

State of use of Electronic Data Capture (EDC) tools in randomized controlled trials in India: results from a survey

Santam Chakraborty (✉ drsantam@gmail.com)

Tata Medical Center, Kolkata, India <https://orcid.org/0000-0003-3580-5979>

Indranil Mallick (✉ imallick@gmail.com)

Tata Medical Center, Kolkata, India <https://orcid.org/0000-0002-5567-9204>

Tapesh Bhattacharyya

Tata Medical Center, Kolkata, India <https://orcid.org/0000-0002-3714-1195>

Moses Arunsingh S

Tata Medical Center, Kolkata, India <https://orcid.org/0000-0002-4003-3259>

Rimpa Basu Achari

Tata Medical Center, Kolkata, India <https://orcid.org/0000-0002-1709-255X>

Sanjoy Chatterjee

Tata Medical Center, Kolkata, India <https://orcid.org/0000-0001-7402-8645>

Research Article

Keywords: Electronic data capture, Randomized Controlled Trials, India

Posted Date: June 7th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-596078/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Introduction

Electronic data capture (EDC) tools can improve data quality gathered in clinical trials and may be more cost-effective than paper forms. However, limited data is available on the adoption of EDC tools in randomized controlled trials (RCT) conducted in India.

Methods

We invited investigators of registered randomized controlled trials in India to an online survey. The questionnaire included questions on the use of EDC (or alternative data capture methods) and features available. An EDC sophistication level (ranging from 1–6) was computed from the responses obtained. Respondents were also asked about barriers to the implementation of EDC in their setting. The EDC adoption rate (EAR) was defined as the proportion of clinical trials where EDC with a sophistication level of 2 or more was used. Multivariable logistic regression was used to identify factors that predicted EDC adoption.

Results

Responses were received for 400 trials, with an EAR of 27.5% (95% confidence intervals : 23.4–32.1%, n = 110). The number of sites influenced EDC adoption (odds ratio : 1.26, 95% CI : 1.12–1.47, p = 0.001) on multivariable analysis. EAR did not increase over time. The key barriers identified for not using an EDC were lack of technical support (170, 63.0%) and software cost (132, 48.9%).

Conclusion

The survey shows a low EAR in randomized trials registered in India. The barriers identified in the survey would need systematic solutions to improve the EAR in the future.

Introduction

Data quality is a multi-dimensional concept with five key dimensions, viz. completeness, uniqueness, timeliness, accuracy, validity, and consistency (1). Ensuring data quality in randomized controlled trials (RCT) is essential to ensure the validity of results. Walther et al. have shown that EDC is as accurate as paper case record forms (CRFs) and has reduced missing data (2). In a randomized trial, the use of eCRFs was associated with significant time saving and improved data quality (3). Pavlovic et al. simulated the costs of administering paper-based and electronic CRFs and showed that costs could be reduced by 49–62% when e-CRFs are used (4). The use of e-CRFs has demonstrated similar advantages in both high and low resource settings (5, 6).

Before this study, a search of the literature revealed five studies from India that described the usage (7–9) or design of an EDC system for clinical research (10, 11). We failed to identify a study that had studied the use of EDC for clinical trials in India. Given the importance of EDC systems in improving trial data quality and cost

reduction, it would be essential to understand the current status of such systems in our country. The study aimed to estimate the proportion of open randomized controlled trials which used an EDC in India.

Methods

Study Design

The survey was designed following the methodology adopted by Emam et al. (12). However, we contacted the investigators instead of the clinical trial unit (CTU) coordinators as the contact details of CTU coordinators were not readily available. The invitation emails requested the investigators to provide the details of coordinators if the information related to EDC use was not available to them. The survey protocol (Appendix I) was prospectively approved by the institutional review board of Tata Medical Center (2021/TMC/212/IRB2), and all participants provided informed consent for participation in the study. The informed consent statement comprised details on the purpose of the study, time involved, the data usage policy, and the investigator's contact details.

