Effects of highly active antiretroviral therapy among HIV-infected patients

Results from randomised and observational studies

Ole Kirk

This review has been accepted as a thesis together with seven previously published papers, by the University of Copenhagen, October 14, 2003, and defended on December 19, 2003.

Department of Infectious Diseases and Copenhagen HIV Programme (CHIP), H.S Hvidovre Hospital, Hvidovre.

Correspondence: Ole Kirk, Copenhagen HIV Programme, H.S Hvidovre Hospital, Køttegård Allé 30, DK-2650 Hvidovre.

Official opponents: Bente Klarlund Pedersen, professor, MD, and Henrik Toft Sørensen, professor, MD.


INTRODUCTION

More than 20 years have passed since the first reports of patients with acquired immunodeficiency syndrome (AIDS) – later documented to be due to infection with human immuno deficiency virus type 1 (HIV) – in USA, and shortly thereafter in Europe [Masur et al., 1981; Gottlieb et al., 1981; Gerstoft et al., 1982]. As of 2002, more than 42 million people are living with HIV/AIDS globally, and more than 27 million have died. At least 95% of the people diagnosed with HIV in 2002 are from developing countries, and the number of HIV-infected patients is rapidly growing in Sub-Saharan Africa, Asia, and Eastern Europe. In Western Europe and North America the prevalence of HIV also continues to increase, but substantially more slowly than in the worst affected regions [UNAIDS/WHO, 2002].

The first therapeutic advances within the field of HIV-infection occurred in the late 1980’s with the introduction of the first nucleoside reverse transcriptase inhibitor (NRTI), Zidovudine, which in randomised controlled trials (RCTs) was documented to prolong survival, though the effect later was shown to be of limited duration [Fischl et al., 1987; Fischl et al., 1990; Volberding et al., 1990; Hamilton et al., 1992; Lundgren et al., 1994; Anonymous, 1994]. In the same time period, chemoprophylaxis against opportunistic infections (OIs) also came into use and reduced the number of new OIs as well as prolonged the relapse free time after an initial event [Haverkos, 1987; Jacobson et al., 1988; Phair et al., 1990; Bozzette et al., 1991; Montaner et al., 1991; Jacobson et al., 1993].

Following these initial advances, it was shown that switch to the later licensed NRTIs, Didanosine and Zalcitabine, improved the clinical prognosis for patients treated with Zidovudine, and subsequently combination therapy of two NRTIs was demonstrated to further improve the clinical prognosis [Kahn et al., 1992; Collier et al., 1993; Abrams et al., 1994; Anonymous, 1996; Saravolatz et al., 1996].

The discovery of a new drug class, protease inhibitors (PI) targeting another point of the HIV replication, was a potential breakthrough. Studies of patients receiving PI therapy documented a large turnover rate of HIV virions and infected CD4 receptor positive T lymphocyte cells (CD4 cells), which combined with a high mutation rate of HIV indicated a need of combining several, potent antiretroviral agents [Ho et al., 1995; Perelson et al., 1996]. Such PI-containing combination therapy lead in RCTs to improved survival as well as decreasing HIV-RNA and increasing CD4 cell count; the latter two being established predictive factors for clinical progression [Collier et al., 1996; Mellors et al., 1996; Hammer et al., 1997; Cameron et al., 1998]. Due to the promising results, the concept of highly active antiretroviral therapy – HAART – a regimen of two NRTIs and a PI or alternatively a non-nucleoside reverse transcriptase inhibitor (NNRTI) – was launched.

However, the durability and long-term clinical benefit as well as toxicity were not known due to the quick approval and licensing of these drugs by the regulatory authorities. Since 1996 the accelerated drug approval process has allowed for the introduction of new drugs based on 24 weeks of follow-up of less than 500 patients using changes in HIV-RNA as the primary end-point. This process has saved lives but has also lead to widespread use of drugs with relatively limited information on their long-term efficacy including the clinical efficacy (as per design of these pivotal RCTs) and no information on long-term and rare toxicities [Anonymous, 2002].

It was also unclear to which extent the results from RCTs, often representing idealised settings due to strict inclusion and exclusion criteria, would translate into clinical effects in a broader spectrum of HIV-infected patients including those not traditionally included in the trials [Madge et al., 2000; Gifford et al., 2002].

OBJECTIVES

The objectives of the present work, initiated in 1997, were therefore:

1. to analyse the rate of introduction of HAART in Europe [Kirk et al., 1998],
2. to analyse the longer-term effects of HAART – in terms of changes in common AIDS defining events (ADEs) [Kirk et al., 2000; Kirk et al., 2001c],
3. to analyse the degree of clinical effectiveness of the HAART-induced immune restoration [Kirk et al., 1999b; Kirk et al., 2002b],
4. to compare the virological, immunological and clinical effects of specific HAART regimens [Kirk et al., 1999a; Kirk et al., 2001b],
5. to discuss the analysis of RCTs with virological end-points and the influence of premature switch from randomised therapy [Kirk et al., 2001a; Kirk et al., 2002a], and
6. to discuss the extent to which observational studies provide data for rational clinical decision making.

STUDIES AND METHODOLOGY

THE EuroSIDA STUDY

EuroSIDA is an observational study of 9805 HIV-infected patients at 70 centres in 26 European countries with 6 monthly follow-up using a standardised data collection form (latest version available at http://www.cphiv.dk/eurosida/eurosida_status.htm) [Lundgren et al., 1997; Kirk et al., 1998]. Pre-determined numbers of patients with a scheduled visit have been enrolled from the participating outpatient clinics in an attempt to ensure inclusion of a representative subset of patients followed at the specific centres. Enrolment was performed regularly in pre-determined time periods (cohort I-V) to better reflect the patient population followed at the participating sites and to reduce a potential long-term survivor bias.

Sites participating in the EuroSIDA study are often university-associated and identified based on their commitment to research projects and are thus likely to represent the golden standard for diagnostics and treatment in the countries involved. Caution should therefore be taken when extrapolating results from the EuroSIDA study to patients in other settings, and to patients not attending the clinics regularly. However, one strength of the analyses made within this multi-national study is the large, heterogeneous study population from 70 sites with different treatment policies and pattern of opportunistic pathogens [Lundgren et al., 1997; Mocroft et al., 1998a].

A low rate of loss-to-follow-up is crucial for interpretation of data in a longitudinal study. As of January 2002, 5482 of 8549 patients...
Table 1. Factors associated with no follow-up data after March 2000 within the EuroSIDA study. Results from a multivariable logistic regression model.

<table>
<thead>
<tr>
<th>HIV transmission category</th>
<th>Odds ratio (95%-CI)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intravenous drug use v. other routes</td>
<td>1.45 (1.24-1.69)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>CD4 count</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥200 v. &lt;200 cells/mm²</td>
<td>0.42 (0.36-0.49)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>HAART v. no HAART</td>
<td>0.26 (0.22-0.30)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Region of EuroSIDA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>East v. north</td>
<td>1.02 (0.62-1.67)</td>
<td>0.93</td>
</tr>
<tr>
<td>South v. north</td>
<td>1.49 (1.25-1.78)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Central v. north</td>
<td>1.59 (1.33-1.89)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Inclusion period</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cohort II v. cohort I</td>
<td>0.69 (0.57-0.84)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Cohort III v. cohort I</td>
<td>0.65 (0.55-0.77)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Cohort IV v. cohort I</td>
<td>0.35 (0.27-0.46)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Supplementary multivariable model including HIV-RNA

| HIV-RNA ≥500 v. <500 copies/mL                           | 1.84 (1.56-2.17)    | <0.001  |

CI: confidence interval.

HAART: highly active antiretroviral therapy, defined as ≥3 drugs including at least one protease inhibitor, one non-nucleoside reverse transcriptase inhibitor or Abacavir.

enrolled into EuroSIDA cohort I-IV (cohort V yet without any follow-up data) were under active follow-up (i.e. with follow-up after June 2000). 1910 had died and 1157 did not have any recent follow-up data. Patients without HAART, at low CD4 cell count and high HIV-RNA were at higher risk of having no recent follow-up, and these patients would be expected to have a worse prognosis compared with those who remained under follow-up (Table 1 [Mellors et al., 1996; Mocroft et al., 1998a; Lundgren et al., 2002]). This finding suggests that the assessments of incidences of new ADEs and the mortality within the EuroSIDA study are minimum estimates. In this connection, it is reassuring that the rate of loss-to-follow-up is fairly stable, approximately 2% per year (data not published), which is lower than or at the same level as mono-centric or national cohort studies [Kirk et al., 1999a; Dudley et al., 1995; Hammer et al., 1997; Sudre et al., 2000; Grabar et al., 2000b].

