Effectiveness and cost-effectiveness of a nurse-delivered intervention to improve adherence to treatment for HIV: a pragmatic, multicentre, open-label, randomised clinical trial

Professor Marijn de Bruin, Ph.D. [1,2]

Edwin Oberjé, Ph.D. [2]

Wolfgang Viechtbauer, Ph.D. [3]

Hans-Erik Nobel, Research Nurse [4]

Mickaël Hiligsmann, Ph.D. [5]

Cees van Nieuwkoop, Ph.D., MD [6]

Jan Veenstra, Ph.D., MD [7]

Frank Pijnappel, Research Nurse [4]

Frank Kroon, Ph.D., MD [8]

Laura van Zonneveld, Research Nurse [9]

Paul H. P. Groeneveld, Ph.D., MD [10]

Marjolein van Broekhuizen, Nurse specialist [11]

Professor Silvia M Evers, Ph.D. [5,12]

Professor Jan M Prins, Ph.D., MD [4]

[1] Corresponding author: University of Aberdeen, Institute of Applied Health Sciences, Foresterhill, AB25 2ZD, Scotland, United Kingdom. Email: M.debruin@abdn.ac.uk Phone: +44 1224 438076

[2] Amsterdam School of Communication Research (ASCoR), University of Amsterdam, the Netherlands

[3] Department of Psychiatry and Neuropsychology, Faculty of Health, Medicine, and Life Sciences, Maastricht University, the Netherlands

[4] Department of Internal Medicine, Division of Infectious Diseases, Academic Medical Centre Amsterdam, the Netherlands

[5] Department of Health Services Research, CAPHRI School for Public Health & Primary Care, Maastricht University, the Netherlands

[6] Haga hospital, the Hague, the Netherlands

[7] Sint Lucas Andreas Hospital, Amsterdam, the Netherlands

[8] Leiden University Medical Centre, Leiden, the Netherlands

[9] Erasmus Medical Centre, Rotterdam, the Netherlands

[10] Isala clinics, Zwolle, the Netherlands

[11] Slotervaart Hospital, Amsterdam, the Netherlands

[12] Trimbos Institute, Netherlands Institute of Mental Health and Addiction, Center of Economic Evaluations

Manuscript accepted for publication by *Lancet Infectious Disease*.

ABSTRACT

Background: No high-quality trials have provided evidence of effectiveness and cost-effectiveness of HIV-treatment adherence interventions. The current multi-centre trial examined the effectiveness and cost-effectiveness of the Adherence Improving self-Management Strategy (AIMS). Preparatory studies demonstrated that AIMS is acceptable, feasible to deliver in routine care, and has reproducible effects on medication adherence.

Methods: A multi-centre randomized controlled trial (RCT) in seven academic and non-academic hospitals, comparing AIMS against treatment-as-usual (TAU). AIMS is provided by nurses during routine clinic visits. Treatment-initiating and treatment-experienced patients at-risk for viral rebound were eligible. Plasma viral load collected at months 5, 10, and 15 was the primary effectiveness outcome. Utilizing cohort data from 7347 Dutch HIV-patients to calculate the natural course of illness, a lifetime Markov model was developed to estimate the costs per quality adjusted life-years (QALYs) gained of AIMS from a societal perspective.

Results: The intent-to-treat sample comprised 221 patients. The primary mixed-effects analysis showed that log viral load was 1.26 [1.04-1.52] times higher in the TAU than AIMS arm. Additional viral load analyses of detectable/undetectable viral load (OR=1.89 [0.98-3.65]) and 'treatment failure' (two consecutive detectable viral loads, OR=2.99 [1.21-7.38]) confirmed this finding. The Markov model showed that AIMS was dominant (more effective and less costly) to TAU in all scenarios (base case scenario: 0.034 QALYs gained and €592,- saved per patient).

Interpretation: This carefully-designed RCT and economic model demonstrate that AIMS reduces viral load, increases QALYs, and saves resources. Implementation of AIMS in routine clinical HIV-care is therefore recommended.

Funding: ZonMW, the Netherlands (Grant Number 171002208).

INTRODUCTION

Efficacious drugs for the treatment of HIV/AIDS are widely available in high-income countries since 1996, and increasingly so in low-income countries. The life expectancy of people living with HIV using combination Antiretroviral Therapy (cART) is now almost identical to that of people living without HIV.[1] Moreover, the risk of forward HIV transmission is reduced considerably for successfully treated patients.[2] However, despite a marked reduction in side-effects and complexity of cART regimens over the last two decades, sub-optimal intake of medication (the faulty execution) and premature discontinuation (non-persistence) of cART are two elements of non-adherence that continue to compromise treatment effectiveness.[3] Non-adherence can lead to poor patient outcomes, the development of drug-resistant virus, fewer treatment options due to drug resistance, and elevated onward transmission risks of (resistant) viral strains.[4-9] Hence, supporting patients' adherence is an important objective from a patient and public health perspective, and essential for achieving the UNAIDS 90-90-90 targets.[10]

For the long-term success of cART and its consequent impact on the spread of HIV, it is key to intervene with suboptimal adherence before virological failure occurs. Although meta-regression analyses suggest that the quality of adherence support provided to patients has a large influence on viral suppression rates,[11,12] there is little direct experimental evidence that adherence interventions have a sustained impact on adherence and – more importantly – on viral loads and CD4 cell counts.[13,14] In fact, a recently updated Cochrane review did not identify any low risk-of-bias trial of HIV adherence interventions in high-income countries providing evidence of intervention effects on adherence and clinical outcomes such as viral load. Two such trials were identified for low-income countries.[15] Moreover, there is no evidence that effective HIV-treatment adherence interventions yield benefits for society in terms of cost-effectiveness.[16]

In 2003, we developed the Adherence Improving self-Management Strategy (AIMS), based on empirical literature, behavioural theories, and input from health care professionals and patients.[17] AIMS is a nurse-delivered, 1-on-1 behavioural intervention that incorporates adherence feedback from electronic medication monitors (MEMS-caps) and is designed to fit in routine clinic visits. After a successful pilot-study demonstrating acceptability, feasibility and effects on adherence,[17] a single-centre randomized

controlled trial (RCT) was conducted amongst treatment-experienced patients.[18] Although powered to detect an effect on adherence (primary outcome), this trial also provided tentative evidence of improved viral suppression rates (secondary outcome). However, this RCT was conducted at a single centre, with a homogeneous patient group, and had a short follow-up (7 months). Demonstrating clinically relevant effects on viral load in a high-quality pragmatic trial with a long follow-up, and a heterogeneous group of patients and HIV clinics, could provide conclusive evidence that AIMS is effective. Moreover, demonstrating that AIMS is also cost-effective would be important for policy makers, as well as for adherence intervention research more generally given the very limited evidence of the economic benefit of adherence interventions.

In sum, effective HIV-treatment adherence interventions should benefit patient and public health, and reduce health care expenditures – yet experimental evidence in support of this is lacking. The current study describes the results from a pragmatic, multi-centre randomized controlled trial evaluating the effectiveness of AIMS and the results of a Markov model evaluating the cost-effectiveness of AIMS over a lifetime horizon.

METHODS

The study protocol has been published,[19] and is registered at clinicaltrials.gov (Identifier: NCT01429142). A separate article has been published on the strategies employed for reducing the risk of bias in this trial,[20] and the risk of bias rationale table summarizing this is included in the Appendix (Table 1). We will therefore only succinctly report the methodology in this paper.

Study Setting and Eligibility Criteria

The study was conducted in seven Dutch HIV clinics (academic and non-academic hospitals). Eligible patients were treatment-experienced (≥9 months on cART) and 'at-risk' of viral rebound, or treatment-naïve patients initiating their first cART regimen. 'At-risk' of viral rebound was determined based on having at least one detectable viral load during the previous three years and suboptimal adherence during two months baseline MEMS monitoring (<100% adherence for QD and ≤95% for a BID regimen). These criteria were based on analyses of data from a large HIV-cohort including all registered HIV patients in the Netherlands,[21] and our previous RCT.[18] Exclusion criteria were: age <18 years, severe

psychiatric disorders or other comorbidities precluding compliance with study procedures, pregnancy, plans to interrupt treatment in the next 14 months, life expectancy less than one year, not able to communicate in English or Dutch, viral resistance to three or more antiretroviral drug classes, and about to initiate hepatitis C treatment.

Eligible patients were approached by their treating physician and/or HIV nurse, and given information about the study verbally and in writing. All patients gave written informed consent and the trial was approved by the medical ethical committees of all participating hospitals. Given the absence of any patient safety risks according to the Medical Ethical Committee that approved the trial, there was no Data and Safety Monitoring board.

Patient recruitment started on the 1st of September 2011 and was completed on the 2nd of April 2013. The last patient completed the study on the 16th of June, 2014.

Randomization and Masking

Consenting patients were randomized to AIMS or treatment-as-usual (TAU) within nurses, since randomizing clinics or nurses was expected to result in recruitment bias. The resulting risk of contamination was kept low because key intervention elements – such as MEMS-feedback and all other intervention materials (see Panel 1) - only appeared on the website when the MEMS-cap of an intervention patient was downloaded (see [20] and Appendix Table 1). Randomization was stratified by treatment experience (experienced versus naïve). Block randomization (with randomly ordered blocks of size four, six, and eight to avoid predictability of assignment) was used to balance intervention and control patients over nurses. The randomization table was computer-generated by a statistician and treatment assignment was done automatically by software after nurses entered the details of consenting patients on a study website. As blinding to treatment assignment is not possible given the nature of the intervention, we developed a 'distraction' strategy for drawing patient and health care provider attention away from the primary study aims. Specifically, we included a second research objective in the study (i.e., 'To examine the content of, and patient satisfaction with, nursing care provided to patients treated for HIV'), and the regular questionnaires nurses and patients completed during the trial focused on this study aim, rather than on the comparison of AIMS versus TAU.[20]

Study Design and Measurements

HIV-nurses (n=21) from the seven participating clinics received a training (3 times 6 hours) on AIMS and using the Medication Event Monitoring System (MEMS-caps, an electronic pill-bottle cap that registers date and time of pill bottle opening). A 1.5-hour booster session was delivered at each HIV-clinic (2-3 nurses per session) after each nurse had seen 2-3 patients. The first author delivered the training and booster sessions. There was no additional support or advice in relation to the delivery of the intervention.

Patient demographics and treatment details were collected at baseline. Plasma viral load and CD4 cell counts were assessed at baseline (Time 1) and at three follow-up time points (Time 2, 3, and 4) as part of routine care. For treatment-initiating patient Time 2 measurement was planned at 5-6 months, to allow patients to become undetectable. Treatment-experienced patients followed the usual 4-5 months visit interval. The observed times of outcome measurement of treatment-experienced versus treatment-naïve patients (mean (SD) number of days) since randomisation were 125 (44) versus 177 (54); 270 (76) versus 306 (69); and 447 (87) and 454 (83) for Time 2, 3 and 4 respectively. The viral load assays used were COBAS AmpliPrep/COBAS TaqMan HIV-1 Test, v2.0 (Roche), Abbott m2000 RealTime HIV-1, and NucliSENS Easy Q HIV-1 v2.0 (Biomerieux), with lower detection limits varying from 20 to 75 copies/ml.

The study was overpowered for detecting an effect on adherence. To avoid unnecessary study burden, we measured MEMS-adherence in a randomly selected 50% of the control group patients. Since a subset of patients prefers using their own medication bottles over the MEMS-caps bottles (especially if MEMS-caps are used for monitoring only, as in the TAU arm),[18,22] and because adherence is a secondary outcome, if randomised patients preferred further trial participation without MEMS-monitoring, they were allowed to do so (for procedure see Appendix).

