Appendix II Electronic Data Capture Systems in India Survey

## Participation Statistics

The following table shows the statistics related to participation in the survey. The number of investigators who were contacted is different from the number of trials as for a trial more than one investigator may have been contacted.

*Table 1 : Participation Statistics*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Investigators (N)** | **Investigators who opened email (N)** | **Investigator open rate (%)** | **Trials (N)** | **Trials whose email was opened (N)** | **Trial open rate (%)** | **Participating Trials (N)** | **Participation rate (%)** |
| 2,890 | 1,523 | 52.7 | 1,909 | 1,141 | 59.8 | 400 | 21 |

## Included trial characteristics

First we take a look at the trial characteristics of all trials were were in the sampling frame.

*Table 2 : Trial Characteristics*

|  |  |
| --- | --- |
| **Characteristic** | **N = 1,9091** |
| **Sponsor Type** |  |
| Governmental | 111 (5.8%) |
| Industry | 302 (16%) |
| Institutional | 1,053 (55%) |
| Other | 443 (23%) |
| **Industry Funded Trial** |  |
| No | 1,607 (84%) |
| Yes | 302 (16%) |
| **Type of Trial** |  |
|  | 38 (2.0%) |
| Interventional | 1,841 (96%) |
| Observational | 26 (1.4%) |
| PMS | 4 (0.2%) |
| **Trial Phase** |  |
| Phase 1 - 2 | 299 (16%) |
| Phase 2 - 3 | 489 (26%) |
| Phase 4 | 333 (17%) |
| Unknown | 788 (41%) |
| **Duration of Trial (Days)** | 545 (365, 730) |
| Unknown | 6 |
| **Sample Size (Total)** | 84 (54, 150) |
| **Number of sites** | 1 (1, 1) |
| **Multicentric Trial** |  |
| No | 1,554 (81%) |
| Yes | 355 (19%) |
| **Country of Recruitment** |  |
| Indian | 1,736 (91%) |
| Multinational | 173 (9.1%) |
| **condition\_type** |  |
| Accidents and Injuries | 51 (2.7%) |
| Chronic non-communicable diseases | 100 (5.2%) |
| Diseases of Circulatory System | 96 (5.0%) |
| Diseases of digestive system | 189 (9.9%) |
| Diseases of Eye | 39 (2.0%) |
| Diseases of genitourinary system | 58 (3.0%) |
| Diseases of respiratory system | 55 (2.9%) |
| Endocrine disease | 155 (8.1%) |
| Infective Diseases | 165 (8.6%) |
| Mental Behavioural Disorders and nervous system disease | 122 (6.4%) |
| Neoplasms | 190 (10.0%) |
| Normal healthy volunteers | 43 (2.3%) |
| Others | 646 (34%) |
| 1n (%); Median (IQR) | |

Next we compare the trial characteristics of those studies which have responded versus those which have not.

*Table 3 : Trial characteristics compared between trials which participated in the survey and those that did not.*