To ensure a high data quality, an extensive data assurance program has been implemented. This includes use of standardised instructions for completion of the forms as well as diagnostic criteria for OIs and malignancies, and circulation of newsletters dealing with key information with respect to the data collection (http://www.cphiv.dk/eurosida/newsletter/newsletter.htm). Standardised control of data entry is performed at the co-ordinating centre, and the validity of the data provided by the centres is further checked at site visits where all clinical events as well as randomly selected patients are reviewed [Lundgren et al., 1997; Kirk et al., 2001c].

**DANISH PROTEASE INHIBITOR STUDY**

The Danish PI study is a randomised five-centre study of 622 patients within Denmark, which was initiated in October 1996 to compare the efficacy and toxicity of specific first-line HAART regimens. The three initial study regimens compared included 2 NRTIs combined with one or two of the PIs initially licensed in Denmark; Indinavir, Ritonavir or Ritonavir/Saquinavir hard gel capsules. Patients were randomised to one of these treatment regimens or two later added treatment options until January 2001, and follow-up is still ongoing. Data is collected in a standardised manner at structured follow-up visits regardless of eventual switch from randomised therapy [Kirk et al., 1999a].

The study intended to reflect clinical practice in Denmark and only had few exclusion criteria; in particular there were no restrictions with respect to HIV transmission category, co-infection with hepatitis B and C, or prior NRTI therapy. However, as in most other studies from this period, no data has been collected on how many of the patients who were eligible for the study were actually included [Begg et al., 1996; Altman et al., 2001].

**DANISH COHORT OF PATIENTS INITIATING THEIR FIRST HAART REGIMEN**

The study population included 505 Danish patients within the EuroSIDA study who started an initial HAART regimen, for whom additional information on the most likely reason for potential switch from the PI component of the initial HAART regimen was retrospectively extracted from the patient records. Due to the design, these patients were representative for patients receiving a HAART regimen and being seen at major Danish HIV-clinics, whereas the limitations include the retrospective design of the study, the restriction of the analysis to switch from the PI components of the HAART regimen and collection of only the principal reason for switching therapy as depicted from the patient notes [Kirk et al., 2001a].

**DANISH COHORT OF PATIENTS DISCONTINUING CHEMOPROPHYLAXIS**

This prospective study was initiated in December 1997 and included patients stopping chemoprophylaxis to protect against the development of OIs caused by cytomegalovirus (CMV), Pneumocystis carinii, Toxoplasma gondii, Mycobacterium avium complex (MAC) and Cryptococcus neoformans according to guidelines developed by the study group. Patients were followed after interruption of chemoprophylaxis and chart review was performed for all patients before analysis.

Similar to what have been reported from RCTs within this research area, data is not available for whether discontinuation of chemoprophylaxis was used for all patients with similar history of HAART and CD4 cell increases, or whether other additional criteria such as regularity of follow-up, adherence to antiretroviral therapy (ART), the degree of immune restoration as well as social situation also influenced the decision [Kirk et al., 1999b; Mussini et al., 2000; Lopez Bernaldo de Quiros et al., 2001].

**A JOINT DATA SET OF PATIENTS INTERRUPTING MAINTENANCE THERAPY**

The data set included patients who interrupted secondary prophylaxis or maintenance therapy (MT) against infections caused by CMV, Toxoplasma gondii, MAC and Cryptococcus neoformans, and was created by merging data from seven existing HIV-cohorts in Europe. Patients identified in the cohorts fulfilled pre-determined criteria in terms of the prophylactic drugs used and interruption of MT at a CD4 cell count above 50 cells/mm³.

A potential drawback is that data regarding reasons for interrupting MT was not available. Patients may interrupt MT according to a variety of criteria, as described above, and the decision to stop the treatment was not done based on commonly agreed well-defined guidelines [Kirk et al., 2002b].

**UPTAKE OF HAART IN CLINICAL PRACTICE**

In North America and Europe there was a marked increase in use of HAART in the mid 1990s, predominantly during 1996, thus indicating a close temporal relation to the first preliminary reports from RCTs reporting the improvement in clinical, immunological and virological status of patients treated with HAART [Collier et al., 1996; Hammer et al., 1997; Gulick et al., 1997; Palella et al., 1998; Kirk et al., 1998; Cameron et al., 1998; Gehrhardt et al., 1998; Spira et al., 1998; Mocroft et al., 1999d; Sackoff et al., 2000; CASCADE Collaboration, 2000].

As of late 1997, the odds of being on a PI-containing regimen were fairly similar across the western part of Europe when adjusting
for regional differences in demographics and immune status. However, before then there had been pronounced regional differences in the uptake of combination therapy, PI therapy as well as usage of specific NTRIs and PIs, suggesting that other factors than results of RCTs also influenced the uptake of therapy in clinical practice including access to new drugs after marketing approval by the European Commission, local traditions and socio-economic factors [Lundgren et al., 1997; Kirk et al., 1998].

Within EuroSIDA, patients of European and non-European origin initiated HAART at similar CD4 cell levels, whereas other studies have reported divergent influences of the ethnicity on uptake of HAART, probably reflecting differences in study population and study design. These studies do, however, not address whether specific ethnic groups may not benefit from therapy because they do not have contact with HIV-clinics [Blaxhult et al., 1999; Sackoff et al., 2000; Hsu et al., 2001].

Patients infected by intravenous drug use were less likely to start HAART, though once started, their response to HAART did not differ from that of other groups, suggesting that clinicians were able to identify the intravenous drug users who could adhere to the complex HAART regimens. Further, patients with lower level of education, homeless patients, and patients who had not previously received ART were less likely to start HAART, as were, not surprisingly, patients with higher CD4 cell count, lower HIV-RNA and without prior diagnosis of AIDS, whereas gender and age in general were not associated with chance of starting HAART. These relations were based on the period immediately after the introduction of HAART, and may have changed or may change due to an increasing understanding of the whole HAART area, increased drug options, and other improvements in patient care [Junghans et al., 1999; Mocroft et al., 1999c; Sackoff et al., 2000; CASCADE Collaboration, 2000; Mocroft et al., 2000a; Hsu et al., 2001]. Whereas the decline in the proportion of patients followed at HIV-clinics without treatment suggests a change in the indication of initiating ART, the optimal time of initiating therapy for the individual patient still remains unaddressed in a RCT, which, if performed, should follow several thousands of patients for many years [Kirk et al., 1998]. Large prospective observational studies have recently provided evidence for similar virological, immunological and clinical responses to HAART as long as it is initiated at a CD4 cell count above 200 cells/mm3, and recent international treatment guidelines have been modified according to these results [Cozzi et al., 2001; Phillips et al., 2001b; Hogg et al., 2001; Anonymous, 2001b; Yeni et al., 2002; Egger et al., 2002].

An updated analysis within EuroSIDA illustrates that as of September 2001 approximately 80% received HAART, and an increasing proportion were treated with 4 drugs or more (Figure 1). Of note, this represents cross sectional assessments of regimens used in consecutive time intervals without any adjustment for previous use of ART, and whereas the results for 1996-1997 may well represent the first (or second) HAART regimen, the results in more recent years include increasing numbers of patients who have switched to second-line or even salvage regimens [Mocroft et al., 2001a; Mocroft et al., 2002b]. Further, the analysis does not differentiate between 4 active antiretroviral drugs and Ritonavir-boosted PI-regimens (described further in the section on 'Comparison of different HAART regimens'), and therefore does not per se document a switch towards starting a quadruple regimen as the initial HAART regimen.

CONCLUSION

Marked changes in use of ART of North American and European HIV-infected patients occurred in the second half of the 1990's closely related to the presentation of RCTs, but some subgroups of patients who were less likely to receive HAART have been identified. Future analyses should address whether results from recent RCTs have led to similar changes in choice of components of the initial and second-line HAART regimens in clinical practice, and evaluate the use of drugs for which the clinical effect has not yet been sufficiently evaluated in RCTs, for example immune stimulating drugs such as Interleukin-2 [Staszewski et al., 1999; Staszewski et al., 2001; Hammer et al., 2002; Emery et al., 2002].