Treatment-As-Usual Provided and the Adherence Improving self-Management Strategy

The quality and quantity of TAU adherence support provided to control groups in adherence trials varies between trials and impacts on effect sizes.[11,12] We developed a minimally intrusive method for collecting TAU data from participating nurses,[23] and found that TAU in participating clinics ranged from medium to

high quality when compared with meta-analyses on this topic.[11,12] Note that TAU was not standardised for the purpose of this trial, and reflects what patients receive in routine clinical care in the Netherlands. AIMS and TAU are described in Panel 1.

Panel 1. Treatment-as-usual versus the Adherence Improving self-Management Strategy

Both TAU and AIMS were delivered as part of routine care by trained HIV/AIDS nurses. The Table below summarises the Materials used and Procedures for TAU and AIMS. A more comprehensive table also including the behaviour change objectives and techniques is included in the Appendix. AIMS intervention materials can be requested from the lead author.

TAU	AIMS
Materials used	
1. Patient information leaflet	1. Easy-to-remember graph explaining how drug levels vary with (non)adherence patterns, and impact on treatment outcomes 2. Seven example adherence reports from electronic monitors ranging from excellent to poor adherence 3. Drop-down lists with common reasons other patients have for achieving high levels of adherence 4. MEMS (view)-cap to monitor own adherence and obtain printed personal adherence reports 5. Templates for action plans and coping plans 6. Drop-down lists with common reasons for non-adherence and effective solutions for dealing with these problems (e.g., electronic reminders, social support, planning ahead for holidays) 7. Ruler (1-10 scale) to score own confidence in improving adherence 8. For treatment initiating patients only: score sheet of 5 reasons for, and 5 concerns about, initiating treatment
Procedures: the activities done and how they relate to the materials described	
When the physician, nurse, and patient agreed treatment should be initiated, typically the following activities were done to support adherence: 1. Patients are verbally explained how the medication works and what the relation is between adherence, viral replication, and treatment outcomes. This includes risks (e.g., viral resistance) and benefits (e.g., healthy immune system, less infectious) of (non) adherence. Information leaflet provided (Material 1). 2. Patients are explained how, how often, and in what dose the medication should be taken	Here we explain AIMS for treatment-experienced patients. Visit 1 is slightly different for treatment naïve patients, which is explained in the Appendix. Prior to the first AIMS intervention visit, patients used an electronic medication monitor for 4-8 weeks. Data were downloaded and a website guided patients and nurses through the steps below. Tailoring of the intervention to the needs and abilities of each individual patient, was a core component of each step. 1. The same educational activities as for TAU step 1, except that Material 1 was used to facilitate discussion and enhance information storage in long-term memory (Material 1). 2. Nurse explains seven exemplar MEMS-reports (Material 2) while linking (non-)adherence patterns to the

3. Nurse and patient discuss when it is best for each individual patient to take their medication (at what time, where, linking intake to daily routines or using reminder devices that can serve as cues) 4. Patient are given a phone number to call in case of difficulties (e.g., side effects, adherence) During follow-up visits (this also applies to treatment-experienced patients) 5. Patient and nurse discuss self-reported adherence (problems) and try to identify solutions that would work for that patient 6. Nurse/Physician ask about any side-effects and discuss how to deal with them (if severe, change of regimen is considered) 7. Nurses provide viral load and CD4 cell count feedback. If results are positive, this serves to reinforce adherence. If results are negative, adherence problems or other causes (e.g., drug resistance, drug interactions) are explored (Objective 6) *Note that Steps 4, 6, and 7 were also delivered to AIMS patients as part of their routine care.*	adherence-outcome information discussed in Step 1. Patient selects one adherence report reflecting how they would like to take their medication ('Desired adherence level') and explain why this is important to them personally/in the long-run (Material 3). 3. Patients' own MEMS-report is printed ('Actual adherence level', Material 4) and compared with their 'Desired adherence level'. Includes reinforcement of good adherence and highlighting discrepancies (i.e., where actual adherence is lower than desired). 4. Patient MEMS-report is used to identify any non-adherence patterns, causes, and solutions. These are written down in coping plans (if-then format) (Materials 4, 5 and 6). 5. Patient selects an adherence goal for the next visit (using Material 2) and scores confidence (Material 7) in their ability to accomplish that. If confidence is low, the nurse explores whether important adherence barriers have been unaddressed and/or if their adherence goal should be approached incrementally. 6. The patient is offered a MEMS-view cap with a display showing how often the bottle has been opened that day (direct behavioural feedback, Material 4). Patient is given printed adherence report and coping plan. Subsequent intervention sessions are mainly repetitions of the activities described under Steps 3, 4, and 5. The aim is that patients reach their desired level of adherence during the first ±5 months of the intervention, strive for behavioural maintenance during the next ±5 months, followed by a follow-up period of another ±5 months. Patients with many adherence difficulties can be seen more frequently.

Outcome Measure, Statistical Analyses, and Sample Size for the Effectiveness Analyses

The study was powered to detect an effect on plasma viral load, measured at three consecutive time points (Time points 2, 3, and 4), while controlling for baseline viral load. A sample of 230 randomized patients was required to obtain 80% power to detect a significant intervention effect on viral load for at least one of three time points with alpha = .05 (two-sided), using a Bonferroni correction and assuming a maximum dropout of 10%.

The primary effectiveness outcome was defined as log_{10}-transformed viral load (copies/ml) across the three follow-up time points. The secondary effectiveness outcome was percentage adherence. Post-hoc outcomes were (1) 'Treatment failure', defined as having a detectable viral load on two consecutive follow-up measurements; (2) CD4 cell count (cells/mm3); and (3) Detectable versus undetectable viral load. The latter measure was to be used as the primary outcome instead of log_{10}- viral load, if the skewed distribution of log_{10} viral load data would lead to violation of statistical model assumptions. As model assumptions were not violated, this analysis is reported as post-hoc.

The primary intent-to-treat analysis for log_{10} viral load used a mixed-effects (multilevel) model .[24,25] A factor for time point (3 levels: Time points 2, 3, 4), group (2 levels), and their interaction (testing for a between-group change during follow-up) were the primary variables of interest. In the absence of a time-by-group interaction, the overall intervention effect can be estimated by a between-group (marginal) contrast across the three follow-up time points. Baseline viral load and the stratification variable (treatment-experienced versus treatment-naïve) were added to the model as covariates; as well as a four-level factor for ethnicity (Caucasian, Sub-Saharan African, Caribbean, and Others patients), as this is an important prognostic covariate.[11,20] The viral load results were exponentiated (with base 10) for easier interpretation. Undetectable viral loads (e.g., <40 copies/ml) were replaced by the corresponding detection limit.

We also conducted: (1) A mixed-effects logistic regression model,[25] using detectable versus undetectable viral load (based on the detection limit of each respective clinic). The detection limit value of each viral load test was added as an additional covariate. (2) A mixed-effects logistic regression model examining 'treatment failure', using the same covariates. (3) A mixed-effects model examining the effects of the intervention on CD4 cell count, using the same model as for the primary viral load analysis, but with viral load replaced by CD4 values.

Based on the fitted models, we also obtained marginal estimates of the AIMS and TAU group-specific means (viral load and CD4 analyses) and risks (detectable viral load and treatment failure analyses), using the median value at baseline for continuous covariates (i.e., baseline viral load and detection limit)

and the observed proportions at baseline for categorical covariates (i.e., treatment-experienced versus treatment-naïve, ethnicity, and detection status at baseline).

No statistical analyses were conducted on the secondary outcome adherence, because of considerable differences in the uptake of the MEMS-monitoring between the study arms (e.g., 91% (52/57) of the treatment-naïve patients randomized to AIMS versus 54% (15/28) in the TAU arm started the use of MEMS after randomisation).

Analyses were carried out in R (version 3.1.2) using the *nlme* package,[24] and Stata (version 13.1) using functions *mixed* and *meqrlogit*. Additional details on the sample size calculation and statistical analyses are in the Appendix.

Cost-effectiveness Analysis

The primary outcome for the *model-based* economic evaluation reported here is lifetime societal costs (including health care costs, productivity loss, HIV transmission costs, and intervention cost) per quality-adjusted life-years (QALYs) of AIMS versus TAU.[19] A *trial-based* economic evaluation, which examines the short-term economic outcomes observed during the follow-up period of the trial and therefore has another primary outcome (see clinicaltrials.gov), will be published separately.

A Markov model was developed based on the Dutch guideline for health economic evaluations and international guidelines for modeling.[26] In a Markov model, a cohort of patients is assumed to transit between health states. Based on the literature and input from clinicians in the participating clinics, 13 health states were identified: three CD4-cell count categories (0-200, 201-500, and >500) combined with 4 viral load categories (0-50, 51-200, 201-1000, and >1000 copies/ml), and death. These health states capture the key changes in viral load and CD4 cell count associated with changes in costs, HIV transmission risk, and quality of life. A 6-month cycle length was used, meaning that patients can change between health states every 6 months. All transitions between health states are possible except when a patient died. Hence, the Markov model is a matrix existing of 13 rows (current health status) and 13 columns (the health state patients move to; see Appendix Table 3).

Next, we calculated the 6-months transition probabilities of TAU patients moving between these health states (the natural course of illness), and the health care consumption in each health state over a 6-month period. For that, we obtained a longitudinal dataset (2008 to 2015) from the HIV Monitoring Foundation, the Netherlands. We used data from all registered Dutch HIV patients (N = 7347) who were on treatment for ≥12 months (two 6-month cycles), and had at least one detectable RNA viral load measurement (>50 copies/ml) in the last 3 years (excluding the first 12 months of treatment), to approximate the inclusion criteria for treatment experienced patients in the trial. Excess mortality per health state was also derived from this cohort. Utility data (i.e., quality of life) per health state were based on CD4 cell count and obtained from another cohort study.[27] HIV transmission probabilities per health state based on viral load data, were estimated by the lead author of an HIV transmission modeling study,[8] and multiplied by the lifetime treatment costs for an HIV-infected patient.[28] For the societal perspective, the model also included productivity losses per health state based on 600 questionnaires completed by 195 patients during the current multi-centre trial. Tables with these transition probabilities, costs (health care costs, HIV transmission costs, and productivity loss), and utilities per health state are included in the online Appendix (Tables 2-3).

To assess the cost-effectiveness of AIMS, data are required on the intervention cost, as well as on the effects of AIMS on the transition probabilities during and after the intervention period. These effects were calculated from the trial data and expressed in relative risks (AIMS versus TAU; Appendix Table 4). For the AIMS intervention, these relative risks were then applied to the natural course of illness (Appendix, Table 3) over three 6-month cycles, which is the approximate duration of the trial. The cohort of patients receiving AIMS has therefore different probabilities of moving between health states than patients receiving TUA, and therefore costs and outcomes will be different.

To define the relative risks of AIMS, a base case and two additional scenarios were conducted. The base case (scenario 1) included all relative risks (AIMS versus TAU) where at least 5 transitions occurred in the trial (see Appendix Table 4). Scenario 2 included all available relative risks irrespective of the number of transitions, whereas the more conservative scenario 3 included only relative risks with at least 10 transitions. Within these 3 scenarios, we varied our assumptions about how long the effects of AIMS would sustain if delivery would be discontinued after the initial 18 months: (1) a linear decrease of the

effects of AIMS to zero 18 months after intervention discontinuation; (2) no effect after AIMS discontinuation; (3) AIMS effects fully sustained for another 18 months, and then to zero. A total of 9 scenarios were therefore tested. Sensitivity analyses were further performed for a healthcare perspective (i.e., excluding productivity losses) and a time horizon of 10 years instead of lifetime. For each scenario and sensitivity analysis, we estimated the societal costs and QALYs of AIMS compared with TAU, and calculate the incremental cost-effectiveness ratio (ICER) between AIMS and TAU. The ICER expresses the additional cost of AIMS compared with TAU to obtain one additional QALY. When an intervention is more effective *and* less costly, the intervention is said to be cost-saving.

