Difficulties in assigning AIDS endpoints in clinical trials in resource-limited settings: lessons from DART

M. Borok[1], A. Palfreeman[2], P. Munderi[3], F. Ssali[4], A. Reid[1], F. Lutwama[5], H. Wilkes[2], D. Winogron[6], R. Tembo[1], H. Musana[4], G. Musoro[1], K. Taylor[2], D.M. Gibb[2], TEA Peto[7], A. Burke[2], and the DART trial team

1UNIVERSITY OF ZIMBABWE, HARARE, ZIMBABWE, 2MRC CLINICAL TRIALS UNIT, LONDON, UNITED KINGDOM, 3MRC/UVRI UGANDA RESEARCH UNIT ON AIDS, ENTEBBE, UGANDA, 4JCRC, KAMPALA, UGANDA, 5INFECTIOUS DISEASES INSTITUTE, MAKERERE UNIVERSITY, MULAGO, UGANDA, 6ROYAL FREE AND UNIVERSITY COLLEGE MEDICAL SCHOOL, LONDON, UNITED KINGDOM, 7JOHN RADCLIFFE HOSPITAL, OXFORD, UNITED KINGDOM

www.ctu.mrc.ac.uk/dart
Email:aub@mrc.ctu.ac.uk

ISCRTN 13968779

TUPEB098

ABSTRACT

Background:

Stage 3/4 who criteria were designed to provide tools for HIV surveillance and staging of disease severity in the pre-HAART era. In 2002, the WHO clinical staging criteria were adapted to define clinical endpoints in the DART trial (2003-2008).

Methods:

DART was designed to compare clinical outcomes under two management strategies in Sub-Saharan Africa. 3316 patients were recruited and followed for up to 6 years accruing 1021 reported WHO 4 events (excluding deaths) to December 2008. The protocol included definitive and presumptive criteria defining each WHO 4 endpoint. Where specific diagnoses could not be allocated to reported events, an independent endpoint review committee developed a "syndromic" categorisation scheme, including "brain syndrome" and "lung syndrome". Syndromes were classified as severe (severity comparable to WHO 4 events) or mild. We compared the rates of rejection of disease endpoints in DART and the last large pre-HAART RCT (DELTA), completed 12 years previously.

Results:

To December 2008, at independent review, 215 (21%) reported events were rejected in DART as not fulfilling the protocol criteria for trial endpoints, more than in DELTA. Lack of diagnostic investigative capacity was a major challenge in ascertaining clinical endpoints. On review of the rejected DART events it was possible to assign them into "syndromic" categorisations in 130/215 (60%) cases. In brain syndromes the lack of access to a CT scanner was the main reason for rejection of the presumptive and almost certainly correct diagnosis of toxoplasmosis in 20/43(47%) rejected events. In lung syndromes the lack of a typical chest X-ray with a documented response to treatment for PCP lead to most rejections (17/35 (49%) rejected endpoints).

Conclusions:

Lack of diagnostic investigations is a challenge for clinical endpoint studies in resource-limited settings. Adaptation of criteria to include a syndromic categorisation for severe HIV-related events, and use of independent endpoint review committees, is recommended.

BACKGROUND

- DART was one of the first large randomised clinical trials of ART in Africa, comparing management strategies in 3316 adults initiating first-line ART in Uganda/Zimbabwe[2].
- Participants were randomised into two arms; a Clinically Driven Monitoring (CDM) arm and a Laboratory and Clinical Monitoring (LCM) arm. Enrolment occurred during 2003/4 and follow-up was until December 2008.
- A large number of endpoints were expected given the relatively advanced HIV disease in those recruited, and thus a protocol to screen reported endpoints to detect those where the diagnosis clearly fulfilled the definite/presumptive WHO criteria was devised.
  - All reported endpoints (WHO 4 events and deaths) were reviewed by the screening committee (Peto, Palfreeman). Only those where the diagnosis was in doubt (not meeting definite or presumptive WHO criteria), or where a drug cause was implicated in a death, were referred to the full Endpoint Review Committee.
- It rapidly became apparent that strict application of WHO criteria led to a large number of events being rejected when it was clear to both the Endpoint Review Committee and the clinicians caring for the patients that a clinical event of significance had occurred
  - Sometimes the lack of appropriate investigational resources at the site or a lack of response to treatment meant that the WHO 4 criteria were not met.
  - Sometimes classification is difficult because more than one WHO 4 diagnosis is equally possible.