THE CLINICAL PROGNOSIS FOR HIV-INFECTED PATIENTS AFTER INTRODUCTION OF HAART

MORTALITY

Use of HAART regimens studied in the pivotal RCTs led to a marked and significant decrease in mortality, and a similar effect was reported in several large cohorts of unselected patients [Hammer et al., 1997; Mouton et al., 1997; Gulick et al., 1997; Egger et al., 1997; Palella et al., 1998; Cameron et al., 1998; Mocroft et al., 1998a; CASCADE Collaboration, 2000; Stellbrink et al., 2000]. Overall, the mortality decreased from approximately 20 deaths/100 person-years of follow-up (PYF) in 1994-95 to less than 5 deaths/100 PYF in 1997-98, and the effect of HAART was most pronounced when initiated at low CD4 cell counts, though observed in all CD4 strata [Hammer et al., 1997; Egger et al., 1997; Cameron et al., 1998; Mocroft et al., 1998a]. In the observational studies there were statistically significant associations between the declining mortality and the

![Figure 1](image-url)
concomitant steep increase in use of combination therapy, in particular HAART, and differences in the survival rate across Europe were associated with regional differences in the uptake of HAART [Egger et al., 1997; Palella et al., 1998; Mocroft et al., 1999]

Of note, the mortality remained at a very low level within the EuroSIDA study based on data until January 2002 (Figure 2a). For example, after March 2001 the incidence of deaths was 1.4 (95%-confidence interval (CI): 1.0-1.8/100 PYF), and thus less than 18.8 (17.4-20.3) observed in patients followed before September 1995.

**MORBIDITY**

**General changes**

Concurrently, decreases in morbidity and rates of admission to hospital together with resolution of ongoing OIs were reported among patients who started HAART [Mouton et al., 1997; Murphy et al., 1997; Palella et al., 1998; Benfield et al., 1998; Carr et al., 1998a; Mocroft et al., 1999a; Ledergerber et al., 1999c; Mocroft et al., 2000b]. The overall incidence of new ADEs decreased from around 3 to 5 events/100 PYF from 1994 to 1998, and decreases were initially documented for common ADEs such as Pneumocystis carinii pneumonia (PCP), CMV chorioretinitis and disseminated MAC infection, all predominantly diagnosed at severe immune deficiency, but later also for other and less common diseases [Mouton et al., 1997; Palella et al., 1998; Mocroft et al., 1999a; de Gaetano et al., 2000; Jones et al., 2000; Baril et al., 2000; Girardi et al., 2000; Mocroft et al., 2000b; Tumbarello et al., 2001; Sacktor et al., 2001; Rossi et al., 2001; Abgrall et al., 2001; Badri et al., 2002]. Figure 2b illustrates and extents these changes until the beginning of 2002 within the EuroSIDA study.

As a consequence of differences in the declines of the individual ADEs, the relative pattern of ADEs changed compared with the pre-HAART era [Mocroft et al., 1998b; Blaxhult et al., 2000; Mocroft et al., 2000b; Blaxhult et al., 2002]. For example, representing very different parts of the spectrum of ADEs, the incidence of MAC infection was more pronouncedly affected by initiation of HAART compared with infection due to Mycobacterium tuberculosis. As a result, Mycobacterium tuberculosis has now generally become the most common mycobacterial agent among HIV-infected patients in Europe, although with marked regional differences [Kirk et al., 2000].

**Mycobacterioses**

The decreasing incidence of disease due to Mycobacterium tuberculosis from 1994 to 1999 within the EuroSIDA study was readily explained by changes in use of ART and in changes induced thereby, in particular the CD4 cell count [Kirk et al., 2000]. However, other factors – for example implementation of tuberculosis control programs including preventive therapy as suggested in an American surveillance study - may also have played some role, though it did not seem to be of substantial importance in Europe [Kirk et al., 2000; Jones et al., 2000].

For disseminated MAC infection, some, but not all, of the decline over calendar time could also be explained by increasing use of ART.
and changes in CD4 cell count. Increased use of primary prophylaxis (PP) against MAC, primarily among patients with severe immune deficiency, may also explain some of the decrease [Kirk et al., 2000].

Non Hodgkin lymphoma
Interestingly, the relative proportion of patients diagnosed with a relatively common HIV-related malignancy, non-Hodgkin lymphoma (NHL), also increased from 4% of all ADEs diagnosed in 1994 to 16% in 1998 [Mocroft et al., 2000b]. As longevity increased and the risk of other competing ADEs was markedly diminished, there were concerns of unchanged or even increasing incidences of malignancies, and in particular NHL [Jacobson et al., 1999]. The biological hypothesis was that a malign transformation and development of NHL is a multifactorial process not readily reversible by immune restoration (i.e. increasing CD4 cell count), as was believed to be the case for OIs and malignancies with a documented infectious etiology [Chang et al., 1994; Murphy et al., 1997; Benfield et al., 1998; Carr et al., 1998a].

Within the Swiss HIV Cohort Study there was no significant decline in the incidence of NHL from 1992-1994 to 1997-1998, and the incidence of NHL did not decrease significantly within the first 12 months of HAART, contrasting the significant decreases observed for all other ADEs analysed [Ledergerber et al., 1999a; Ledergerber et al., 1999c]. However, a large multi-cohort study later reported a decrease in the incidence of NHL from 1992: 1996 to 1997: 1999. This decrease was observed for primary brain lymphoma and immunoblastic lymphoma, but not for Burkitt's lymphoma [Anonymous, 2000]. With longer follow-up among patients on HAART, the EuroSIDA study found significant decreases from 1994 to late 2000 in all NHL subtypes for which data was collected: Burkitt, immunoblastic, primary brain lymphoma and other/unknown histology [Kirk et al., 2001c]. The decline was most pronounced for primary brain lymphoma, and all findings were confirmed in an updated analysis including follow-up till January 2002 (Figure 3a). There are two caveats for the interpretation of these subtype analysis: namely a change in the subtype classification of NHL (‘diffuse large B-cell’ is now used rather than ‘B-immunoblastic’), and a not ignorable number of patients with NHL being categorised as ‘other/unknown histology’, thus more detailed evaluation of the changes in the histological pattern of NHL being categorised as ‘other/unknown histology’ has not been possible so far [Harris et al., 1994; Vilchez et al., 2002].

To address this issue, the categories of NHL collected within the EuroSIDA study is currently being revised based on the latest classification system [Jaffe et al., 2001].

Further, among patients starting HAART, the incidence of NHL decreased significantly from the first year after starting HAART to more than one year after starting HAART [Kirk et al., 2001c]. The result held true in an updated analysis with longer follow-up, which allowed for further subdividing the time period on HAART (Figure 3b).

Taken together, these results suggest that NHL behaves as an OI, though the protective effect of HAART seems to arise after a substantially longer period on HAART compared with the traditional OIs [Ledergerber et al., 1999a; Kirk et al., 2001c]. This pattern may however differ between the subtypes of NHL, and lymphomas associated with Epstein-Barr virus and human herpes virus type 8 such as primary brain lymphoma and body cavity-based lymphoma, may well behave more like a traditional OIs [Pedersen et al., 1991; Nador et al., 1996; Hansen et al., 2000].

Interestingly, the incidence of NHL among patients starting HAART early in the era of HAART (i.e. before June 1997) was higher than that of patients starting HAART more recently, and this difference was at least in part explained by sub-optimal CD4 and HIV-RNA responses to HAART in the early period compared with later [Kirk et al., 2001c].

Inflammatory immune response
Early reports suggested a ‘HAART-induced inflammatory immune response’, namely development of OIs with an altered clinical presentation at relatively high CD4 cell counts shortly after initiation of HAART [Jacobson et al., 1997; Race et al., 1998; Jacobson and French, 1998; Sepkowitz, 1998; French et al., 2000]. However, when stratifying by latest CD4 count, the incidence of OIs still declined suggesting that for a given CD4 count, the risk of ADE was lower in the HAART era compared with the pre-HAART era, and the increasing CD4 cell count at diagnosis of an ADE thus represented a cohort effect [Kirk et al., 2000; Mocroft et al., 2000b; Tumbarello et al., 2001; Abgrall et al., 2001].

More detailed characterisation of potentially new presentations of known clinical entities based on prospectively collected, standardised data is warranted.
New diseases

The decline in conventional ADEs, which was described in the pre-HAART epidemic, might not capture all the diseases seen in HAART treated patients. There have been some indications of a changing pattern of death causes and terminal diseases; relatively more patients now die due to non-HIV-related diseases including cardiovascular disease and hepatitis C and correspondingly fewer following an ADE [Valdez et al., 2001; Louie et al., 2002; Mocroft et al., 2002a].

Further analysis of prospectively collected and standardised data on severe diseases as well as death causes is required to address whether these changes reflect that HIV-infected patients now live longer and consequently acquire and die of non-HIV-related diseases or whether the changes are explained by new HIV-related diseases or even side effects of the ART [Friis-Møller et al., 2003].

Overall, continuous surveillance of the patterns of morbidity and mortality remains a high priority in EuroSIDA and other large-size cohort studies (Figure 2a and 2b).