Role of the funding source

This trial was funded from public money by the Netherlands Organisation for Health Research and Development (ZonMW) (Grant Number 171002208). This funding source had no role in study design, data collection, analysis, interpretation, or writing/revising the report.

RESULTS

Recruitment and randomization

224 patients were randomized: slightly below the target of 230 but dropout was lower than anticipated (4.5% (10/224) instead of 10%). The intent-to-treat sample comprised 221 patients: one patient who was not planning to start with cART was accidentally randomized, and two eligible patients (one in each arm) did not provide any outcome data, because soon after randomization one died of a cardiovascular event, and the other was incarcerated in another country. As these reasons were unrelated to group assignment or the dependent variable, team members (MdB, WV, and JMP) blinded to group assignment concluded these were valid reasons for exclusion (Cochrane handbook 8.13 and 16.2).[29]

Figure 1 shows the flow of participants through the study and the most frequent patient-reported reasons for study refusal. In a logistic regression analysis, treatment experience (treatment-naïve patients were more likely to participate), but not gender, age, ethnicity, CD4, or viral load predicted study participation.

Sample descriptives and intervention fidelity

The majority of the intent-to-treat sample was male (185/221, 84%), Caucasian (143/221, 65%), with an average age of 44 years (*SD*=10.9). The majority had a low to medium educational level. About half of

the participants were treatment-experienced and 34% (37/109) of those had a detectable viral load at baseline, confirming that the 'at-risk' selection criteria were useful (viral suppression rate in the general treatment-experienced population in the Netherlands is 91%).[30] See Table 1 for further sample descriptives.

The mean follow-up of study participants was 14.6 months (SD=2.7). The mean number of TAU and AIMS visits were 3.2 (SD=1.6) and 3.2 (SD=1.7), respectively. The delivery of AIMS took an average of 10.3 additional minutes/visit (35 minutes in total during follow-up). Intervention patients received on average 85% of all planned intervention visits, during which 65% of all the intervention elements were delivered (recorded on the intervention website). The main reason recorded for not delivering all intervention elements, was adherence having improved during follow-up sessions, without additional issues to address; or because the action/coping plans made during the previous intervention session remained relevant and did not need to be completed again.

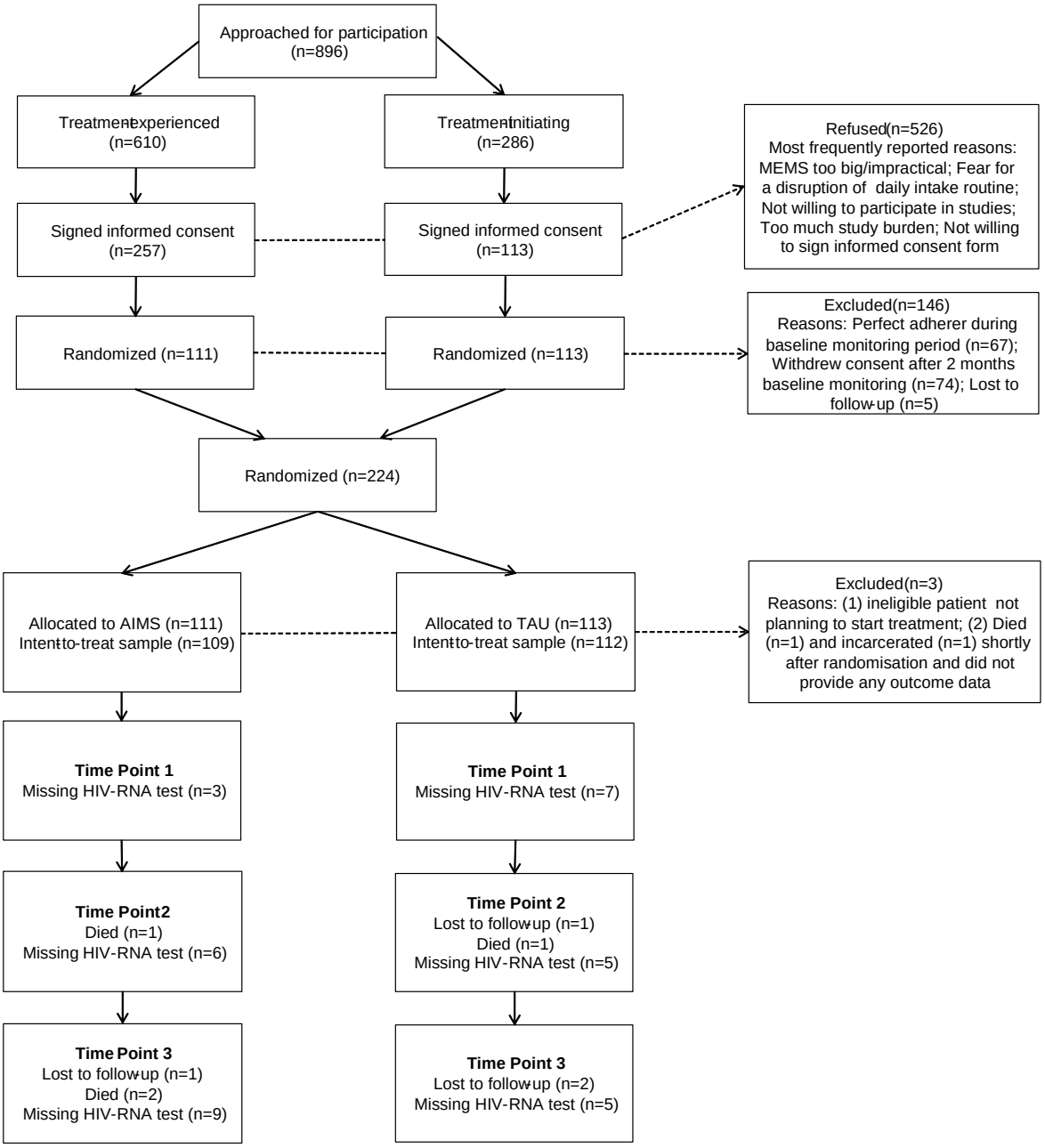


Figure 1. CONSORT flow diagram

Table 1: Baseline Characteristics

Characteristic	*Intervention group (N = 109)*	*Control group (N = 112)*
Female, *n* (%)	14 (12.8%)	22 (19.6%)
Age, years, mean (*SD*)	45.4 (11.0)	43.3 (10.8)
Ethnicity, *n* (%)		
Caucasian	81 (74.3%)	62 (55.3%)
Sub-Saharan African	16 (14.7%)	21 (18.8%)
Caribbean[a]	9 (8.2%)	21 (18.8%)
Other	3 (2.8%)	8 (7.1%)
Education,[b] *n* (%)		
Low	47 (43.1%)	45 (40.2%)
Medium	40 (36.7%)	39 (34.8%)
High	22 (20.2%)	28 (25.0%)
Sexual orientation, *n* (%)		
Homosexual	56 (51.4%)	63 (56.3%)
Bisexual	11 (10.1%)	11 (9.8%)
Heterosexual	42 (38.5%)	38 (33.9%)
Treatment status, *n* (%)		
Treatment-experienced	52 (47.7%)	57 (50.9%)
Treatment-naïve	57 (52.3%)	55 (49.1%)
CD4+ cell count, cells/mm^3, mean (*SD*)		
Treatment-experienced	520.6 (212.9)	535.1 (226.4)
Treatment-naïve	379.1 (239.5)	431.8 (200.5)
Plasma HIV-RNA, $mean_{log}$ (*SD*)		
Treatment-experienced	1.74 (0.61)	1.83 (0.83)
Treatment-naïve	4.83 (0.71)	4.30 (1.01)

[a] Surinamese, Latin American and Antillean. [b] Categorization was based on the Dutch education system, ranging from (a) low: (less than) primary education, lower secondary education; (b) medium: higher secondary education, lower vocational education; (c) high: higher vocational education, university.

Handling missing data

There were 634/663 (95,6%) completed follow-up viral load measurements and 29/663 (4.4%) missing values, which were not associated with group assignment or viral load values at other time points in logistic regression models. Missing data were assumed to be missing at random, except for two patients who dropped out of care, discontinued medication after Time point 3, and did not provide a viral load at Time point 4. As AIMS should reduce such non-persistence (i.e., a stage of non-adherence),[3] and non-persistence affects the dependent variable, these data cannot be treated as missing at random. Based on clinical advice, the two missing values were replaced by the median baseline viral load (50,123 copies/ml) and CD4 count (400 cells/mm^3) of the treatment-naive patients participating in the study. These decisions were based on consensus between team members (MdB, WV, and JMP) blinded to group assignment.

Since all 221 patients provided data at least one follow-up measure, the mixed-effects analyses include the full intent-to-treat sample. The statistician (WV) who conducted the analyses was blinded to group assignment. The main treatment effects are described here (for the results on the covariates and exploratory subgroup analyses, please see the Appendix).

Primary Effectiveness Analysis

The three-level mixed-effects regression model showed that there was no indication of a change in the intervention effect across the three follow-up time points (time-by-group interaction ($F(2,409)=0.75$, $p=.47$)). We therefore examined the between-group contrast across the three follow-up time points, which showed that the intervention was effective ($F(1,196)=6.40$, $p=.012$), while controlling for baseline viral load, treatment experience, and ethnicity. Patients in the control group had viral loads that were on average 1.26 times (95%CI: 1.04 to 1.52) higher than those in the intervention group. There was no significant variability of the treatment effect across nurses ($p = .14$).

Post-hoc Effectiveness Analyses

The three-level mixed-effects logistic regression model with detectable versus undetectable viral loads showed the same pattern ($\chi^2(df=1)=3.66$, $p=.056$). Overall, patients in the control group had a 1.89 times higher odds of a detectable viral load across the three time points (95%CI: 0.98 to 3.65).

The two-level logistic regression model of treatment failure indicated a significant group difference ($\chi^2(df=1)=5.61$, $p=.012$). The odds of treatment failure were 2.99 times higher in the control group (95% CI: 1.21 to 7.38).

The model examining the effects on CD4 cell count did reveal a significant time-by-group interaction ($F(2,398)=3.09$, $p=0.047$). We therefore examined the group difference for each follow-up time point separately. At the first follow-up (Time 2), there was a non-significant increase in CD4 cell count in the intervention compared with the control arm (31 cells/mm^3, 95%CI: -8.37 to 70.37); at Time 3 the control group caught up (-6.55 cells/mm^3, 95%CI: -46.03 to 32.92); and at Time 4 CD4 cell counts continued to rise in the intervention but not in the control arm, and the difference was significant (39.39 cells/mm^3, 95%CI: 0.10 to 78.67). Marginal group means and risks for these analyses are shown in Table 2.