|  |  |  |
| --- | --- | --- |
| **Characteristic** | **No, N = 1,5091** | **Yes, N = 4001** |
| **Sponsor Type** |  |  |
| Governmental | 86 (5.7%) | 25 (6.2%) |
| Industry | 266 (18%) | 36 (9.0%) |
| Institutional | 812 (54%) | 241 (60%) |
| Other | 345 (23%) | 98 (24%) |
| **Industry Funded Trial** |  |  |
| No | 1,243 (82%) | 364 (91%) |
| Yes | 266 (18%) | 36 (9.0%) |
| **Type of Trial** |  |  |
|  | 37 (2.5%) | 1 (0.2%) |
| Interventional | 1,450 (96%) | 391 (98%) |
| Observational | 19 (1.3%) | 7 (1.8%) |
| PMS | 3 (0.2%) | 1 (0.2%) |
| **Trial Phase** |  |  |
| Phase 1 - 2 | 228 (15%) | 71 (18%) |
| Phase 2 - 3 | 382 (25%) | 107 (27%) |
| Phase 4 | 268 (18%) | 65 (16%) |
| Unknown | 631 (42%) | 157 (39%) |
| **Duration of Trial (Days)** | 545 (365, 730) | 545 (365, 730) |
| Unknown | 6 | 0 |
| **Sample Size (Total)** | 83 (50, 150) | 90 (60, 160) |
| **Number of sites** | 1 (1, 1) | 1 (1, 1) |
| **Multicentric Trial** |  |  |
| No | 1,215 (81%) | 339 (85%) |
| Yes | 294 (19%) | 61 (15%) |
| **Country of Recruitment** |  |  |
| Indian | 1,364 (90%) | 372 (93%) |
| Multinational | 145 (9.6%) | 28 (7.0%) |
| **condition\_type** |  |  |
| Accidents and Injuries | 43 (2.8%) | 8 (2.0%) |
| Chronic non-communicable diseases | 72 (4.8%) | 28 (7.0%) |
| Diseases of Circulatory System | 74 (4.9%) | 22 (5.5%) |
| Diseases of digestive system | 151 (10%) | 38 (9.5%) |
| Diseases of Eye | 35 (2.3%) | 4 (1.0%) |
| Diseases of genitourinary system | 49 (3.2%) | 9 (2.2%) |
| Diseases of respiratory system | 42 (2.8%) | 13 (3.2%) |
| Endocrine disease | 123 (8.2%) | 32 (8.0%) |
| Infective Diseases | 145 (9.6%) | 20 (5.0%) |
| Mental Behavioural Disorders and nervous system disease | 92 (6.1%) | 30 (7.5%) |
| Neoplasms | 129 (8.5%) | 61 (15%) |
| Normal healthy volunteers | 34 (2.3%) | 9 (2.2%) |
| Others | 520 (34%) | 126 (32%) |
| 1n (%); Median (IQR) | | |

## EDC Adoption Rate

**EDC Adoption Rate (EAR)**: The primary outcome measure is EAR. This will be defined as the ratio of the number of CTRI registered trials that use an EDC with sophistication level 2 or more to that of the participating trials (unique CTRI registered trials for which investigators agreed to participate in the study. The proportion and the binomial 95% confidence intervals of the same will be reported.

The **EDC sophistication level** is defined as follows:

* **Level 1:** There is a unique account and password for each user to access the online system.
* **Level 2:** Sites enter subject visit data through a Web interface into electronic case report forms (eCRFs). The completion status of each eCRF for each subject can be tracked automatically online. The system provides an audit trail for all data entry and data modification
* **Level 3:** Data validation happens automatically when data are entered into the eCRF. The system will automatically log the user off after a period of inactivity.
* **Level 4:** Subjects are randomized automatically
* **Level 5:** Subject recruitment can be tracked online for each site
* **Level 6:** The system allows tracking of medication inventory at the sites.

For a level to be considered complete, **all the questions** should be marked as **Yes**. If one of the questions is marked as No and a higher level is marked Yes then the **higher level** will be taken. For each unique trial we will therefore calculate the highest EDC sophistication level. If EDC is not used then sophistication level will be marked as missing.

The following table shows the EDC adoption rate and the different levels in the trials for which responses were received in the survey.

*Table 4 : EDC use and adoption rate with EDC sophistication levels among responding studies*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variable** | **Total** | **Yes** | **Percentage** | **95% CI** |
| EDC ADOPTION | 400 | 110 | 27.5 | ( 23.4 - 32.1 ) |
| EDC USE | 400 | 130 | 32.5 | ( 28.1 - 37.2 ) |
| LEVEL 1 | 400 | 106 | 26.5 | ( 22.4 - 31 ) |
| LEVEL 2 | 400 | 76 | 19.0 | ( 15.5 - 23.1 ) |
| LEVEL 3 | 400 | 65 | 16.2 | ( 13 - 20.2 ) |
| LEVEL 4 | 400 | 75 | 18.8 | ( 15.2 - 22.9 ) |
| LEVEL 5 | 400 | 83 | 20.8 | ( 17.1 - 25 ) |
| LEVEL 6 | 400 | 64 | 16.0 | ( 12.7 - 19.9 ) |

The following table shows the breakdown of key trial characteristics by EDC adoption status. Comparison between groups has been done using Chi-square test for categorical variables and Wilcox rank sum test for continuous variables.