Survey Instrument

Questions related to EDC usage and EDC sophistication levels were taken from the previous survey conducted by Emam et al. (12). If EDCs were not used in the RCT, users were asked how they collected data using a separate question. Conditional logic was used to reduce the response burden. Additionally, inquiries on the perceived barriers toward EDC use in the trial and the publication status of the trial were asked. Before rolling out the survey the navigation and skip logic (conditional logic) was checked manually by four users who also provided feedback on the time taken in navigating the survey.

Clinical Trial selection

Investigators of open (status: open or not yet recruiting) randomized controlled trials registered in the Clinical Trial Registry of India (CTRI) and the ClinicalTrials.gov were invited to participate in a short online survey. Given our prior experience with response to online surveys, random sampling was not done. Investigator details were obtained from the information available on the trial registry website. The list of trials was retrieved using a broad search which included the keywords "randomized", "randomised", and "trial". In addition to the contact details of the investigators, we also retrieved the following data :

1. Type of trial - Interventional (I), Observational (O), Bioavailability and Bioequivalence (BA/BE), and post-marketing surveillance (PMS) studies

1. Recruitment Status
2. Health condition studied
3. Intervention(s) and Comparator(s) as applicable
4. Sponsor Type
5. Trial design and phase
6. The planned sample size for India
7. Date enrollment started and total trial duration.
8. Multinational / Indian Study

The list of trials was then filtered to retain only studies that have reported the use of randomization in their design. We used a simple title matching technique to see if the record exists in the CTRI search records to eliminate duplicate trials between two registries. For those studies that did not have a perfect match of the title, manual curation was done assisted by using a fuzzy word distance-based matching technique to determine if the registered study was unique (13). The same information regarding the trial was also retrieved from the ClinicalTrials.gov site.

Recruitment

Emails with personalized links to the survey were sent to each investigator using a transactional email service (Gmass, Inc). The customized email had the CTRI or NCI trial ID as a pre-populated query string in a read-only field. All emails had an unsubscribe link. No advertisements were done for the survey, and the survey link was not available on a website. The survey emails were sent from 16th March 2021 till 15th May in batches. Four reminders were sent at intervals of a few days to people who had not responded. The transactional email service automatically removed any invalid email addresses, unsubscribes, and bounced email addresses. The survey was closed to participation on 29th May 2021.

Survey Administration

The survey was deployed on a commercial website (<https://lakxya.com>), and the survey instrument was designed using the Drupal Webform module (<https://www.drupal.org/project/webform>). While the survey was open to all visitors, a direct link was not available on the website. The IP address of the respondent was not recorded. All questions were marked as required, and automatic validation ensured that questions had been answered before submission. Questionnaires were displayed in 6 pages with a variable number of questions on each page. While navigating between the pages, a back button was provided. Questions related to EDC sophistication were only asked when users responded that they had used an EDC in the clinical trial. No incentives were provided for participation.

Access to survey data was limited to an authorized user who had to log in with a username and password to access the survey data.

Analysis

The campaign tracking feature of the transactional email service was used to calculate the total number of emails delivered and the email open rate. The survey participation rate was defined as the proportion of clinical trials where at least one of the investigators agreed to participate in the survey by clicking the checkbox in the consent form provided on the first page. The survey completion rate was defined as the proportion of clinical trials where at least one investigator completed the survey.

All responses which do not have a CTRI or NCI trial ID in the read-only field were discarded. We retained the first response for trials where more than one respondent has answered, which states that EDC was used. The total number of trials with multiple responses was reported.

An EDC sophistication level was computed from each response using data from available features in the EDC used (12). Six levels could be assigned where higher levels represented higher levels of sophistication (Appendix I). The primary outcome was EDC adoption rate (EAR) which was defined as the ratio of the number of CTRI registered trials that use an EDC with a sophistication level of 2 or more to that of the number of participating

trials. Clinical trial data may be entered into a standalone database with password protection, corresponding to sophistication level 1. Such a system can be used to enter data from paper CRFs manually. Therefore only software with features other than password protection was considered as an EDC. This is consistent with the definition used by Emam et al. (12).

Proportion and binomial confidence intervals were reported. Additional analyses focussed on evaluating clinical trial-related factors associated with the EAR (using logistic regression). Additionally, time trends of EAR after 2010 were graphically explored.

