

Measuring the Quality of Data Collection in a Large Observational Cohort of HIV and AIDS

Mariska Hillebregt^{*1}, Elly de Lange-de Klerk², Dirk Knol², Frank de Wolf^{1,3} and Colette Smit¹

¹HIV Monitoring Foundation, Amsterdam, The Netherlands

²Department of Epidemiology and Biostatistics, VU University Medical Center Amsterdam, The Netherlands

³Imperial College, London, UK

Abstract: The aim of this study was to examine the quality of data collection by studying the validity of collected data. Data were extracted from the clinic charts of two anonymous outpatients by 38 data collectors. A standard for the data to be collected was determined (168 items). The validity was measured by comparing the collected items with the standard; in this way, the percentages of the collected items that were 'correct' could be calculated. The percentage 'correct' was higher for clinic chart 1 (mean: 83% correct, SD 7%) than for clinic chart 2 (mean: 78% correct, SD 8%). All categories contained incorrectly collected data. These data were divided into missing data, incorrect start-stop dates, and surplus collected data. Almost all start-stop dates would change into 'correct' if 'monthyear' was considered correct (instead of the standard 'daymonthyear'). Not all data collectors used specific protocols, and sources other than the written comments were not always checked. This study shows that a high proportion of data was correctly collected. However, the collection of start-stop dates was not optimal, and the collected data included surplus and missing data. Data collectors should be more knowledgeable about HIV disease and trained in the use of difficult protocols, so that they can better recognize what data to collect and how it should be collected. Among physicians, there should be more agreement about what information to record in the charts, to facilitate data extraction for data collectors.

Keywords: Database, manual data entry, quality of data collection, HIV/AIDS.

INTRODUCTION

From a public health perspective, it is important to monitor the outcomes of HIV-infected individuals on a national level [1]. In the Netherlands, the HIV Monitoring Foundation (HMF) performs the prospective collection of data on HIV-infected persons, which is part of routine health care for these patients. The mission of the HMF is to expand the knowledge and understanding of the epidemiology and course of both treated and untreated HIV infection. To achieve this mission, the HMF engages in activities such as collecting and processing data and making them available to researchers. By 1 June 2009, the HMF had collected and stored data in the ATHENA (AIDS Therapy Evaluation in the Netherlands) database from 16,129 HIV-infected persons registered in one of the 25 HIV treatment centres in the Netherlands. In the treatment centres, data collectors obtain data directly from the patients' medical files. Data collectors are divided into core function data collectors and additional job data collectors (HIV consultants and medical secretaries). Data collectors (DCs), supervised by treating physicians, enter the data online into the national ATHENA database, using specific protocols to standardize the data collection and to minimize errors [2, 3]. This database was developed in Oracle Clinical[®] (OC), a system specifically

designed for the management of clinical trial data that complies with the guidelines of the International Conference on Harmonisation (Good Clinical Practice) and the U.S. Food and Drug Administration. The OC database provides cross-checking options and discrepancy checks, which increases the efficiency and accuracy of data entry [2, 4-7]. Data monitors (DMs) supervise the quality of the data by checking their accuracy and completeness. The data are then compared to those in the source documents, such as medical files and electronic laboratory results. In addition, the DMs check for the correct use of specific protocols and privacy regulations, and they train the DCs in complete and accurate collection. Regular feedback and training maximises the reliability of the data [2, 8]. Controlling the quality of data obtained from patients is crucial for all clinical research, and source data verification (SDV) is the preferred approach [2-4, 8, 9]. However, given the large population size, 100% SDV is not feasible. Therefore, the HMF has implemented customized procedures for improving the quality of data by restricting SDV to all end-points essential for key data analysis, such as death. In this way, SDV can be used even for data sets that include large numbers of patients. In addition, HMF develops strategies to replace the manual collection of laboratory results by 'direct uploading' from hospital laboratory databases ('lablink') [10], since direct uploading of the laboratory results improves the quality of data [10]. In this study, we determined whether the DCs collected the information that should be collected. We studied the validity of the collected data in all other categories ('adverse events', 'AIDS-defining events (CDC)',

*Address correspondence to this author at the HIV Monitoring Foundation, Academic Medical Center University of Amsterdam, HVA-A.3.10, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands; Tel: 020-56 66473; Fax: 020-56 69189; E-mail: m.m.hillebregt@amc.uva.nl

‘Antiretroviral Therapy (ART)’, ‘co-medication’ and ‘once-only data’) to provide further insight into the optimisation of procedures for manual data entry [10].

