making its use somewhat problematic in regions where anaemia from malaria, intestinal parasitoses and repeated pregnancies is common.

These studies, taken collectively, however, suffer from a variety of limitations and the overall GRADE rating was very low. Five of the six studies were open-label, several studies compared AZT+3TC to d4T+ddI potentially obscuring the head-to-head comparison of AZT and d4T and several others used protease inhibitors as a third drug, studying combinations that will not be widely available in the developing world. Additionally, several of the studies were small, leading to imprecision and potential for type II error; only one study (which involved four arms) enrolled more than 300 patients, and few enrolled patients from the developing world.

Interestingly the cleanest comparisons will likely come from the two ongoing studies in Africa, the Tshepo study in Botswana and the Optimizing Pediatric HIV-Treatments in Infants with Prophylactic Exposure to Nevirapine in Kenya. Both of these will provide direct comparisons of standard therapeutic combinations of AZT+3TC+NVP or EFV and d4T+3TC+NVP or EFZ. Additional analysis of the more recent large cohort studies from low- and middle-income countries may provide additional data to

**EFV vs. NVP**

Six separate randomised controlled trials and three other experimental studies contribute to this literature. The six trials enrolled 1,352 participants and found there to be no critical differences between EFV and NVP with EFV slightly less likely than NVP to be associated with development of antiretroviral resistance. The quality of the body of literature that supports this is moderate to high, with the exception of drug resistance, which was examined in only a single study. This literature is dominated by the landmark 2NN study, which found no difference between EFV and NVP in a non-inferiority randomised open-label, industry-funded, four-arm study. Overall the 2NN study accounted for 787 (58%) of the 1,352 patients randomised.

The three ongoing studies we identified will add substantially to this literature. Specifically ANRS 12146 and the d4T/ddi backboned study in India will examine the use of EFV and NVP in patients with tuberculosis and hopefully extend the findings of the 2NR study. The DAYANA study (ARNS 12115) being conducted in Senegal and Cameroon will examine the use of EFV and NVP in a more contemporary TDF+FTC backbone as well as compare the NNRTI-containing regimens to one using LPV/r.

Even, however, without these studies, this is a rich, compelling literature, which clearly finds the clinical equivalence of EFV and NVP. Interestingly, the observational studies from low- and middle-income countries were unable to find the superiority of EFV reported from high-income countries. Recommendations will need to move beyond efficacy to examine specific toxicities and risk of resistance.
TDF vs. d4T or AZT

The side effects that make d4T and AZT somewhat difficult to use in low- and middle-income countries have led to consideration of TDF as a first-line NRTI, either in combination with 3TC or FTC. Only three studies, all industry sponsored, allow for direct comparison of d4T or AZT-containing regimens compared to TDF-containing regimens. Two of these studies (Gallant 2004; Rey 2009) use 3TC as the second NRTI and allow for direct comparisons; the third (Arribas 2008) compares AZT+3TC with TDF+FTC. The two large studies (Arribas 2008; Gallant 2004), each of which enrolled more than 500 patients, had equivalent findings: that AZT+3TC+EFV and d4T+3TC+EFV had similar efficacy, tolerability and safety as TDF+FTC+EFV and TDF+3TC+EFV, respectively. Of note, however, is the much smaller DAUFIN study (Rey 2009) that prematurely discontinued the study after failure of the TDF+3TC+NVP arm to suppress viral replication in eight of 35 participants. This same problem was not seen in the TDF+3TC+EFV arm reported by Gallant (2004).

Additionally, none of the observational studies examined these combinations in low- and middle-income settings. Although they found consistent evidence of the superiority of TDF-containing regimens to those containing AZT or d4T, there was also a finding of slower CD4 cell recovery (Mocroft 2006; Wolbers 2007). An industry-sponsored cost-effectiveness analysis comparing TDF/FTC to AZT/3TC random combinations found that TDF/FTC was both cost-saving and more effective over a two-year period (Sanchez-de-la-Rosa 2008); these findings were confirmed by another study that suggested that TDF+3TC would cost less than AZT+3TC over a three-year period (Fernandez Lison 2005).