In the survey reported by Emam et al., 41% of the users in Canada had reported using an EDC. We hypothesized that EAR in India might lie between 20–40%. The sample size to estimate the EAR ranged between 246 to 369 with a 95% confidence interval and a margin of error of 5%. Assuming a response rate of 30%, we needed to survey about 1230 studies. All analysis was done using R version 4.0.3 and packages gtsummary, tidyverse, and ggplot2 (14–17). The complete analysis with code and results is available online (https://rpubs.com/santam/edc_appendix).

Results

A total of 1,909 open randomized controlled trials with investigator emails were identified. Contact details of 3,173 investigators were available. Several investigators had conducted more than one trial. Hence we could contact a total of 2,890 unique emails. 1,523 investigators (52.7%) opened at least one email. Emails regarding 1,141 trials were opened (email open rate of 59.8%). Google analytics data indicated that the survey landing page was viewed 772 times by 622 users from India during this period. Investigators of 400 unique trials consented to participate in the survey (participation rate of 21%). All participants completed the survey (completion rate 100%). We received multiple responses for eight trials.

An EDC was used in 130 (32.5%, 95% CI : 28.1% – 37.2%) trials. Of these, 110 trials had an EDC with a sophistication level of 2 or more. EDC sophistication score could not be computed for twenty studies as all features were absent. The EAR was 27.5% (95% confidence intervals : 23.4–32.1%). In the trials which did not use an EDC (n = 270), the most common method of data collection was using a spreadsheet (260, 96%) (Appendix II Table 10). Email and fax were used to send data to a central data collection unit in 46 (17%) and 9 (3.3%) trials, respectively.

Table 1 shows the results of the univariate analysis of factors that influenced the EAR. Trials that were industry-funded, multicentric, and multinational were more likely to have adopted EDCs. Multivariable analysis using logistic regression showed that the number of sites was a statistically significant variable influencing EDC adoption (odds ratio: 1.26, 95% CI: 1.12–1.47, p = 0.001). The odds ratio of adoption of EDC in industry-funded trials was 2.46 (95% CI: 0.86–5.13, p = 0.09) though it was not statistically significant (Table 2).

Among the trials which had used EDC with a sophistication level of 1 or more (n = 123), 93 (76%) of the trials reported using EDCs that had a sophistication level of 5 or more (Appendix II Table 7). Trial characteristics associated with EDC use with lower sophistication scores were longer duration and larger sample size (Table 3). Additionally, trials conducted in institutes with access to CTU tended to have EDCs with lower sophistication scores.

EAR time trends did not show a notable trend towards increasing adoption of EDC. EAR between 2019–2015 was 28% compared to 27% between 2014 – 2010 (Fig. 1 and Appendix II: Table 9).

Respondents identified lack of technical support (n = 170, 63.0%, 95% CI : 57.1% – 68.5%) and software cost (n = 132, 48.9%, 95% CI: 43.0% – 54.8%) as the most reasons for not using an EDC in their clinical trial (Appendix II Table 11). The analysis of free-text responses showed that lack of budgetary support and technical support were the primary impediments identified by respondents in adopting an EDC (Appendix II Table 12). Some respondents also mentioned that they were unaware of any EDC systems.

Discussion

The results of this survey suggest that EAR in randomized controlled trials in India is low, with an upper bound of 95% confidence interval of 32.1%. This is in contrast to the situation in most developed countries where EDC use is now more widespread. Emam et al.'s survey identified that as far back as 2007, nearly 41% of the trials were using an EDC (12). In 2010, a study of the European Clinical Research Infrastructure Network (2010) identified that nearly 90% of the centers use a clinical data management system (CDMS) - which enables the capture of trial data using electronic case report forms (18). However, they also identified a wide heterogeneity in the systems used across centers and adherence to standards.