SURVIVAL AFTER DIAGNOSIS OF AN ADE

Several studies have reported an improved prognosis after the diagnosis of most of the individual ADEs in recent years, though most of these studies had quite low power due to the low number of new ADEs [Miralles et al., 1998; De Luca et al., 2000; Tumbarcelo et al., 2001; Girardi et al., 2001; Rossi et al., 2001; Fordyce et al., 2002; Dore et al., 2002]. For NHL, however, the improvement seemed to be more modest, if present. Most studies have reported a better prognosis, primarily for systemic lymphoma, but a few also for primary brain lymphoma [Aviles and Halabe, 1999; Levine et al., 2000; Thiessard et al., 2000; Navarro et al., 2001; Ratner et al., 2001; Vacher et al., 2001; Antinori et al., 2001; Hoffmann et al., 2001; Besson et al., 2001; Tam et al., 2002; Gerard et al., 2002]. Of note, the survival for primary brain lymphoma or systemic NHL did not improve in several recently published large cohort studies. This is of concern, though lack of detailed data on chemotherapeutic therapy and staging of disease in these and several other studies should be acknowledged as potential confounding factors [Kirk et al., 2001c; Fordyce et al., 2002; Dore et al., 2002].

Further analyses of large numbers of patients and longer follow-up as well as more detailed information on the subtype of NHL and data of chemotherapy are warranted to assess whether new promising, better targeted treatment options (e.g. monoclonal antibody directed against the CD20 antigen, rituximab) with less toxicity and interactions combined with a high efficacy for NHL-patients in general, will also translate into more pronounced changes for HIV-infected patients [Czuczman et al., 1999; Hainsworth et al., 2000].

RISK FACTORS FOR CLINICAL PROGRESSION

In clinical practice, the ability to identify patients at high risk for clinical progression is essential. As in the pre-HAART era, the CD4 cell count, HIV-RNA, age and a diagnosis of AIDS at initiation of HAART are associated with risk of clinical progression in terms of death or new ADE, as is the immunological and virological response to HAART [Goedert et al., 1987; Goedert et al., 1989; Mellors et al., 1996; O’Brien et al., 1996; Ledergerber et al., 1999a; Graber et al., 2000a; Sterling et al., 2001; Egger et al., 2002]. Further, presence of anaemia remains a strong, independent predictive factor for death among patients with and without HAART, and hepatitis C co-infection has been shown to deteriorate the clinical prognosis for HIV-infected patients [Sash et al., 1994; Mocroft et al., 1999b; Greub et al., 2000; De Luca et al., 2002].

Among patients who initiated HAART within EuroSIDA, the most recently measured CD4 count, HIV-RNA and haemoglobin as well as a previous diagnosis of severe AIDS (i.e. NHL or progressive multifocal leukoencephalopathy) were all independently related to the risk of clinical progression, and the CD4 cell count was the strongest predictor [Sterling et al., 2001; Lundgren et al., 2002].

Taking the latest measurements into account, the CD4 cell count/HIV-RNA at initiation of HAART and the nadir CD4 cell count were not significantly associated with clinical progression, indicating that present rather than past immune deficiency affects the risk of clinical progression [Kirk et al., 2001c; Lundgren et al., 2002].

Future studies within large observational studies should include analyses on the prognostic influence of hepatitis B and C co-infections as well as the protective role of long-term toxicities, in particular dyslipidaemia and lipodystrophy (see section on toxicity). Such analyses should be carefully balanced between the wish to include large numbers of parameters and the accuracy and validity of the collected data.

DURABILITY OF EFFECT OF HAART

Due to the limited experience with HAART, the durability of the virological, immunological and clinical effects of HAART remains to be determined. A suggestion for a model of the durability of HAART on a population level is shown in Figure 4. The model is based on past experience and hence the patients starting HAART in this time frame. Importantly, many of these patients had already been exposed to NRTIs prior to commencing HAART.

Associations between the factors in this model have been documented, and the sequential associations between decreased HIV-RNA level, increased CD4 cell counts and lowered risk of clinical progression among patients who start HAART (the left hand side of the figure), have been described [Ho et al., 1995; Mellors et al., 1996; O’Brien et al., 1996; Hammer et al., 1997; Gulick et al., 1997; Egger et al., 1997; Paleia et al., 1998; Pakker et al., 1998; Mocroft et al., 1998a; Notermans et al., 1999; Gulick et al., 2000].

Conversely, virological failure of HAART may start with the development of resistance, though there are other reasons to virological rebound such as poor patient adherence, sub-optimal virological efficacy of the antiretroviral regimen used and non-therapeutic plasma concentrations of the antiretroviral drugs [Havlir et al., 2000; Paredes et al., 2000; Paterson et al., 2000; Nieuwkerk et al., 2001; Baxter et al., 2002].

Development of drug resistance is associated with poor outcome to therapy. Firstly, resistance to Zidovudine as monotherapy was shown to be associated with more rapid clinical progression, and more recently, an association between drug resistance for components of combination therapy and lower virological response has been reported in prospective observational studies. Finally, resistance testing has been shown to improve the virological response of
HAART among treatment experienced patients in RCTs [D’Aquila et al., 1995; Durant et al., 1999; DeGruttola et al., 2000; Baxter et al., 2002].

The subsequent sequence of increasing HIV-RNA, decreasing CD4 cell count and clinical progression is fairly well described in individual patients and well known in the HIV-clinics. However, within the presently available studies, the CD4 cell count and the rate of clinical progression remained stable for patients on HAART with virological failure [Ledergerber et al., 1999b; Deeks et al., 2000]. In fact, there remained a beneficial virological and immunological effect of continuing HAART after loss of virological control as well as a clinical effect of HAART among patients with a low CD4 cell count [Mocroft et al., 2000b; Deeks et al., 2001; Kirk et al., 2001c; Müller et al., 2002].

Further, on a population level – within the EuroSIDA study – there was as of 2001/2002 no sign of a time limitation in the effect of HAART on a population level. The mortality and incidence of new ADEs overall as well as of the individual ADEs remained at a stable low level within the EuroSIDA study in the period 1998-2002 (Figure 2a and 2b). The CD4 cell count for the entire population, in terms of the median value as well as the percentage below 200 cells/mm², was also relatively stable. However, if focus is entirely on median values, smaller fractions of the population coming at risk of progression will be missed. Interestingly, the 5th percentiles of the CD4 cell count also remained relatively stable. Further, among patients on HAART, the CD4 cell count as well as the percentage with HIV-RNA ≤500 copies/mL did not shown any signs of a failing effect (Figure 5a and Figure 5b), nor did the clinical score (i.e. the short-term risk of clinical progression remained unchanged) with longer time on HAART (EuroSIDA – unpublished data) [Lundgren et al., 2002]. It was interesting to see that in spite of only 55-65% of the patients who started HAART experienced virological suppression at a given time point, the 5th percentile of the CD4 cell count did not decrease over time, thus suggesting a lag time of more than 4 years between lack of virological suppression and immunological failure (Figure 5b).

Large numbers of patients have now been exposed to many antiretroviral drugs from all classes, and the proportion of patients followed within EuroSIDA who received a regimen of at least 5 drugs increased from 1% in 1996 to 10% in 2001. The median number of antiretroviral drugs a patient had ever been exposed for increased from 2 (range 0–6) in January 1996 to 6 (0–16) in January 2001, and the proportion of patients exposed to 10 drugs or more increased from 0% to 9% in the same period (EuroSIDA, unpublished data) [Mocroft et al., 2001a; Mocroft et al., 2002b]. This may reflect drug resistance, tolerability or simply changed fashion of treatment. The implications of this extensive exposure for development of drug resistance are essentially unknown, though there are some indications from studies reporting a higher risk of virological rebound among patients who have been exposed to NRTIs before starting HAART compared with patients being treatment naïve at time of starting HAART. Further, a recent study reported an increasing risk of virological rebound with longer duration of NRTI experience before initiation of HAART [Ledergerber et al., 1999b; Paredes et al., 2000; Phillips et al., 2002].

As of 2002, it is not yet possible to assess the durability of HAART on a population level. It is important to underline that the right side of Figure 4 is a worst-case scenario. New antiretroviral drugs, within the existing drug classes as well as drugs targeting other steps of the HIV life cycle (e.g. fusion inhibitors, integrase inhibitors), and immune stimulating drugs, for example Interleukin-2, may influence the model considerably.

Therefore, long-term monitoring of large patient populations is required for early detection of changes in the clinical prognosis for HIV-infected patients receiving HAART, and for better understanding the relations between development of resistance, loss of virological, immunological and clinical response. In particular, the relation between drug resistance and clinical outcome has not yet been documented sufficiently.