*Table 5 : Comparision of trial characteristics between trials with adopted an EDC and those that did not*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Characteristic** | **N** | **Overall, N = 4001** | **No, N = 2901** | **95% CI** | **Yes, N = 1101** | **95% CI** | **p-value2** |
| **Sponsor Type** | 400 |  |  |  |  |  | <0.001 |
| Governmental |  | 25 (6.2%) | 16 (5.5%) | 3.4% - 8.8% | 9 (8.2%) | 4.3% - 15% |  |
| Industry |  | 36 (9.0%) | 14 (4.8%) | 2.9% - 8.0% | 22 (20%) | 14% - 29% |  |
| Institutional |  | 241 (60%) | 184 (63%) | 58% - 69% | 57 (52%) | 42% - 61% |  |
| Other |  | 98 (24%) | 76 (26%) | 21% - 32% | 22 (20%) | 14% - 29% |  |
| **Industry Funded Trial** | 400 |  |  |  |  |  | <0.001 |
| No |  | 364 (91%) | 276 (95%) | 92% - 97% | 88 (80%) | 71% - 86% |  |
| Yes |  | 36 (9.0%) | 14 (4.8%) | 2.9% - 8.0% | 22 (20%) | 14% - 29% |  |
| **Type of Trial** | 400 |  |  |  |  |  | 0.7 |
|  |  | 1 (0.2%) | 1 (0.3%) | <0.1% - 2.4% | 0 (0%) |  |  |
| Interventional |  | 391 (98%) | 284 (98%) | 95% - 99% | 107 (97%) | 92% - 99% |  |
| Observational |  | 7 (1.8%) | 4 (1.4%) | 0.5% - 3.6% | 3 (2.7%) | 0.9% - 8.2% |  |
| PMS |  | 1 (0.2%) | 1 (0.3%) | <0.1% - 2.4% | 0 (0%) |  |  |
| **Trial Phase** | 400 |  |  |  |  |  | 0.076 |
| Phase 1 - 2 |  | 71 (18%) | 53 (18%) | 14% - 23% | 18 (16%) | 11% - 25% |  |
| Phase 2 - 3 |  | 107 (27%) | 72 (25%) | 20% - 30% | 35 (32%) | 24% - 41% |  |
| Phase 4 |  | 65 (16%) | 55 (19%) | 15% - 24% | 10 (9.1%) | 4.9% - 16% |  |
| Unknown |  | 157 (39%) | 110 (38%) | 33% - 44% | 47 (43%) | 34% - 52% |  |
| **Duration of Trial (Days)** | 400 | 545 (365, 730) | 545 (365, 730) | 640 - 766 | 545 (365, 865) | 637 - 916 | 0.5 |
| **Sample Size (Total)** | 400 | 90 (60, 160) | 86 (60, 149) | 119 - 178 | 100 (52, 204) | -405 - 3,638 | 0.2 |
| **Number of sites** | 400 | 1 (1, 1) | 1 (1, 1) | 1.1 - 1.4 | 1 (1, 3) | 2.6 - 5.7 | <0.001 |
| **Multicentric Trial** | 400 |  |  |  |  |  | <0.001 |
| No |  | 339 (85%) | 268 (92%) | 89% - 95% | 71 (65%) | 55% - 73% |  |
| Yes |  | 61 (15%) | 22 (7.6%) | 5.0% - 11% | 39 (35%) | 27% - 45% |  |
| **Country of Recruitment** | 400 |  |  |  |  |  | <0.001 |
| Indian |  | 372 (93%) | 282 (97%) | 95% - 99% | 90 (82%) | 73% - 88% |  |
| Multinational |  | 28 (7.0%) | 8 (2.8%) | 1.4% - 5.4% | 20 (18%) | 12% - 27% |  |
| **Access to Institutional CTU** | 400 |  |  |  |  |  | 0.3 |
| No |  | 195 (49%) | 146 (50%) | 45% - 56% | 49 (45%) | 35% - 54% |  |
| Yes |  | 205 (51%) | 144 (50%) | 44% - 55% | 61 (55%) | 46% - 65% |  |
| 1n (%); Median (IQR) | | | | | | | |
| 2Pearson's Chi-squared test; Fisher's exact test; Wilcoxon rank sum test | | | | | | | |

