Original Article

Reliability Analysis of a Newly Developed Questionnaire for Quality Control of Follow-up Visits in PolyIran Study

Gholamreza Roshandel MD PhD[1,2], Mohammad Reza Ostovaneh MD[2], Hossein Poustchi MD PhD•[2], Fatemeh Malekzadeh MD[2], Sadaf Ghajarieh Sepanlou MD PhD[2], Mohammad Reza Honarvar MD PhD[3], Shahryar Semnani MD[1], Shahin Merat MD[2], Reza Malekzadeh MD[2]

Abstract

Background: The PolyIran study is a large-scale pragmatic cluster randomized controlled trial of fixed-dose combination therapy (Polypill) for prevention of cardiovascular diseases (CVD) in Iran. The PolyIran Quality Control Program (PIQCP) including a new questionnaire was developed to assess the quality of data collection during follow-up visits. The aim of this study was to assess the inter-rater reliability of PIQCP questionnaire.

Methods: The study was conducted in 26 (11%) randomly selected clusters (from a total of 236 PolyIran clusters). All participants within these 26 clusters were enrolled. The quality scores were measured according to the PIQCP guidelines by two independent raters. The intraclass correlation coefficients (ICC) were measured. In addition, the quality scores were categorized into good (³70%) and poor (<70%). The kappa coefficient was used to assess inter-rater agreement for this categorical quality scores.

Results: A total number of 945 PolyIran participants were enrolled of which, 501 (53%) were from intervention arm. In 934 participants (98.8%), the quality score could be successfully identified by both raters. The ICC (95%CI) of the overall quality scores was 0.985 (0.983–0.987). It was 0.976 (0.972–0.980) and 0.988 (0.986–0.990) in intervention and control arms, respectively. We found excellent agreement between the two raters in identifying participants with good and poor quality scores (kappa = 0.988, $P < 0.001$). The kappa values were 0.972 ($P < 0.001$) and 1.000 ($P < 0.001$) in intervention and control arms, respectively.

Discussion: Our results suggested that the PIQCP questionnaire is a reliable tool for assessing quality of data collection in PolyIran follow-up visits. Using this measure will help us in efficient monitoring of the PolyIran follow-ups and may ensure high quality data.

Keywords: Follow-up visit, PolyIran study, quality control, reliability

Cite this article as: Roshandel G, Ostovaneh MR, Poustchi H, Malekzadeh F, Sepanlou SG, Honarvar MR, Semnani S, Merat S, Malekzadeh R. Reliability Analysis of a Newly Developed Questionnaire for Quality Control of Follow-up Visits in PolyIran Study. *Arch Iran Med*. 2016; **19(8):** 551 – 555.

Introduction

Cardiovascular diseases (CVD) are leading causes of death worldwide.[1–3] A fixed-dose combination of established drugs (polypill) was suggested as an effective strategy for controlling CVD.[4] This approach may reduce costs of treatment and increase participants' adherence and consequently will reduce CVD incidence and mortality. Regarding the high rates of CVD mortality in Golestan Cohort Study (GCS),[5] it was suggested to consider the polypill strategy for primary and secondary prevention of CVD in this population. The GCS is a prospective population-based long-term study with 50,000 participants which started 10 years ago, primarily to study esophageal cancer which was exceptionally common in Northeastern Iran (eastern part of the Golestan province) with 99% follow-up success rate. CVD was found to be the etiology of death in 50% of all mortalities among GCS participants. The results of the pilot phase of polypill study suggested that it is feasible and even necessary to conduct a large-scale trial on GCS population.[6] Therefore, the main phase of polypill study, called "PolyIran study", was designed. The PolyIran study is a pragmatic cohort multiple cluster randomized trial (cmRCT) to assess the safety and efficacy of a single fixed-dose combination of cardiovascular medications intervention for primary and secondary prevention of major cardiovascular events in middle-aged and elderly Iranians. The details of the PolyIran study have been described elsewhere.[7] The enrolment phase started in February 2011 and was completed by April 2013, with each village considered a cluster. The follow-up phase of the PolyIran project began in May 2011. The intervention arm clusters were followed up at months 1, 2, 3, 6, and then at 6-month intervals whereas the control arm clusters were followed up at months 3, 6, and subsequently at 6-month intervals.

