Drug Level Monitoring in a Double-Blind Multicenter Trial: False-Positive Zidovudine Measurements in AIDS Clinical Trials Group Protocol 019

DONALD J. KROGSTAD,1†† MICHAEL R. EVEKLAND,1 LYNETTE L.-Y. LIM,2 PAUL A. VOLBERDING,3,4 AND BRIAN M. SADLER5

Departments of Medicine and Pathology, Washington University School of Medicine, St. Louis, Missouri 63110; Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts 02115; University of California, San Francisco, San Francisco, California 94110; San Francisco General Hospital, San Francisco, California 94110; and Research Triangle Institute, Research Triangle Park, North Carolina 27709

Received 10 September 1990/Accepted 26 March 1991

Twenty-three different laboratories using four different assay methods reported zidovudine (ZDV; azidothymidine) measurements in a double-blind trial of ZDV for asymptomatic human immunodeficiency virus-infected patients (AIDS Clinical Trials Group Protocol 019). The risk of false-positive ZDV measurements was defined with coded specimens containing no ZDV in a quality control testing program. This testing identified six problem laboratories which reported ZDV levels of ≥100 ng/ml for specimens with no ZDV; all of these laboratories used high-performance liquid chromatography. Six laboratories reported a disproportionately high fraction of positive assays for subjects randomized to the placebo group (31% for these 6 laboratories versus 4% for the other 17 laboratories; P < 0.0001). The high number of false-positive ZDV results reported by these six laboratories suggested that many of the positive results that they reported for patient specimens were also false-positive results. This hypothesis was examined by retesting specimens from patients in the placebo group that had been reported as positive by these laboratories. Ninety percent (19 of 21) of these specimens were negative on retesting at the reference laboratory. These results confirm the hypothesis; they demonstrate the need for quality control testing to avoid the misinterpretation of multicenter trials because of incorrect laboratory data.

The interpretation of quantitative laboratory data may be confounded by at least four sources of error: false-positives, false-negatives, imprecision, and inaccuracy. Each of these affects the quality and interpretation of drug level data in clinical trials (11). False-positive results produce overestimates of surreptitious drug use among subjects randomized to placebo groups and overestimates of compliance among subjects randomized to treatment groups. Conversely, false-negative results produce underestimates of surreptitious drug use among subjects randomized to placebo groups and underestimates of compliance among subjects randomized to treatment groups. Imprecision (unreliability) and inaccuracy confound efforts to define level-related toxicity or efficacy.

These problems are compounded in multicenter clinical trials. Although many clinical trials have used drug assays to estimate compliance (1, 2, 9), there are few or no previous analyses of these issues or of their impact on the interpretation of clinical trials. These sources of error may be reduced by the use of uniform commercially available reagents (10) or by reliance on a single reference laboratory (2, 5, 9, 14). However, neither of these strategies was feasible when the first clinical trials of zidovudine (ZDV) began. At that time, there were no commercially available reagents or assays to measure ZDV, no high-performance liquid chromatography (HPLC) methods had been published, and no one laboratory within the AIDS Clinical Trials Group (ACTG) had sufficient staff and equipment to handle the volume of drug level determinations required for multicenter clinical trials with several thousand patients on various doses of ZDV.

To evaluate ZDV concentration data from the 23 different ACTG pharmacology laboratories, a quality testing program was established. The results of this testing were used to compare performance among laboratories and to identify laboratories likely to have increased prevalences of false-positive or false-negative results or reduced precision or accuracy. This paper describes the use of quality control testing to identify laboratories likely to report false-positive ZDV results for patient specimens, the assay methods associated with false-positive results, and the effect of those results on estimates of compliance in ACTG Protocol 019 (13).

(This research was presented in part at the Eighth Meeting of the AIDS Clinical Trials Group, Bethesda, Md., March 1990.)

MATERIALS AND METHODS

Quality control testing program. Quality control testing for drug level measurements was organized by the Pharmacology Committee of the ACTG. Since July 1988, a set of six quality control specimens has been sent to participating laboratories 1 to 2 months before each meeting so that the results of a new round of testing can be presented and discussed at every ACTG meeting (three times per year). Results are presented in both tabular (Table 1) and graphic (Fig. 1) form. Graphs are coded so that differences in results associated with specific methods are immediately obvious (Fig. 1).