## Influence of trial parameters on EAR

Influence of trial parameters on EAR

To determine the influence of the trial parameters on EAR, we will use a logistic regression model where the dependent variable will be EDC adoption with EDC sophistication level 2 or more (modeled as Yes or No). Independent variables will be:  
1. Trial sponsor: Industry or Investigator-Initiated. In studies where the primary sponsor is a pharmaceutical company or device manufacturer, the user will be considered industry-sponsored, and the rest will be considered investigator-initiated.  
2. Trial sample size: Total trial sample size will be modeled as a continuous variable. To relax the linearity assumption, this will be expanded using a restricted cubic spline with 3 knots.  
3. Trial sites: The number of sites will also be modeled as a continuous variable. Again to relax the linearity assumptions, the model term will be expanded using a restricted cubic spline with three knots.

Interactions will be testing in an omnibus model containing all interaction terms. Wald test will be used for determining the significance of any interaction. Odds ratios with 95% confidence intervals will be reported.

## Wald Statistics Response: edc\_adoption   
##   
## Factor Chi-Square d.f.  
## sample\_size (Factor+Higher Order Factors) 6.30 6   
## All Interactions 2.62 4   
## Nonlinear (Factor+Higher Order Factors) 2.75 3   
## sites (Factor+Higher Order Factors) 11.31 3   
## All Interactions 1.39 2   
## industry\_funded (Factor+Higher Order Factors) 1.83 3   
## All Interactions 0.91 2   
## sample\_size \* sites (Factor+Higher Order Factors) 1.39 2   
## Nonlinear 0.46 1   
## Nonlinear Interaction : f(A,B) vs. AB 0.46 1   
## sample\_size \* industry\_funded (Factor+Higher Order Factors) 0.91 2   
## Nonlinear 0.81 1   
## Nonlinear Interaction : f(A,B) vs. AB 0.81 1   
## TOTAL NONLINEAR 2.75 3   
## TOTAL INTERACTION 2.62 4   
## TOTAL NONLINEAR + INTERACTION 3.65 5   
## TOTAL 24.95 8   
## P   
## 0.3905  
## 0.6233  
## 0.4315  
## 0.0101  
## 0.4994  
## 0.6095  
## 0.6335  
## 0.4994  
## 0.4958  
## 0.4958  
## 0.6335  
## 0.3674  
## 0.3674  
## 0.4315  
## 0.6233  
## 0.6008  
## 0.0016

As the results of the above ANOVA show, the Wald test for non-linear terms as well as interactions is not significant. Hence we show the simplified model without the interaction terms as well as without the non-linear assumption. The table below shows the results of the logistic regression analysis.