Taken together, this literature, while all industry-sponsored, is of moderate quality, with two large studies with 144-week follow up adding to its precision and at least some patients enrolled from Latin American countries. It supports the superiority of TDF and fixed-dose TDF/FTC over more conventional d4T and fixed-dose AZT/3TC in NRTI-backbones for initial therapy of HIV infection in two important respects: tolerance and retention in studies and development of drug resistance. Again, the one trial that examined AZT and TDF in combination with 3TC and NVP (as opposed to EFV, which was used in Studies 903 and 934), ended early because of high rates of early virologic failure in the TDF+3TC+NVP arm. This was similarly seen in other studies that were employing a TDF+3TC backbone to compare NVP and ATV/r (Lapadula 2008) where early failure only occurred in the NPV arm or examined TDF+3TC+NVP in a single-arm study (Towner 2004). As TDF+3TC+NVP is one of the current WHO recommendations for first-line therapy, additional data from cohort studies is needed to judge the effectiveness of this regimen. Of note is the PEARLS trial (ACTG A5175), being conducted in poor communities in Africa, Asia, Haiti, South America and the United States of America, that is planning on enrolling 1 574 patients in a three-arm comparison of AZT+3TC+EFV, FTC+ddI+EFV+ATV and TDF+FTC+EFV. This study will add substantially to this literature and will complement Arribas 2008 by providing a direct comparison of AZT and TDF in dual NRTI backbones with EFV.

TDF vs. ABC

Two randomised controlled trials have compared ABC-3TC and TDF-FTC in combination with either EFV or a RTV-boosted protease inhibitor and a non-randomised study that compared open-label ABC-3TC and TDF-FTC with either EFV or a protease inhibitor; another two studies are continuing these evaluations. We found no studies that compared ABC with TDF without FTC. The three studies with published results had findings in the same direction, primarily that ABC-3TC and TDF-FTC were therapeutically equivalent and well tolerated nucleoside
backbones for triple-drug therapy. In fact there were no statistically significant differences in meta-analysis of the three studies for six major outcomes.

The finding that ABC-3TC was less potent and was associated with a more rapid time to Grade 3/4 adverse events than TDF-FTC in patients with HIV-1 RNA plasma loads $\geq 100,000$ copies/mL in ACTG 5202 is, however, troubling (Sax 2008). A meta-analysis (Pappa 2008) of 2,940 ART-naive patients beginning ABC+3TC-containing arms in six other trials found that by 48 weeks subjects with HIV-1 RNA levels of $<100,000$ and $\geq 100,000$ had similar experiences and that between 87% and 95% did not experience virological failure, as defined by ACTG5202, by 48 weeks. Additionally, safety endpoints were similar in both strata.

The two large studies, ACTG 5202 and HEAT, will together contribute more than 1,700 subjects, and about 530 patients from the remainder of ACTG 5202 should be randomised to receive EFV-containing regimens. While the findings of the HEAT Study are certainly encouraging (Smith 2009), especially when taken along with the meta-analysis of ABC+3TC arms from other studies (Pappa 2008), it is unlikely that RTV-boosted protease inhibitors will be widely likely in low- and middle-income countries in the near future. Unfortunately, aside from very small numbers of patients who received EFV in Manfredi’s (2008b) non-randomised study, there is essentially no published experience comparing ABC+3TC+EFV to TDF+3TC+EFV, which is likely what would be used in resource-constrained settings. The ASSERT study, which is projected to be completed in October 2009, may be able to provide additional data on this comparison but is primarily a study of renal and other metabolite endpoints. Moreover, we were unable to identify any observational studies that would bear on this question. Thus, while data collection is ongoing in ACTG 5202 and ASSERT, there is no basis for making a recommendation to use either of these two regimens.