During follow-up visits, different types of information were collected including participants' demographic data, pill counts, reasons for non-adherence, blood pressure measurement, new medications, symptoms, referrals and cardiovascular events. The full list of information collected in PolyIran follow-up visits is shown in Table 1.

Quality assurance and quality control constitute an important part of any clinical trial.[8,9] Because of the importance and pivotal role of follow-up visits in the PolyIran study, the PolyIran Quality Control Program (PIQCP) was developed to assess the quality of data collection during PolyIran follow-up visits. It contained two questionnaires, one for the intervention group, and the other for

Authors' affiliations: [1]Golestan Research Center of Gastroenterology and Hepatology, Golestan University of Medical Sciences, Gorgan, Iran. [2]Digestive Disease Research Center, Digestive Disease Research Institute, Shariati Hospital, Tehran University of Medical Sciences, Tehran, Iran. [3]Department of Health, Golestan University of Medical Sciences, Gorgan, Iran.
•**Corresponding author and reprints:** Hossein Poustchi MD, Digestive Disease Research Center, Digestive Disease Research Institute, Tehran University of Medical Sciences, Shariati Hospital, 1411713135, Tehran, Iran. Telefax: +982182415104, E-mail. h.poustchi@gmail.com
Accepted for publication: 15 May 2016

Table 1. Calculating activity-specific and overall quality scores for each participant in the PolyIran Quality Control Program (PIQCP).

Activity code	Activity details	Expected (maximum) score (S_E)		Observed (calculated) score (S_O)	Quality score
		Intervention arm	Control arm		
A1	To record participants' demographic data	10	10	S_O	$(S_O*100) / S_E$
A2	To calculate and record number of Polypill tablets have been taken/ had to be taken by participants	35	—	S_O	$(S_O*100) / S_E$
A3	To record the reasons for non-adherence	15	—	S_O	$(S_O*100) / S_E$
A4	To assess and record blood pressure	35	50	S_O	$(S_O*100) / S_E$
A5	To refer hypertensive participants to health system physician	15	25	S_O	$(S_O*100) / S_E$
A6	To obtain and record data on newly occurred cardiovascular events	15	25	S_O	$(S_O*100) / S_E$
A7	To record drug history (Aspirin, Anti-lipid, Anti-hypertensive)	—	25	S_O	$(S_O*100) / S_E$
A8	To record new symptoms	30	40	S_O	$(S_O*100) / S_E$
A9	To record the participant's follow-up status and reasons for loss to follow up	10	25	S_O	$(S_O*100) / S_E$
A10	To make and record final decision on drug continuation for the next 6 months	20	—	S_O	$(S_O*100) / S_E$
A11	To refer participants with adverse events (or any other problems) for physician visit	35	35	S_O	$(S_O*100) / S_E$
A12	To give the Polypill tablets to participants	35	—	S_O	$(S_O*100) / S_E$
A13	To prepare the list of participants needed to be visited by the study physician	15	15	S_O	$(S_O*100) / S_E$
A14	To give the participants information on lifestyle modification	30	50	S_O	$(S_O*100) / S_E$
Overall quality score		$\sum S_E=300$	$\sum S_E=300$	$\sum S_O$	$(\sum S_O*100) / \sum S_E$

the control group asking about the quality of data collection on all of the follow-up activities (Table 1). The questionnaires were completed by an experienced staff after reviewing the PolyIran follow-up data collection forms and by telephone interviews with followed up participants. A PIQCP guideline was also developed, in which the methods of scoring the quality of data were described in detail.

Because the PIQCP questionnaire was a new questionnaire developed for the PolyIran, we decided to validate this quality control measure. Validation studies mainly assess the validity and reliability of questionnaires. The different aspects of validity of the PIQCP questionnaire (content validity, face validity and construct validity) were assessed and approved by experts and experienced individuals in this field. In the present study, we aimed to determine the inter-rater reliability (agreement) for the PIQCP questionnaire in assessing quality of data collection in PolyIran follow-up visits.

Materials and Methods

Study population

This study was done on the PolyIran study participants during May and September 2015. A total of 8410 participants were enrolled in PolyIran study, including 4233 participants (120 clusters) and 4177 (116 clusters) in the intervention and control arms, respectively.

Sample size

We aimed to enroll about 10% of the PolyIran participants (800 participants) in this validation study with the intervention/control ratio of 1/1.