LONG-TERM TOXICITY OF HAART

As experience with HAART has grown, a range of long-term, late-onset side effects, potentially affecting the long-term treatment of HIV-infected patients and their prognosis have been documented. These include lipodystrophy, hyperlactataemia, osteopenia, diabetes mellitus, dyslipidemia and potential risk of cardiovascular disease [Carr et al., 1998b; Carr et al., 1999; Stephens et al., 1999; Gerard et al., 2000; Carr and Cooper, 2000; Fellay et al., 2001; Friis-Moller et al., 2003].

Though still debated, mitochondrial toxicity, induced by use of NRTIs, has been proposed as a common pathogenesis for several of the side effects [Brinkman et al., 1998; Brinkman et al., 1999].

Lipodystrophy or fat abnormalities – loss of peripheral subcutaneous fat located to arms, legs, buttocks, and buccal region – and/or central fat accumulation – breast hypertrophy, increased intra-abdominal fat and buffalo hump – is associated with use of ART. In particular, fat loss has over the last 3–4 years been associated with...
use of Stavudine in several observational studies, and this relation was recently further supported by results from a RCT (Carr et al., 1998b; Saint-Marc et al., 1999; Carr, 2000b; Saves et al., 2002; Dube et al., 2002; Worm et al., 2002).

Approximately, half of the patients on HAART have developed symptoms or signs of fat abnormalities, depending on the diagnostic criteria used, the patient population studied and the follow-up period available [Carr et al., 1998b; Carr et al., 1999; Martinez et al., 2001; Worm et al., 2002]. So far, no curative therapy for the fat abnormalities has been identified, and the reversibility, being tested in RCTs, is still debated and if at all, occurs slowly [Martinez et al., 1999; Carr and Cooper, 2000; Ruiz et al., 2001; Carr et al., 2002].

For future analyses it is important to use uniform objective diagnostic criteria for the lipodystrophy syndrome. Such criteria are currently being developed [Carr et al., 2003].

Of even greater concern is the potential impact of dyslipidaemia, insulin resistance and the clinical fat abnormalities on the risk of cardiovascular disease. This postulated association is currently being analysed in a large multi-cohort collaboration, the DAD study, following more than 23,000 patients (Vittecoq et al., 1998; Friis-Møller et al., 2003).

It is worth emphasising that the prevalence and incidence of these severe and potentially life-threatening side effects should be evaluated in the light of the pronounced declines in mortality and morbidity antiretroviral regimens have resulted in. Of note, these side effects have all been reported in observational studies, whereas RCTs powered to analyse the beneficial effect of HAART due to their limited patient sample and length of follow-up have been of limited value.

CONCLUSION

The introduction of HAART has led to a marked reduction in mortality and morbidity among HIV-infected patients, initially documented in RCTs and later substantiated in large observational studies with long-term follow-up. The latter have also provided data on changes in incidences of individual ADEs as well as survival thereafter and identified predictive factors for clinical progression in general and for individual ADEs. Further, the durability of the effect of HAART is being monitored in the same observational studies, and on a population level the immunological and clinical effects appear to last at least 5-6 years.

In the coming years there are important tasks for both RCTs and observational studies; the former to evaluate causal relations, for example between identified side effects and specific drugs or regimens, and the latter to monitor durability of HAART, as well as to analyse the influence of drug resistance, and not least to detect new and emerging side effects.

THE CLINICAL EFFECTIVENESS OF THE HAART-INDUCED IMMUNE RESTORATION

Before the introduction of HAART, chemoprophylaxis against common OIs was recommended for patients with severe immune deficiency in terms of CD4 cell counts below threshold levels at which the risk of OIs was considered sufficiently high to warrant usage of prophylactic medication. It was plausible to extend these recommendations and threshold levels for usage of chemoprophylaxis to patients who started HAART, though patients on HAART generally are at lower risk of clinical progression and specific OIs compared with patients not receiving HAART at a similar CD4 cell level (Phair et al., 1990; Kaplan et al., 1998; Mocroft et al., 1998b; Kirk et al., 2000; Mocroft et al., 2000b; Kirk et al., 2001c; Miller et al., 2002; Kaplan et al., 2002b).

However, for the majority of patients, HAART has led to immune reconstitution as determined by increasing numbers of memory and later also naïve CD4 cells, improved CD4 cell proliferation against specific antigens such as CMV antigen and tuberculin, and resolving OIs for which no specific treatment was available (Autran et al., 1997; Murphy et al., 1997; Li et al., 1998; Benfield et al., 1998; Carr et al., 1998a).

It was therefore questioned whether the predictive value of the CD4 cell level for the relative risk of developing ADEs was similar among patients on HAART relative to those not on HAART. That is, whether chemoprophylaxis against OIs could be interrupted after CD4 cell counts had increased to above the threshold levels prompting initiation of similar medication among patients with declining CD4 cell counts during the natural history of HIV; thereby basing the indication for continuation of chemoprophylaxis on the latest CD4 cell count rather than the nadir. In this respect, it should be noted that also in the pre-HAART era some patients (although few) did develop OIs at CD4 levels above those used as thresholds for when to start chemoprophylaxis [Phair et al., 1990; Mocroft et al., 1998b].

Though this should ideally be addressed in RCTs, preliminary reports indicated a low incidence of OIs following interruption of PP (i.e. prevention of initial events of OIs) as well as secondary prophylaxis or MT (i.e. prevention of recurrence of OIs), and therefore only large RCTs with a substantial follow-up time would be able to definitively address this question. A randomised superiority trial addressing the safety of interrupting a specific type of PP or MT against an OI should probably include at least 300-1500 patients, based on assumptions of event/relapse rates of 0-1% among patients on chemoprophylaxis and 3-5% among those without chemoprophylaxis, a statistical power of 80% and a significance level of 5% [Trikalinos and Ioannidis, 2001; Kirk et al., 2002b].

In the absence of conclusive results from RCTs, data from European cohort studies of patients interrupting PP for PCP and also cerebral toxoplasmosis provided compelling evidence that this type of PP could be safely interrupted after sustained CD4 increases to above 200 cells/mm³ (Schneider et al., 1999; Wewering et al., 1999; Furrer et al., 1999; Kirk et al., 1999b; Furrer et al., 2000a). Subsequently, these results were confirmed in two sufficiently powered RCTs, and in a recent metaanalysis of the studies published so far, the incidence of PCP after interruption of PP against PCP was as low as 0.24 (95%-CI: 0.10-0.49) events/100 PYF [Mussini et al., 2000; Lopez Bernaldo de Quiros et al., 2001; Trikalinos and Ioannidis, 2001].

For another common opportunistic pathogen, MAC, two large RCTs as well as an observational study documented that also for this pathogen it was safe to interrupt MT after CD4 cell count increased (Currier et al., 2000; El Sadr et al., 2000; Furrer et al., 2000b).

Extrapolation of the results on PP to MT was not straightforward. Patients with a history of an OI had sufficient degrees of impairment of their immune system to allow the OI to develop, as more than half of the patients in the pre-HAART era experienced a relapse within 6-12 months after stopping specific treatment for the OI if MT was not provided. Patients interrupting MT against a specific OI were thus at a substantially higher risk of experiencing a relapse of the OI compared with patients who interrupted PP against the same OI [Haverkos, 1987; Jacobson et al., 1988; Bozzette et al., 1991; Montaner et al., 1991; Jacobson et al., 1993; Katlama et al., 1996; Mocroft et al., 1998b; Kaplan et al., 2002b].

A collaboration of 8 large European cohorts found no relapses of PCP within 374 PYF after MT against PCP was interrupted after HAART had led to an increase of the CD4 count to above 200 cells/mL, and thereby provided strong evidence on the safety of the interruption of this type of MT [Ledergerber et al., 2001]. This finding was supported by a RCT, in which 60 patients interrupted MT against PCP, though this trial was not of a sufficient size to base recommendations on [Lopez Bernaldo de Quiros et al., 2001]. Accumulating all published series of interruption of MT against PCP, 1 relapse has been reported in 501 PYF, resulting in an incidence of 0.20 (0.01-1.11) relapse of PCP/100 PYF [Trikalinos and Ioannidis, 2001].

The strategy of safe interruption of MT against other common
OIs - CMV end organ disease, disseminated MAC infection, cerebral infection with Toxoplasma gondii, and extrapulmonary Cryptococcus neoformans - has been reported on in patient series of varying sizes and inclusion criteria, though based on substantially lower number of patients and PYF than for PCP [Whitcup et al., 1997; Abeg et al., 1996; Whitcup et al., 1999; Kirk et al., 1999b; Jouan et al., 2001; Zeller et al., 2002; Shafran et al., 2002].