Given the demonstrated utility of EDC in ensuring data quality(19) while improving efficiency (20) and reducing the costs of clinical trials (21), it is interesting to note that software costs were considered as a critical barrier towards implementation. This is not surprising considering the high initial capital investment (server and software purchase) required for setting up such systems. Only 51% of the randomized trials in this survey have access to an institutional CTU. However, access to a CTU was not associated with a higher EAR. This may indicate that CTUs are not actively investing in and advocating for EDCs.

Unless institutional funding and technical support are available, it may not be feasible for investigators to individually adopt and maintain EDC systems for academic clinical trials (Appendix II Table 11). Additionally, high bandwidth internet access or hospital local area networks are required for most systems. Even among the free-text responses, lack of funding and difficulties in hiring, training, and retaining personnel who can use and manage the EDC systems were prominent. Several investigators were also unaware that such systems existed. The role of an institutional CTU which has the capability to manage an EDC is, therefore, paramount to improve the EAR.

Data from our survey also suggests that EAR has not changed over the past decade as 95% confidence intervals of EAR tend to overlap for all years. This is not the case in the West, which has seen a steady increase in EAR over the past decade (21). Increasingly direct electronic health record (EHR) to eCRF data collection is being adopted in the West (22). The lack of change in EAR India is likely to be related to systematic barriers that have remained consistent over time.

Between 2013–2015, the Central Drugs Standard Control Organization (CDSCO) instituted several regulations like a requirement for audio-video consenting, mandatory registration of ethics committees as well as the specification of compensation norms (23). This may have resulted in fewer industry-sponsored trials being registered (24). Analysis of our data suggests, that the proportion of industry-sponsored trials fell from 37.7% of

the total trials in 2012 to 11.9% of total trials in 2013 (Appendix II, Fig. 3). The decline in industry-sponsored trials may have affected the EAR across time.

The categorization of EDC sophistication levels used in the present study was derived from Emam et al. to maintain comparability across study results (12). These levels can be modeled using a Guttman scale, which is suitable for cumulative functionality levels (25). EDC systems which have implemented level 6 are likely to have implemented level 5 also. Noteworthy level 6 sophistication level corresponds to the ability to track medication inventory at the site, and level 5 is subject recruitment that can be tracked online at each location. Both of these requirements are important for multicentric industry-sponsored studies. Thus it is not surprising that industry-sponsored studies tended to use more sophisticated forms of EDC. However, only 15 of 32 (46.8%) industry-sponsored trials had used an EDC with a sophistication level of 6.

Setting up EDC systems to be available institute-wide requires substantial infrastructure, equipment, and human resources. While commercial EDC systems exist, they are expensive and may be out of reach for most academic investigators. Free and feature-rich, web-based EDC systems like REDCap are available, but increased awareness is necessary (26). Our institute has maintained a REDCap EDC, which has hosted over 300 projects over the past decade. The REDCap consortium website shows that 96 centers in India are using REDCap (27). Cloud-based EDCs may ease the adoption barrier in centers with limited IT support. However, the complex regulatory and patient privacy requirements would mandate that a system is available at a national level with stringent protections for patient privacy and intellectual property.

Meanwhile, developments in the software have resulted in improved functionality achieved by EDC. Some examples are the ability to abstract data from electronic health records through the integration of natural language processing and artificial intelligence, including decision support systems, integration with electronic patient-reported outcomes, and wearables (21). In resource-constrained settings like India, it may make more sense to develop and maintain such advanced EDC systems as a web-based application available to all researchers across the country.

Limitations

The study included only registered randomized controlled trials where contact details were available. However, since RCTs provide the highest level of evidence regarding efficacy (and therefore need high-quality data to be interpretable and usable), the use of EDC will be most important in these studies. We also contacted the investigators instead of the study coordinators as the contact details of the study investigators were only available. Contact details of contract research organizations (CROs) were also inconsistently available on CTRI. Contacting the CROs may have improved the EAR as industry-sponsored studies are more likely to have funds to use an EDC. As shown by the results of this survey, only half of the studies had a clinical trial unit at the center. However, the definition of a clinical trial was not specified in the survey instrument. Finally, random sampling was not used in the study. However, key trial characteristics viz. sponsorship type, number of sites, countries of recruitment, and sample size seem to have a similar distribution in the trials for which we received a response versus those we did not (Appendix II: Table 3).