Sampling

Considering a median cluster size of about 30 participant,[7] we aimed to enroll 26 PolyIran clusters to achieve a total sample size of about 800 for the present study. According to the timetable of the PolyIran follow-up, 180 clusters (90 clusters from the control and 90 clusters from the intervention arms) were followed between May and September 2015. Of these followed clusters, 13 intervention and 13 control clusters were randomly selected for this validation study. We aimed to enroll all PolyIran participants within these cluster.

Data collection

For all participants within selected clusters, the quality scores were assessed (using the PIQCP questionnaire) by two independent raters within one week after the end of follow-up. The quality scores were measured by the PIQCP questionnaire according to the PIQCP guidelines (Table 1). Thus, the following process was considered to measure the quality score for each participant.

At first, the quality score of each activity (activity-specific quality score) was calculated (Table 1). For each of the PolyIran follow-up activities, a maximum score was earned if the data were correctly and completely collected according to the protocol of the PolyIran follow-up and there was no error. Otherwise, part of the score was earned (based on the level of incompleteness, errors, etc). For example, according to the protocol of the PolyIran follow-up, the section of "demographic data" included 5 items (first name, last name, participants ID, village code, date). So, if there was an error in one of these items, 80% of the maximum score would be earned for this activity, and so on. Then, the sum

of scores of all activities (overall observed score) was calculated. Finally, the "overall quality score" was calculated by dividing the "overall observed score" by sum of maximum scores of all activities (overall expected score) (Table 1). Each rater was blinded to the results reported by the other.

Analysis of data

The activity-specific quality scores as well as the overall quality score were calculated for each participant using the above-mentioned formula (Table 1). The intraclass correlation coefficient (ICC) and its 95% confidence intervals (CI) were measured to assess the level of agreement between the raters. Bland-Altman graph[10] was also used to assess the concordance of quality scores reported by rater 1 and rater 2.

In addition, the quality of data collection was categorized as good (quality score ³ 70%) or poor (quality score < 70%). The kappa coefficient was used to assess inter-rater agreement for this categorical quality scores. *P* values of less than 0.05 were considered significant.

Results

A total of 945 PolyIran participants were enrolled, of which, 472 (49.9%) were male and the mean (SD) age of participants was 60.01 (7.14) years. A total of 501 participants (53%) were recruited from the intervention arm. The number of male participants in the intervention and control arms was 246 (49.1%) and 226 (50.9%), respectively (*P* value = 0.58). The mean (SD) of participants' age was 59.88 (7.21) and 60.27 (7.07) years in the intervention and control arms, respectively (*P* value = 0.40).

In 934 participants (98.8%), the quality score could be successfully identified by both raters. The mean (SD) of overall quality scores reported by raters 1 and 2 was 94.8% (13.9) and 95.1% (14.3), respectively. Table 2 shows the means of the overall quality scores in the intervention and control arms.

The quality of data was rated as good in 95.1% and 95.2% of participants by raters 1 and 2, respectively. The number (proportion) of participants with good and poor quality scores reported by raters 1 and 2 in the intervention and control arms is shown in Table 2.

The ICC (95% CI) of the overall quality scores reported by raters 1 and 2 was 0.985 (0.983–0.987). The ICCs (95% CI) of the overall quality scores in the intervention and control arms were 0.976 (0.972–0.980) and 0.988 (0.986–0.990), respectively. Table 3 shows the ICCs (95%CIs) of activity-specific quality scores reported by the two raters in the intervention and control arms.

Figure 1 shows the concordance of two raters in assessing quality scores in PolyIran participants.

Our results showed an excellent agreement between the two raters in identifying participants with good and poor quality scores (kappa = 0.988, $P < 0.001$). The kappa values were 0.972 ($P < 0.001$) and 1.000 ($P < 0.001$) in the intervention and control arms, respectively.

Discussion

The PolyIran study is a large-scale pragmatic randomized controlled trial.[7] It was conducted at the community level and large number of healthcare workers (*behvarzes*) were involved in the process of data collection and follow-up. It was believed that follow-up visits were an important phase of the PolyIran study and quality of data collection in this phase could strongly affect the study results. To ensure high-quality data acquisition and reporting during PolyIran follow-up visits, a quality assurance and monitoring program (PIQCP) was designed for the PolyIran study. A new PIQCP-specific questionnaire was developed and the quality of data collection was regularly assessed according to the PIQCP guideline during each round of follow-up. The PIQCP questionnaire was a newly developed questionnaire and needed to be validated in the study population. The aim of this study was to assess the inter-rater reliability of PIQCP questionnaire in the PolyIran study.