A recent European multi-cohort study was by far the largest of these studies and in addition had the strength of analysing four common OIs, and the consistent overall findings could thereby provide support to the experience with OIs with the previously reported shortest follow-up after interruption of MT, such as toxoplasmosis and cryptococcosis. A total of 358 patients on HAART interrupted MT at a CD4 count above 50 cells/mm², and 5 relapses were diagnosed within 781 PYF, yielding relapse incidences with upper limit of the 95%-CI below 5-6 relapses/100 PYF. These incidences were very low compared with relapse rates in the pre-HAART era, and the risk of a recurrence was well out-weighted by costs, and side effects of the existing drugs used for MT [Kaplan et al., 2002a; Kirk et al., 2002b].

In addition, the incidence of bacterial infections, in particular bacterial pneumonia, was low among patients who interrupted chemoprophylaxis, and therefore was not a argument for continuing prophylactic regimens with antibacterial effect [Weverling et al., 1999; El Sadr et al., 2000; Ledergerber et al., 2001; Koletar et al., 2001].

Of note, recurrences of OIs often occur shortly after interrupting MT, and some patients experience several recurrences, suggesting that these patients may also have had some specific immune deficiencies despite experiencing CD4 cell increases [Muray et al., 2001; Johnson et al., 2001; Kirk et al., 2002b]. Such deficiencies have now been reported, as have methods for identifying high risk patients for experiencing recurrent CMV end organ disease before it is decided to interrupt the MT by assessing the CMV-specific CD4 cell frequency, though not yet available for broader clinical practice [Komanduri et al., 2001; Piccinini et al., 2001].

One caveat for all the studies reported, including the RCTs, is that information on those who do not interrupt chemoprophylaxis are generally lacking, and the participating patients may not be representative of all patients who experience HAART-induced CD4 cell increases. In particular, the median CD4 cell count at discontinuing chemoprophylaxis was for most studies far above the threshold of 200 CD4 cells/mm² [Kaplan et al., 2002a]. The cohort studies have, however, due to their design the possibility of analysing differences between patients who remain on and interrupt MT, and future studies should further elucidate this question [Yust et al., 2001]. Future analyses should also include a more refined evaluation of the level of CD4 cell at which it is safe to interrupt chemoprophylaxis for patients with increasing CD4 cell count and when to restart chemoprophylaxis when the CD4 cell count starts to decline again. Until further information becomes available, it seems reasonable to restart at the same criteria as for starting it initially [Kaplan et al., 2002a; Kaplan et al., 2002b].

Presently, patients with detectable HIV-RNA do not seem to be at a substantially higher risk of an OI after discontinuation of chemoprophylaxis compared with patients with a similar CD4 cell count and undetectable HIV-RNA, though the numbers of patients with detectable HIV-RNA have been relatively limited [Schneider et al., 1999; Weverling et al., 1999; Kirk et al., 1999b; Currier et al., 2000; El Sadr et al., 2000; Furrer et al., 2001; Kirk et al., 2002b]. Consistent with this, the CD4 cell count seems to be a much better marker of risk of clinical disease than viral load [Sterling et al., 2001; Lundgren et al., 2002].

Most patients in the published series of chemoprophylaxis have been treated with PI-based HAART regimens, which have been reported to have in vivo activity against Pneumocystis carinii, and it remains to be determined whether NNRTI-based or entirely NRTI-based HAART regimens provide a similar degree of protection, although this is likely the case [Atzori et al., 2000].

Finally, extrapolation to interruption of chemoprophylaxis against other less common OIs such as Histoplasma capsulatum, Coccidioides immitis, and Salmonella species seems sensible, though documentation is less clear. The much lower incidences of these OIs makes it, however, difficult to monitor increases or decreases in incidence [Kaplan et al., 2002b].

CONCLUSION
Based on results from RCTs and especially observational studies there is now sufficient evidence to conclude that it is safe to discontinue PP and MT for the most common OIs in HIV-patients after sustained HAART-induced CD4 cell increases [Kaplan et al., 2002b].

Items to be addressed include refining thresholds for when to interrupt and restart chemoprophylaxis for patients with increasing and decreasing CD4 cell count, the influence of continuous HIV-replication and of specific components of the HAART regimens, as well as the long-term monitoring. These topics must be addressed in large populations due to the very low number of events after discontinuation of chemoprophylaxis.

COMPARISON OF DIFFERENT HAART REGIMENS

RCTS
Generally, the effect of HAART regimens have been documented in RCTs comparing a regimen of two NRTIs with and without the new component, being a PI or NNRTI [Collier et al., 1996; D'Aquila et al., 1996; Hammer et al., 1997; Montaner et al., 1998; Cameron et al., 1998; Stellbrink et al., 2000; Saag et al., 2001; Gartland, 2001]. Only few RCTs comparing various HAART regimens head-to-head have been conducted. This is unfortunate since optimising the treatment remains of immense importance as a large proportion of patients do not experience fully suppressed viral replication [Fakhenheuer et al., 1997; Ledergerber et al., 1999b; Paredes et al., 2000] (Figure 5b).

The Danish PI study was one of the first of these studies and compared regimens of either Indinavir 800 mg TID, Ritonavir 600 mg BID or Ritonavir 400 mg BID plus Saquinavir hard gel 400 mg BID, all combined with 2 NRTIs. In the latter regimen Ritonavir inhibited the metabolism of Saquinavir hard gel, resulting in higher plasma concentrations of this PI and thereby 2 virologically active PIs [Merry et al., 1997; Kirk et al., 1999a].

An analysis of the first 269 patients followed for 24 weeks suggested a better short-term virological response of the regimen containing Ritonavir plus Saquinavir hard gel, whereas the discontinuation rate was significantly higher among patients who started a Ritonavir only based regimen. In a pre-determined subgroup analysis the virological response differed significantly across the treatment regimens among treatment naïve patients, but not among experienced patients (Figure 6) [Kirk et al., 1999a]. The primary strength of these findings was the consistency in the findings between the various definitions of virological response - proportion with HIV-RNA ≤200 copies/mL or HIV-RNA ≤20 copies/mL at week 24, as well as decrease of viral load within the first 24 weeks - but the possibility of a type I error should not be ignored, due to a relatively limited number of patients in the analysis, in particular when analysing treatment naïve and experienced patients separately. We attempted to apply a conservative approach to this three-armed trial by only performing pair wise comparisons if the overall p-value was less than 0.05, although the power calculations were based on a comparison of two drugs.

In a latter analysis with more patients and longer follow-up, patients starting Ritonavir and Saquinavir hard gel still tended to have better virological responses, though the results were no longer statistically significant [Katzenstein et al., 2000].

However, as of 2002, this dual PI-regimen is no longer a first-

DANISH MEDICAL BULLETIN VOL. 51 NO. 1 / FEBRUARY 2004

71
choice regimen, primarily due to common side effects and concern of long-term toxicity. Instead, 'baby-dose' Ritonavir boosted PI-containing HAART regimens have come into use: a small dosage (100 mg BID) of Ritonavir without intrinsic antiretroviral effect ensures a stable plasma concentration of the other PI by inhibiting its metabolism. This results in a high virological efficacy combined with an improved tolerability [Anonymous, 2001b; Walmsley et al., 2002; Yeni et al., 2002; Dragstedt et al., 2003].

Other RCTs have compared the virological response after 24-48 weeks among patients starting different initial HAART regimens [Cohen et al., 1999; Staszewski et al., 1999; Moyo et al., 2000; Eron, Jr. et al., 2000; Squires et al., 2000; Carr et al., 2000a; Staszewski et al., 2001; Walmsley et al., 2002; Nunez et al., 2002]. In general, they found few differences in short-term virological response, and several of these studies seemed inadequately powered [Mohr et al., 1994; Altman and Bland, 1995]. The notable exception from this pattern was the superior virological effect of Efavirenz, Zidovudine and Lamivudine compared with Indinavir, Zidovudine and Lamivudine in treatment naïve patients, although there were caveats in the design of that trial (see 'Design and analysis of RCTs with virological end-point') [Staszewski et al., 1999].

POTENTIAL ROLE OF OBSERVATIONAL STUDIES

Though being the golden standard, there are several issues inadequately addressed (and likely to remain unaddressed) in the RCTs of HAART regimens: i) the longer-term response to specific HAART regimens including clinical outcome, and ii) the lack of comparisons; i.e. the inability to perform all theoretical comparisons of HAART regimens, as 18 antiretroviral drugs have been licensed.

A study of 2203 mainly treatment experienced patients recently reported a significantly better virological response of Efavirenz and Nevirapine to be reported in 2003 [van Leeuwen and The ZNN Study Group, 2002].

