Randomized, Placebo-Controlled Trial or Post Hoc Subgroup Analysis: The Importance of Standardized and Comprehensive Reporting

To the Editor—In a recent issue of this journal, Hatano et al [1] presented results of a study investigating the effect of raltegravir intensification on 2–long terminal repeat (2-LTR) circles as a proxy for low-level viral replication in treated patients with human immunodeficiency virus (HIV) infection [1]. Although the authors tackle an important issue in HIV research, we have serious concerns about a possible misinterpretation of the level of evidence associated with this study.

The study is presented as a randomized placebo-controlled trial, but this characterization is not supported by the information provided by the authors. After reviewing the clinical trial registration number given in the article (NCT00843713), we believe it is obvious that the presented study (including 31 participants) differs in sample size from the study registered, which was published in 2012 (including 56 participants) [2]. This leads to the assumption that the present study is a non-specified subgroup analysis of the original randomized controlled trial (RCT). The criteria used to select this subgroup from the original study population (HIV RNA level <40 copies/mL and CD4+ T-cell count ≥350 cells/µL for ≥1 year) are presented in the article as inclusion criteria before randomization. This hides from the reader how the process of randomization was embedded in the design of the present study.

For a number of reasons, post hoc subgroup analyses of RCTs generally do not fulfill the criteria of an RCT, as described by Pocock et al [3]. Because “randomized controlled trials are . . . regarded as the most scientifically rigorous method of hypothesis testing” and are often referred to in the development of clinical guidelines, this term should be used with great care [4]. Therefore, we suggest changing the study description to a “post-hoc subgroup analysis of data from a randomized controlled trial” and ask Hatano et al to interpret the results accordingly.

Many factors that may negatively affect the level of evidence in subgroup analyses are present in this study [1, 3]. We are concerned that the study groups might not have been balanced at baseline. Although they seem balanced with respect to numbers of patients and some of the baseline characteristics, there is a relevant difference between groups in duration since diagnosis of HIV infection and duration of highly active antiretroviral therapy suppression that was not present in the baseline measurements of the original trial [2]. The interpretation of the corresponding nonsignificant P values should take into account the low power they are based on (n = 31). Moreover, a large interindividual heterogeneity in 2-LTR values at baseline was shown in the article’s Figure 1 [1, p. 1439]. Only 2 individuals in the treatment group but 7 in the control group have a 2-LTR value >0 at baseline, so the treatment groups were not balanced with respect to the outcome variable. Because 2-LTR levels of 0 can only increase or remain at 0, where higher levels could also decrease, this imbalance between study groups, if not accounted for, might have biased the estimates of the presented effect. Hatano et al should report whether and how this issue has been addressed in the analysis.

Several other points discussed by Pocock et al [3] could not be appraised because they are not reported appropriately by Hatano et al [3]. For example, there is a lack of information on the handling of missing data (with 3 missing values in the placebo group and 1 in the treatment group) and the maintenance of observer blinding with regard to already published results [2, 5, 6]. Moreover, the article includes no information about power or sample size calculation. The study may have been powered for the outcomes reported in the main analysis, which should be stated in the methods.

To ensure that readers can correctly evaluate the quality of RCTs, we believe it is important to report studies according to the CONSORT guidelines, as advised in the “instructions to authors” section of this journal’s Web site [5, 6]. Unfortunately, these guidelines were not rigorously followed for the current study.

Notes

Financial support. This work was supported by the German Centre for Infection Research (Deutsches Zentrum für Infektionsforschung: DZIF; clinical leave stipend to B.K.).

Potential conflicts of interest. All authors: No reported conflicts.

All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

André Karch,1 Ralf Krumkamp,2 and Benno Kreuels2,3

1Research Group Epidemiological and Statistical Methods (ESME), Helmholtz Center for Infection Research, Braunschweig. 2Research Group Infectious Disease Epidemiology, Bernhard Nocht Institute for Tropical Medicine, Hamburg, and 3Division of Tropical Medicine, I. Department of Internal Medicine, University Medical Centre Hamburg Eppendorf, Hamburg, Germany

References


Reply to Karch et al

To the Editor—We appreciate the interest by Karch et al [1] in our study of raltegravir intensification in treated human immunodeficiency virus (HIV)–infected patients [2]. Our study showed that intensification resulted in a rapid and transient increase in the level of 2–long terminal repeat (2-LTR) circles in a proportion of treated HIV-infected subjects, suggesting that low-level viral replication persists in some individuals even after long-term antiretroviral therapy. Intensification also reduced the level of D-dimer, a coagulation biomarker that has been shown to be predictive of morbidity and mortality among treated HIV-infected individuals [3].