Table 2. Estimated marginal means, estimated risks, and 95% confidence intervals for viral load and CD4 values in the TAU and AIMS groups

	AIMS	TAU
Viral load (copies/ml)	35.4 (29.9 to 42.0)	44.5 (35.5 to 55.9)
Viral load (% detectable)	9.6 (3.8 to 15.4)	16.7 (8.2 to 25.3)
Treatment failure (%)	9.0 (2.4 to 15.7)	22.8 (11.7 to 34.0)
CD4 Time 1 (cells/mm^3)	550.9 (520.4 to 581.4)	519.9 (489.3 to 550.5)
CD4 Time 2 (cells/mm^3)	562.5 (531.7 to 593.3),	569.0 (538.7 to 599.4)
CD4 Time 3 (cells/mm^3)	597.8 (567.1 to 628.5)	558.4 (528.2 to 588.6)

Note: For CD4 cell count, analysis were conducted per time-point given the significant time-by-group interaction during the 3 follow-up measures (i.e., effects were different at different follow-up time points)

The Cost-Effectiveness Analyses

In the base-case analysis, the Markov model estimated that AIMS reduces lifetime societal costs by €592 per patient and increases QALY by 0,034 per patient. AIMS was therefore cost-saving (i.e., more

QALYs and less costs) in the base case. Results were comparable for the other scenarios and for the sensitivity analyses with a health care perspective, and a 10-year time horizon (Table 3).

Table 3. Lifetime costs per patient, QALYS, and incremental cost-effectiveness ratio of AIMS compared with TUA: base case and sensitivity analyses

	Lifetime costs	Lifetime QALYs	ICERs
Offset: linear decrease of AIMS effect over 18 months			
Scenario 1 (base case)	€-592	0,034	AIMS dominant
Scenario 2	€-843	0,036	AIMS dominant
Scenario 3	€-412	0,025	AIMS dominant
Offset: effect AIMS maintained over another 18 months			
Scenario 1	€-793	0,046	AIMS dominant
Scenario 2	€-1117	0,049	AIMS dominant
Scenario 3	€-599	0,035	AIMS dominant
Offset: no effect after stopping AIMS			
Scenario 1	€-375	0,023	AIMS dominant
Scenario 2	€-546	0,024	AIMS dominant
Scenario 3	€-221	0,016	AIMS dominant
Sensitivity analyses (base case)			
Healthcare perspective	€-597	0,034	AIMS dominant
10 year time horizon	€-643	0,028	AIMS dominant

ICER = Incremental Cost-Effectiveness Ratio; Scenario 1: all relative risks where at least 5 transitions occurred; Scenario 2: all available transition probabilities irrespective of the number of transitions; Scenario 3: only relative risks with at least 10 transitions in total.

CONCLUSIONS

To our knowledge, this is the first randomized controlled trial of an HIV treatment adherence intervention that demonstrates a clinically meaningful effect on viral load as well as cost-effectiveness. Importantly, the economic model shows that AIMS is dominant to TAU: both cheaper and more effective, regardless of the time horizon (life time or 10 year) and perspective (health care or societal). These results have been obtained in a heterogeneous sample of HIV-infected patients and clinics, where AIMS was delivered by nurses as part of routine care.

A recent Cochrane review did not identify RCTs with a low risk of bias that demonstrated an impact of HIV treatment adherence interventions on adherence and clinical outcomes in high-income settings.[15] Short follow-up periods (<6 months) and a high risk of bias were important reasons for excluding many trials from these analyses. In the design of the current study, we tried to overcome these and additional challenges by designing a study with a long follow-up period (15 months), extensive efforts to minimize the risk of bias (which is particularly challenging in behavioural trials as blinding to treatment assignment is typically not possible), and the detailed reporting of TAU provided to control participants.[11,12,20,23] Although one limitation of the current study was the low uptake of MEMS-monitoring in the TAU arm, precluding meaningful secondary adherence analyses, the effects of AIMS on adherence had already been demonstrated in two earlier studies.[17,18] Moreover, as there is no plausible other pathway to improved viral loads in the AIMS-arm than through improved adherence, this limitation does not influence the overall study conclusions. Secondly, although dropout rates were very low, a 60% study refusal rate may limit generalizability of the findings. We did, however, not find demographic or clinical differences between participants and those refusing study participation. Relevant to note here is that most reasons for study refusal (see Figure 1) are unlikely to be a barrier to the uptake of AIMS in routine care. Specifically with regards to patients' willingness to utilize an electronic adherence monitor, we expect substantially fewer issues when AIMS is implemented in routine care, since patients know they will receive AIMS and the feedback, AIMS can now be presented as evidence-based care, and ongoing technological developments should make more used-friendly devices available shortly. Indeed, in a pharmacy-based HIV-treatment adherence clinic in Lausanne (Switzerland) that uses MEMS-monitoring in routine care, refusal of MEMS-monitoring is rare (personal communication with Dr. M.P. Schneider).

Panel 2. Research in context

Evidence before this study

Two systematic reviews synthesizing the evidence on the (cost)effectiveness of medication adherence interventions up to January and July 2013, did not identify any adherence interventions that were effective and cost-effective. We conducted two updated searches to identify the most recent evidence on (cost)effectiveness of HIV treatment adherence interventions. We searched for effectiveness and cost-effectiveness evidence from RCTs conducted in high-income countries, with at least 12 months follow-up, including a clinical outcome, and focusing on adult HIV infected patients. Interventions had to promote autonomous behaviour (i.e., directly observed therapy interventions were excluded) and treatment simplification studies (e.g., once versus twice daily medication) were excluded.

Effectiveness**:**. Search terms: ((HIV or HAART or cART or Antiretroviral) and (adherence or compliance or persistence or concordance) and (viral load or virologic failure or CD4) in title or abstract) and ((random* or clinical trial) in all text) and (("2013" or "2014" or "2015" or "2016") in year). Databases: MEDLINE, PsycINFO, Embase. Search dates: January 2013 to October 2016. Search results: 529 unique titles were obtained of which 27 evaluated an adherence intervention. Twenty-six were excluded as they were short-term and small-scale (pilot)studies, and/or conducted in low-resource settings, and/or focused on youth, and/or did not have an RCT design. The one eligible RCT evaluated the Managed Problem Solving (MaPS) intervention by Gross and colleagues which – similarly to the AIMS intervention – utilises electronic monitoring feedback in an interpersonal (1-on-1) intervention. MaPS was delivered outside of routine clinical care by specially trained staff and is more labor intensive than AIMS: approximately 250 minutes of face-to-face contact plus 22 telephone calls per patient over a 12-month period. MaPS improved adherence (primary outcome; 1.78 (95% CI,1.07-2.96)), and the effect on detectable/undetectable viral load (secondary outcome) was on the border of significance (OR = 1.48, 95% CI: 0.94-2.31). A particular strength of the trial was the high consent rate, which may reflect participants' positive perception of the trial and intervention, but may in part also be due to the financial incentives offered for completing study measures. Possible trial weaknesses were (differential) attrition (36% in MaPS and 26% in controls) and – according to the

Cochrane risk of bias tool – a missing data imputation method that 'can lead to serious bias' (i.e., missing equals treatment failure). No cost-effectiveness analysis was reported.

Cost-effectiveness: Search terms ((HIV or HAART or cART or Antiretroviral) and (adherence or compliance or persistence or concordance) and (Cost Analysis or Cost Effectiveness or Cost Benefit or Cost Utility or Cost Minimi#ation or Economic Evaluation) in title or abstract); and (2013 or 2014 or 2015 or 2016) in year. Databases: MEDLINE, PsycINFO, Embase. Search dates: January 2013 to October 2016. Search results: 137 unique titles/abstracts were scanned and 6 studies were examined in more detail. Five studies were directly excluded as they were only examining costs (not cost-effectiveness), and/or were conducted in low-resource settings, and/or were a conference abstract (so quality could not be assessed). The one eligible study by Ownby and colleagues (2013) reported the cost-effectiveness of a computer-delivered intervention to promote adherence to HIV medication (Florida, United States). This evaluation was, however, based on effectiveness data from a subgroup analysis in a short-term intervention feasibility study. Further limitations were that the effectiveness data was derived from self-reported adherence and did not line-up with the effectiveness input in the economic model (i.e., CD4 counts were used to define health states); hence, the authors had to make assumptions about the relationship between their self-reported adherence measure and CD4 counts, which were not supported by empirical data– and in fact opposed by some studies.

Hence, these (updated) searches did not identify any adherence interventions from high-quality, long-term trials and economic evaluations that provided evidence of effectiveness and cost-effectiveness.

Added value of this study

To our knowledge, this carefully designed multi-center, randomized controlled trial and economic model are the first to demonstrate that an adherence intervention can produce meaningful effects on viral load and be cost-effective in a high-resource setting. In fact, this study shows that (HIV) treatment adherence interventions can increase QALYs while saving resources, even when compared against medium-to-high quality treatment-as-usual. Moreover, AIMS requires few resources as it has been adapted to fit in routine HIV clinic services, which should facilitate implementation in routine care.

Implications of all the available evidence

HIV treatment adherence interventions can benefit patients, even in high-resource settings, and lead to gains in QALYs while saving resources. AIMS seems at present to be the only adherence intervention of which the effects have been replicated in consecutive trials. The current economic evaluation also provides robust evidence on cost-effectiveness. Implementation of AIMS in routine clinical care is therefore recommended.

Similarly, a recent systematic review identified a lack of evidence on the cost-effectiveness of HIV treatment adherence interventions, as it identified only one cost-effective HIV treatment adherence intervention evaluated in an RCT and subjected to a high-quality economic evaluation.[16] However, the paper did not report evidence of intervention effectiveness, or the content of the control and experimental interventions, so that generalizability, replicability, and scalability of the intervention (effects) are unclear.[16] Our aim was to collect and report this information, and conduct a similarly high-quality economic evaluation. Given the absence of a suitable and up-to-date Markov model for that purpose, a new model was developed using ISPOR-SMDM guidelines.[26] Up-to-date cohort data (2008-2015) from all registered HIV patients in the Netherlands meeting our inclusion criteria were used to describe the natural course of illness. Besides effects on costs (health care and productivity) and quality of life, the model also incorporated HIV transmissions avoided given the evidence that lower viral loads reduce transmission risk.[2,8] Although a limitation of the current model was the absence of trial data to populate the full health state transition matrix, the finding that AIMS is more effective and saves resources was robust as all scenarios and sensitivity analyses produced the same result.

The cumulative results of the current multi-centre RCT and the previous pilot study and single-centre RCT, show that AIMS requires few resources, is feasible to deliver in routine care, and is acceptable to health care providers and patients (although more patient-friendly electronic monitoring devices are desirable). Moreover, they demonstrate relevant and replicable effects of AIMS on adherence (in the pilot study and single-centre RCT) and viral load (in the single-centre RCT and multi-centre RCT).[17,18] On average patients receiving TAU had a 1.26 higher log viral load than AIMS patients, and AIMS reduced the risk of treatment failure (2 consecutive detectable viral loads) by 61% (22.8% versus 9.0%).

These effects were comparable for treatment naïve and treatment experienced patients at-risk for viral rebound (see subgroup analyses in the Appendix), and despite some risk of contamination and the medium to high-quality TAU adherence support provided to the control group. The economic analysis showed that AIMS is dominant and that when the intervention is provided to 10,000 patients over a period of 18 months, the approximate savings would be 5,300,300 euro while 340 QALYs would be gained. As these results have been obtained in a heterogeneous sample of patients and clinics, we would expect at least similar effects if AIMS was to be rolled out nationally in the Netherlands, and in other countries were HIV care is organized in a similar manner (i.e., western Europe). Nation-wide training of health care professionals, reimbursement of electronic monitors, and adoption of AIMS in national HIV-treatment guidelines in the Netherlands is currently being negotiated, as a first step.

