

**Sonia Khatawkar,
Arun Bhatt, Rutika Shetty,
Patricia Dsilva**

*Clininvent Research Private Limited,
Mumbai, Maharashtra, India*

Address for correspondence:

Dr. Arun Bhatt,
Clininvent Research Private Limited.,
A-302, Everest Chambers, Marol Naka,
Andheri-Kurla Road, Andheri (E),
Mumbai - 400 059, Maharashtra, India.
E-mail: arunbhatt@clininvent.com

Abstract

Analysis of data query as parameter of quality

Introduction: Data query (DQ) rate per case record form (CRF) page is an index of quality of clinical trial data, which can be affected by the phase of clinical trial, the therapeutic area, and the country, where investigator site is located. **Objective:** To compare DQ rate per page by countries, phases, and therapeutic areas. **Materials and Methods:** Data from 19 paper CRF clinical trials conducted at 352 sites, in 5,610 patients were entered into clinical data management system by double data entry method, and DQs were generated. The DQ rate per page was compared for the phase, therapeutic area, and the country, by parametric analysis of variance (ANOVA) and nonparametric test—Kruskal-Wallis and median test. Multiple comparisons test was conducted for each category using Tukey's Studentized Range Test. **Results:** The total number of DQs from 108,599 CRF pages was 33,177 (0.306/CRF page). The studies included 12 countries, all phases, and seven therapeutic areas. Comparison of DQ rate per page showed a significant difference between phases, countries, and therapeutic areas. However, multiple comparisons showed that the DQ rate per page was significantly different between Phase I and other phases, and oncology, cardiology, endocrinology, and dermatology. The DQ categories were: Missing 21.8%, confirmatory 50.1%, and clarification 28.2%. The most common DQs by CRF field were: Lab data (23.2%), physical examination (17.4%), study assessment (17.4%), and concomitant medication (12%). There was no correlation between the number of CRF pages and DQ rate per page per study. **Conclusion:** The phase of study and therapeutic area could impact the data quality as measured by DQ rates.

Key words: Case record form, country, data query, data query rate per page, phase, therapeutic area, quality

INTRODUCTION

Since last several years, the clinical trial industry has become global with inclusion of new emerging countries such as India. One of the major reasons for inclusion of emerging countries is potential for fast recruitment. During the same period, the protocols have become more

complex, with an increase in number of selection criteria, number of protocol procedures, and number of case record form (CRF) pages.^[1] These changes have resulted in increasing the burden of work at investigator sites, which in turn increased the risk of adversely affecting the quality and performance of the team at the investigator site. As most sites in emerging countries are new in the field of clinical trials, there is a concern about quality of data collected from these new regions. Data query (DQ) rate per CRF page is one index of quality, used to assess quality of clinical trial data from some of the western countries and Russia.^[2]

A recent analysis of data queries from 26 large phase II/III trials comparing DQ rate per page from North America or western Europe and other emerging regions

Access this article online

Quick Response Code:



Website:

www.picronline.org

DOI:

10.4103/2229-3485.134312

failed to detect any statistically significant differences.^[3] However, in an analysis of 8,317 protocols, Getz *et al.*, found wide differences in protocol complexity and work burden across phases and therapeutic areas.^[4] Hence, it would be important to consider the impact of therapeutic area whilst analyzing DQ rates. We report here comparison of DQs by countries, phases, and therapeutic areas.

MATERIALS AND METHODS

Data generated from 19 clinical trials, conducted by 352 sites in 5,610 patients, using paper CRF were included for analysis of DQs.

A cumulative total of 108,599 CRF pages were entered into Oracle Clinical Data Management System using double data entry method. The CRF pages were then validated using edit checks, logic checks, and 100% manual review to generate per page DQs. Multiple DQ for each page were counted as two separate DQs.

DQ rate per page was calculated by dividing total number of DQs raised by total number of CRF pages entered.

Statistical analysis was carried out to:

- Compare statistically, the DQ per page among countries, phases of clinical trial, and therapeutic areas using parametric analysis of variance (ANOVA), nonparametric Kruskal-Wallis and median test. If the distribution of DQ rate per page was significant, multiple comparisons were performed for each category–country, phase, and therapeutic area by Tukey’s Studentized Range Test
- Assess correlation between number of CRF pages/study and average number of DQs per CRF page/study
- Distribution of category of various types of DQs: Missing data, confirmatory queries, and clarification
 - Missing data queries are queries raised when data is not provided
 - Confirmation queries are queries raised to confirm the data not recorded/provided as per CRF entry criteria
 - Clarification queries are queries raised to get more clarification on the data provided/recorded on CRF pages but lacks clarity.
- Distribution of DQs per CRF page by data fields on CRF page such as:
 - Laboratory data
 - Physical examination
 - Study assessment
 - Concomitant medication
 - Administrative
 - Adverse event
 - Medical history
 - Investigational product

- Administrative, and
- Electrocardiogram.