The specimens used for the first round of testing were
kindly provided by Steven Good of Burroughs-Wellcome and had been prepared by adding ZDV and zidovudine glucuronide (GZDV) to bovine serum. Beginning with round 2 in March 1989, human serum was used. Beginning with round 4 in October 1989, serum from human immunodeficiency virus (HIV)-infected human subjects receiving ZDV was used. These sera contained both ZDV and GZDV. Quality control specimens are prepared by mixing patient specimens containing known ZDV concentrations to produce pooled sera with the desired ZDV concentrations. After preparation, each sample is tested by HPLC, radioimmunoassay (RIA), and fluorescence polarization immunoassay (FPIA) at the reference laboratory to ensure that the results obtained are consistent with all three methods. As an additional control, the reference laboratory uses spectroscopy at 267 nm to define ZDV concentrations with the molar extinction coefficient (ε = 7,750) (3). The target values generated by these procedures are consistently within 5 to 10% of the median values obtained by the participating laboratories. To prevent deterioration of the samples during shipping, each set of serum specimens is packed in 3 to 4 lb (ca. 1.4 to 1.8 kg) of dry ice and sent to participating laboratories by overnight express delivery on Monday or Tuesday of a week without holidays.

The following persons and institutions are members of the Pharmacology Committee of the ACTG or participants in the quality control testing program: Columbia University College of Physicians and Surgeons—Thomas Cooper and R. F. Suckow; Harvard Medical School—Donald W. Kufe and David A. Schoenfeld; Hershey Medical Center—Daike Li; Indiana University—D. Craig Brater; Johns Hopkins University—David M. Kornhauser and Paul S. Littman; Memorial Sloan-Kettering Cancer Center—William Tong; Mount Sinai School of Medicine—John Roboz; National Institutes of Health (Division of AIDS, National Institute of Allergy and Infectious Diseases)—Richard Hafner and Jonathan M. Kagan; New York University—Howard Hochster and Leonard Liebes; Stanford University—Trence F. Blachek; Tulane University—William J. George, June Z. Lertora, C. D. Arvind B. Rege; University of California, San Diego—James D. Connor; University of California, San Francisco—John Gambertoglio; University of Cincinnati—Amadeo Pesce and Brian Stretchers; University of Massachusetts, Worcester—Sarah Cheeseman; University of Miami—Ram Agarwal; University of Minnesota—Carolyn Beatty, Henry H. Balfour, Jr., and Courtney V. Fletcher; University of Pittsburgh—Richard Ptachinski; University of Rochester—Gene D. Morse (State University of New York, Buffalo); University of Southern California—Vasilios I. Avramis; University of Washington—Kent Opheim and Teresa Tartaglione; George Washington University Medical Center—Ti Li Loo; and Washington University—Donald J. Krogstad.

**Heat inactivation.** When quality control testing was begun, it was not known whether heating to inactivate HIV (56°C for 120 min) (8) would interfere with the ZDV assays. This question was examined in the first and second rounds of quality control testing.

**Testing methods.** The 23 ACTG laboratories participating in quality control testing used four different assays to measure ZDV: RIA, FPIA, and fluorescence immunoassay (FIA) (Table 2).

**HPLC.** Fifteen laboratories measured ZDV by HPLC using protocols similar or identical to the original procedure of Good et al. (3). Most (12 of 15) of the laboratories used an isocratic elution procedure with acetonitrile in ammonium phosphate (20 mM, pH 3.0). Most (11 of 15) of the laboratories also used the optical isomer of ZDV (A22U) as the internal standard and measured ZDV by comparing its UV absorbance at 267 nm with that of the internal standard. In this assay, GZDV elutes earlier than ZDV and thus may be confused with nonspecific serum peaks, especially in patients with renal or hepatic failure.