Conclusion

The results of this survey suggest a low EAR in randomized controlled trials in India. This is likely due to systematic issues as costs, and lack of technical support were the key barriers to EDC use. Therefore, improving EAR would be contingent on providing adequate funding and technical support, especially for academic clinical trials. Given the technological developments in this space, it may make more sense to fund a nationwide EDC solution available for all academic investigators across the country.

Declarations

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

1. Wand Y, Wang RY. Anchoring data quality dimensions in ontological foundations. *Commun ACM* [Internet]. 1996 Nov 1;39(11):86–95. Available from: <https://doi.org/10.1145/240455.240479>
2. Walther B, Hossin S, Townend J, Abernethy N, Parker D, Jeffries D. Comparison of electronic data capture (EDC) with the standard data capture method for clinical trial data. *PLoS One* [Internet]. 2011 Sep 23;6(9):e25348. Available from: <http://dx.doi.org/10.1371/journal.pone.0025348>
3. Fleischmann R, Decker A-M, Kraft A, Mai K, Schmidt S. Mobile electronic versus paper case report forms in clinical trials: a randomized controlled trial. *BMC Med Res Methodol* [Internet]. 2017 Dec 1;17(1):153. Available from: <http://dx.doi.org/10.1186/s12874-017-0429-y>
4. Pavlović I, Kern T, Miklavcic D. Comparison of paper-based and electronic data collection process in clinical trials: costs simulation study. *Contemp Clin Trials* [Internet]. 2009 Jul;30(4):300–16. Available from: <http://dx.doi.org/10.1016/j.cct.2009.03.008>
5. Jennifer YY, Goldberg T, Lao N, Feldman BM, Goh YI. Electronic Forms for Patient Reported Outcome Measures (PROMs) are an Effective, Time-Efficient, and Cost-Minimizing Alternative to Paper Forms. 2020; Available from: <https://www.researchsquare.com/article/rs-36967/latest.pdf>
6. Dillon DG, Pirie F, Rice S, Pomilla C, Sandhu MS, Motala AA, et al. Open-source electronic data capture system offered increased accuracy and cost-effectiveness compared with paper methods in Africa. *J Clin Epidemiol* [Internet]. 2014 Dec;67(12):1358–63. Available from: <http://dx.doi.org/10.1016/j.jclinepi.2014.06.012>
7. Bassi A, John O, Gallagher M, Kotwal S, Joshi R, Essue B, et al. Methodological challenges to collecting clinical and economic outcome data: Lessons from the pilot dialysis outcomes India study. *Nephrology* [Internet]. 2019 Apr;24(4):445–9. Available from: <http://dx.doi.org/10.1111/nep.13257>
8. Gupta PC. Survey of sociodemographic characteristics of tobacco use among 99,598 individuals in Bombay, India using handheld computers. *Tob Control* [Internet]. 1996 Summer;5(2):114–20. Available from: <http://dx.doi.org/10.1136/tc.5.2.114>
9. Kumar AMV, Naik B, Guddemane DK, Bhat P, Wilson N, Sreenivas AN, et al. Efficient, quality-assured data capture in operational research through innovative use of open-access technology. *Public Health Action* [Internet]. 2013 Mar 21;3(1):60–2. Available from: <http://dx.doi.org/10.5588/pha.13.0004>