Our findings suggested high inter-rater reliability values for the PIQCP questionnaire. Both quantitative and qualitative analyses showed high concordance for assessing the overall quality scores between the two raters (Table 2 and Figure 1). In addition, table 3 shows high ICCs for all activity-specific quality scores. Therefore, the PIQCP questionnaire was a reliable tool for assessing quality of data collection in PolyIran follow-up visits.

The effects of polypill strategy have been studied on CVD outcomes in different populations.[11–15] But, similar to many other researches,[16] little information was reported on measures of quality control and quality assurance for follow-up data collection in these studies. Quality control is one of the most important issues of clinical trials, especially those with large number of study sites and many different interviewers.[17] Data of poor quality may have strong negative effects on the power of study[18,19]. Many different factors may affect the quality of data in clinical researches, including study design, sample size, complexity of data collection, etc. Therefore, different methods may be considered to ensure high quality data in various studies.[20,21] In addition, an individual study may need different quality control measures during each phase of the study. Marinez *et al*. considered an integrated data management system to ensure high quality data.[22] In a study

Table 2. The mean (SD) of the overall quality scores as well as the number (proportion) of participants with good (quality score ≥ 70%) and poor (quality score<70%) quality scores reported by raters 1 and 2 in intervention and control arms of the PolyIran study.

Statistics	Raters		Intervention arm (n = 492)	Control arm (n = 442)	*P*-Value
Mean (SD) of quality score	Rater 1		95.7 (10.3)	93.8 (15.1)	0.06
	Rater 2		95.9 (10.2)	94.3 (15.7)	0.09
Number (%) of participants with good and poor quality scores	Rater 1	Good quality	473 (96.1)	415 (93.9)	0.11
		Poor quality	19 (3.9)	27 (6.1)	
	Rater 2	Good quality	474 (96.3)	415 (93.9)	0.08
		Poor quality	18 (3.7)	27 (6.1)	

Table 3. The levels of agreement between raters 1 and 2 for assessing the quality scores for each activity of the PolyIran quality control program in intervention and control arms.

Activity code	Activity details	Intervention arm, ICC (95%CI)	Control arm, ICC (95%CI)
A1	To record participants' demographic data	0.772 (0.727–.0809)	0.822 (0.806–0.836)
A2	To calculate and record number of Polypill tablets have been taken/ had to be taken by participants	0.874 (0.843–919)	N/A
A3	To record the reasons for non-adherence	0.969 (0.963–0.974)	N/A
A4	To assess and record blood pressure	0.864 (0.838–0.868)	0.907 (0.888–0.923)
A5	To refer hypertensive participants to health system physician	0.962 (0.954 – 0.968)	0.972 (0.966–0.977)
A6	To obtain and record data on newly occurred cardiovascular events	0.953 (0.944–0.960)	0.996 (0.995–0.996)
A7	To record drug history (Aspirin, Anti-lipid, Anti-hypertensive)	N/A	0.959 (0.951–0.966)
A8	To record new symptoms	0.960 (0.952–0.966)	0.995 (0.994–0.996)
A9	To record the participant's follow-up status and reasons for loss to follow up	0.986 (0.983–0.988)	0.981 (0.977–0.984)
A10	To make and record final decision on drug continuation for the next 6 months	0.918 (0.902–0.932)	N/A
A11	To refer participants with adverse events (or any other problems) for physician visit	0.931 (0.917–0.942)	0.971 (0.965–0.976)
A12	To give the Polypill tablets to participants	0.900 (0.853–0.933)	N/A
A13	To prepare the list of participants needed to be visited by the study physician	0.795 (0.698–0.861)	0.823 (0.743–0.879)
A14	To give the participants information on lifestyle modification	0.873 (0.813–0.914)	0.855 (0.745–0.932)

ICC = intraclass correlation coefficient; N/A= not applicable.