Like all meta-analyses, our findings are limited by the completeness of our search and the literature that is available. Larger studies that more directly study the questions posed by WHO could change our results. However, at the present we find essentially no evidence from randomised controlled trials, non-randomised trials or observational studies from low- and middle-income countries that clearly indicate the superiority of d4T over AZT, EFV over NVP, TDF over AZT or d4T or TDF over ABC in triple-drug antiretroviral regimens for treatment-naive patients.
REFERENCES

General references

Barroso 2000

Bendavid 2009

Brokek 2008

Bunnell 2008

Cu-Uvin 2000

Division of AIDS, NIAID 2004
Division of AIDS, National Institute of Allergy and Infectious Diseases, National Institutes of Health. Table for Grading the Severity of Adult and Pediatric Adverse Events. Bethesda, Maryland: National Institutes of Health, December 2004.

Garcia 1999

Gilks 2006

Gupta 1997

Guyatt 2008
Higgens 2008

Holtgrave 2005

Kaplan 1996

Musicco 1994

Operskalski 1997

Palella 1998

Pedrazza 1999

Porco 2004

Quinn 2000

Ragni 1998

UNAIDS 2008

**UNAIDS/WHO**


**Vernazza 1997a**


**Vernazza 1997b**


**Vernazza 2000**


**Wells 2009**


**WHO 2002**


**WHO 2003**


**WHO 2004**


**WHO 2006a**


**WHO 2006b**

**WHO 2006c**


**WHO 2006d**


**WHO 2008**

d4T vs. AZT

**Amin 2003**

**Anastos 2007**

**Antiretroviral Therapy Cohort Collaboration 2006**

**ART-LINC Collaboration of International Databases to Evaluate AIDS**

**Benetucci 2000**
Benetucci J, Oliva S, Redini L, Pegliese D, Cantaravoro R, Maranzana A HAART in the daily practice. Comparison of the potency between two PI including regimens (Gadis Cohort) [Abstract ThPeB5003]; International Conference on AIDS, Durban, South Africa, 9-14 July 2000.

**Blanco 2003**

**Bogner 2001**

**Bolton-Moore 2007**

**Boubaker 2001**
Boullé 2007


Bussman 2009


Carbonell 2004


Carr 2000


Carr 1999a


Carr 1999b


Carr 2009


Castelnuovo 2009


Chen 2004a


Chen 2004b

D:A:D Study Group 2008

De Wit 2008

Dokekias 2008

Domingo 2003

Dragovic 2003

Dubé 2004

Dubé 2007a

Dubé 2007b
Dubé MP, Parker MA, Mulligan K, Tebas P, Robbins GK, Roubenoff R, Grinspoon SK. Effects of potent antiretroviral therapy on free testosterone levels and fat-free mass in men in a

Dubé 2005

Eiros Bouza

Enanoria 2004

Eron 1998a

Eron 1998b

Eron 2000

Fellay 2001

Fisher 1998
Fisher M, Stoehr A, Podzamczer D, Angarano G, Silleni M, Ledeine JM, Dunkle LM. A randomized double-blind study of d4T+ddl vs ZDV+ddl as initial treatment in subjects with CD4 count less than or equal to 500 cells/mm³ [Abstract 204]. 5th Conference on Retroviruses and Opportunistic Infections, Chicago, USA, 1-5 February 1998.

French 2002

Funk 1999


Gandhi 2006


Gathe 2002


Geijo Martínez 2006


George 2009


Gulick 1998


Gulick 1998


Harrigan 2004


Havlir 1998

Hill 2009

Hogg 1998

Hudson 1998

Hulgan 2005

Hulgan 2006

Jahnke 1999

Jimenez-Nacher 1998

Justice 2000

Justice 2004
Kallianpur 2006

Kamya 2007

Kumar 2003

Kumar 2004

Kumar 2006

Kumarasamy 2006

Kumarasamy 2008

Lauenroth-Mai 2002

Laurent 2008
nevirapine and lamivudine plus zidovudine or stavudine in Cameroon. AIDS Res Hum Retroviruses 2008; 24:393-9.