Another example of the role of observational studies is that in the absence of RCTs evaluating Saquinavir hard gel relative to other PI-containing HAART regimens, several cohort studies have reported an inferior virological effect of Saquinavir containing HAART regimens (mono PI-regimens), which is biologically plausible due to the low bioavailability of Saquinavir hard gel [Merry et al., 1997; Jensen-Fang et al., 1999; Ledergerber et al., 1999b; Paredes et al., 2000; Grabar et al., 2000b]. In extension of these findings, a significantly higher risk of clinical progression was documented among patients starting a Saquinavir hard gel-based HAART regimen; being consistent with and in the same direction as inferior virological and immunological responses in this group compared with those starting regimens based on Indinavir or Ritonavir [Kirk et al., 2001b]. The consequences beyond the first few years of starting a virologically sub-optimal regimen including the development of resistance and the influence thereof remain to be evaluated.

Interpretation of drug comparisons within observational studies should be done with caution, as several biases may influence the results. In an often-referenced paper, consistent results on the relative difference in clinical outcome among patients who initiated one or two NRTIs were very similar in three large observational studies and in three large RCTs. This suggests that comparisons of treatment regimens under certain circumstances can be performed within observational studies in the field of HIV where risk factors for clinical outcome are fairly well established, allowing for adjustment to eliminate or at least reduce some of the biases inherent to observational studies [Phillips et al., 1999].

Other HIV- and non-HIV-related studies have compared the assessment of treatment effect in observational and randomised studies without pronounced inconsistencies between the two approaches, though some variability has been reported between RCTs and observational studies, as is also the case for the RCTs and observational studies, as is also the case for the RCTs and observational studies internally. Though the statistical methods and the quality of the design and data collection within observational studies have improved, there are still concerns of the biases in the latter, in particular the inherent selection bias or 'confounding by indication' (i.e. patients are not randomly allocated to a treatment regimen) [Kunz and Oxman, 1998; McKee et al., 1999; Pocock and Elbourne, 2000; Concato et al., 2000; Benson and Hartz, 2000; Barton, 2000; Ioannidis et al., 2001b; Dunn et al., 2002].

In general, for treatment comparisons within observational studies, the more pronounced difference, the larger number of patients, the more homogeneous patient populations in the different treatment arms to be compared, and the more robust/stable the findings are in sensitivity analyses, the more likely is the finding to hold true, especially if independently confirmed in other cohorts [Phillips et al., 1999; Pocock and Elbourne, 2000; Barton, 2000; Sabin and Phillips, 2001a; Sabin and Phillips, 2001b].

Observational studies should not replace results from RCTs or reduce the willingness to perform RCT; the two types of study should rather complement each other. First of all, observational data is useful for evaluating the generalisability of the results obtained in RCTs with extensive exclusion criteria to a broader spectrum of patients in clinical practice. Further, results from observational studies should serve as encouragement to perform RCTs and potential sources of identifying drugs to compare in future RCTs, sources of knowledge in cases where RCTs for ethical or practical reasons are not feasible, and sources for long-term clinical data for drugs originally tested in short-term virological RCTs [Black, 1996; Pocock and Elbourne, 2000; Barton, 2000; Ioannidis et al., 2001a].

CONCLUSION

Few comparative RCTs of HAART regimens have been performed,
and the recommendations on first-choice initial HAART regimens are generally based on relatively short-term virological results from randomised settings, on toxicity data as well as considerations of future treatment options [Anonymous, 2001b; Yeni et al., 2002].

More comparisons and trials with longer follow-up including clinical outcome are warranted to facilitate the choice of optimal components of the (initial) HAART regimen and treatment strategy. RCTs remain the golden standard for evaluating treatment regimens, and should be performed whenever possible, but observational studies may in several situations be able to provide important information, especially if confirmed in several independent cohorts.

DESIGN AND ANALYSIS OF RCTS WITH VIROLOGICAL END-POINT

Whereas clinical outcome was the primary outcome in early RCTs, the regulatory authorities such as the Food and Drug Administration, FDA, in USA and the European Medical Evaluation Agency, EMEA, in Europe currently approve new drugs based on virological data of 24 weeks of follow-up (accelerated approval) or 48 weeks (standard approval), due to the low number of clinical events [Anonymous, 2002]. This procedure allows for a quick access to new therapy of benefit for the present patients, but as a consequence data on long-term follow-up including clinical outcome are not available [Anonymous, 2001b; Yeni et al., 2002].

In these RCTs with virological response as primary endpoint it has become widely accepted to categorise switches from randomised therapy (i.e. discontinuation of randomised therapy and likely initiation of another HAART regimen) as virological failures. These failures were classified with patients who do not obtain a HIV-RNA below a specified level of quantification, ITT/s=f. In addition, patients who die are lost to follow-up, or do not have any HIV-RNA measurement at a given time point, are also categorised as virological failures. The reasoning for this approach is a wish to assess the effectiveness of a treatment regimen by combining the virological and safety responses. Contrasting the ITT/s=f approach, a true intention-to-treat analysis requires continuous follow-up on all patients whether on or off randomised therapy, ITT/s incl, and the virological response is assessed based on the actual HIV-RNA measurement regardless of premature switch from randomised therapy [Anonymous, 1999; Kirk et al., 2002a].

Many RCTs including most recent pivotal trials discontinue follow-up permanently at switch from randomised therapy, thereby limiting the possibilities of performing complementary analyses [Staszewski et al., 1999; Gilbert et al., 2001; Staszewski et al., 2001].

As one of few trials, the Danish PI Study continued to follow patients after premature switch from randomised therapy, and it was interesting to note that the choice of analysis, ITT/s=f or ITT/s incl did matter; the result of a regimen-to-regimen comparison was significant in one and not significant in the other analysis for two of the three regimen-to-regimen comparisons in this study [Katzstein et al., 2000; Kirk et al., 2002a].

Based on data from two RCTs with complete follow-up and as well as observational data from EuroSIDA, the difference between the two analytic approaches has been shown – as expected – to increase with increasing switch rate, and the response to a second-line regimen at week 48 varies across the regimens used (Figure 7). Therefore, the ITT/s=f approach neglects that a given drug or regimen may have impact on the virological response and the risk of toxicity after the patient has switched to other therapy. For example, among patients randomised to initiate Ritonavir in the Danish Protease Inhibitor Study, 55% switched from randomised therapy within 48 weeks, and 29% and 48% had a HIV-RNA below level of quantification after 48 weeks according to the ITT/s=f and ITT/s incl, respectively [Kirk et al., 2002a].

As the focus of RCT has changed towards an evaluation of treatment strategies rather than effect of specific regimen, follow-up data after premature switch from therapy is of increasing importance, and several ongoing studies have specified sequential regimens after each of the randomised regimens in the protocol, and continue to follow patients after premature interruption of randomised therapy to assess the long-term efficacy. Results from such studies on treatment strategies rather than specific regimens will be more generalisable to the handling of patients in clinical practice [Smeaton et al., 2001; MacArthur et al., 2001; Anonymous, 2001a; Emery et al., 2002; Kirk et al., 2002a].

SWITCH OF COMPONENTS OF HAART

In analyses based on the ITT/s=f approach, the virological response rate is closely linked to the switch rate. More than one fourth of patients starting a first HAART regimen switched components of their HAART regimen within the first year in a randomised as well as an observational setting with large variations according to the drugs used [Kirk et al., 1999a; d’Arminio et al., 2000; Kirk et al., 2001a; Kirk et al., 2001b; M christoff et al., 2001b; Kirk et al., 2002a].

Interestingly, the most common reason for switching from components of HAART regimens were side effects, predominantly of gastrointestinal or neurological nature, though with pronounced differences according to the HAART regimen used. In contrast, virological failure was only considered to be the principal reason for switch from initial HAART therapy in few patients [d’Arminio et al., 2000; Kirk et al., 2001a; Kirk et al., 2002a].

In general, patients who initiated a Ritonavir-containing regimen were at highest risk of treatment discontinuation within the first 12 months of HAART [Kirk et al., 1999a; d’Arminio et al., 2000; Bonfanti et al., 2000; Kirk et al., 2001a; Kirk et al., 2001b]. Other predictive factors for switches of a HAART regimen varied across studies, depending on the patient populations studied, whether the analysis focused on discontinuation of one specific component or any component of the HAART regimen, and whether the analysis included all switches or those related to toxicity. The risk factors seemed relatively similar in randomised and observational settings [d’Arminio et al., 2000; Kirk et al., 2001a; M christoff et al., 2001b].

Predictive factors for treatment switch in general were previous usage of ART before starting HAART, type of HAART regimen initiated, body weight or gender; the later primarily explained by differences in body weight [Lucas et al., 1999; Ferrer et al., 1999; Kirk et al., 2001a]. Further, the risk of treatment-limiting side effects of PI
therapy increased with decreasing body weight, thus suggesting the need of individualisation of dosages of antiretroviral drugs according to body weight and the potential use of therapeutic drug monitoring for safety reasons in clinical practice (Figure 8) [Gatti et al., 1999; Kirk et al., 2001a].