METHODS

Design

The outpatient clinic charts from two anonymous patients were selected. The first clinic chart consisted of data from the first visit to the outpatient clinic of an HIV treatment centre (once-only data) and three follow-up visits (longitudinal data). The second clinic chart consisted of nine follow-up visits. A pilot study was performed by the four DMs, who collected the data from the two charts. The purpose of that study was to examine the representativeness of the charts and to determine the standard for the present study. On the basis of the DMs' experience, all four considered the charts to be representative. Because the DMs normally check the data's accuracy and completeness, and check for the correct use of specific protocols, they were able to determine the standard for the collected information. The specific protocols were used to determine this standard.

Study Population

Of the 42 DCs eligible to participate in this study, 38 took part; four DCs were not able to participate (another job (2), sick leave (1), too busy (1)). All 38 DCs collected the data from the first chart; 36 DCs collected the data from the second chart. Two DCs could not collect the data from the second chart (time shortage at last working day and just started as a DC).

Data Collection

A letter was sent to all DCs inviting their participation, and a presentation was given regarding the research. To guarantee the anonymity of the DCs, we used new log in codes for the OC database, and the data were reversibly anonymised. From April to October 2007, the DCs participated with the Oracle Clinical® database in their own workplace. After the data was collected from each chart, the DCs were interviewed by a researcher using a standardized questionnaire. The purpose of this questionnaire was to gather information about the use of protocols and the characteristics of the DCs (main topics of these questions were: function (HIV consultants, medical secretaries, core function DCs), education (university (of applied sciences) and university-preparatory school, post-secondary and vocational education), full-time equivalent (1-16 hours a week, 17-40 hours a week), duration of the data collection (hours), year of employment), as well as to determine the representativeness of the charts as ascertained by the DCs (lay-out of the charts, handwriting, abbreviations, availability of information).

Statistical Analysis

The items collected by the DCs were compared to the standard, which consisted of 168 items, and they were coded as ‘correct’ or ‘incorrect’. The percentages of correct and incorrectly collected data of the individual charts (overall), both charts together (overall), and of all categories (‘adverse

events’, ‘CDC events’, ‘ART’, ‘co-medication’ and ‘once-only data’) were calculated (by data collector). The statistical tests used for the normally distributed variables were Student's t-test and univariate linear regression analysis; those used for the non-normally distributed variables were Wilcoxon rank sum test, Spearman's rank correlation coefficient, and Fisher's exact test.

All statistical analyses were performed with SAS® version 9.1 (Statistical Analysis Software).

RESULTS

Representativity Chart 1 (N=38)

No difficulties in collecting data were mentioned by 86% of the DCs. These DCs found chart 1 no different from charts they usually handle, although the type of information in this chart occurs less frequently during usual data collection. Of all the DCs, 84% evaluated the lay-out of the chart as clearer than or similar to the charts they usually handle. The handwriting was assessed as clearer or similar by 92% of the DCs; 92% also found the abbreviations clearer or similar, and 63% usually had equal or less information available.

Representativity Chart 2 (N=36):

No difficulties in collecting data were mentioned by 77% of the DCs. These DCs found chart 2, in general, not different from usual, except there were more switches of treatment regimen during data collection. Of all the DCs, 92% evaluated the lay-out of the chart as clearer than or similar to the charts they normally handle. The handwriting was assessed as clearer or similar by 86% of the DCs, 88% found the abbreviations clearer or similar, and 72% usually had equal or less information available.

Percentages ‘Correct’ Chart 1:

The collected items from chart 1 were grouped into six categories (Table 1a); ‘data collection first visit’ (29% of all items to collect from chart 1), ‘co-medication’ (19%), ‘ART’ (18%), ‘clinical visits’ (18%), ‘adverse events’ (8%) and ‘CDC events’ (8%). Overall, the mean score of chart 1 was 83% ‘correct’ (SD 7%). The categories ‘adverse events’ and ‘CDC events’ showed the lowest mean percentages ‘correct’ (respectively, mean 56%, SD 31% and mean 60%, SD 13%). The categories of ‘clinical visits’ and ‘ART’ had the highest mean percentages ‘correct’ (respectively, mean 96%, SD 7% and mean 92%, SD 11%). There was no significant difference in the percentages ‘correct’ when the protocols were used and when they were not used.