Law 2003

Li 2008

Lodwick 2008

Lonergan 2001

Matthews 2000

Mercier 2009

Mocroft 2006

Mocroft 2005

Moyle 2004

Mulligan 2006

Njoroge 2009


Nolan 2003


Pavia 2002

Pavia A, Olson JS, Mauney J, Wills B. Differences in hematological complications in the selection of thymidine analog regimen therapy trials (START I and START II) [AbstractTuPeB4499]. International Conference on AIDS, Barcelona, Spain, 7-12 July 2002.

Pazare 2008


Picard 2001


Pujari 2004


Pujari 2005


Robbins 2003


Saint-Marc 1999

Shafer 2003


Shah 2006


Shet 2009


Shlay 2008


Siegfried 2006


Smeaton 2001


Squires 2000a


Squires 1998


Squires 2000b


Squires 1999

Sulkowski 2005

Torriani 2008

Toure 2008

Van Griensven 2007

Wamalwa 2007

Wester 2007

Wood 2006

Zhou 2007
**EFV vs. NVP**

**Ananworanich 2005**


**Anastos 2007**


**Annan 2005**


**Annan 2009**


**Antiretroviral Therapy Cohort Collaboration**


**Aranzabal 2005**


**Arranz-Caso 2004**


**Arribas 2001**


**ART-LINC Collaboration of International Databases to Evaluate AIDS (IeDEA) 2008**

Aurpibul 2007

Ayala Gaytan 2004
Ayala Gaytan JJ, de la Garza ERZ, Garcia MC, Chavez SBV. Nevirapine or efavirenz in combination with two nucleoside analogues in HIV infected antiretroviral naïve patients. Med Intern Mex 2004; 20:24

Bannister 2008

Basu 2006

Berenguer 2008

Boulle 2007

Boulle 2008

Braithwaite 2007

Bruck 2008

Bussman 2009

Chen 2004

Chung 2002
Chung U, Fichtenbaum CJ. A minority of patients achieve long-term durable viral suppression with potent antiviral therapy [Abstract TuPeB4452]. XIV International Conference on AIDS, Barcelona, Spain, 7-12 July 2002.

Cozzi-Lepri 2002

de Beaudrap 2008

Ding 2008

Dokekias

Ena 2003

Forna 2007

George 2009

Han 2005

Hartmann 2005

Hartmann 2005

Kamya 2007

Kappelhoff 2005a

Kappelhoff 2005b

Keiser 2002

Keiser 2003

Kumarasamy 2006

**Kumarasamy 2008**


**Lapphra 2008**


**Lima 2009**


**Lonca 2002**


**MacArthur 2001**


**Madec 2007**


**Maleewong 2008**


**Manfredi 2004**


**Manfredi 2005**


**Manfredi 2006**
Manfredi R, Calza L. Nevirapine versus efavirenz in 742 patients: no link of liver toxicity with female sex, and a baseline CD4 cell count greater than 250 cells/microl. AIDS 2006; 20:2233-6.

Manosuthi 2004

Manosuthi 2008

Manosuthi 2009

Martín-Carbonero 2003

Matthews 2003

Matthews 2000

Matthews 2002

McConnell 2009

Mocroft 2005

**Moyle 2000**


**Moyle 2003**


**Mugavero 2008**


**Nachega 2007**


**Nachega 2008**


**Nakanjako 2008**


**Neuwelt 2003**


**Núñez 2002**


**Palmon 2002**


**Patel 2006**

**Phillips 2004**


**Phillips 2001a**


**Phillips 2001b**


**Potard 2007**


**Sanne 2005**


**Sathia 2008**


**Sheran 2005**


**Shipton 2009**


**Siegfried 2006**

Skowron 2001

Sow 2006

Srasuebkul 2007

Stern 2004

Sulkowski 2002

Sungkanuparph 2006

Torre 2001

Trein 2005

van den Berg-Wolf 2006

van den Berg-Wolf 2008
van den Berg-Wolf M, Hullsiek KH, Peng G, Koziel MJ, Novak RM, Chen L, Crane LR, Macarthur RD; CPCRA 058 Study Team, the Terry Beirn Community Programs for Clinical