OBJECTIVE

To devise a set of reproducible reliable criteria to supplement WHO 4 criteria to evaluate endpoints in clinical trials

METHODS

- Reported endpoints were reviewed using a standardised form, blinded to randomised arm, and classified using WHO criteria for stage 3 and 4 events (as modified in 2006[3]). Cause of death was determined using a standard form based on the CoDe[4] system.
- 1021 WHO 4 events were reported by sites after enrolment and before 31 December 2008 (the cutoff for the primary analysis) to the Endpoint Review Committee
  - 215 (21%) were rejected by the Endpoint Review Committee using standard WHO 4 criteria in the DART protocol.

We therefore developed a classification to capture these events and to document them appropriately as "DART syndromes". A DART syndrome is defined as an event in which:

❖ **The clinical criteria of a WHO 4 presumptive event are met, but clinical details available, including response to treatment and investigations, are unable to confirm or refute the event.**

Brain syndrome
Fulfils clinical criteria for at least one of the following
- Cerebral toxoplasmosis
- Cryptococcal meningitis
- TB meningitis
- PML

Mild brain syndrome: responds to outpatient treatment only or to first line inpatient treatment with no severe residual disability.
Severe brain syndrome: Fail to respond to first line treatment or results in severe residual disability.

Lung syndrome
Fulfils clinical criteria for at least one of the following
- PCP
- Extra pulmonary TB
- Pulmonary KS

Mild lung syndrome: responds to out patient treatment only or to first line inpatient treatment
Severe lung syndrome: Requires hospital admission and did not respond to first line inpatient treatment.

Mouth Syndrome
Fulfils clinical criteria only for at least one of the following
- Oesophageal candidiasis
- CMV oesophagitis
- HSV oesophagitis

Other syndromes
Fulfils clinical criteria only for at least one of the following
- Extrapulmonary TB
- Disseminated KS
- Non-Hodgkins Lymphoma
- Malaria and other conditions - where HIV was excluded as a cause or another unrelated diagnosis became apparent over time

RESULTS

	DELTA (Europe/Australia) Accepted / Reported (% accepted)		DART (Uganda/Zimbabwe) Accepted / Reported (% accepted)	
All events	1017/1100	(92%)	806/1021	(79%)
HIV wasting	35/45	(88%)	25/34	(74%)
PCP	148/162	(91%)	29/64	(45%)
Toxoplasmosis	49/53	(92%)	13/56	(23%)
Cryptosporidia	49/46	(93%)	34/34	(100%)
Cryptococcosis	26/28	(93%)	158/180	(88%)
CMV	107/114	(94%)	9/12	(75%)
Herpes simplex, mucocutaneous	9/16	(56%)	52/63	(83%)
Herpes simplex, visceral	2/3	(67%)	5/9	(56%)
PML	12/25	(84%)	0/1	(0%)
Oesophageal candidiasis	254/264	(96%)	284/311	(91%)
MAI/MAC	77/82	(94%)	0/2	(0%)
Recurrent salmonella infection	0		0/2	(0%)
Tuberculosis, extra-pulmonary	3/2	(67%)	150/198	(76%)
Lymphoma	45/46	(96%)	15/17	(88%)
KS	184/186	(99%)	32/33	(97%)
Encephalopathy	17/28	(61%)	0/5	(0%)

➢ **A higher percentage of reported endpoints were rejected in resource-limited than resource-rich studies.**

- The main reasons for rejection were lack of investigations to confirm the diagnosis (see below), e.g a typical CXR appearance, or failure to fulfil other criteria such as a response to anti-PCP treatment when this did not occur
- In severely ill patients, lack of clinical response to therapy was also a major factor
- There was little doubt in the minds of the reviewers that the diagnoses were AIDS-related but a failure to meet the criteria required meant that many of the events could not be confirmed as definite or presumptive and thus had to be rejected
- **This was mainly in "Lung syndrome" and "Brain syndrome"**
  - For "lung syndrome", the criteria needed to define PCP included an appropriate clinical presentation, a typical CXR together with appropriate response to treatment for PCP
  - **In DELTA only 9% of PCP diagnoses were rejected vs. 55 % in DART**
  - For "brain syndrome", the CT scanner needed to verify a toxoplasmosis diagnosis was often not available. The diagnosis then rested on a clinical response to treatment, which is not always apparent in patients with severe disease, so the event could not be verified in 13/55 as presumptive although in many cases this was very likely and thus these patients were classified as brain syndrome.
  - **In DELTA only 8% of toxoplasmosis diagnoses were rejected vs 77% in DART**

There were 21 serious and 15 mild lung syndromes and 34 serious and 23 mild brain syndromes

In addition there were 17 other serious conditions of comparable severity to WHO 4 (see below)

Other comments:

- It was very rare for a definitive microbiological diagnosis to be made
- In patients with suspected meningitis, a clear-cut definitive diagnosis was easy to confirm from a CRAG test if the cause was cryptococcal, with 88% accepted. The few that were rejected were mostly where there was serological evidence of cyptococcal infection but no clinical evidence to support a diagnosis.