In conclusion, the current pragmatic, randomized controlled trial and the economic model demonstrates that AIMS is feasible to deliver in routine are, reduces viral load, increases QALY, and saves resources. To our knowledge this is the first HIV treatment adherence intervention for which such an evidence-base has been established. The AIMS intervention should thus be scalable and the results generalizable to the wider population of patients and HIV clinics – at least in high-income settings. Implementation of AIMS in routine HIV clinical care is therefore strongly recommended.

CONFLICT OF INTERESTS

We declare that we have no conflicts of interest.

CONTRIBUTORS

MdB, JMP, SME, and WV designed the study and obtained project funding. All authors were involved in defining inclusion/exclusion criteria, measures to protect against bias, and data collection procedures. All authors except WV and SME were involved in data collection for the effectiveness and/or cost-effectiveness analyses. MdB, EO, JMP and WV were primarily responsible for the effectiveness analyses. MH, EO, and MdB were primarily responsible for the cost-effectiveness analyses. All other authors critically examined the analyses and findings. MdB, EO, JMP, and MH drafted the manuscript. All other authors critically read and commented on draft versions of the manuscript, and approved of the final version.

ACKNOWLEDGMENTS

This trial was funded from public money by the Netherlands Organisation for Health Research and Development (ZonMW) (Grant Number 171002208). Aardex ltd provided support on the development of the study website and had no influence on the data collection, analysis, interpretation and writing of the report. We thank all the HIV-nurses and physicians (those who are not included as co-authors here) from the seven HIV-clinics involved in the AIMS-study for their input and collaboration (Academic Medical Centre, Slotervaart hospital, and St. Lucas-Andreas hospital, all in Amsterdam; the Leiden University Medical Centre, Leiden; HAGA hospital, Den Haag; Erasmus Medical Centre, Rotterdam; and Isala clinic, Zwolle.. We also would like to express our gratitude to the study participants. We would like to thank the Stichting HIV Monitoring (SHM) for their support in accessing the SHM database for identifying patient inclusion criteria and developing the Markov model. Finally, we would like to thank and remember Professor Herman Schaalma (†) for his valuable contribution to the study design and grant application.

.

REFERENCES

1 Nakagawa F, Lodwick RK, Smith CJ, et al. Projected life expectancy of people living with HIV according to timing of diagnosis. *AIDS* 2012; **26**: 335-43.

2 Cohen MS, Chen YQ, McCauley M, et al. Prevention of HIV-1 Infection with Early Antiretroviral Therapy. *N Engl J Med* 2011; **365**: 493-505.

3 Blaschke TF, Osterberg L, Vrijens B, Urquhart J. Adherence to medications: Insight arising from studies on the unreliable link between prescribed and actual dosing histories. *Annu Rev Pharmacol Toxicol* 2012; **52**: 275-301.

4 Howard AA, Arnsten JH, Lo Y, et al. A prospective study of adherence and viral load in a large multi-centre cohort of HIV-infected women. *AIDS* 2002; **16**: 2175-82.

5 Lima VD, Harrigan R, Bangsberg DR, et al. The combined effect of modern highly active antiretroviral therapy regimens and adherence on mortality over time. *J Acquir Immune Defic Syndr* 2009; **50**: 529-36.

6 Mannheimer S, Matts J, Telzak E, et al. Quality of life in HIV-infected individuals receiving antiretroviral therapy is related to adherence. *AIDS Care* 2005; **17**: 10-22.

7 Das M, Chu PL, Santos G, et al. Decreases in community viral load are accompanied by reductions in new HIV infections in San Francisco. *PLoS One* 2010; 5: e11068.

8 Wilson DP, Law MG, Grulich AE, Cooper DA, Kaldor JM. Relation between HIV viral load and infectiousness: a model-based analysis. *Lancet* 2008; **372**: 314-20.

9 Pasternak AO, de Bruin M, Jurriaans S, et al. Modest nonadherence to antiretroviral therapy promotes residual HIV-1 replication in the absence of virological rebound in plasma. *J Infect Dis* 2012; **206**: 1443-52.

10 UNAIDS. 90-90-90 An ambitious treatment target to help end the AIDS epidemic. Geneva: UNAIDS, 2014. URL: http://www.unaids.org/sites/default/files/media_asset/90-90-90_en_0.pdf

11 de Bruin M, Viechtbauer W, Schaalma HP, Kok G, Abraham C, Hospers HJ. Standard care impacts on effects of highly active antiretroviral therapy adherence interventions: a meta-analysis of randomized controlled trials. *Arch Intern Med* 2010; **170**: 240-50.

12 de Bruin M, Viechtbauer W, Hospers HJ, Schaalma HP, Kok G. Standard care quality determines treatment outcomes in control groups of HAART-adherence intervention studies: implications for the interpretation and comparison of intervention effects. *Health Psychol* 2009; **28**: 668-74.

13 Amico KR, Harman JJ, Johnson BT. Efficacy of antiretroviral therapy adherence interventions: a research synthesis of trials, 1996 to 2004. *J Acquir Immune Defic Syndr* 2006; **41**: 285-97.

14 Simoni JM, Amico KR, Pearson CR, Malow R. Strategies for promoting adherence to antiretroviral therapy: a review of the literature. *Curr Infect Dis Rep* 2008; **10**: 515-21.

15 Nieuwlaat R, Wilczynski N, Navarro T, et al. Interventions for enhancing medication adherence. *Cochrane Database Syst Rev* 2014; **11**: CD000011.

16 Oberjé EJ, de Kinderen RJ, Evers SM, van Woerkum CM, de Bruin M. Cost effectiveness of medication adherence-enhancing interventions: a systematic review of trial-based economic evaluations. *Pharmacoeconomics* 2013; **31**: 1155-68.

17 de Bruin M, Hospers HJ, Borne HW, Kok G, Prins JM. Theory- and evidence-based intervention to improve adherence among HIV-infected patients in The Netherlands: a pilot study. *AIDS Patient Care STDs* 2005; **19**: 384-94.

18 de Bruin M, Hospers HJ, van Breukelen GJ, Kok G, Koevoets WM, Prins JM. Electronic monitoring-based counseling to enhance adherence among HIV-infected patients: a randomized controlled trial. *Health Psychol* 2010; **29**: 421-8.

19 Oberjé E, de Bruin M, Evers S, et al. Cost-effectiveness of a nurse-based intervention (AIMS) to improve adherence among HIV-infected patients: design of a multi-centre randomised controlled trial. *BMC Health Serv Res* 2013; **13**: 274.

20 de Bruin M, McCambridge J, Prins JM. Reducing the risk of bias in health behaviour change trials: improving trial design, reporting or bias assessment criteria? A review and case study. *Psychol Health* 2015; **30**: 8-34.

21 Gras L, Smit C, de Bruin M, et al. Risk of HIV viremia depends on length of prior suppression <50 copies/ml and is increased in sub Saharan African patients living in the Netherlands. Poster presented at the Conference on Retroviruses and Opportunistic Infections, Seattle, USA. 2012. URL: http://www.hiv-monitoring.nl/files/5413/3363/4143/Poster_CROI_2012_Luuk_Gras.pdf.

22 de Bruin M. Does managed problem solving work and in what setting? *JAMA Intern Med* 2013; **173**: 1474.

23 Oberjé EJ, Dima AL, Pijnappal FJ, Prins JM, de Bruin M. Assessing treatment-as-usual provided to control groups in adherence trials: exploring the use of an open-ended questionnaire for identifying behaviour change techniques. *Psychol Health* 2015; **30**: 897-910.

24 Pinheiro J, Bates D. Mixed-effects models in S and S-PLUS. Springer: New York; 2000.

25 Molenberghs G, Verbeke G. Models for discrete longitudinal data. Springer: New York; 2005.

26 Siebert U, Alagoz O, Bayoumi AM, et al. State-transition modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-3. *Value Health* 2012; **15**: 812-20.

27 Schackman BR, Goldie SJ, Freedberg KA, Losina E, Brazier J, Weinstein MC. Comparison of health state utilities using community and patient preference weights derived from a survey of patients with HIV/AIDS. *Med Decis Making* 2002; **22**: 27-38.

28 Schackman BR, Gebo KA, Walensky RP, et al. The lifetime cost of current human immunodeficiency virus care in the United States. *Med Care* 2006; **44**: 990-7.

29 Higgins JP, Green S. Cochrane handbook for systematic reviews of interventions version 5.1.0 [update March 2011]. URL: www.cochrane-handbook.org. The Cochrane Collaboration; 2011.

30 Stichting HIV Monitoring the Netherlands. Annual report 2014. Amsterdam; 2014.

Appendix

***Part I* – Table 1. The Risk of Bias Justification Table (RATIONALE) for the AIMS cost-effectiveness trial**

Type of bias	**Common strategies for reducing risk of bias**	**Common strategies applied in trial (protocol)? Explain**	**Additional/alternative strategies applied? Explain**
Selection bias	Random sequence generation	Yes: Computer random number generator.	
	Concealment of allocation until assigned	Yes: Randomly permuted block sizes and researchers/ personnel have no access to randomization table.	Blinding of research personnel responsible for including patients (i.e., nurses) to randomization strategy.
Recruitment bias	Include participants before randomization *[or]* Blind recruiters for cluster assignment	*Not applicable: Randomization of individual patients rather than clinics.*	
Baseline imbalance (chance bias)	Include sufficiently large sample size (>100-200)	Yes: Large number of individuals randomized (instead of a small number of clusters).	
	Control for key prognostic covariates	Yes: Stratification, block randomization, and measurement of prognostic covariates.	

Performance bias	Blind personnel for treatment assignment *[or]* Use a strict protocol for participant contact	No: Impossible to blind personnel for treatment assignment. No: A strict protocol for provider-participant contact incompatible with delivering usual care/the intervention.	Distraction strategy to draw attention away from primary research hypotheses. Concealment of non-adherence in control group by randomising adherent patients to TAU arm.
Detection bias	Objective outcome measure *[or]* Blinding outcome assessors to group assignment	Yes: Viral load is the primary outcome variable. Yes: The nurses drawing the blood and the lab-technicians analysing the blood sample are unaware that this sample belongs to a patient who participates in a trial.	
Attrition bias	Analyse participants as randomized	Yes: participants are included in analyses as randomized, regardless of intervention exposure.	Minimize risk of attrition/missing data through reducing study burden, stepwise withdrawal protocol, and building clinic/patient commitment to participation ('informed consent').
	Advanced data imputation procedures	Yes: (a) missing data was missing at random, and mixed-effects regression analyses were conducted including all patients, (b) 3 decision makers blinded to	

		group assignment imputed 1 data point for 2 patients not missing at random.	
Reporting bias	Online registration trial protocol	Yes: Trial protocol registered online prior to start inclusion and published in BMC health services research.	
	CONSORT guideline reporting	Yes: Trial report will be CONSORT compliant.	
Contamination bias	Blinding for treatment assignment *[or]* Cluster randomization *[or]* Restrict access to intervention materials	No: See explanation on performance bias. No: Benefits of individual randomization outweigh the advantages of cluster randomization in this trial. Yes: All intervention materials can only be accessed online for intervention patients. Moreover, key strategies like MEMS-data feedback are not transferrable.	See explanation on performance bias.
	Control for contamination in the analyses	No: No data available on contamination, but risk of contamination was very low (see above).	
Inappropriate administration	Promote accurate program delivery	Yes: Training strategy (initial training and booster) and supportive materials (protocol, website) offered.	
	Control for variable program delivery	Yes: Intervention completeness (the number of planned AIMS modules actually delivered) assessed and will be	

		reported in follow-up paper. Initial analyses suggest that AIMS effects are larger if completeness was better.	
Stop early /continue for benefit	Report sample size computation in the study protocol	Yes: Sample size computations given in study protocol.	
	Report planned interim analyses in the study protocol and apply appropriate analyses	*Not applicable: No interim analyses planned.*	
Scientific misconduct	Be transparent about study methods before, during and after trial	Yes: Trial protocol published, reporting adheres to CONSORT guidelines, statistical outputs will be published as online appendices, and RATIONALE table published.	
	Monitoring conduct by team or board	Yes: All trial decisions (design, management, data analyses) are discussed openly with a research team from different institutions. Ambiguous decisions taken by multiple team members blinded to group assignment.	
	Minimize effects of vested interest	Yes: The commercial company (AARDEX) had no access to data or influence on analyses. A statistician with minimal trial involvement and blinded to group assignment ran the analyses.	