**RIA.** Five laboratories measured ZDV by RIA using commercially available reagents (Incstar, Stillwater, Minn.) (12). In this procedure, 200 μl of a diluted patient serum sample (diluted 1:21 in buffer) is mixed with 100 μl of 125I-labeled ZDV and 100 μl of rabbit antiserum to ZDV, and the mixture is incubated at room temperature for 2 h. Next, 500 μl of goat antirabbit serum with polyethylene glycol is added, and the tubes are mixed and incubated at room temperature for 30 min prior to centrifugation and counting of the pellet. Cross-reactivity with GZDV is <0.01%. The GZDV concentration may be determined by performing the RIA before and after treatment with β-glucuronidase (4). The β-glucuronidase procedure was used to measure GZDV

**TABLE 1.** Effect of adjusting for false-positive ZDV determinations on the variability of the data

| Specimen | ZDV concn (ng/ml) | Target | Group mean ± SD (n = 17) | Adjusted (n = 14) |
|----------|------------------|--------|------------------------|------------------|
| A        | 0                | 12 ± 34 (287) | 0.8 ± 2.3 (287)   |                  |
| B        | 10               | 126 ± 401 (319) | 16.3 ± 12.8 (79) |                  |
| C        | 30               | 147 ± 415 (282) | 28.6 ± 11.8 (41) |                  |
| D        | 300              | 343 ± 204 (60)  | 313.3 ± 57.7 (18) |                  |
| E        | 40               | 74 ± 162 (219)  | 36.9 ± 16.0 (43)  |                  |
| F        | 40               | 88 ± 214 (243)  | 38.5 ± 17.1 (44)  |                  |

*a Numbers in parentheses are coefficients of variation (percentages). Specimens A to D were mixtures of sera from HIV-infected patients taking ZDV. Specimens E and F were produced by a 1:10 dilution of ZDV-containing serum with serum from renal or hepatic failure patients, respectively. When data from three laboratories with unacceptable results were deleted, the group means were closer to the targets and the coefficients of variation were markedly reduced (adjusted). Unacceptable results were defined as false-positive ZDV concentrations of ≥100 ng/ml for blank specimens and concentrations more than twice the targets for three or more of the six specimens.
TABLE 2. ZDV levels reported for a blank serum specimen

| Laboratory | Method | ZDV (ng/ml) |
|------------|--------|-------------|
| A          | HPLC   | 0           |
| B          | HPLC   | 4           |
| C          | HPLC   | 5           |
| D          | HPLC   | 20          |
| E          | HPLC   | 21          |
| F          | HPLC   | 30          |
| G          | HPLC   | 33          |
| H          | HPLC   | 60          |
| I          | HPLC   | 87          |
| J          | HPLC   | 100         |
| K          | HPLC   | 110         |
| L          | HPLC   | 125         |
| M          | HPLC   | 140         |
| N          | HPLC   | 217         |
| O          | HPLC   | 350         |
| P          | RIA    | 0           |
| Q          | RIA    | 4           |
| R          | RIA    | 6           |
| S          | RIA    | 6           |
| T          | FPIA   | 3           |
| U          | FPIA   | 5           |

* Four of the five laboratories using RIA and two of the three laboratories using FPIA participated in this round of quality control testing. The laboratory that used FIA used HPLC in this round of quality control testing. One laboratory reported both HPLC and RIA results (accounting for 24 sets of results from 23 laboratories). Carryover from specimens with high ZDV concentrations has now been shown to cause low-level false-positive results with FPIA (7).

by the reference laboratory and by three of the five laboratories using the RIA.

**FPIA.** Three laboratories measured ZDV by FPIA using commercially available reagents (Sigma, St. Louis, Mo.). In this assay, the protein in the serum specimen is first precipitated with an acidified methanol extraction reagent (1:1.1). The treated serum specimen (15 or 160 μl) is then mixed with rabbit antibody to ZDV (25 μl), fluorescein-labelled ZDV (25 μl), and Tris base (25 μl of 1 M). The ZDV concentrations of unknown specimens are determined by comparing the fluorescence polarization of unknown specimens with that of known controls by using excitation and emission wavelengths of 485 and 525 nm (4). GZDV may be determined by subtracting the ZDV measurement obtained prior to treatment with β-glucuronidase from that obtained after treatment with β-glucuronidase (4). This procedure was performed by the reference laboratory and by two of the three laboratories using the FPIA.