*Table 6 : Multivariable analysis of factors influencing EDC use*

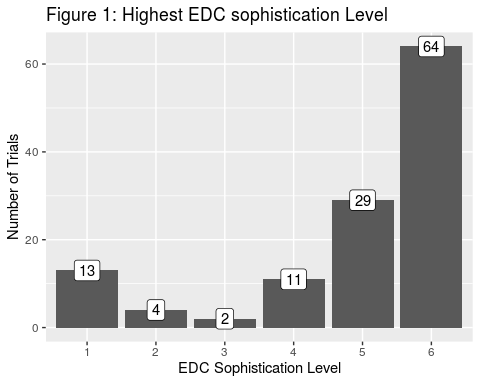
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Characteristic** | **N** | **Event N** | **OR1** | **95% CI1** | **p-value** |
| **(Intercept)** | 400 | 110 | 0.20 | 0.14, 0.27 | <0.001 |
| **sample\_size** | 400 | 110 | 1.00 | 1.00, 1.00 | 0.10 |
| **sites** | 400 | 110 | 1.26 | 1.12, 1.47 | <0.001 |
| **industry\_funded** | 400 | 110 |  |  |  |
| No |  |  | — | — |  |
| Yes |  |  | 2.14 | 0.86, 5.13 | 0.090 |
| 1OR = Odds Ratio, CI = Confidence Interval | | | | | |

## EDC Sophistication Level

We will provide data on the median EDC sophistication levels as well as a plot showing the proportion of CTRI registered trials with different levels of EDC sophistication. Further visualization and analysis will also explore the association between trial sample size, number of trial sites, and type of trial sponsorship with EDC sophistication.

*Table 7 : Highest level of EDC sophistication*

|  |  |
| --- | --- |
| **Characteristic** | **N = 4001** |
| Highest EDC sophistication Level |  |
| 1 | 13 (11%) |
| 2 | 4 (3.3%) |
| 3 | 2 (1.6%) |
| 4 | 11 (8.9%) |
| 5 | 29 (24%) |
| 6 | 64 (52%) |
| Unknown | 277 |
| 1n (%) | |



In the following table we will show the univariable analysis of the factors which influenced EDC sophistication level. We will dichotomize the level into two categories (score 6 or score 1-5).

*Table 8 : Univariate analysis of factors associated with EDC sophistication level*

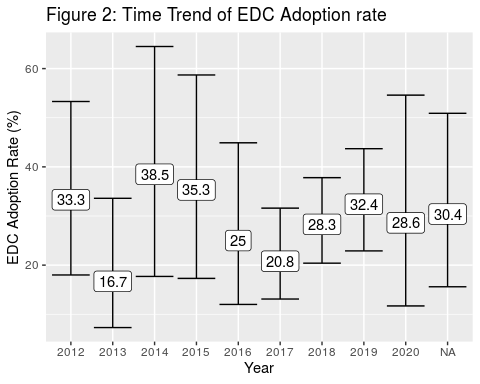
|  |  |  |  |
| --- | --- | --- | --- |
| **Characteristic** | **Level 1 - 5, N = 591** | **Level 6, N = 641** | **p-value2** |
| **Sponsor Type** |  |  | 0.12 |
| Governmental | 7 (12%) | 2 (3.1%) |  |
| Industry | 7 (12%) | 15 (23%) |  |
| Institutional | 30 (51%) | 34 (53%) |  |
| Other | 15 (25%) | 13 (20%) |  |
| **Industry Funded Trial** |  |  | 0.094 |
| No | 52 (88%) | 49 (77%) |  |
| Yes | 7 (12%) | 15 (23%) |  |
| **Type of Trial** |  |  | >0.9 |
| Interventional | 58 (98%) | 62 (97%) |  |
| Observational | 1 (1.7%) | 2 (3.1%) |  |
| **Trial Phase** |  |  | 0.9 |
| Phase 1 - 2 | 9 (15%) | 12 (19%) |  |
| Phase 2 - 3 | 20 (34%) | 18 (28%) |  |
| Phase 4 | 5 (8.5%) | 7 (11%) |  |
| Unknown | 25 (42%) | 27 (42%) |  |
| **Duration of Trial (Days)** | 730 (365, 1,095) | 411 (240, 730) | <0.001 |
| **Sample Size (Total)** | 153 (63, 293) | 64 (40, 128) | <0.001 |
| **Number of sites** | 1 (1, 2) | 1 (1, 2) | >0.9 |
| **Country of Recruitment** |  |  | 0.4 |
| Indian | 51 (86%) | 52 (81%) |  |
| Multinational | 8 (14%) | 12 (19%) |  |
| **Multicentric Trial** |  |  | 0.8 |
| No | 39 (66%) | 44 (69%) |  |
| Yes | 20 (34%) | 20 (31%) |  |
| **Access to Institutional CTU** |  |  | 0.071 |
| No | 20 (34%) | 32 (50%) |  |
| Yes | 39 (66%) | 32 (50%) |  |
| 1n (%); Median (IQR) | | | |
| 2Fisher's exact test; Pearson's Chi-squared test; Wilcoxon rank sum test | | | |