We performed a randomized, double-blind, placebo-controlled trial as described [2], making it the first such study to show that intensification led to an increase in the level of 2-LTR circles and a decrease in D-dimer. Our findings were consistent with those of a previous open-label, randomized study of raltegravir intensification, which showed that intensification led to an increase in the level of 2-LTR circles [4].

Karch et al [1] correctly noted a lack of clarity in our clinicaltrials.gov entry; we have edited it to clarify the study structure. We performed 2 independent, randomized, double-blind, placebo-controlled intensification studies (one in “immunologic nonresponders” [n = 30] [5] and the other in “immunologic responders” [n = 31] [2]). All subjects from both studies were offered coenrollment into a cardiovascular study before randomization to evaluate whether raltegravir intensification led to an improvement in endothelial function; the cardiovascular outcomes from both intensification studies have been published elsewhere [6].

The impact of raltegravir intensification on the dynamics of 2-LTR circles has been controversial. Other intensification studies that measured 2-LTR circles at later time points have failed to show changes in 2-LTR circles, perhaps because they were measured too late after intensification was initiated (at 4 weeks [7] and 12 weeks [8] after intensification). Given the discrepant results of these previous studies [4, 7, 8], our second intensification study in immunologic responders was specifically designed to obtain samples from very early time points (weeks 1 and 2) [2]; our first intensification study in immunologic nonresponders had already begun enrollment, so 2-LTR data were not collected or analyzed for that study [5]. The 2-LTR data in our recent article [2] were, therefore, obtained only from our second intensification study in immunologic responders and hence do not represent a “nonspecified subgroup.”

Our most recently published raltegravir intensification study was an independent, randomized, double-blind, placebo-controlled study, with enrollees required to meet inclusion and exclusion criteria, as described before randomization [2]. Thus, any real or perceived imbalances between the raltegravir and placebo groups at baseline were due to random chance. The primary virologic end point of the study was an increase in 2-LTR circles at week 1 or 2. This end point was determined a priori (ie, during study design and before any patient enrollment occurred), and specimens were, accordingly and specifically, obtained at early time points (weeks 0, 1, 2, and 8) for measurement of potential changes in 2-LTR circles. All study samples (including those for 2-LTR analysis) were collected and analyzed, and all statistical analyses were completed, before the study code was unlocked. We analyzed all of the 2-LTR data, as shown in our table and figures [2], without imputing values for missing data. Power is widely acknowledged to be irrelevant for interpreting completed studies; the estimates and confidence intervals given in our article permit much more direct and reliable interpretation of our study results than any reasoning involving power calculations [9–11]. The findings of our randomized, double-blind, placebo-controlled study suggest that low-level viral replication may persist in the setting of otherwise effective antiretroviral therapy and may be a modifiable factor in future cure and treatment strategies.

Notes

Financial support. This work was supported by the National Institute of Allergy and Infectious Diseases (grants R01 AI087145, K24 AI069994, AI080193, AI69432, AI047745, and AI174621), the Delaney AIDS Research Enterprise (DARE; grant U19 AI0961090), the Collaboratory of AIDS Researchers for Eradication (CARE; grant U19 AI0966113), the National Heart Lung and Blood Institute (grant R01 HL095130), the American Foundation for AIDS Research (grant 106710–40–RGRL), the University of California, San Francisco/Gladstone Institute of Virology & Immunology Center for AIDS Research (CFAR; grant P30 AI027763), the University of California, San Diego CFAR (grant AI306214), the University of California, San Francisco Clinical and Translational Research Institute Clinical Research Center (grant UL1 RR024131), the Center for AIDS Prevention Studies (grant P30 MH62246), the CFAR Network of Integrated Systems (grant R24 AI067039), and the Department of Veterans Affairs.

Potential conflicts of interest. H. H. and S. G. D. have received research support from Merck. The study drug was provided by Merck.