

Table 8: Big Data - How Are Analytical Teams Addressing This, and Automation in Biopharmaceutical Development- An Answer or the Answer?

SCOPE:

Biological structures, such as cells, tissues or whole organs are one of the most complex systems on earth. Cellular responses to ligand interaction can lead to many different effects. For example, the binding of the rather simple molecule insulin to its receptor introduces at least twelve different pathways [1]. Since all these effects do not stand alone but may influence each other, it could be of great benefit if all derived measurement data would be analyzed in context. However, if a certain amount of data is exceeded, it becomes very difficult for humans to analyze it. Could the concept of big data be the answer to this?

QUESTIONS FOR DISCUSSION:

1. What do you think about big data? Will it have the praised benefits or is it overhyped?
2. What is your experience with chemo informatics? Do you already use chemometrics or other types of big data in your Lab? What are the main difficulties to implement it?
3. What decisions may be taken by AI in the future and which under no circumstances?
4. How would you rate the degree of automation in your field? Where do you see potentials for more automation?

[1] De Meyts P. The Insulin Receptor and Its Signal Transduction Network. [Updated 2016 Apr 27]. In: Feingold KR, Anawalt B, Boyce A, et al., editors. Endotext [Internet]. South Dartmouth (MA): MDText.com, Inc.; 2000-. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK378978/>

DISCUSSION NOTES:

Only one attendee participated in this table so the questions were skipped and some interesting points were highlighted

- For measure certain ligand binding assays, it seems that the published ligand binding data (for example dissociation constants) are quite uncertain, and there are many reasons for this matter.
With a normal portfolio of analytical techniques is feasible to address these issues but it is far easier when we use algorithms to sort it, for example to classify which factors are more important to a particular assay.
- The way specific data is gathered and storage has improved a lot over the years. Without a consistent data storage there was not a system in place for identification and registration of proteins for example. This evolution allowed to systematic connect the assay data with

basic things like proteins sequences and composition. Which served as a leverage to tap into machine learning approaches

- In general, the topic of big data and algorithms do make sense for biologics products, for process and product characterization. According to the parameters you want to assess, a toolbox of machine learning can be of extreme help.