RESULTS - REJECTED EVENTS

Rejected endpoints in DART	**215**	**(100%)**
Site decided on a different diagnosis	12	(6%)
No response to treatment	7	(3%)
Tests negative	14	(7%)
Too ill to respond	16	(7%)
Died before tests	16	(7%)
No tests done	76	(35%)
Negative test, but poor sample	2	(1%)
Duration not to WHO 4 level	13	(6%)
Severity not at WHO 4 level	19	(9%)
Other plausible diagnosis	35	(16%)
Other reasons	5	(2%)

RESULTS - DART SYNDROMES

	Total
Rejected events	215
Adjudicated to be serious brain syndrome	34
Originally reported as:	
PCP	*1*
toxoplasmosis	*17*
cryptococcosis	*4*
CMV	*1*
PML	*1*
tuberculosis, extra-pulmonary	*4*
lymphoma	*2*
encephalopathy	*3*
serious adverse event	*1*
Adjudicated to be serious lung syndrome	21
Originally reported as:	
PCP	*17*
extra-pulmonary TB	*4*
Adjudicated to be other serious condition*	17

*** abdominal, ano-genital, cardiac failure, chronic myeloid leukaemia, gastrointestinal bleed, heart syndrome of unknown origin, overwhelming sepsis, pancreatic tumour, pelvic inflammatory disease, peritonitis, possible extra-pumonary TB, pyrexia of unknown origin, jaundics, vaginal herpes/ulcers**

References

1. Monitoring HIV treatment in the developing world N Kumarasamy, Timothy P Flanigan, Anish P Mahajan, Charles CJ Carpenter, Kenneth H Mayer, Sunil Solomon. The Lancet Infectious Diseases - Vol. 2, Issue 11, November 2002, Pages 656-657
2. DART Development of Antiretroviral therapy in Africa. www.ctu.mrc.ac.uk/dart/summary.asp
3. WHO case definitions of HIV for surveillance and revised clinical staging and immunological classification of HIV-related disease in adults and children. Geneva: World Health Organization, 2007. www.who.int/hiv/pub/guidelines/hivstaging/en/index.html
4. The CoDe ("Coding of Death in HIV") Project www.cphiv.dk/Portals/_default/pdf_folder/code_protocol_ver_1.0.pdf
5. Delta Coordinating Committee. Delta: a randomised double-blind placebo-controlled trial comparing combinations of zidovudine plus didanosine or zalcitabine with zidovudine alone in HIV-infected individuals. Lancet 348(9023):283-291, 1996.

DISCUSSION

- Most HIV clinical trials published and presented in recent years have used surrogate markers as endpoints. The change from using clinical to surrogate endpoints has been driven by the reduction in events due to the effectiveness of therapy, and by the economy of scale and cost given the timescales needed for clinical endpoint trials. The availability of effective treatment and monitoring has also made it likely that treatment is modified or changed before clinical events occur.
- The vast majority of these studies have been in populations in resource-rich settings with good access to antiretroviral drugs and to laboratory and radiological diagnostic facilities.
- Most of the 32 million people in the world who are living with HIV do not have access either to ART or to the laboratory measurements[1] on which trials in resource-rich settings depend.
- As treatment becomes increasingly available in resource-limited settings there is an opportunity to undertake trials based on clinical outcomes to answer crucial strategic questions about HIV management.
- Endpoint Review Committees have often been used to independently evaluate outcomes in a standardised manner, and are often blinded to treatment arm.
  - This should mean they compare like with like across different populations, but the clinical endpoints measured in trials may differ either due to variations in the geographical prevalence of indicator diseases or due to difference in availability of investigations to detect them.
- Here we present 2 different trials where the classified and defined clinical endpoints were very similar, but where there were differences in access to confirmatory investigations. The differences in rates may have been due to the different populations studied but may also have been due to differences in availability of investigations to confirm clinical endpoints.
- **The dilemma for Endpoint Review Committees is: should they adhere strictly to standard agreed criteria for judging clinical events and risk underestimating the number of events due to lack of facilities in some sites, or should they adopt different standards of evidence for different sites at the risk of over estimating progression of disease events?**