10. Bose A, Das S. Trial analytics—a tool for clinical trial management. *Acta Pol Pharm* [Internet]. 2012 May;69(3):523–33. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/22594267>
11. Ruth CJ, Huey SL, Krisher JT, Fothergill A, Gannon BM, Jones CE, et al. An Electronic Data Capture Framework (ConnEDCt) for Global and Public Health Research: Design and Implementation. *J Med Internet Res* [Internet]. 2020 Aug 13;22(8):e18580. Available from: <http://dx.doi.org/10.2196/18580>
12. El Emam K, Jonker E, Sampson M, Krleza-Jerić K, Neisa A. The use of electronic data capture tools in clinical trials: Web-survey of 259 Canadian trials. *J Med Internet Res* [Internet]. 2009 Mar 9;11(1):e8. Available from: <http://dx.doi.org/10.2196/jmir.1120>
13. Kumari S, Mohan A, Saberwal G. Hidden duplicates: 10s or 100s of Indian trials, registered with ClinicalTrials.gov, have not been registered in India, as required by law. *PLoS One* [Internet]. 2020 Jun 19;15(6):e0234925. Available from: <http://dx.doi.org/10.1371/journal.pone.0234925>
14. R Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2020. Available from: <https://www.R-project.org/>
15. Sjoberg DD, Curry M, Hannum M, Larmarange J, Whiting K, Zabor EC. gtsummary: Presentation-Ready Data Summary and Analytic Result Tables. 2021.
16. Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, et al. Welcome to the tidyverse. *Journal of Open Source Software* [Internet]. 2019;4(43):1686. Available from: <http://dx.doi.org/10.21105/joss.01686>
17. Wickham H. ggplot2: Elegant Graphics for Data Analysis [Internet]. Springer-Verlag New York; 2016. Available from: <https://ggplot2.tidyverse.org>
18. Kuchinke W, Ohmann C, Yang Q, Salas N, Lauritsen J, Gueyffier F, et al. Heterogeneity prevails: the state of clinical trial data management in Europe - results of a survey of ECRIN centres. *Trials* [Internet]. 2010 Jul 21;11:79. Available from: <http://dx.doi.org/10.1186/1745-6215-11-79>
19. Nahm ML, Pieper CF, Cunningham MM. Quantifying data quality for clinical trials using electronic data capture. *PLoS One* [Internet]. 2008 Aug 25;3(8):e3049. Available from: <http://dx.doi.org/10.1371/journal.pone.0003049>
20. Litchfield J, Freeman J, Schou H, Elsley M, Fuller R, Chubb B. Is the future for clinical trials internet-based? A cluster randomized clinical trial. *Clin Trials* [Internet]. 2005;2(1):72–9. Available from: <http://dx.doi.org/10.1191/1740774505cn069oa>
21. Eisenstein E, Zozus MN, Sanns W. Beyond EDC. Winter 2021 [Internet]. 2021 Mar 12;1(1). Available from: <http://www.jscdm.org/article/id/33/>
22. Wilkinson M, Young R, Harper B, Machion B, Getz K. Baseline Assessment of the Evolving 2017 eClinical Landscape. *Ther Innov Regul Sci* [Internet]. 2019 Jan;53(1):71–80. Available from: <http://dx.doi.org/10.1177/2168479018769292>
23. Gogtay NJ, Ravi R, Thatte UM. Regulatory requirements for clinical trials in India: What academicians need to know. *Indian J Anaesth* [Internet]. 2017 Mar;61(3):192–9. Available from: http://dx.doi.org/10.4103/ija.IJA_143_17
24. Bhave A, Menon S. Regulatory environment for clinical research: Recent past and expected future. *Perspect Clin Res* [Internet]. 2017 Jan;8(1):11–6. Available from: <http://dx.doi.org/10.4103/2229-3485.198551>

25. Guttman L. A Basis for Scaling Qualitative Data. *Am Sociol Rev* [Internet]. 1944;9(2):139–50. Available from: <http://www.jstor.org/stable/2086306>
26. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* [Internet]. 2009 Apr;42(2):377–81. Available from: <http://dx.doi.org/10.1016/j.jbi.2008.08.010>
27. World Map of Member Institutions in REDCap consortium [Internet]. REDCap. [cited 2021 May 29]. Available from: https://redcap.vanderbilt.edu/consortium/map_fullscreen.php