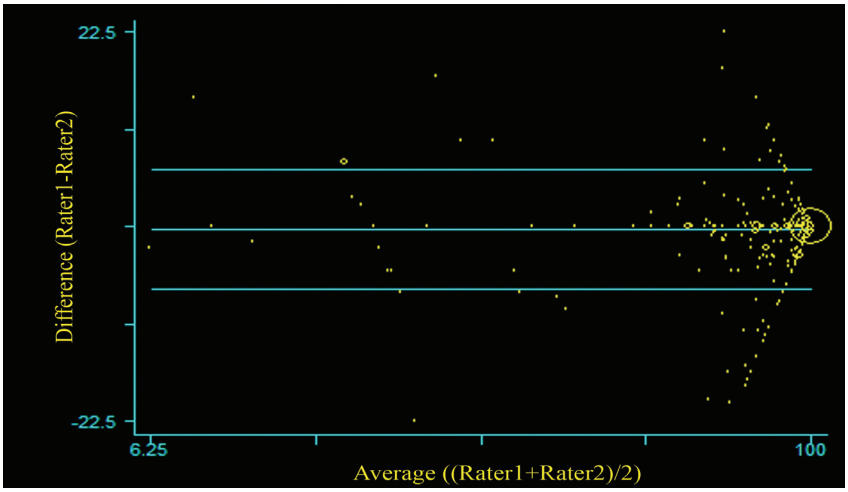

Figure 1. Bland-Altman graph for assessing the concordance of two raters for measuring quality of data in PolyIran follow-up visits.

from Malawi, using a local quality assessment tool could result in improvements in data quality.[23,24] In the Diabetes Control and Complications Trial and the follow-up Epidemiology of Diabetes Interventions and Complications study a quality assurance program was implemented to identify trends, data inconsistencies and process variability of results over time.[17] Therefore, it is important to design and implement an appropriate quality control program for different phases of clinical researches, especially large-scale clinical trial.[18,25]

In conclusion, we found high inter-rater accordance for assessing the overall and activity-specific quality scores. Our results suggested that the PIQCP questionnaire is a reliable tool for assessing quality of data collection in PolyIran follow-up visits. Using this measure will help us in efficient monitoring of the PolyIran follow-ups and will consequently ensure high quality data in this large-scale pragmatic randomized controlled trial.

Acknowledgment

The authors would like to thank the PolyIran staff for their cooperation in data collection. The efforts of local healthcare workers (behvarzes) in the PolyIran study are also appreciated.

