BEACONCURE

# RESULTS OF SURVEY INTO STATISTICAL PROGRAMMING PRACTICES

Orit Veinerman,
Vice President of Marketing, Beaconcure

# INTRODUCTION

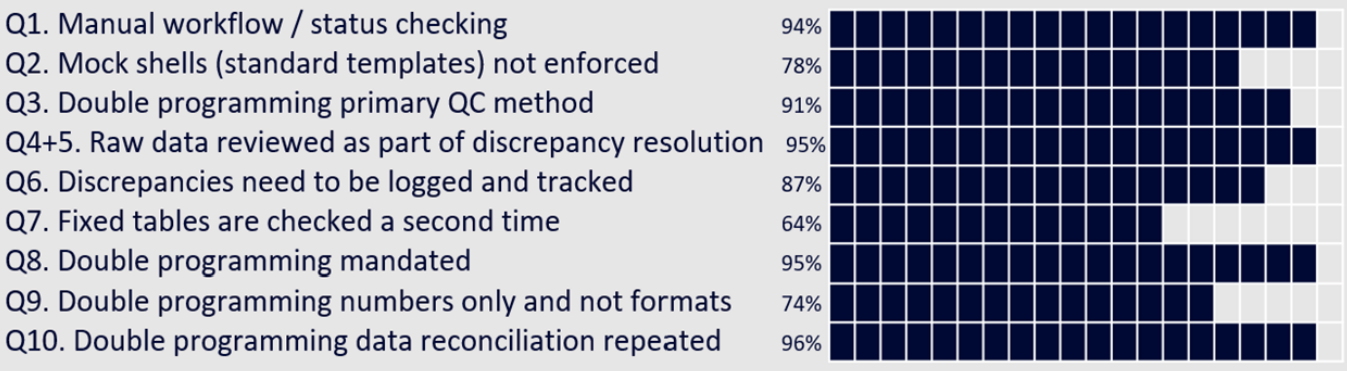
A 10-question survey was sent to 45 pharmaceutical companies and CROs in January 2021. The objectives were to discover the current methods used in the production and validation of programs developed to produce tables for use in clinical study reports and, based on the results, to determine how automation, including machine learning and artificial intelligence, might be used to increase efficiency, reduce costs and increase quality. Twenty-five responses were received from 22 companies. Sixteen responses were from pharmaceutical companies and nine from CROs. Respondents were mostly from senior level leaders in the Statistical Programming function. All responses were anonymous and therefore there was no follow up to clarify answers or to ask further questions.

## RESULTS

The key findings are shown in Table 1.

### INDUSTRY SURVEY

25 responses from 22 companies: 64% sponsor organizations, 36% from CROs



## Q 1

For managing the production of output, one company relies on email, all others depend on manual tracking in Excel or some other tool. There appears to be no automation of this process, with direct links to the statistical computing environment.

## Q 2

In the majority of cases, the mock shell tables are not enforced, and changes can be made in all but one case. However, in twenty cases the mock shells are updated to reflect changes in the table format. In only four cases do the final mock shells not reflect the final tables. This synchronization is achieved manually.

## Q 3

Results are discussed in Q8 .

## Q 4 - 5

All but one respondent reported that the raw data must be reviewed in at least some studies to resolve discrepancies in output, with nine reporting that raw data has to be looked at in all studies. Some sponsors do not review the raw data if discrepancies are found, instead requiring the CRO to do the investigation. If the data that was responsible for the discrepancy could be automatically displayed during the review of discrepancies, the resolution process would be greatly improved. Currently reviewers have to access the raw data or listings manually and locate the data that caused the discrepancy.

## Q 6

In all but three cases, discrepancies need to be logged and tracked through to resolution. This process is typically manual with no linkage between the table and the tracking tool to facilitate the tracking of the process. Such a connection would significantly accelerate the process.

## Q 7

In eight cases, the double programming validation is not rerun after the original program has been fixed. The automation of the checking of the output, would mitigate the risk these companies are taking and streamline the process for those that currently do validate programs once a discrepancy is resolved.

Q8

# DOUBLE PROGRAMMING

**(Including Q3, Q9, Q10)**

Twenty-three of the companies use double programming to validate output before submissions. Eleven also use manual validation. Two companies use some form of automated checking, but the scope and nature of this automation is unknown. There is significant room to increase the automation of the QC process. Only one company does not use double programming at some point in the production of output (one company uses double programming in the early stages of program development but not just before a submission).

Seventeen respondents use it for all tables and seven for a subset of tables. In nearly three quarters of cases, only the numbers are validated by double programming; format, wording and footnotes are not. In one case, as mentioned above, the double programming is only run during the development of the program. In the remaining 24 instances, the double programming process is repeated if the data changes. In nine of these cases double programming is also repeated if the format, wording, and footnotes change. Double programming is time-consuming and error prone. Alternative, automated, processes to eliminate the need for double programming would be beneficial.

# CONCLUSION

There is considerable uniformity of the processes used in the production of output used in the analysis of clinical trials. In all cases, the standard process would benefit from automation and integration of the processes. Double programming could be eliminated, relevant data discrepancies could be automatically provided to the reviewer, the current labor-intensive management of table production and discrepancy resolution could be replaced by an integrated automated process, and the consistency between the mock shells could be enforced by technology, removing the need for reliance on a manual process.

# WHO WE ARE

Established in 2016, Beaconcure is an Israeli high-tech startup located in Israel with an office in the US.
We developed an ML/NLP-driven analytics platform that is designed to support the life science industry. Our first-to-market product 'Verify Outputs' converts any clinical data format and layout into a semantic & dynamic database, enables to applies any required check and detects 99.7% of discrepancies within hours.
Our technology helps our clients to significantly improve the quality of their outputs by identifying programming issues, specification deficiencies, and data inconsistencies. It also reduces costs, and accelerates decision making and time to market.

**For more information on implementing a state-of-the-art validation solution, visit Beaconcure today**
**https://beaconcure.com/ hello@beaconcure.com**