**FIA.** One laboratory measured ZDV using an FIA performed with the FPIA reagents. In this assay, serum (100 μl) is mixed with fourfold dilutions of rabbit antibody to ZDV (100 μl) and of fluorescein-labelled ZDV (100 μl) in assay buffer (0.1 M sodium phosphate [pH 7.4] with 0.01% bovine gamma globulin and 0.1% NaCl). After incubation at 37°C for 60 min, goat antirabbit serum (50 μl) is added. After an additional 30 min of incubation at room temperature, 2 ml of 3% polyethylene glycol is added and the specimen is centrifuged at 2,000 × g and 4°C for 15 min. After the supernatant is aspirated, 2% sodium dodecyl sulfate in 0.1 N NaOH is added, the tubes are mixed by vortexing, and fluorescence is measured after 6 to 24 h of incubation in the dark at room temperature by using excitation and emission wavelengths of 494 and 520 nm, respectively.

ACTG Protocol 019. ACTG Protocol 019 was a double-blind study of ZDV in patients with asymptomatic HIV infections. Participants were randomized to placebo, low-dose ZDV (500 mg/day), or high-dose ZDV (1,500 mg/day) groups. They were then monitored prospectively for neutropenia, anemia, CD4-positive cell counts, p24 antigen levels, and opportunistic infections or other evidence of progression to AIDS-related complex or AIDS (13).

**RESULTS**

Heat inactivation. Results from the first two rounds of testing established that heat inactivation did not affect the measurement of ZDV by HPLC, RIA, or FPIA. The mean ZDV values reported by the 8 laboratories that did not heat inactivate were not different from those of the 13 laboratories that did (P > 0.30). No information is available for the FIA. Heat inactivation prior to testing for ZDV has been recommended for all patient specimens since that time (February 1989).

**Blank specimens containing no ZDV.** Blank specimens containing no ZDV were included in rounds 3 to 5 to estimate the prevalence of false-positive results among participating laboratories. The highest false-positive ZDV concentrations were reported by laboratories using HPLC. Six of the 15 laboratories using HPLC reported false-positive ZDV concentrations of ≥100 ng/ml versus none of 8 laboratories using either RIA or FPIA (P = 0.0496; Fisher’s exact test). The results reported for these samples in round 3 (when patient serum samples were still being analyzed for ZDV in ACTG Protocol 019; Table 2) demonstrated that the ZDV concentrations reported for blank specimens varied widely among laboratories.

**False-positives among patients randomized to the placebo group.** If false-positive ZDV results for blank quality control specimens (Table 2) are associated with false-positive assay results for patient specimens, the prevalence of reportedly positive ZDV assay results for the placebo group should be greater in laboratories with high false-positive ZDV results in quality control testing. This hypothesis was tested and confirmed with data from ACTG Protocol 019. The prevalence of reportedly positive ZDV assay results among patients in the placebo group was eight- to ninefold greater among laboratories that reported ZDV concentrations of ≥100 ng/ml for the blank quality control specimens in round 3 in July 1989 (23 of 75 versus 4 of 108; P < 0.0001).

**GZDV.** Similar problems were also observed with GZDV. High false-positive GZDV determinations were concentrated among laboratories that used HPLC. However, GZDV levels in patient specimens were not reported to the ACTG Data Base. Therefore, it was not possible to examine the effect of false-positive GZDV determinations on the prevalence of reportedly positive GZDV assay results for patient specimens.

**Identification of potential problem laboratories and initial assessment of their impact on the ACTG Data Base.** Six laboratories reported false-positive ZDV concentrations of ≥100 ng/ml (0.375 μM) for blank quality control specimens. ZDV concentrations of 10 to 100 ng/ml reported by these laboratories were both uninterpretable (because these laboratories reported false-positive concentrations of ≥100 ng/ml) and within the expected range of levels in serum (because of the short elimination half-life of ZDV) (6). Therefore, the prevalence of positive ZDV assay results in ACTG Protocol 019 was recalculated after deletion of the results from these six problem laboratories (Table 3). After
this adjustment, the prevalence of positive ZDV assay results for the placebo group was reduced from 15 to 4%.