Figures

Figure 2: Time Trend of EDC Adoption rate

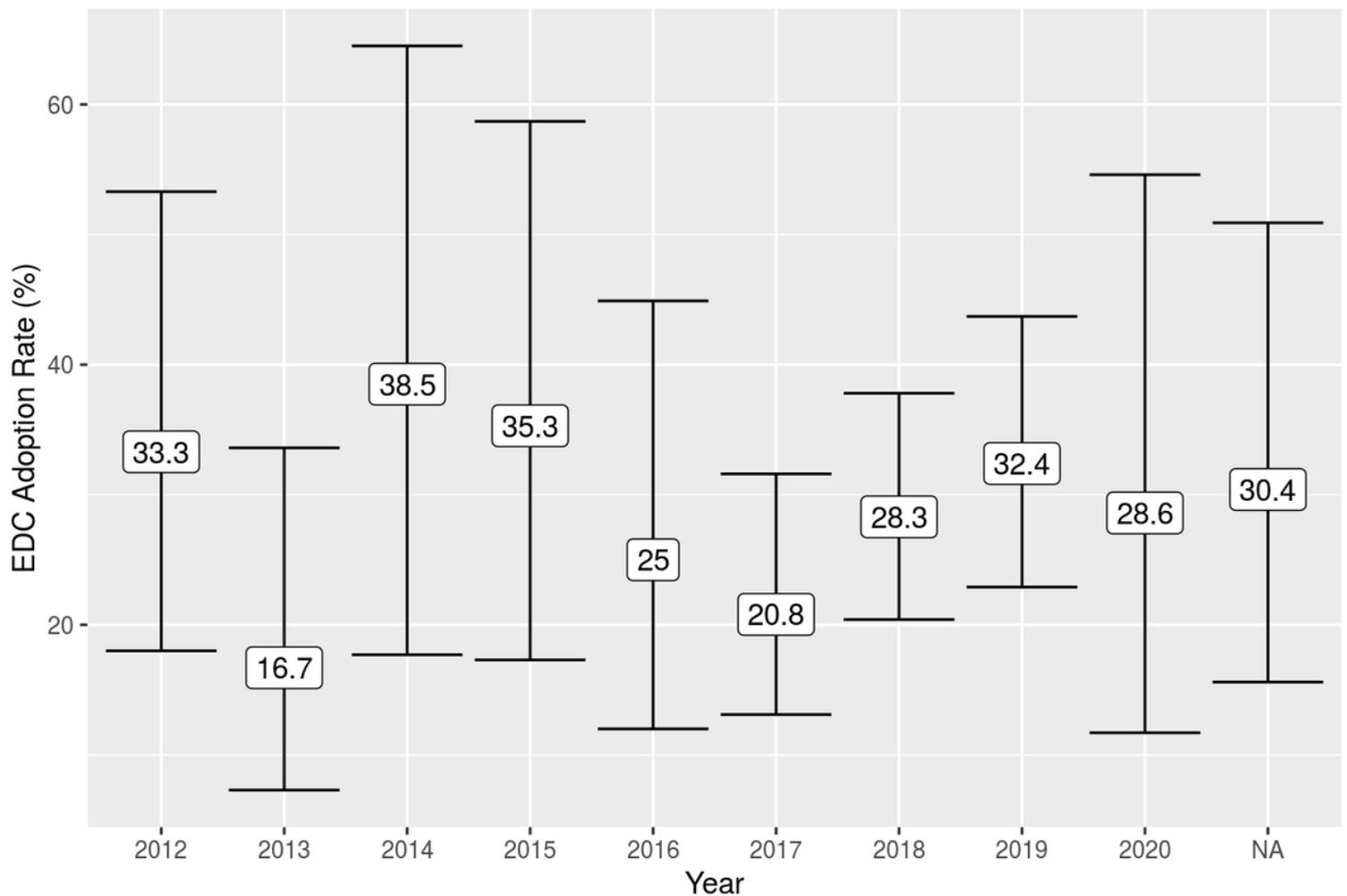


Figure 1

Shows the time trends in adoption of EDC across each year from 2012 - 2020. Numbers show the EDC rate in each year, and error bars represent the 95% confidence intervals. The year was obtained from the date the trial was registered in the clinical trial registry. If the date of registration was not available, then the year was marked as NA.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [AppendixIProtocolofSurveyofuseofElectronicDataCapturetoolsinrandomizedcontrolledtrialsinIndiaEDC2.pdf](#)
- [AppendixIIEDCSurveyPaper.docx](#)