Percentages ‘Correct’ Chart 2

The collected items from chart 2 were grouped into five categories (Table 1b); ‘ART’ (58% of all items to collect from chart 2), ‘clinical visits’ (28%), ‘adverse events’ (9%), ‘co-medication’ (4%) and ‘CDC events’ (1%). Overall, the mean score of chart 2 was 78% ‘correct’ (SD 8%). The category of ‘co-medication’ showed the lowest mean percentage ‘correct’ (mean 65%, SD 18%). The category of ‘clinical visits’ had the highest mean percentage ‘correct’ (mean 91%, SD 8%). There was no significant difference in

Table 1a. Percentages ‘Correct’ Collected Categories, Chart 1

Chart 1 N=38 data collectors	Percentages ‘Correct’				Using Protocol Mean (N)	Not Using Protocol Mean (N)	P-Value
	Mean	SD***	Minimum	Maximum			
Overall:							
correct chart 1 (%)	83	7	65	94	85 (N=16)	82 (N=22)	0.16 *
Categories:							
correct data, only collected at first visit (%)	84	8	68	100	85 ^a (N=16)	84 ^a (N=22)	0.67 **
correct AIDS-defining events (%)	60	13	33	80	60 ^a (N=34)	65 ^a (N=4)	0.46 **
correct adverse events (%)	56	31	0	100	57 ^a (N=22)	55 ^a (N=16)	0.93 **
correct Antiretroviral Therapy (%)	92	11	63	100	86 ^a (N=26)	88 ^a (N=12)	0.31 **
correct co medication (%)	84	14	50	100	85 ^a (N=36)	66 ^a (N=2)	0.09 **
correct visits (%) ****	96	7	75	100	-	-	-

*Student’s t-test.

**Wilcoxon rank sum test.

***Standard deviation.

****No protocol.

^aThe ‘mean’ is given instead of the ‘mean rank score’ to better understand the differences.**Table 1b. Percentages ‘Correct’ Collected Categories, Chart 2**

Chart 2 N=36 Data Collectors	Percentages ‘Correct’				Using Protocol Mean (N)	Not Using Protocol Mean (N)	P-Value
	Mean	SD***	Minimum	Maximum			
Overall							
correct chart 2 (%)	78	8	58	91	81 (N=6)	77 (N=30)	0.40 *
Categories:							
<i>correct AIDS-defining events (%) ****</i>	<i>100</i>	<i>0</i>	<i>100</i>	<i>100</i>	-	-	-
correct adverse events (%)	75	20	14	100	74 ^a (N=16)	75 ^a (N=20)	0.90 **
correct Antiretroviral Therapy (%)	72	11	47	92	73 ^a (N=29)	69 ^a (N=7)	0.31 **
correct co medication (%)	65	18	33	100	67 ^a (N=30)	57 ^a (N=6)	0.23 **
correct visits (%) *****	91	8	76	100	-	-	-

*Student’s t-test.

**Wilcoxon rank sum test.

***Standard deviation.

****No AIDS-defining events had to be collected.

*****No protocol.

^aThe ‘mean’ is given instead of the ‘mean rank score’ to better understand the differences.

the percentages ‘correct’ when the protocols were used and when they were not used.

Characteristics of the Charts and Percentages ‘Correct’

Tables 2a and 2b show the results by characteristics of the collected items in charts 1 and 2. The tables consist of missing items, correct start and stop dates (according to the standard), and surplus collected items. The start and stop dates are divided into ‘ddmmyy’ (standard in most cases), only ‘mmyy’, and only ‘yy’. When we looked only at ‘mmyy’ and ‘yy’, the percentages ‘correct’ increased substantially. The surplus collected data were partially false

and partially unnecessarily collected (data not shown). All second stop reasons for ‘ART’ (Table 2b) showed low percentages ‘correct’.

Correlation Between Variables

The correlation of seven variables of chart 1 and chart 2 was assessed (duration of data collection, ‘adverse events’, ‘ART’, ‘co-medication’, ‘ART’ with ‘co-medication’, duration with experience (by chart)). Of these variables, only the duration of the data collection from chart 1 correlated with the duration of the data collection from chart 2 (Spearman R= 0.6820, P<0.001).