References

1. Hartley A, Marshall DC, Salciccioli JD, Sikkel MB, Maruthappu M, Shalhoub J. Trends in Mortality from Ischaemic Heart Disease and Cerebrovascular Disease in Europe: 1980-2009. *Circulation*. 2016; 133(20): 1916 – 1926.
2. Yusuf S, Reddy S, Ounpuu S, Anand S. Global burden of cardiovascular diseases: part I: general considerations, the epidemiologic transition, risk factors, and impact of urbanization. *Circulation*. 2001; 104(22): 2746 – 2753.
3. Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*. 2012; 380(9859): 2224 – 2260.
4. Wald NJ, Law MR. A strategy to reduce cardiovascular disease by more than 80%. *BMJ*. 2003; 326(7404): 1419.
5. Pourshams A, Khademi H, Malekshah AF, Islami F, Nouraei M, Sadjadi AR, et al. Cohort Profile: The Golestan Cohort Study--a prospective study of oesophageal cancer in northern Iran. *Int J Epidemiol*. 2010; 39(1): 52 – 59.
6. Malekzadeh F, Marshall T, Pourshams A, Gharravi M, Aslani A, Nateghi A, et al. A pilot double-blind randomised placebo-controlled trial of the effects of fixed-dose combination therapy ('polypill') on cardiovascular risk factors. *Int J Clin Pract*. 2010; 64(9): 1220 – 1227.
7. Ostovaneh MR, Poustchi H, Hemming K, Marjani H, Pourshams A, Nateghi A, et al. Polypill for the prevention of cardiovascular disease (PolyIran): study design and rationale for a pragmatic cluster randomized controlled trial. *Eur J Prev Cardiol*. 2015; 22(12): 1609 – 1617.
8. Gassman JJ, Owen WW, Kuntz TE, Martin JP, Amoroso WP. Data quality assurance, monitoring, and reporting. *Control Clin Trials*. 1995; 16(2 Suppl): 104s – 136s.
9. Groom N, Wilson E, Lyn E, Faivre-Finn C. Is pre-trial quality assurance necessary? Experiences of the CONVERT Phase III randomized trial for good performance status patients with limited-stage small-cell lung cancer. *Br J Radiol*. 2014; 87(1037): 20130653.
10. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986; 1(8476): 307 – 310.
11. Bautista LE, Vera-Cala LM, Ferrante D, Herrera VM, Miranda JJ, Pichardo R, et al. A 'polypill' aimed at preventing cardiovascular disease could prove highly cost-effective for use in Latin America. *Health Aff (Millwood)*. 2013; 32(1): 155 – 164.
12. Yusuf S, Pais P, Afzal R, Xavier D, Teo K, Eikelboom J, et al. Effects of a polypill (Polycap) on risk factors in middle-aged individuals without cardiovascular disease (TIPS): a phase II, double-blind, randomised trial. *Lancet*. 2009; 373(9672): 1341 – 1351.
13. Soliman EZ, Mendis S, Dissanayake WP, Somasundaram NP, Gunaratne PS, Jayasingne IK, et al. A Polypill for primary prevention of cardiovascular disease: a feasibility study of the World Health Organization. *Trials*. 2011; 12: 3.
14. Thom S, Poulter N, Field J, Patel A, Prabhakaran D, Stanton A, et al. Effects of a fixed-dose combination strategy on adherence and risk factors in patients with or at high risk of CVD: the UMPIRE randomized clinical trial. *JAMA*. 2013; 310(1): 918 – 929.
15. Wald DS, Morris JK, Wald NJ. Randomized Polypill crossover trial in people aged 50 and over. *PLoS One*. 2012; 7(7): e41297.
16. Rosa C, Campbell A, Kleppinger C, Sampson R, Tyson C, Mamay-Gentilin S. Quality assurance of research protocols conducted in the community: the National Institute on Drug Abuse Clinical Trials Network experience. *Clin Trials*. 2009; 6(2): 151 – 161.
17. Lorenzi GM, Braffett BH, Arends VL, Danis RP, Diminick L, Klumpp KA, et al. Quality Control Measures over 30 Years in a Multicenter Clinical Study: Results from the Diabetes Control and Complications Trial / Epidemiology of Diabetes Interventions and Complications (DCCT/EDIC) Study. *PLoS One*. 2015; 10(11): e0141286.
18. Whitney CW, Lind BK, Wahl PW. Quality assurance and quality control in longitudinal studies. *Epidemiol Rev*. 1998; 20(1): 71 – 80.
19. Severe JB, Schooler NR, Lee JH, Haas G, Mueser KT, Rosen P, et al. Ensuring data quality in a multicenter clinical trial: remote site data entry, central coordination and feedback. *Psychopharmacol Bull*. 1989; 25(3): 488 – 490.
20. Prud'homme GJ, Canner PL, Cutler JA. Quality assurance and monitoring in the Hypertension Prevention Trial. Hypertension Prevention Trial Research Group. *Control Clin Trials*. 1989; 10(3 Suppl): 84s – 94s.
21. Chalmers TC, Smith Jr H, Blackburn B, Silverman B, Schroeder B, Reitman D, et al. A method for assessing the quality of a randomized control trial. *Control Clin Trials*. 1981; 2(1): 31 – 49.
22. Marinez YN, McMahan CA, Barnwell GM, Wigodsky HS. Ensuring data quality in medical research through an integrated data management system. *Stat Med*. 1984; 3(2): 101 – 111.
23. Admon AJ, Bazile J, Makungwa H, Chingoli MA, Hirschhorn LR, Peckarsky M, et al. Assessing and improving data quality from community health workers: a successful intervention in Neno, Malawi. *Public Health Action*. 2013; 3(1): 56 – 59.
24. Hedt-Gauthier BL, Tenthani L, Mitchell S, Chimbwandira FM, Makombe S, Chirwa Z, et al. Improving data quality and supervision of antiretroviral therapy sites in Malawi: an application of Lot Quality Assurance Sampling. *BMC Health Serv Res*. 2012; 12: 196.
25. Fukushima M. Quality control in clinical trials [in Japanese]. *Gan To Kagaku Ryoho*. 1996; 23(2): 172 – 182.