Retesting of specimens reported as positive for ZDV by problem laboratories. To determine whether these six laboratories also reported false-positive results for patient specimens, 21 specimens from patients in the placebo group that had been reported as positive for ZDV were retested at the reference laboratory. Nineteen (90%) of these 21 specimens were found to contain no ZDV.

DISCUSSION

Blank quality control specimens and the risk of false-positive ZDV results for patient specimens. Blank quality control specimens are a straightforward way to estimate the prevalence of false-positive results for each laboratory. The significance of these observations is that they suggest that laboratories that have false-positive ZDV determinations for quality control specimens have increased prevalences of false-positive ZDV results for patient specimens. Thus, unadjusted drug level data substantially overestimate the surreptitious use of ZDV among subjects randomized to the placebo group of ACTG Protocol 019. False-positive ZDV results for patients randomized to the treatment group are also affected (Table 3) and produce an overestimate of compliance. Rigorous analysis of those results is beyond the scope of this paper because it requires an additional adjustment for false-negative ZDV determinations resulting from the rapid elimination half-life of ZDV (6).

The relationship between false-positive results for blank quality control specimens and the prevalence of reportedly positive assay results for the placebo group supports the findings of the quality control testing program. To our knowledge, this is the first time that the validity of any quality control testing program has been tested in this explicit fashion.

Retesting of specimens reported as positive by problem laboratories. The retesting of reportedly positive specimens from the placebo group indicated that many (≥90%) of these specimens represented false-positive assays. Thus, these data suggested that the prevalence of surreptitious ZDV ingestion in the placebo group in ACTG Protocol 019 was actually ≤3 to 4% (not 31%).

Methodological differences. With ZDV concentrations of less than 50 ng/ml, a number of the laboratories using HPLC had substantial difficulty (Table 1 and Fig. 1). However, this was not a universal phenomenon. Three laboratories produced excellent results (consistently within 1 standard deviation of both the group means and target values) using HPLC alone. On the basis of a survey of the participating laboratories (data not shown), there were no differences in extraction procedures, columns, mobile or elution phases, or internal standards between laboratories that experienced problems with HPLC and those that did not. In contrast, the eight laboratories using immunoassays (five using RIA and three using FPIA) produced results that were all within 1 standard deviation of the group means and target values at both low and high ZDV concentrations. These data suggest that the commercially available immunoassays (RIA and FPIA) produce consistently accurate results more readily than does HPLC.

Drug level data and the interpretation of multicenter clinical trials. The impact of false-positive assay results for quality control specimens on the estimate of surreptitious drug use among the placebo group in ACTG Protocol 019 demonstrates that multicenter clinical trials must have quality control programs to ensure the validity of laboratory data and clinical conclusions.

However, the most important aspect of these studies is their potential applicability to trials involving other diseases and agents. A lack of quality control testing would have similar effects on the interpretation of trials involving antihypertensive, antiarrhythmic, or antineoplastic agents. In addition, the same considerations pertain to the selection of an assay. Assay data used to estimate compliance must be obtained by a sensitive method with few false-positives, although the method may be a qualitative assay. Conversely, data used to define early pharmacokinetics may be obtained by less sensitive methods but must be quantitative. For these reasons, the RIA and FPIA were, in retrospect, better methods than HPLC was for obtaining data on patient compliance in ACTG Protocol 019.

ACKNOWLEDGMENTS

We thank Vasilios I. Avramis, D. Craig Brater, Jonathan M. Kagan, David M. Kornhauser, Jack H. Ladenson, Steven W. Lagakos, Juan J. Lertora, Paul S. Lietman, Gerald Medoff, Gene D. Morse, Amadeo Pesce, William G. Powderly, Richard Pichichicini, Lee J. Ratner, Brian Stretcher, and Teresa Tartaglione for reviews of the manuscript and Steven S. Good for providing A22U and the first set of quality control specimens. This study was supported by a grant to Washington University (AI 25903) and by contracts with the University of California, San Francisco (AI 62541), the Harvard School of Public Health (AI 95030), and the Research Triangle Institute (AI 62561) from the Division of AIDS of the National Institute of Allergy and Infectious Diseases.