Table 2a. Most Relevant Results of the Data Collection, Chart 1, N=38 Data Collectors

Categories	Characteristics of the Chart	Missing Items	Correct Start Date (If Collected)			Missing Stop Date	Correct Stop Date (if Collected)			Surplus Collected Data
			ddmmyy ¹	mmyy ²	yy ³		ddmmyy ¹	mmyy ²	yy ³	
Demography	Male (%)	0	-	-	-	-	-	-	-	-
Transmission	Heterosexual (%)	0	-	-	-	-	-	-	-	-
History of HIV infection	HIV-1 infection (%)	0	34.2	94.7	100	-	-	-	-	15 HIV-2 dates
Data collection	21-07-2004 (%)	0	50.0	55.2	100	-	-	-	-	-
AIDS-defining events	Toxoplasmosis (%)	0	15.7	100	100	52.6	27.7	50	94.4	7 AIDS-defining events
Adverse events	Hospitalization (%)	18.4	29.2	97.5	100	0	75.6	75.6	100	27 adverse events
Antiretroviral Therapy	Combivir (%)	2.5	86.4	100	100	-	-	-	-	1 ART
	Kaletra (%)	0	86.8	100	100	-	-	-	-	
Co-medication	Cotrim (%)	2.5	54.0	94.5	100	5.4	91.8	91.8	91.8	11 co medications
	Folinezuur (%)	0	71.0	94.7	100	-	-	-	-	
	Pyrimethamine (%)	0	63.1	94.7	100	-	-	-	-	
	Sulfadiazine (%)	0	86.8	100	100	13.1	90.9	90.9	100	
Visits	4 visits (%)	3.2	98.6	100	100	-	-	-	-	2 visits

¹Daymonthyear, ²Monthyear, ³Year.

Requests for Supplementary Data

DCs had the opportunity to request supplementary data from the treating physician (by means of a handwritten note). To examine the association between 'a request' and 'no request' and the percentage 'correct' for this question, Fisher's exact test was used. The Fisher's exact test showed only one association, and that was between the question 'Is the patient still using omeprazol?' and the percentage 'correct' of stop date omeprazol ($P=0.0341$). DCs who asked this question collected the stop date incorrectly.

Characteristics of DCs

Table 3 shows the relationship between characteristics of the DCs (function, education, full-time equivalent, duration of data collection, year of employment, employment experience) and percentage 'correct' for chart 1. In the univariate linear regression analysis, only function is associated with the percentage 'correct' for chart 1 and for combined charts 1 and 2 (data not shown). HIV consultants had significant lower overall score for percentage 'correct' than the core function DCs had for chart 1 (80% vs 86%) and for combined charts 1 and 2 (79% vs 85%).

HIV Consultants vs Core Function DCs

Since function was associated with percentage 'correct' for chart 1 and for combined charts 1 and 2, the Wilcoxon rank-sum test was used to compare the percentages 'correct' for HIV consultants with those of core function DCs (Table 4). A significant difference in percentages 'correct' was seen for 'adverse events', 'ART' and 'data collected at first visit only'. The HIV consultants had lower percentages 'correct' for those categories than core function DCs had (respectively 57% vs 74%, 70% vs 80%, 81% vs 87%).

DISCUSSION

The aim of this study was to examine the quality of data collection of HIV infected patients by studying the validity of collected data. Data were extracted from the clinic charts of two anonymous outpatients by data collectors.

Policy, guidelines, and recommendations for further research are based on the results of research that depend on the quality of the collected data [4, 8]. Improving the validity of data is essential to ensure valid results [4]. In this study, the validity was measured by comparing the collected items with the standard, so that the percentages 'correct' of the collected items could be calculated.

To collect accurate data, it is necessary that the charts contain complete data. If information is missing in the charts, data collection will be incomplete. In this study, all collected data was compared to the standard. Therefore, if the charts consisted of incomplete data (care was provided but not recorded), the standard consisted of incomplete data as well. Incomplete information in the charts does not explain missing and incorrectly collected data in this study.