CONCLUSION

- **The WHO clinical staging criteria were originally designed to provide tools for HIV surveillance and staging of disease severity in the pre-HAART era. Definitions were sufficiently precise to avoid clinical events which could commonly occur in HIV uninfected immuno-competent individuals. The effect of this was not to include a number of events which were probably but not definitively HIV-related.**
- **Criteria for clinical endpoints in HIV trials need to be adapted beyond WHO staging, particularly to allow for limited access to laboratory diagnostic facilities**

We thank all the patients and staff from all the centres participating in the DART trial. University of Zimbabwe, Harare, Zimbabwe: A Latif, J Hakim, V Robertson, A Reid, E Chidziva, R Bulaya-Tembo, G Musoro, F Taziwa, C Chimbetete, L Chakonza, A Mawora, C Muvirimi, G Tinago, P Svovanapasis, M Simango, O Chirema, J Machingura, S Mutsai, M Phiri, T Bafana, M Chirara, L Muchabaiwa, M Muzambi. MRC Programme on AIDS/Uganda Virus Research Institute, Entebbe, Uganda: H Grosskurth, P Munderi, G Kabuye, D Nsibambi, R Kasirye, E Zalwango, M Nakazibwe, B Kikaire, G Nassuna, R Massa, K Fadhiru, M Namyalo, A Zalwango, L Generous, P Khauka, N Rutikarayo, W Nakahima, A Mugisha, J Todd, J Levin, S Muyingo, A Ruberantwari, P Kaleebu, D Yirrell, N Ndembi, F Lyagoba, P Hughes, M Aber, A Medina Lara, S Foster, J Amurwon, B Nyanzi Wakholi. Joint Clinical Research Centre, Kampala, Uganda: P Mugyenyi, C Kityo, F Ssali, D Tumukunde, T Otim, J Kabanda, H Musana, J Akao, , H Kyomugisha, A Byamukama, J Sabiiti, J Komugyena, P Wavamunno, S Mukiibi, A Drasiku, R Byaruhanga, O Labeja, P Katundu, S Tugume, P Awio, A Namazzi, GT Bakeinyaga, H Katabira, D Abaine,J Tukamushaba, W Anywar, W Ojiambo, E Angweng,S Murungi , W Haguma, S Atwiine,J Kigozi. Infectious Diseases Institute (formerly the Academic Alliance) Makerere University, Mulago, Uganda: E Katabira, A Ronald, A Kambungu, F Lutwama, A Nanfuka, J Walusimbi, E Nabankema, R Nalumenya, T Namuli, R Kulume, I Namata, L Nyachwo, A Florence, A Kusiima, E Lubwama, R Nairuba, F Oketta, E Buluma, R Waita, H Ojiambo, F Sadik, J Wanyama, P Nabongo. The AIDS Support Organisation (TASO), Uganda: R Ochai, D Muhweezi. Imperial College , London, UK: C Gilks, K Boocock, C Puddephatt, D Winogron, J Bohannon. MRC Clinical Trials Unit, London, UK: J Darbyshire, DM Gibb, A Burke, D Bray, A Babiker, AS Walker, H Wilkes, M Rauchenberger, S Sheehan, L Peto, K Taylor, M Spyer, A Ferrier, B Naidoo, D Dunn, R Goodall. DART Virology Group: P Kaleebu (Co-Chair), D Pillay (Co-Chair), V Robertson, D Yirrell, S Tugume, M Chirara, P Katundu, N Ndembi, F Lyagoba, D Dunn, R Goodall, A McCormick. DART Health Economics Group: A Medina Lara (Chair), S Foster, J Amurwon, B Nyanzi Wakholi, J Kigozi, L Muchabaiwa, M Muzambi. Independent DART Trial Monitors: R Nanfuka, C Mufuka-Kapuya. Trial Steering Committee: I Weller (Chair), A Babiker (Trial Statistician), S Bahendeka, M Bassett, A Chogo Wapakhabulo, J Darbyshire, B Gazzard, C Gilks, H Grosskurth, J Hakim, A Latif, C Mapuchere, O Mugurungi, P Mugyenyi; Observers: C Burke, S Jones, C Newland, S Rahim, J Rooney, M Smith, W Snowden, J-M Steens. Data and Safety Monitoring Committee: A Breckenridge (Chair), A McLaren (Chair-deceased), C Hill, J Matenga, A Pozniak, D Serwadda. Endpoint Review Committee: T Peto (Chair), A Palfreeman, M Borok, E Katabira. Funding: DART is funded by the UK Medical Research Council, the UK Department for International Development (DFID), and the Rockefeller Foundation. GlaxoSmithKline, Gilead and Boehringer-Ingelheim donated first-line drugs for DART.