High HIV-RNA and to some extent also low CD4 cell count seemed to be associated with a higher risk of switch, though low HIV-RNA was associated with a higher risk of switch among patients participating in a RCT [Kirk et al., 2001a]. The latter probably reflects that patients followed in this RCT who were at low risk of clinical progression were less likely to accept side effects or dietary restrictions due to the regimen initiated, as was also the case with increasing calendar time [d’Arminio et al., 2000; Kirk et al., 2001a; Mocroft et al., 2001b].

CONCLUSION
Most RCTs do not routinely follow patients after premature switch of the randomised therapy, and this compromises the possibilities of several additional approaches. Emphasising the relevance of continuous follow-up, such treatment switches are predominantly due to side effects, whereas virological failure is only seen in few patients. Predictive factors for switch of therapy have been identified.

Future studies should ensure that patients remain under follow-up in RCTs even after premature switch from first-line randomised therapy and subsequent treatment options should be specified in the study protocols, as this is important for analysis of virological effect and assessment of long-term consequences of the treatment to be evaluated in the trial.

GENERAL CONCLUSIONS AND PERSPECTIVES
In summary, the introduction of antiretroviral combination therapy in the western world has been associated with marked improvements in the clinical prognosis for HIV-infected patients in terms of decreased mortality, morbidity and need of chemoprophylaxis for opportunistic pathogens. However, many unsolved questions remain with respect to HIV-therapy – in terms of developing new antiretroviral drugs and ideally a vaccine, but also in terms of optimal usage of the therapy already available for the clinical handling of HIV-infected patients. These questions include when to initiate ART, choice of optimal treatment regimen, how to treat (role of immune stimulating drugs, structured treatment interruptions and therapeutic drug monitoring as well as risk of toxicity), when to switch therapy (components of salvage therapy, role of resistance testing and identification of patients at high risk of clinical progression), and when to interrupt and restart chemoprophylaxis for OIs.

The present series of studies document that observational studies have been an important supplement to RCTs in studying the effects of HAART among HIV-infected patients [Kirk et al., 1998; Kirk et al., 1999a; Kirk et al., 1999b; Kirk et al., 2000; Kirk et al., 2001a; Kirk et al., 2001b; Kirk et al., 2001c; Kirk et al., 2002a; Kirk et al., 2002b]. Clearly, RCTs are the golden standard and should be performed whenever possible [Pocock and Elbourne, 2000; Barton, 2000]. However, there are a number of the issues mentioned above, which RCTs cannot readily address for practical or ethical reasons. In this situation, observational studies may provide useful information as an important supplement to expert opinions. Several of the publications involving the EuroSIDA study has led to revisions of public health guidelines in USA and Europe [Weverling et al., 1999; Ledergerber et al., 2001; Phillips et al., 2001b; Anonymous, 2001b; Yeni et al., 2002; Egger et al., 2002; Kaplan et al., 2002b].

Observational studies play an important role in the continuous evaluation and assessment of the longer-term effect of HAART, i.e. the durability of HAART, identification of patients at high risk of developing a specific ADE or dying, monitoring of changes in incidences of individual ADEs and thereby in the pattern of ADEs, as well as detection of new emergent HIV-related diseases. Longer-term side effects are also poorly elucidated and require continuous follow-up in large-scale studies, as antiretroviral drugs are being approved in a relatively fast manner based on short-term RCTs with virological rather than clinical end-points [Anonymous, 2001b; Anonymous, 2002].

It remains important to assess the treatment response in clinical practice, outside the often idealised setting of a RCT including selected groups of patients and being performed at selected clinics; i.e. how do results from RCTs translate into daily clinical practice among patient populations commonly not included in RCTs, such as intravenous drug users and patients co-infected with hepatitis B or C. Ideally, RCTs should include these subgroups.

With respect to comparisons of specific HAART regimens, the role of observational studies is debated, and results of drug comparisons in cohort studies should be interpreted with caution. Ideally, findings in observational studies should be tested in RCTs and the observational studies may thus help to identify objectives for future RCTs, though the feasibility of such trials will naturally depend on the size of the effect observed in the observational studies [Ioannidis et al., 2001a]. Observational studies may still provide important information on relative effect of different HAART regimens not readily available from RCTs, in particular on long-term effect.

To conclude, there are several areas where observational studies in the coming years can provide essential information, which may never be available from RCTs. Standardised data collection, data monitoring, and standardised analyses can all help to improve the quality of an observational study and provide more reliable results. Regardless of the nature of the study, well designed studies will better address unsolved questions, and it could be argued that large well-designed high-quality observational studies may provide more reliable information than small, sub-optimally designed randomised studies of selected patient populations [Fenéstein, 1984]. Further, even large-scale observational studies may for some analyses not be large enough. For this reason and due to the importance of the reproducibility of findings in other settings, collaboration between the existing and the many new HIV cohorts is essential for the scientific outcome of these studies in the coming years.

SUMMARY
The thesis includes nine previously published articles and a review of the literature. This work was carried out in the period 1997-2002.
The aim of this series of studies was to analyse the effects of potent antiretroviral combination therapy among HIV-infected patients and to discuss to which extent results obtained from observational studies can support randomised controlled trials.

Since the introduction of highly active antiretroviral therapy, HAART, in Europe, the mortality and incidences of all AIDS defining events have decreased substantially, and the decrease was most pronounced for AIDS defining events diagnosed at severe immune suppression, as for example infection with Mycobacterium avium complex compared with Mycobacterium tuberculosis, which is diagnosed over a broader spectrum of immune deficiency. In contrast to what had been feared, the risk of non-Hodgkin lymphoma also decreased with longer time on HAART, though more slowly than most other AIDS defining diseases. Patients, with sub-optimal CD4 and HIV-RNA responses to HAART, were at highest risk of NHL.

In the absence of data from randomised trials, observational data from European cohorts including a Danish study has documented that it is safe to interrupt primary prophylaxis and maintenance therapy against most common opportunistic infections after HAART-induced CD4 lymphocyte increases.

Few randomised trials have compared the virological effect of different HAART regimens: a Danish trial suggested a better short-term virological effect of regimens based on Ritonavir/Saquinar hard gel compared with regimens containing either Indinavir or Ritonavir, whereas the immunological and clinical effects did not vary significantly. In addition, we found within a large population with substantial follow-up, consistency between inferior virological and immunological responses and the clinical effect of HAART-regimens, which contained Saquinavir hard gel compared with Indinavir.

Most current randomised trials of virological response to HAART do not include virological data after premature interruption of randomised therapy. Instead patients are categorised as virological failures from the time of switching therapy, though the most common reason for the switch is side effects and the risk of switch varies widely according to the therapy initiated and the patient population. Based on results from both observational and randomised settings, it was demonstrated that such analyses might lead to an imprecise and incomplete assessment of the treatment response whereas complete follow-up allows for several analytic approaches.

To conclude, the introduction of antiretroviral combination therapy has been associated with marked improvements in the clinical prognosis for HIV-infected patients in Europe. Observational studies have proven an essential supplement to randomised trials in the analyses of the effects of HAART, and will be ideally placed to address future important questions, which otherwise might not be addressed sufficiently.

ABBREVIATIONS

ADE AIDS defining event
AIDS acquired immunodeficiency syndrome
ART antiretroviral therapy
BID twice daily
CD4 cell CD4 receptor positive T lymphocyte cell (plasma level)
CMV cytomegalovirus
HAART highly active antiretroviral therapy
HIV human immunodeficiency virus type 1
HIV-RNA HIV-RNA plasma level
ITT/s incl intention-to-treat switches included
ITT/s ef intention-to-treat switches equal failures
MAC Mycobacterium avium complex
MT maintenance therapy
NHL non-Hodgkin lymphoma
NRTI nucleoside reverse transcriptase inhibitor
NNRTI non-nucleoside reverse transcriptase inhibitor
OI opportunistic infection
PCP Pneumocystis carinii pneumonia
PYF person-years of follow-up
PI protease inhibitor
PP primary prophylaxis
RCT randomised controlled trial
TID three times daily

REFERENCES


Baxter JD, Merigan TC, Wentworth DN, Neaton JD, Hoover ML, Hoetelmans RM, Piscitelli SC, Verbiest WH, Meyers DL. (2002) Both
baseline HIV-1 drug resistance and antiretroviral drug levels are associated with short-term virologic responses to salvage therapy. AIDS 16:1131-1138.


Johnson SC, Bensa C, Johnson DW, Weinberg A. (2001a) Recurrences of cytomegalovirus retinitis in a human immunodeficiency virus-infected patient, despite potent antiretroviral therapy and apparent immune re-