A chart consists of medical letters, written comments, diagnostic information, and laboratory results. All of these sources can contain information that has to be collected. The written comments are the guideline for collection, but all other sources need to be checked for additional information. Chart 1 contained more data sources than chart 2, which mainly pertained to 'the data collected at the first visit only' ('once-only data'), 'adverse events', and 'CDC events'. The sources for 'once-only data' were medical letters, written comments, and diagnostic information. The sources for events ('adverse events' and 'CDC events') were mainly diagnostic information and medical letters. If the

Table 2b. Most Relevant Results of the Data Collection, Chart 2, N=36 Data Collectors

Categories	Characteristics of the Chart	Missing Items	Correct Start Date (If Collected) ddmmyy ¹ mmyy ² yy ³			Missing Stop Date	Correct Stop Date (If Collected) ddmmyy ¹ mmyy ² yy ³			Correct First Stop Reason Therapy ⁴	Correct Second Stop Reason Therapy ⁴	Surplus Collected Data
AIDS-defining events	No event	-	-	-	-	-	-	-	-	-	-	-
Adverse events	Groggy/ dizzy (%)	2.7	51.1	60.4	86.0	-	72.0	72.0	100	-	-	12 adverse events
	Diarrhoea (%)	5.5	41.1	73.5	94.1	2.9	72.7	72.7	96.9	-	-	
	Lipodystrophy (%)	0	-	-	-	-	-	-	-	-	-	
Antiretroviral Therapy	Start:											1 ART
	Efavirenz (%)	0	69.4	100	100	-	-	-	-	-	-	
	DDI (%)	0	69.4	100	100	-	-	-	-	-	-	
	TDF (%)	0	69.4	100	100	-	-	-	-	-	-	
	Kaletra (%)	0	80.5	100	100	-	-	-	-	-	-	
	Emtricitabine (%)	0	77.7	100	100	-	-	-	-	-	-	
	DDI (%)	2.7	77.1	100	100	-	-	-	-	-	-	
	Lamivudine (%) ⁵	69.5	-	45.4	100	-	-	-	-	-	-	
	DDI (%)	27.8	84.6	100	100	-	-	-	-	-	-	
	Abacavir (%)	2.7	80.0	100	100	-	-	-	-	-	-	
	Stop:											
	Lamivudine (%)	0	-	-	-	0	77.7	97.2	100	83.3	27.7	
	Stavudine (%)	0	-	-	-	0	77.7	97.2	100	91.7	27.7	
	Kaletra (%)	0	-	-	-	0	77.7	97.2	100	91.7	27.7	
	Efavirenz (%)	0	-	-	-	0	97.2	100	100	94.4	30.5	
	DDI (%)	0	-	-	-	0	100	100	100	83.3	-	
	Emtricitabine (%)	0	-	-	-	0	-	13.8	100	86.1	13.8	
	Lamivudine (%) ⁵	69.5	-	-	-	0	-	9.0	100	63.6	27.2	
	DDI (%)	27.8	-	-	-	0	100	100	100	46.1	-	
TDF (%)	5.5	-	-	-	0	100	100	100	27.7	-		
Co-medication	Omeprazol (%)	0	33.3	52.7	91.6	52.7	64.7	64.7	100	-	-	4 co medications
Visits	9 visits (%)	7.0	90.8	100	100	-	-	-	-	-	-	24 visits

¹Daymonthyear, ²Monthyear, ³Year, ⁴Antiretroviral Therapy, ⁵The standard was 'mmyy'.

characteristics of the patient contained 'once-only data' and/or events, sources other than written comments had to be checked. The sources that deviated from the written comments in chart 1 then might be an explanation for missing and incorrectly collected data.

Many start-stop dates would have been classified as 'correct' if 'mmyy' had been considered correct, instead of 'ddmmyy'. However, specific start-stop dates are essential for all researchers who study, for example, short-term effects of ART, toxicity reactions, ART switch/stop reasons, and clinical progression. Therefore, start-stop dates should be correctly collected as 'ddmmyy', when information about dates is available.

The HIV consultants achieved a significant lower correct score than core function DCs in the specific categories of 'adverse events', 'ART' and 'once-only data'. An explanation for the lower correct score is that the HIV consultants used the specific protocols less frequently than

did the core function DCs (data not shown). Another explanation for the lower correct score is that the consultants usually treat HIV patients themselves and know (mostly) their state of health from memory. In this study, they had to collect data from unknown patients. Therefore, this could have been a disadvantage. However, to reduce errors, DCs should always use the clinical charts, not memory, as the source for data collection.