RESULTS

The total number of DQs from 108,599 CRF pages was 33,177 giving a DQ rate per page of 0.306/CRF page. The studies were from 12 countries, covering 352 sites, all phases, and seven therapeutic areas.

The phase wise distribution of number of subjects, CRF pages and DQ per CRF page is presented in Table 1.

The DQ rate per page for Phase IV was lowest (0.308) and was highest for Phase I (1.043) [Table 1].

Table 2 describes country-wise distribution of number of subjects, CRF pages, and DQ per CRF page.

The DQ rate per page across countries ranged from 0.035 for Ukraine to 1.149 for Canada [Table 2].

Table 3 describes therapeutic area wise distribution of number of subjects, CRF pages and DQ per CRF page.

DQ rate per page was lowest for ophthalmology (0.075) and highest for oncology 0.649 [Table 3].

Table 1: Distribution of DQ rate per page by phase of study

Phase	No. of projects	No. of sites	No. of patients	Total no. of CRF pages	Total no. of queries	DQ/ page
I	3	7	121	6,215	6,482	1.043
II	2	7	130	14,664	5,906	0.403
III	8	57	1,405	41,646	6,586	0.158
IV	6	281	3,954	46,074	14,203	0.308
Total	19	352	5610	108,599	33,177	0.306

DQ=Data query, CRF=Case record form

Table 2: Distribution of DQ rate per page by country

Country	No. of sites	No. of patients	Total no. of CRF pages	Total no. of DQs	DQ rate
India	107	2161	69,822	22,162	0.317
Poland	19	695	15,546	1,656	0.107
United states	58	510	5,454	2,311	0.424
Indonesia	66	563	4,474	704	0.157
Canada	2	35	3,333	3,828	1.149
Korea	22	596	2,980	316	0.106
Vietnam	16	300	2,100	1,212	0.577
Malaysia	42	304	1,520	238	0.157
Philippines	14	268	1,340	246	0.184
Taiwan	2	107	1,315	469	0.357
Ukraine	1	36	540	19	0.035
Brunei	3	35	175	16	0.091
Total	352	5,610	108,599	33,177	0.306

DQ=Data query, CRF=Case record form

Comparison of DQ rate per page from all phases showed a significant difference between countries and therapeutic areas [Table 4].

The DQ rate per page showed a significant difference between phases, countries, and therapeutic areas. The multiple comparison tests showed:

Phases

- DQ rate per page from phase I studies were significantly different from Phase II, Phase III, and Phase IV
- DQ rate per page from Phase II, Phase III, and Phase IV was not significantly different from each other.

Countries

- There were no significant differences between countries.

Therapeutic areas

- DQ rate per page from oncology studies were significantly different from cardiology, endocrinology, and dermatology. All other comparisons did not show any significant differences.

Number of CRF pages ranged from two pages for a Phase IV study to 155 pages for a Phase I study [Table 5].

There was no correlation between the number of CRF pages and DQ rate per page.

The distribution of DQ categories was: Missing 21.8%, confirmatory 50.1%, and clarification 28.2%.

The most common DQs [Table 6] by CRF field were: Lab data (23.2%), physical examination (17.4%), study assessment (17.4), and concomitant medication (12.0%).

Table 3: Distribution of DQ rate per page therapeutic area

Therapeutic area	No. of trials	No. of sites	No. of patients	No. of CRF pages	No. of DQs	DQ rate per page/CRF
Endocrinology	5	95	1,813	34,877	10,294	0.295
Dermatology	6	75	1,206	20,269	2,594	0.128
Cardiovascular	3	149	2,164	19,716	5,335	0.271
Oncology	2	9	120	16,848	10,936	0.649
Nephrology	1	7	125	7,623	2,499	0.328
Ophthalmology	1	8	110	6,759	505	0.075
Urology	1	9	72	2,507	1,014	0.404
Total	19	352	5,610	108,599	33,177	0.306

DQ=Data query, CRF=Case record form

Table 4: Statistical comparison of DQ rate per page

Tests	ANOVA P value	Kruskal-Wallis P value	Median P value
Phase	<0.0001	<0.0034	<0.0360
Country	<0.0001	<0.0001	<0.0001
Therapeutic area	<0.0001	<0.0001	<0.0001

DQ=Data query, ANOVA=Analysis of variance

DISCUSSION

Our study showed that the average DQ rate per page was 0.306/CRF page or 1 DQ in 3 CRF pages. DQ rate per page was highest for Phase I or oncology study. Getz has reported that Phase I protocols were the most complex and the most burdensome.^[4] The study also reported that anti-infectives, immunomodulation, CNS, and oncology studies are consistently the most complex and burdensome to execute.^[4]

Getz *et al.*, study has revealed that protocol complexity and work for all phases in diverse therapeutic areas have grown between 2000 and 2007. Hence, comparison of DQ rate per page for countries needs to consider the effect of phase or therapeutic area in the analysis.