van Leth 2005a

van Leth 2004a
van Leth F, Andrews S, Grinsztejn B, Wilkins E, Lazanas MK, Lange JM, Montaner J; 2NN study group. Virologic failure in antiretroviral therapy naive patients is only determined by extreme low values of CD4+ cells or high values of HIV-1 RNA concentration, not by choice of treatment with nevirapine or efavirenz [Abstract 550]. 11th Conference on Retroviruses and Opportunistic Infections. San Francisco, USA, 8-11 February 2004.

van Leth 2004b

van Leth 2004c

van Leth 2004d

van Leth 2006a
van Leth F, Hall DB, Lange JM, Reiss P. Plasma lipid concentrations after 1.5 years of exposure to nevirapine or efavirenz together with stavudine and lamivudine. HIV Med 2006; 7:347-50.

van Leth 2005b

van Leth 2006b

van Leth 2006c

van Leth 2006d

Varma 2009

Wamalwa 2007

Wester 2005

Wit 2007

Wolff 2006

Zhang 2009

Zhang 2007

Zhou 2006

**d4T or AZT vs. TDF**

**Anonymous 2005**

**Arribas 2008**

**Carr 2009**

**Cassetti 2007**

**Clotet 2008**

**Fernandez 2005**

**Frampton 2006**

**Frampton 2005**

**Gallant 2003**

**Gallant 2002**
Gallant J, Staszewski S, Pozniak A, Lu B, Miller MD, Coakley DF, Cheng A. Favorable lipid and mitochondrial (mt) DNA profile for tenofovir disoproxil fumarate (TDF) compared to stavudine (d4T) in combination with lamivudine (3TC) and efavirenz (EFV) in antiretroviral therapy (ART)

Gallant 2004a

Gallant 2006a

Gallant 2006b

Gallant 2004b

Gallant 2008

Havlir 2005

Izzedine 1995

Jones 2005

Laesig

Lapadula 2008

**Lewis 2004**


**Lodwick 2008**


**Makinson 2008**


**Manfredi 2006**


**Margot 2006**


**Mascho 2007**


**McGowan 2001**


**Mocroft 2006**


**Parienti**


**Pozniak 2006**

Pozniak AL, Gallant JE, DeJesus E, Arribas JR, Gazzard B, Campo RE, Chen SS, McCall D, Enejosa J, Toole JJ, Cheng AK. Tenofovir disoproxil fumarate, emtricitabine, and efavirenz

Pozniak 2003

Rey 2009

Sanchez-de-la-Rosa 2008

Staszewski 2002
Staszewski S, Gallant J, Pozniak A, Suleiman JM, DeJesus E, Koening E, Coleman S, Lu B, Cheng AK, Coakley DF; the 903 Study Team. Efficacy and safety of tenofovir disoproxil fumarate (TDF) versus stavudine (d4T) when used in combination with lamivudine (3TC) and efavirenz (EFV) in HIV-1 infected patients naive to antiretroviral therapy (ART): 48-week interim results [Abstract LbOr17]. 14th International Conference on AIDS, Barcelona, Spain, 7-12 July 2002.

Staszewski 2003a

Staszewski 2003b

Towner 2004

Willig 2008

Wolbers 2007

**TDF vs. ABC**

**Anastos 2007**


**Hill 2009**


**Manfredi 2008a**


**Manfredi 2008b**


**Pappa 2008**

Pappa K, Hernandez J, Ha B, Shaefer M, Brothers C, Liao Q. Abacavir/lamivudine (ABC/3TC) shows robust virologic responses in ART-naïve patients for baseline (BL) viral loads (VL) of ≥100,000 c/mL and <100,000 c/mL by endpoint used in ACTG5202 [Abstract THAB0304]. XVII International Conference on AIDS, Mexico City, August 3-8, 2008.

**Sax 2008**


**Smith 2009**