All collected categories contained incorrect data. The incorrect data were divided into missing data, incorrect start-stop dates, and surplus collected data. The surplus collected data consisted of partially false and partially unnecessary information (data not shown). To reduce data errors, DCs should use protocols [2, 11, 12]. Protocols should be easy to follow, with clear descriptions of what data to collect [3, 11]. During the data collection in this study, not all DCs used specific protocols. Contrary to our expectations, we found that, although start-stop dates of all medication, 'adverse

Table 3. Univariate linear Regression Analysis on Percentage ‘Correct’ for Clinic Chart 1, N=38 Data Collectors

Independent Variable	Characteristics of the DCs	Sample Size	B (% Correct)	SE	P-Value ^a	Overall P-Value
Function	HIV consultants	15	80	2.2	0.015	0.049
	Medical secretaries	4	84	3.5		
	Core function (constant)	19	86	1.5		
Education	University (of Applied Sciences), University- preparatory School	18	82	2.2	0.34	
	Post-secondary and Vocational education (constant)	20	84	1.5		
Full-time equivalent	1-16 hours a week	21	82	2.2	0.12	
	17-40 hours a week (constant)	17	85	1.6		
Duration of the data collection	Hours	38	0.01448	0.02442	0.56	
	Constant		0.81603	0.03274		
Year of employment	After 2003	22	83	2.3	0.39	
	Before 2003 (constant)	16	85	1.7		
Year of employment	After 2005	13	84	2.4	0.94	
	Before 2005 (constant)	25	83	1.4		
Employment experience	Years	38	0.03506	0.03121	0.27	
	Constant		0.82268	0.01514		

^aP values <0.05 were considered significant and are represented in bold.

Table 4. HIV Consultants vs Core Function Data Collectors

Variable	HIV Consultant (N=14)	Core Function (N=18)	P-Value ^{a b}
% Correct	Mean ^c	Mean ^c	
Adverse events chart 1 + chart 2	57	74	0.013
Antiretroviral Therapy chart 1 + chart 2	70	80	0.006
Data collected at first visit only, chart 1	81	87	0.049
Co-medication chart 1 + chart 2	79	81	0.64
AIDS-defining events chart 1 + chart 2	66	67	0.75

^aP-values <0.05 were considered significant and represented in bold.

^bWilcoxon rank sum test.

^cThe ‘mean’ is given instead of the ‘mean rank score’ to better understand the differences.

events’, and ‘CDC events’ were defined in specific protocols, there was no significant difference in the percentages ‘correct’ when the protocols were used and when they were not used. A plausible explanation is that the protocols of the HMF were too difficult to follow, and therefore, errors were made. The protocols are difficult because of the complexity of the HIV disease and because the content of charts varies amongst physicians. The protocols of the HMF try to include all variations in chart content to help the DCs with data extraction.

There were several limitations to this study. First, the small sample size resulted in a lack of power; because of this, we found fewer predictors of data quality than expected. Nevertheless, the sample size was representative of the population of interest; nearly all data collectors were included. There were four non-responders for chart 1 and six non-responders for chart 2. If all data collectors had been included, this would not have changed the proportion of HIV consultants vs that of core function data collectors.

Multivariate regression analyses could not be performed because of the small sample size, and thus we could not correct for confounding factors. Second, multiple analyses were performed in the study, so we expected no confounding factors, but unmeasured confounding will always exist. Because of multiple testing, there is chance of a type 1 error. However, we do not think the significant lower correct score of consultants was a result of a type 1 error. These data errors can be explained by the consultants' collecting data from patients they did not know and by the less frequent use of specific protocols.