## EAR Time trends

For unique responding CTRI registered trials, we will create a subset containing trials registered on or after 1st January 2010. From this subset, we will then aggregate the EAR for each year based on the methodology for calculating EAR as above. This will be graphically demonstrated using a bar plot or a dot plot with a bar for each year. Note that as each trial is independent of each others, we will not use a line plot for the visualization. EAR will be compared between two time periods: period 1 from 1st January 2015 to 31st December 2019 and period 2 from 1st January 2010 to 31st December 2014. Given that most randomized trials will be completed by 10 years, we expect to have few open clinical trials available for analysis that was registered before 2010. However, if more than 30 trials are available, we will also analyze an earlier time point, i.e., between 1st January 2005 and 31st December 2009.

*Table 9 : EAR across time period*

|  |  |  |  |
| --- | --- | --- | --- |
| **Characteristic** | **Period 1 (2015 - 2019), N = 2861** | **Period 2 (2010 - 2014), N = 751** | **Pre 2010, N = 21** |
| EDC Adoption | 79 (28%) | 19 (25%) | 1 (50%) |
| 1n (%) | | | |



# Additional Analyses

Additionally the survey collected data on alternative methods for data collection used in the trial as well a single item question on the key perceived barriers towards adoption of EDC in their trial.

*Table 10 : Data collection methods used when EDC was not used*

|  |  |  |
| --- | --- | --- |
| **Characteristic** | **N** | **N = 2701** |
| **Spreadsheet** | 270 | 260 (96%) |
| **Data sent by Email** | 270 | 46 (17%) |
| **Data sent by Fax** | 270 | 9 (3.3%) |
| 1n (%) | | |

*Table 11 : Reason for not using EDC*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Reason** | **Total** | **Yes** | **Percentage** | **95% CI** |
| Lack of Technical Support | 270 | 170 | 63.0 | ( 57.1 - 68.5 ) |
| Software Cost | 270 | 132 | 48.9 | ( 43 - 54.8 ) |
| CTU Staff Motivation to implement EDC | 270 | 76 | 28.1 | ( 23.1 - 33.8 ) |
| Lack of user friendly Software | 270 | 71 | 26.3 | ( 21.4 - 31.9 ) |
| Complex Regulatory Requirements | 270 | 62 | 23.0 | ( 18.3 - 28.3 ) |

Other reasons identified for not using EDC were:

*Table 12 : Free text responses to reasons for not using EDC in trial*

|  |
| --- |
| **Reasons** |
| Small sample size |
| Data recording on paper has been the conventional method and it needs time and effort to make a change. All data collection personnel may not be comfortable using EDC and their training will be needed. |
| Lack of budget for procuring such software for the project because this is a single institution intra murally funded study which does not allocate funds for any personnel for datda recording/management |
| Standardized assessment tools can not be applied and analysed |
| Not aware of free software and concern regarding data safety |
| can not comment about my institution |
| Unawareness |
| Cost. No funding for trials |
| scholars can't afford for their research |
| Unaware about edc |
| I was unaware of this EDC system before this |
| Not heard of it when trial was conducted |
| unaware of EDC |
| Cost effectivness when volume of data collected is less and external sponsorship is not available. |
| Lack of knowledge about EDC |
| Not being used in our area so didn’t have access to it. |
| Certain complexity of design made it difficult to go with EDC |
| Study was not started at our site due to premature termination. Hence we could not access the EDC systems for the study |
| learnt about them much later after the start of the trial |
| EDC was not under dicussion / consideration at thta time |
| Was not aware. |
| Not funded by the institution |
| We have not thought about the possibility of an electronic software. As the department of cardiology and the hospital are next door, and since it is a single centre study, we thought of collecting data and write in paper |
| this was an academic trial. We did not have the resources available in 2012 to use EDC for an academic study. |
| At that time lack of information that how effective this could have been |
| Not aware |
| Lack of awareness of it and proper institutional software. |
| not enough hardware with internet connection in the premises -so difficult to implement online EDC. Also hospital does not have a ready to use software which can be used across trials. found it too cumbersome to set it up for one trial |
| We have the competency to use EDC,which we are using for some of the Phase 2 ongoing studies. But for this particular study Sponsor,has provided us SOW of only paper CRF study |
| I did not know about it |
| Lack of awareness |
| Sponsor did not inform us or provide us with any EDC software or platform. |

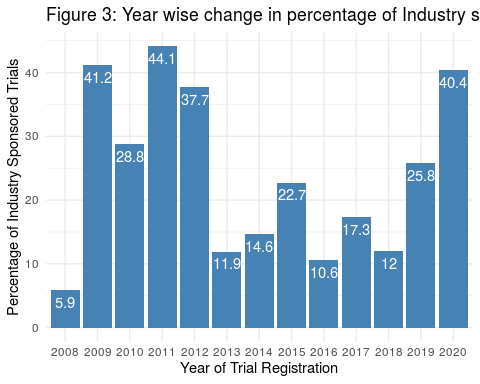
Finally two additional questions were asked about the trial center weather they had access to a CTU and an IRB. We will evaluate the data in relation to EDC use.

*Table 13 : EDC adoption by CTU and IRB availability*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Characteristic** | **N** | **No, N = 2901** | **Yes, N = 1101** | **p-value2** |
| **resource** | 400 |  |  | 0.2 |
| Both |  | 135 (47%) | 54 (49%) |  |
| None |  | 146 (50%) | 49 (45%) |  |
| Only CTU |  | 9 (3.1%) | 7 (6.4%) |  |
| 1n (%) | | | | |
| 2Fisher's exact test | | | | |

# Industry Sponsored trials

The percentage of industry sponsored trials by each year of registration is shown in the figure below.



# Packages used

1. R : R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL [https://www.R-project.org/.](https://www.r-project.org/.)
2. Tidyverse : Wickham et al., (2019). Welcome to the tidyverse. Journal of Open Source Software, 4(43), 1686, <https://doi.org/10.21105/joss.01686>
3. gtsummary : Daniel D. Sjoberg, Michael Curry, Margie Hannum, Joseph Larmarange, Karissa Whiting and Emily C. Zabor (2021). gtsummary: Presentation-Ready Data Summary and Analytic Result Tables. [<https://github.com/ddsjoberg/gtsummary>,](https://github.com/ddsjoberg/gtsummary,) <http://www.danieldsjoberg.com/gtsummary/.>
4. Hmisc : Frank E Harrell Jr, with contributions from Charles Dupont and many others. (2021). Hmisc: Harrell Miscellaneous. <https://hbiostat.org/R/Hmisc/,> <https://github.com/harrelfe/Hmisc/>
5. flextable : flextable: Functions for Tabular Reporting. <https://ardata-fr.github.io/flextable-book/,> <https://davidgohel.github.io/flextable/.>
6. rms : Frank E Harrell Jr (2021). rms: Regression Modeling Strategies. <https://hbiostat.org/R/rms/,> <https://github.com/harrelfe/rms.>
7. ggplot2: H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.
8. Lubridate: Garrett Grolemund, Hadley Wickham (2011). Dates and Times Made Easy with lubridate. Journal of Statistical Software, 40(3), 1-25. URL <https://www.jstatsoft.org/v40/i03/.>