**See also reference 26 in main manuscript. Some minor difference between RATIONALE table in [26] and here are related to (a) have run out of resources (e.g., to also measure TAU after the trial) and (b) not being able to report all the planned analyses in the main manuscript.*

Part II – Details on the strategy for MEMS-use

Patients in the Netherlands often prefer to use their own medication box over the MEMS-cap bottle, due to its design [23, 24]. All treatment-experienced patients had to agree with baseline MEMS-monitoring, however, since this was used to determine trial eligibility. For patients who are then randomized to the intervention, the use of MEMS-caps has clear advantages: it offers a detailed overview of adherence to discuss with the health care professionals. Patients value that information highly [23]. For patients in the TAU arm, there are no such advantages of using the MEMS-caps. In both arms study participation was prioritized over the use of the MEMS-caps. So patients who indicated that they would not want (to continue) participation in the trial because of the MEMS-monitoring were allowed to discontinue MEMS-monitoring. As MEMS-data feedback is an important element of the AIMS intervention, if AIMS patients expressed wanting to discontinue MEMS-use, nurses would first discuss with patients the intermittent use of MEMS-caps (i.e., only the 4 weeks prior to the next visit) before agreeing with discontinuation of MEMS-monitoring.

The main reasons for a subgroup of patients not wanting to use the MEMS-caps in this study were (a) it is bulky and looks like a medication bottle, so hard to discretely carry with them when they are not at home; and (b) it has only a single compartment. These designs issues are currently being, or have already been, addressed by various companies.

Part III – TAU and AIMS description

Panel 1. Treatment-as-usual versus the Adherence Improving self-Management Strategy

Both TAU and AIMS were delivered as part of routine care. Trained HIV/AIDS nurses delivered the adherence support in the HIV clinic during routine clinic visits. TAU and AIMS share several objectives, as well as so-called 'behaviour change techniques' (i.e., the active ingredients of behaviour change interventions) for realising these objectives. For example, both aim for patients being informed about the role of adherence for long-term treatment success, that patients are motivated, and have a medication intake plan that fits in their daily life. Key differences are that AIMS has more objectives (i.e., addresses more determinants of adherence), and uses more and more advanced behaviour change techniques to accomplish these objectives. Additionally, whereas TAU relies on patients' ability and willingness to self-report (barriers to) adherence, AIMS uses electronic medication monitors that produce objective, detailed, longitudinal adherence reports. Finally, AIMS uses several carefully developed materials and information resources in each step of the intervention.

The Table below summarises the Rationale (objectives), Materials, Procedures, and Behaviour Change Techniques for TAU and AIMS. Under Procedures, we link the activities described to the objectives and materials. AIMS intervention materials can be requested from the lead author.

TAU	AIMS
Rationale: behaviour change objectives	
1. Enhancing understanding of the relationship between (non) adherence and outcomes 2. Evoking motivation for high adherence 3. Facilitating the translation of intention for high adherence into action, and the overcoming of adherence barriers for achieving/maintaining good adherence 4. Offering continuous professional support 5 Addressing side effects (an adherence barrier) 6. Feeding back viral load and CD4 cell counts	1. Ensuring understanding and storage of information in long-term memory of relation (non) adherence and outcomes 2. Evoking internal motivation for high adherence 3. Raising awareness of own adherence and barriers/facilitators 4. Evoking motivation for improving adherence 5. Facilitating the translation of intention for high(er) adherence into action, and the overcoming of adherence barriers for achieving/maintaining good adherence 6. Enhancing self-efficacy for changing suboptimal, or maintaining good adherence 7. Consolidating/protecting motivation and self-efficacy for high adherence 8. For treatment-initiating patients only: Strengthening/introducing beliefs about the perceived needs for, and reducing concerns about, antiretroviral treatment *Note that TAU objectives 4, 5, and 6 were also delivered to AIMS patients as part of routine care.*
How: Materials used	

1. Patient information leaflet	1. Easy-to-remember graph explaining how drug levels vary with (non)adherence patterns, and impact on treatment outcomes 2. Seven example adherence reports from electronic monitors ranging from excellent to poor adherence 3. Drop-down lists with common reasons other patients have for achieving high levels of adherence 4. MEMS (view)-cap to monitor own adherence and printed personal adherence reports 5. Templates for action plans and coping plans 6. Drop-down lists with common reasons for non-adherence and effective solutions for dealing with these problems (e.g., electronic reminders, social support, planning ahead for holidays) 7. Ruler (1-10 scale) to score own confidence in improving adherence 8. For treatment initiating patients only: score sheet of 5 reasons for, and 5 concerns about, initiating treatment
How: Procedures (activities, and how they relate to the objectives and materials described)	
When the physician, nurse, and patient agreed treatment should be initiated, typically the following activities were done to support adherence: 1. Patients are verbally explained how the medication works and what the relation is between adherence, viral replication, and treatment outcomes. This includes risks (e.g., viral resistance) and benefits (e.g., healthy immune system, less infectious) of (non) adherence. Information leaflet provided (Objectives 1 and 2, Material 1). 2. Patients are explained how, how often, and in what dose the medication should be taken (Objective 3) 3. Nurse and patient discuss when it is best for each individual patient to take their medication (at what time, where,	We first explain AIMS for treatment-experienced patients, then explain the differences with treatment-initiating patients: Prior to the first AIMS intervention visit, treatment-experienced patients used an electronic medication monitor for 4-8 weeks. Data were downloaded and a website guided patients and nurses through the steps below. Tailoring to needs and abilities of individual patient, was a core component of each step. 1. Started with the same educational activities as for TAU step 1, except that Material 1 was used to facilitate explanation /discussion and increase transfer of information provided to long-term memory (Objective 1 and 2, Material 1). 2. Nurse explains seven exemplar MEMS-reports ranging from perfect to poor adherence, linking this to the adherence-outcome information discussed in Step 1. Patient is asked to select one adherence report reflecting how they would like to take their medication ('Desired adherence level') and explain why this is important to them personally/in the long-run (Objective 2, Material 2 and 3).* 3. Nurse prints patients' own MEMS-report ('Actual adherence level') and a discussion follows which includes reinforcement of (periods of) good adherence, and inducing a state of cognitive dissonance by comparing (periods of) suboptimal adherence with patients' desired adherence level defined in Step 2 (Objectives 3 and 4, Material 4) 4. Patient and nurse use the MEMS-report to identify non-adherence patterns and causes. Patient is encouraged to identify solutions to deal with these barriers, using drop-down list of common adherence barriers and solutions if useful. These are written down in coping plans (if-then format) and – if desired - printed for the patient to take home (Objectives 3 and 5; Materials 4, 5 and 6). 5. The nurse asks the patient to select an adherence goal for the next visit (using Material 2) and to score confidence (10-point scale) in being able to accomplish that goal given their action/coping plans. If patient

linking intake to daily routines or using reminder devices that can serve as cues) (Objective 3) 4. Patient are given a phone number to call in case of difficulties (e.g., side effects, adherence) (Objective 4) During follow-up visits (this also applies to treatment-experienced patients) 5. Patient and nurse discuss self-reported adherence (problems) and try to identify solutions that would work for that patient (Objective 3) 6. Nurse/Physician ask about any side-effects and discuss how to deal with them (if severe, change of regimen is considered) (Objective 5) 7. Nurses provide viral load and CD4 cell count feedback. If results are positive, this serves to reinforce adherence. If results are negative, adherence problems or other causes (e.g., drug resistance, drug interactions) are explored (Objective 6)	confidence is low, the nurse explores whether important adherence barriers have been unaddressed; and if not, if their adherence goal should be tackled incrementally rather than at once (Objective 6, Material 7). 6. The patient is offered a MEMS-view cap, with a display on top showing how often the bottle has been opened that day. This serves as a direct feedback mechanism for missed/late doses, supporting adherence self-monitoring (Objective 3, Material 4). Patient is given printed adherence report and plans. During subsequent intervention sessions… 7. Nurse and patient evaluate whether the action/coping plans were successful, if there were any new barriers, and how the patient dealt with those. 8. Patients' new MEMS-report is printed and discussed. Improvements/good adherence is reinforced and attributed to the persons' efforts/capabilities (Objectives 3 and 7, Materials 4). 9. The causes of suboptimal/disappointing adherence levels are explored, trying to ensure patients attribute 'failure' to controllable/avoidable causes (i.e., a learning experience). Adherence goals and plans are revised, and if needed new/alternative solutions are identified and written down in if-then coping plans. If patients have excellent adherence, the nurse and patient focus on behavioural maintenance by discussing what could be potential barriers disrupting their routine (e.g., a holiday) and identifying coping strategies for those barriers (Objectives 5 and 7; Materials 4, 5, an 6). 10. Replicates Step 5. The aim in this study was that patients reach their desired level of adherence during the first ±5 months of the intervention, strive for behavioural maintenance during the next ±5 months, followed by a follow-up period of another ±5 months. The differences in AIMS for treatment-initiating patients were in Session 1 and related to the patients not having any electronically compiled adherence reports yet (as they still had to initiate their treatment). Step 3 was skipped. Step 4 focused on developing a written medication intake plan (when, where and how to take the medication; linked with daily routines or other cues), an action plan (e.g., to plan organising social support or storing of spare doses of medication at convenient places) and a coping plan (identifying solutions to anticipated adherence barriers). *Note that Steps 4, 6, and 7 were also delivered to AIMS patients as part of their routine care.*
Behaviour Change Techniques (coded with taxonomy https://osf.io/hnyuk/)	
1. Providing general information	1. Provide general information

2. Increasing memory & understanding 3. Feedback of clinical outcomes 4. Risk communication 5. Persuasive arguments 6. Planning of coping responses 7. Develop medication intake schedule 8. Review of adherence goals 9. Use of cues 10. Continuous professional support 11. Cope with side effects	2. Increasing memory & understanding 3. Risk communication 4. Self-monitoring of adherence 5. Electronic monitoring of adherence and delayed feedback on adherence patterns 6. Direct feedback of behaviour 7. Feedback of clinical outcomes 8. Re-evaluation, self-evaluation 9. Persuasive arguments 10. Rewards for behavioural progress 11. Planning of coping responses 12. Setting of graded tasks 13. (Re)attribution of success/failure 14. General intention formation 15. Develop medication intake schedule 16. Specific goal setting 17. Review of adherence goals 18. Use of social support 19. Use of cues 20. Goals for maintenance 21. Relapse prevention 22. Provide supportive materials 23. Continuous professional support 24. Cope with side effects.