This study shows a high proportion of correctly collected data. However, the collection of start-stop dates was not optimal, and the collected data consisted of surplus as well as missing data. The explanation for this is that not all collectors used the specific protocols, although there was no significant difference in the percentages ‘correct’ when the protocols were used and when they were not used. Furthermore, DCs probably did not always check sources

other than the written comments in the charts, which might have resulted in a higher proportion of incorrectly collected data. The HIV consultants scored lower in percentages 'correct' than the core function data collectors, possibly because they were not familiar with these patients and used the specific protocols less frequently than the core function data collectors. The following recommendations make it possible for the HMF to increase the validity of the data. First, the specific protocols should be easy to follow, with clear descriptions [3, 11], but because of the complexity of the HIV disease and the variation in the content of the charts, the protocols remain difficult. Therefore, DCs should increase their knowledge of HIV disease, so they can better recognize what data to collect. DCs should also be trained in the use of difficult protocols to improve their ability to collect data. In addition, there should be more agreement among physicians about what information to record in the charts to facilitate data extraction for the DCs [13]. Finally, all DCs need to be trained to check sources other than written comments, especially if the characteristics of the patient contain 'data collected at first visit only' and/or events ('adverse events' and 'CDC events'). In particular, it must be recognized that recall bias might result from the HIV consultants collecting data partially from memory (knowing their own patients' state of health). Ideally, collected data should be based strictly on information in the clinical chart.

ACKNOWLEDGMENTS

The authors would like to thank all data collectors for their study participation and all data monitors for their participation in the pilot project, Luuk Gras for his assistance with the SAS[®] syntaxes and the HMF for its support in completing this study. We also thank Sally H. Ebeling for editing the manuscript.

ABBREVIATIONS

AIDS	=	Acquired Immunodeficiency Syndrome
ATHENA	=	AIDS Therapy Evaluation in the Netherlands
ART	=	Antiretroviral Therapy
CDC	=	U.S. Centers for Disease Control and Prevention

DC	=	Data collectors
DM	=	Data monitors
HIV	=	Human Immunodeficiency Virus
HMF	=	HIV Monitoring Foundation
OC	=	Oracle Clinical [®]
SAS [®]	=	Statistical Analysis Software
SDV	=	Source Data Verification

REFERENCES

- [1] The UK Collaborative HIV Cohort Steering Committee. The creation of a large UK-based multicentre cohort of HIV-infected individuals: The UK Collaborative HIV Cohort (UK CHIC) Study. *HIV Med* 2004; 5: 115-24.
- [2] Schneider JK, Deenan A. Reducing quantitative data errors: tips for clinical researchers. *Appl Nurs Res* 2004; 17: 125-29.
- [3] Favalli G, Vermorken JB, Vantongelen K, Renard J, Van Oosterom AT, Pecorelli S. Quality control in multicentric clinical trials. An experience of the EORTC Gynecological Cancer Cooperative Group. *Eur J Cancer* 2000; 36: 1125-33.
- [4] Cassidy LD, Marsh GM, Holleran MK, Ruhl LS. Methodology to improve data quality from chart review in the managed care setting. *Am J Manag Care* 2002; 8: 787-93.
- [5] Mullooly JP. The effects of data entry error: an analysis of partial verification. *Comput Biomed Res* 1990; 23: 259-67.
- [6] McGinniss MJ, Chen R, Pratt VM, *et al.* Development of a web-based query tool for quality assurance of clinical molecular genetic test results. *J Mol Diagn* 2007; 9: 95-98.
- [7] Black N. Using clinical databases in practice. *BMJ* 2003; 4: 326-29.
- [8] Freeman JA, Hobart JC, Playford ED, Undy B, Thompson AJ. Evaluating neurorehabilitation: lessons from routine data collection. *J Neurol Neurosurg Psychiatry* 2005; 76: 723-28.
- [9] Allison JJ, Wall TC, Spettell CM, *et al.* The art and science of chart review. *Jt Comm J Qual Improv* 2000; 26: 115-36.
- [10] HIV Monitoring Foundation. Monitoring of Human Immunodeficiency Virus (HIV) Infection in the Netherlands. Report. Amsterdam: Stichting HIV Monitoring/HIV Monitoring Foundation 2005.
- [11] Moody LE, McMillan S. Maintaining data integrity in randomized clinical trials. *Nurs Res* 2002; 51: 129-33.
- [12] Baigent C, Harrell FE, Buyse M, Emberson JR, Altman DG. Ensuring trial validity by data quality assurance and diversification of monitoring methods. *Clin Trials* 2008; 5: 49-55.
- [13] Luck J, Peabody JW, Dresselhaus TR, Lee M, Glassman P. How well does chart abstraction measure quality? A prospective comparison of standardized patients with the medical record. *Am J Med* 2000; 108: 642-49.