REFERENCES

1. Curb, J. D., C. Ford, C. M. Hawkins, E. O. Smith, N. Zimbaldi, B. Carter, and C. Cooper. 1983. A coordinating center in a clinical trial: the hypertension detection and followup program. Controlled Clin. Trials 4:171–186.
2. Diabetes Control and Complications Trial Research Group. 1987. Feasibility of centralized measurements of glycated hemoglobin in the diabetes control and complications trial: a multicenter study. Clin. Chem. 33:2267–2271.
3. Good, S. S., D. J. Reynolds, and P. de Miranda. 1988. Simultaneous quantification of zidovudine and its glucuronide in serum by high-performance liquid chromatography. J. Chromatogr. 431:123–133.
4. Granich, G. G., M. R. Eveland, and D. J. Krogstad. 1989.
Fluorescence polarization immunoassay for zidovudine. Antimicrob. Agents Chemother. 33:1275–1279.
5. Hainline, A., Jr., D. T. Miller, and A. Mather. 1983. The coronary drug project: role and methods of the central laboratory. Controlled Clin. Trials 4:377–387.
6. Klecker, R. W., Jr., J. M. Collins, R. Yarchean, R. Thomas, J. F. Jenkins, S. Broder, and C. E. Myers. 1987. Plasma and cerebrospinal fluid pharmacokinetics of 3'-azido-3'-deoxythymidine: a novel pyrimidine analog with potential application for the treatment of patients with AIDS and related diseases. Clin. Pharmacol. Ther. 41:407–412.
7. Krogstad, D. J., and M. R. Eveland. Unpublished observations.
8. McDougal, J. S., L. S. Martin, S. P. Cort, M. Mozen, C. M. Heldebrant, and B. L. Evatt. 1985. Thermal inactivation of the acquired immunodeficiency syndrome virus, human T-lymphotropic virus III/lymphadenopathy-associated virus, with special reference to antihemophilic factor. J. Clin. Invest. 76:875–877.
9. Meinert, C. L., J. Tonascia, and S. Tonascia. 1989. The hypertension prevention trial: design, methods and baseline results. Controlled Clin. Trials 10(Suppl.):1S–115S.
10. Minne, H. W. 1984. Quality control in parathyroid hormone radioimmunoassays: a multicentre study performed by the European Parathyroid Hormone Study Group. Eur. J. Clin. Invest. 14:16–23.
11. Moers, D. S. 1979. Statistics: concepts and controversies. W. H. Freeman & Co., San Francisco.
12. Puckett, L. R., B. S. Orf, C. R. Hostager, and J. E. Goldensoph. 1988. Characteristics of a two hour RIA for assessment of the levels of AZT in serum. IVth Int. Conf. AIDS, Stockholm, Sweden, abstr. 3642.
13. Volberding, P. A., S. W. Lagakos, M. A. Koch, C. Pettinelli, M. W. Myers, D. K. Booth, H. H. Balfour, Jr., R. C. Reichman, J. A. Bartlett, M. S. Hirsch, R. L. Murphy, W. D. Hardy, R. Soeiro, M. A. Fischl, J. G. Bartlett, T. C. Merigan, N. E. Hyslop, D. D. Richman, F. T. Valentine, L. W. Corey, and the AIDS Clinical Trials Group. 1990. Zidovudine in asymptomatic human immunodeficiency virus infection: a controlled trial in persons with fewer than 500 CD4-positive cells per cubic millimeter. N. Engl. J. Med. 322:941–949.
14. Widdowson, G. M., M. Kuehneman, A. G. DaChene, S. B. Hulley, and G. R. Cooper. 1986. Quality control of biochemical data in the multiple risk factor intervention trial: central laboratory. Controlled Clin. Trials 7(Suppl. 3):17S–33S.