Part IV – Details on the sample size computation

The study was powered to detect an effect on plasma viral load, measured at three consecutive time points (Time points 2, 3, and 4), while controlling for baseline viral load. A sample of 230 randomized patients was required to obtain 80% power to detect a significant intervention effect on viral load for at least one of three time points with alpha = .05 (two-sided), using a Bonferroni correction. The sample size calculation was conservatively based on a dichotomous outcome variable (detectable versus non-detectable viral load) using the following assumptions: (a) 22 nurses deliver the AIMS intervention and treatment-as-usual, (b) a nurse recruits on average 11 patients for the trial, (c) a maximum dropout rate of 10%, (d) 20% of treatment-experienced patients and all treatment-initiating patients have a detectable viral load at baseline, and (e) depending on the nurse, (i) 60% to 80% of the patients receiving TAU achieve an undetectable viral load during follow-up, and (ii) the percentage of undetectable viral loads increases for patients in the intervention condition by 5 to 20 percentage points (12.5% on average, based on the effects observed in the RCT) [24]. This sample size computation took into account that baseline viral load would be used as a covariate to enhance power [24], and that a multilevel model with random intercepts and random treatment effects at the nurse level would be used.

Note that a change in inclusion/exclusions criteria shortly before recruitment initiation led to an adjusted sample size computation. The rationale and details have been reported both on clinicaltrials.gov and in the published trial protocol [22].

Part V – Details on statistical models

The primary intent-to-treat analysis used a mixed-effects (multilevel) model [33,34]. A factor for time point (3 levels), group (2 levels), and their interaction were the primary variables of interest. Baseline viral load and the stratification variable (treatment-experienced versus treatment-naïve) were added to the model as covariates; as well as a four-level factor for ethnicity (Caucasian, Sub-Saharan African, Caribbean, and Others patients), as ethnicity was identified as an important prognostic covariate [15, 26]. Viral load values were $\log_{10}$-transformed before the analysis. Values for undetectable viral loads (e.g., <40 copies/ml) were replaced by the corresponding detection limit. To account for the nesting of the three follow-up measurements within patients and the nesting of patients within nurses, random intercepts at the patient and nurse level were added to the model. Since nurses administered both the control and intervention treatment, and may differ in how well they implement the intervention, a random group effect was added at the nurse level (and allowed to be correlated with the random intercepts at the same level). The model was fitted using restricted maximum likelihood (REML) estimation.

The time point (Time 2, 3, and 4), group, and time-by-group interaction (note that this interaction does not test for an intervention effect from baseline to follow-up, but for a group effect over the 3 follow-up time points) were tested with Wald-type F-tests. In the absence of a time-by-group interaction, the overall intervention effect can be estimated by a between-group (marginal) contrast across the three follow-up time points. For easier interpretation, the estimated group difference (with corresponding 95% CI) was exponentiated (with base 10) and therefore reflects the average viral load ratio of the control versus the intervention group. Relationships between the covariates and follow-up viral loads are also expressed as viral load ratios. To examine the consistency of the intervention effect, the random group effect at the nurse level was tested using a likelihood ratio test. Post model fitting checks included examining the size and distribution of the random effects and residuals and checking for autocorrelation in the residuals.

The primary analysis was supplemented by dichotomizing the viral load values into detectable versus undetectable, based on the viral load detection limit used in each respective clinic. A three-level mixed-effects logistic regression model was then fitted to these data with fixed and random effects as described previously, with the addition of the detection limit of the viral load test as an additional covariate, while baseline viral load was replaced by the detection status at baseline (detectable versus undetectable).

In an additional viral load analysis, data across time points were aggregated by determining which patients had detectable viral loads on two consecutive follow-up measurements (based on the detection limit of the viral load test in the respective clinic), which we defined as treatment failure. This outcome was analysed with a two-level mixed-effects logistic regression model with random effects at the nurse level. The same covariates as in the previous model were included. The logistic regression models were fitted using maximum likelihood (ML) estimation and Wald-type chi-square and z-tests were used to test the significance of the fixed effects in the model.

Finally, intervention effects on CD4 cell counts (secondary analysis) were examined using the same model as for the primary $\log_{10}$-transformed viral load analyses, except that baseline viral load was replaced by the baseline CD4 cell count.

Based on the fitted models, marginal estimates of the group-specific means (viral load and CD4 analyses) and risks (detectable viral load and treatment failure analyses) were obtained, using the median value at baseline for continuous covariates (i.e., baseline viral load and detection limit) and the observed proportions at baseline for categorical covariates (i.e., treatment-experienced versus treatment-naïve, ethnicity, and detection status at baseline).

Analyses were carried out in R (version 3.1.2) using the *nlme* package and Stata (version 13.1) using functions *mixed* and *meqrlogit*. The statistician conducting the analyses was blinded to group assignment.

Part VI – Markov model tables

Table 2: Costs and utilities per health state

Health State		Health care cost	Productivity loss	Transmission costs	Total costs	Utilities
1	CD4 >500; viral load 0-50	€4014	€810	€1470	€6294	0.954
2	CD4 >500; viral load 51-200	€4180	€771	€2490	€7441	0.954
3	CD4 >500; viral load 201-1000	€3770	€55	€4680	€8505	0.954
4	CD4 >500; viral load >1000	€2995	€2140	€15000	€20135	0.954
5	CD4 >201-500; viral load 0-50	€4169	€682	€1470	€6321	0.929
6	CD4 >201-500; viral load 51-200	€4372	€157	€2490	€7019	0.929
7	CD4 >201-500; viral load 201-1000	€3899	€2149	€4680	€10728	0.929
8	CD4 >201-500; viral load >1000	€2913	€2186	€15000	€20099	0.929
9	CD4 >0-200; viral load 0-50	€4647	€0	€1470	€6117	0.863
10	CD4 >0-200; viral load 51-200	€4459	€307	€2490	€7256	0.863
11	CD4 >0-200; viral load 201-1000	€4267	€0	€4680	€8947	0.863
12	CD4 >0-200; viral load >1000	€3576	€1348	€15000	€19924	0.863
13	Dead	€0	€0	€0	€0	0

Note: 6-month intervention costs were based on the costs of training nurses, time of AIMS delivery, and MEMS-caps, and were estimated at €41,50.

Table 3: Transition probabilities general population

Health state	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0,848	0,047	0,011	0,008	0,068	0,005	0,002	0,005	0,001	0,000	0,000	0,000	0,003
2	0,535	0,299	0,050	0,017	0,057	0,020	0,007	0,010	0,001	0,000	0,000	0,001	0,003
3	0,319	0,141	0,373	0,066	0,024	0,014	0,034	0,024	0,000	0,000	0,001	0,001	0,004
4	0,106	0,032	0,050	0,584	0,019	0,007	0,012	0,180	0,000	0,000	0,001	0,006	0,003
5	0,167	0,014	0,003	0,002	0,702	0,053	0,015	0,014	0,016	0,002	0,001	0,006	0,007
6	0,116	0,044	0,009	0,003	0,415	0,292	0,053	0,031	0,012	0,006	0,004	0,008	0,007
7	0,069	0,030	0,035	0,007	0,297	0,129	0,295	0,093	0,009	0,008	0,008	0,012	0,008
8	0,052	0,020	0,010	0,054	0,147	0,060	0,063	0,532	0,008	0,003	0,005	0,042	0,005
9	0,012	0,001	0,000	0,000	0,249	0,027	0,009	0,005	0,568	0,048	0,019	0,032	0,030
10	0,015	0,008	0,000	0,000	0,197	0,075	0,015	0,012	0,309	0,228	0,056	0,065	0,020
11	0,007	0,000	0,002	0,002	0,106	0,037	0,042	0,010	0,214	0,120	0,295	0,155	0,010
12	0,006	0,000	0,001	0,001	0,076	0,042	0,029	0,042	0,110	0,061	0,063	0,540	0,030
13	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	1,000

Table 4: Relative risks AIMS vs. TAU

Health state	1	2	3	4	5	6	7	8	9	10	11	12	13
1	1,156*	0,718*	1,077***		0,249*	0,538***							1,077***
2	0,882*	1,100*			2,000***								
3	2,000***												
4	1,364*	0,545***											
5	0,806*	0,627**			0,940*	1,175**							
6	1,078**	0,288**			1,438*	1,150**	1,438***						
7					4,714**								
8	1,203*	2,647***			1,544*	0,147**	0,882***						
9					1,818**				1,091**				
10									2,000***				
11													
12					0,923*	1,846*			0,923**				
13	0,000*	0,000*	0,000*	0,000*	0,000*	0,000*	0,000*	0,000*	0,000*	0,000*	0,000*	0,000*	1,000*

* Used for all scenarios; ** Only used for scenario 1 and 2; *** Only used for scenario 2

Part VII – Results including covariates and subgroup analyses

Primary Analysis

The mixed-effects multilevel model showed that the intervention was effective across the three time points ($F(1, 196) = 6.40$, $p = .012$). Overall, patients in the control group had viral loads that were on average 1.26 times (95% CI: 1.04 to 1.52) higher than those in the intervention group during follow-up. There was no indication of a change in the intervention effect across the three follow-up time points (time-by-group interaction ($F(2,409) = 0.75$, $p = .47$) or an overall effect of time regardless of group ($F(2, 409) = 0.62$, $p = .54$). There was no significant variability of the treatment effect across nurses ($p = .14$). Estimated marginal viral load was 35.44 copies/ml (95% CI: 29.91 to 42.00) in the intervention group and 44.53 copies/ml (95% CI: 35.47 to 55.89) in the control group.

Viral loads at follow-up were significantly related to baseline viral load, with a one-point increase in $\log_{10}$ baseline viral load (e.g., from 100 to 1000 copies/ml) leading on average to 1.33 times higher viral loads at follow-up (95% CI: 1.19 to 1.48). Viral loads for treatment-experienced patients were on average 2.49 times higher than those of treatment naive patients (95% CI: 1.74 to 3.56). Ethnicity just failed to be a significant predictor when considering the factor as a whole ($F(3,196) = 2.46$, $p = .064$). Nevertheless, compared to Caucasian patients, viral loads were 1.26 times higher for patients with a Caribbean (95% CI: 0.97 to 1.64) and 1.32 times higher for patients with an African ethnicity (95% CI: 1.02 to 1.69). On the other hand, patients with some other ethnic identity had on average lower viral loads, but due to the relatively small size of this group, the viral load ratio (i.e., 0.88) was not estimated precisely (95% CI: 0.59 to 1.31).

Post model fitting checks revealed that a small number of the viral load measurements (11 out of 634) were quite high (above 1,000 copies/ml, with a maximum of 225,014) and led to some noteworthy outliers and an extremely large correlation between the random intercepts and group effects at the nurse level ($r = .93$). Censoring these values at 1,000 copies/ml resolved these issues, but did not alter any of the previous conclusions, except that the ethnicity factor then became significant ($F(3,196) = 3.61$, $p = .014$).

Post-hoc Analyses

The three-level mixed-effects logistic regression model, in which the viral load values were dichotomized into detectable versus undetectable viral loads, generally confirmed the previous results, although the group effect just failed to be significant at α = .05 (χ^2(*df* = 1) = 3.66, *p* = .056). The time-by-group interaction (χ^2(*df* = 2) = 1.38, *p* = .50) and the time effect (χ^2(*df* = 2) = 4.81, *p* = .09) were again not significant. However, the close to significant time effect suggests a possible decrease in the proportion of patients with a detectable viral load over the three follow-up time points, irrespective of group. Overall, patients in the control group had 1.89 times higher odds of having a detectable viral load across the three time points (95% CI: 0.98 to 3.65). The nurse level variability in the treatment effects was again not significant (*p* = .99). Estimated marginal risks of a detectable viral load were 9.6% (95% CI: 3.8% to 15.4%) and 16.7% (95% CI: 8.2% to 25.3%) in the intervention and control group, respectively.