We found a significant difference in DQ rate per page by parametric test ANOVA or non-parametric tests for all phases, all countries, and therapeutic areas.

Table 5: Distribution of CRF pages and DQ rate per page

Phase	Therapeutic area	No. of pages per CRF	Average DQ rate	
I	Oncology	155	1.171	
	Cardiovascular	47	0.819	
	Dermatology	15	0.035	
II	Dermatology	37	0.018	
	Oncology	127	0.467	
III	Dermatology	29	0.150	
	Endocrinology	41	0.158	
	Endocrinology	41	0.145	
	Endocrinology	34	0.223	
	Ophthalmology	67	0.082	
	Cardiovascular	30	0.199	
	Dermatology	20	0.133	
	Urology	35	0.381	
	IV	Dermatology	2	0.149
		Dermatology	33	0.249
Endocrinology		60	0.328	
Cardiovascular		5	0.198	
Endocrinology		12	0.332	
Nephrology		61	0.360	

DQ=Data query, CRF=Case record form

Table 6: Distribution of number of DQs by data field on CRF pages

Data field on CRF page	No. of DQs (%)
Lab Data	7,690 (23.2)
Physical examination	5,779 (17.4)
Study assessment	4668 (14.1)
Concomitant medication	3,979 (12.0)
Administrative	3,128 (9.4)
Adverse Event	2,927 (8.8)
Other (OTH)	1837 (5.5)
Medical history	1,588 (4.8)
Investigational product	1,120 (3.4)
Electrocardiogram	461 (1.4)

DQ=Data query, CRF=Case record form

Our results for the difference DQ rate per page between the countries are in contrast to the study of Desai *et al.*,^[3] who reported that there were no statistically significant differences in the query rates, when developing country regions were compared to North America or Western Europe. However, our study differs from this study in several aspects. Desai *et al.*, compared combined DQ rate per page for Phase II and III studies, whilst our study has compared DQs of all phases separately. Desai's study combined studies of several companies, covering all 10 regions of the globe, whilst we could include eastern Europe, US, India, and Asia Pacific. Also, Desai's study did not consider the effect of therapeutic area on the DQ rate, but only analyzed the data quality of developing countries.

When the multiple comparisons were done for each group, DQ rate per page was significantly different between Phase I and other Phases, but not significantly different between Phase II or III or IV. Amongst therapeutic areas, the DQ rate per page was significantly different between oncology and cardiology, endocrinology and dermatology. However, amongst the countries, there was no significant difference in DQ rate per page, when multiple comparison test was applied. Hence, it appears that the phase of the study or therapeutic area could explain the differences in DQ rate per page between the countries.

The length of CRF could be another factor which could impact DQ rate. However, we did not find any correlation between length of CRF and the DQ rate.

Confirmatory DQs featured as the most common category type highlighting the need for thorough training in protocol and CRF completion for sites. Over 20% DQs were due to missing data, suggesting the need for greater and vigilant monitoring by clinical research associate prior to data retrieval.

Physical examination and lab data pages generated the highest number of DQs again highlighting the need for precise CRF completion guidelines and ongoing site training.

Our study included paper CRFs. It is possible that use of electronic CRF could have reduced the DQ rates. However, Desai's study reported the query rate per page was higher for clinical trials using electronic CRFs.^[3] Hence, it appears that technology alone would not be able reduce the number of DQs.

There were some limitations of our study. The study did not cover countries from western Europe, Latin America, Africa, and China. The number of sites was not similar from all countries for all phases. There was preponderance of Phase III and IV studies and only seven therapeutic areas covered.

In conclusion, the phase of study, and the therapeutic area could impact the data quality as measured by DQ rate per page.

REFERENCES

1. Platonov PG, Varshavsky S, Escandon S. An evaluation of query generation rate per page as an indicator of major factors influencing data quality. *Appl Clin Trials* 2003;12:32-4.
2. Getz K. Protocol design trends and their effect on clinical trial performance. *Raj Pharma* 2008:315-6.
3. Desai PB, Anderson C, Sietsema WK. A Comparison of the quality of data, assessed using query rates, from clinical trials conducted across developed versus emerging global regions. *Drug Inf J* 2012;46:455-63.
4. Getz KA, Campo RA, Kaitin KI. Variability in protocol design complexity by phase and therapeutic area. *Drug Inf J* 2011;45:413-20.

How to cite this article: Khatawkar S, Bhatt A, Shetty R, Dsilva P. Analysis of data query as parameter of quality. *Perspect Clin Res* 2014;5:121-4.

Source of Support: Nil. **Conflict of Interest:** None declared.