Those with a detectable viral load at baseline had 7.67 times higher odds of also having a detectable viral load at follow-up (95% CI: 3.10 to 18.98). Moreover, the odds of a detectable viral load were 7.47 times higher in treatment-experienced versus treatment-naive patients (95% CI: 3.02 to 18.49). Not surprisingly, the odds of a detectable viral load at follow-up was related to the detection limit of the test, with 1.58 times higher odds (95% CI: 1.12 to 2.22) for a 10-point decrease in the detection limit (e.g., from 50 to 40 copies/ml). Finally, in comparison to the primary analyses, ethnicity was clearly not significant (χ^2(*df* = 3) = 1.95, p = .58).

Considering the analysis for treatment failure, 19 out of the 112 patients in the control group (17.0%, 95% CI: 10.8% to 25.5%) and 8 out of the 109 intervention patients (7.3%, 95% CI: 3.5% to 14.4%) had two consecutive detectable viral loads post-randomization. The two-level logistic regression model for treatment failure indicated a significant group difference (χ^2(*df* = 1) = 5.61, *p* = .012). The odds of treatment failure were 2.99 times higher in the control group (95% CI: 1.21 to 7.38), corresponding to estimated risks of treatment failure of 9.0% (95% CI: 2.4% to 15.7%) in the intervention and 22.8% (95% CI: 11.7% to 34.0%) in the control group. Again, no heterogeneity in treatment effects was observed across nurses (*p* = 1.0). The odds of treatment failure were 4.17 times higher for those with a detectable viral load at baseline (95% CI: 0.88 to 19.78) and 6.97 times higher for treatment-experienced patients (95% CI: 1.46 to 33.32). Ethnicity was not a significant predictor (χ^2(*df* = 1) = 1.27, p = .26). Note that ethnicity was dichotomized in this model (Caucasian versus other) to avoid perfect separation (there were no treatment failures in the 11 patients falling into the ‘other’ category).

The mixed-effects multilevel model for CD4 cell-count showed that the intervention was not effective across the three time points ($F(1, 196) = 2.38$, $p = .12$). However, there was a significant time-by-group interaction ($F(2,398) = 3.09$, $p = 0.047$), so that the effect of group cannot be examined with a marginal contrast across the 3 follow-up time points. The per time point analysis suggest an initial advantage of being in the intervention group at Time 1 (mean difference 31.00, 95% CI: -8.37 to 70.37), after which the control group catches up at Time 2 (mean difference 6.55, 95% CI: -46.03 to 32.92), but then the CD4 cell count in the control group remains stable while it further increases in the intervention group, leading to a significant difference at Time point 3 (mean difference 39.39, 95% CI: 0.10 to 78.67). See Figure below, with the reference group being treatment-experienced, Caucasian patients with a median CD4 cell count (i.e., 0) at baseline. Estimated marginal means were 550.89 cells/mm^3 (95%CI: 520.40 to 581.39), 562.46 cells/mm^3 (95%CI: 531.65 to 593.28), and 597.79 cells/mm^3 (95%CI: 567.07 to 628.52) in the intervention group across the three follow-up points, compared to 519.90 cells/mm^3 (95%CI: 489.31 to 550.49), 569.02 cells/mm^3 (95%CI: 538.69 to 599.35), and 558.41 cells/mm^3 (95%CI: 528.21 to 588.61) in the control group.

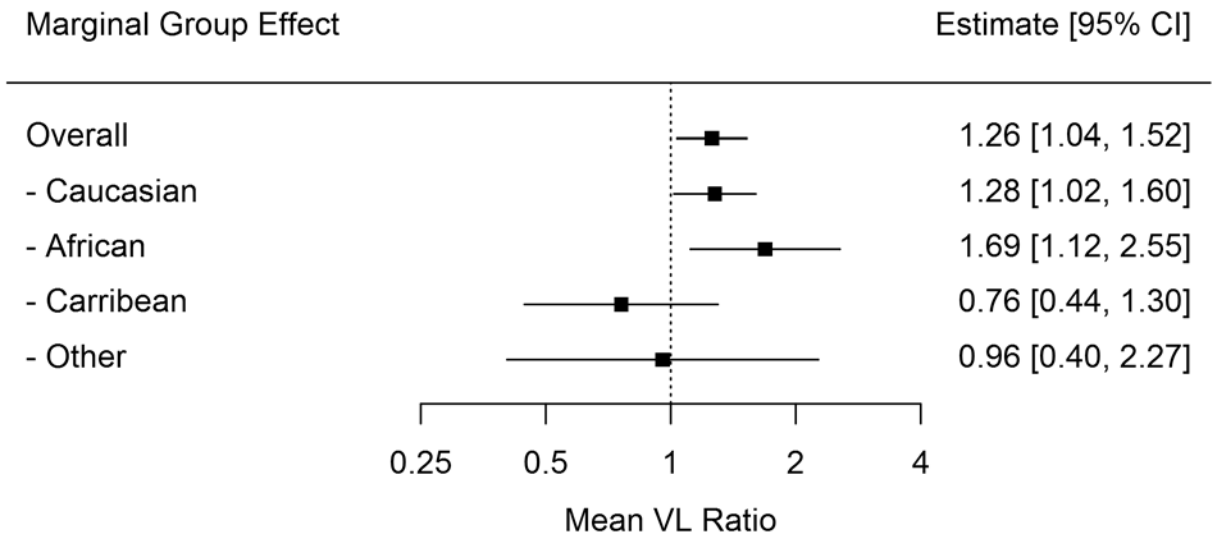
The model also revealed a main effect of time ($F(2,398) = 6.03$, $p = .003$). Both a higher CD4 cell count at baseline (F (1,196) = 470.55, p <.0001) and being treatment-naïve at the start of the study (F (1,196) = 27.59, p <.0001) were strong predictors of CD4 during follow-up. Ethnicity did not predict CD4 values during follow-up (F (3,196) = 1.89, p = .13).

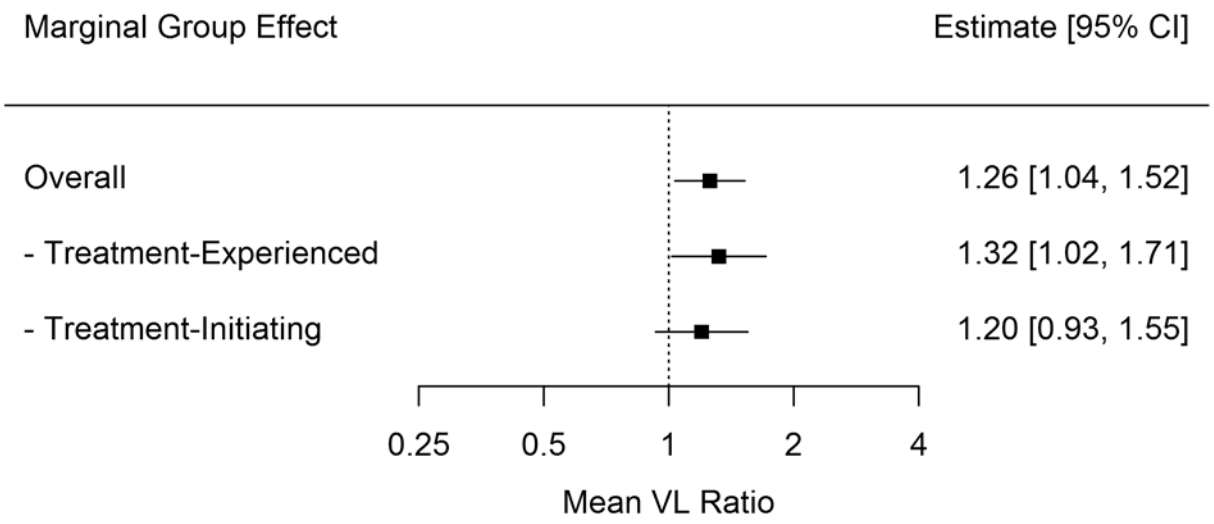
Subgroup analyses for Mean Viral Load

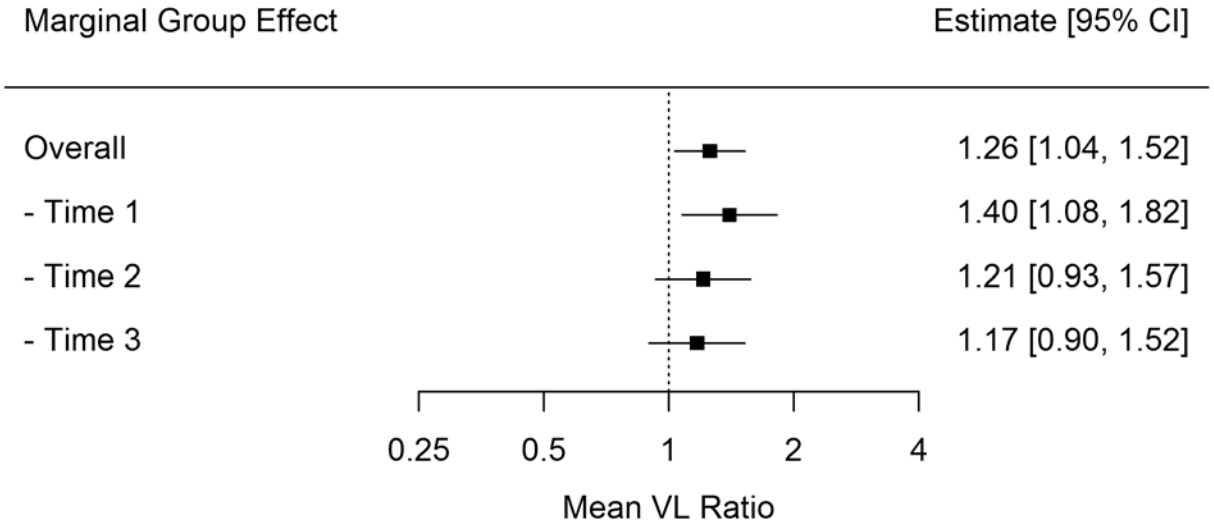
Although the study was not powered to detect a treatment effect for subgroups or per time point, visual inspection of the marginal effects can reveal relevant trends. Figure 1a shows that the effects for treatment naïve and treatment experienced patients are very similar. Figure 1b suggests that Caucasian and sub-Saharan African patients benefit most from the intervention, whereas patients with a Caribbean or Other background do not (these latter groups were small and had high viral suppression rates throughout the trial in both arms). Figure 1c suggests that the intervention has a strong initial effect on viral load, that reduces at follow-up 2 (Time point 3) and then is sustained at follow-up 3 (Time point 4).

INSERT FIGURES 1-A-C ABOUT HERE

Figures 1a-c: Forest plots of marginal group effects for treatment experience, ethnicity, and per time point.




Marginal Group Effect
Estimate [95% CI]
Overall
1.26 [1.04, 1.52]
- Treatment-Experienced
1.32 [1.02, 1.71]
- Treatment-Initiating
1.20 [0.93, 1.55]
0.25
0.5
1
2
4
Mean VL Ratio


Marginal Group Effect
Estimate [95% CI]
Overall
1.26 [1.04, 1.52]
- Time 1
1.40 [1.08, 1.82]
- Time 2
1.21 [0.93, 1.57]
- Time 3
1.17 [0.90, 1.52]
0.25
0.5
1
2
4
Mean VL Ratio