



European Federation of Statisticians in the Pharmaceutical Industry
Representing Statistical Associations in Europe



Organisational Evolution:

The statistical/data science organisation
today and in the future

Justine Rochon

Virtual Statistics Leaders Meeting 2020

1. Which kind of company are you working for?

	Pharmaceutical Company	20
	CRO	2
	Other	2

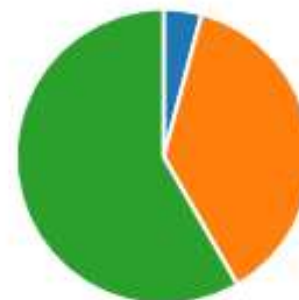


2. If "Other" please specify

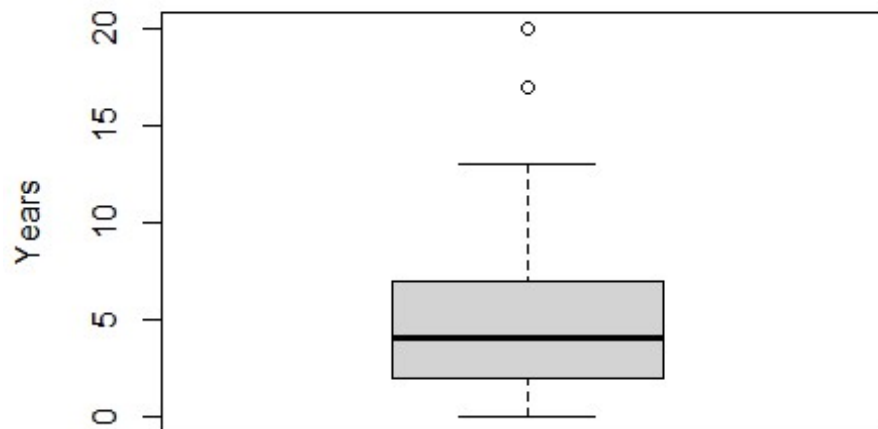
- Nutrition
- FMCG

3. What is the size of your company?

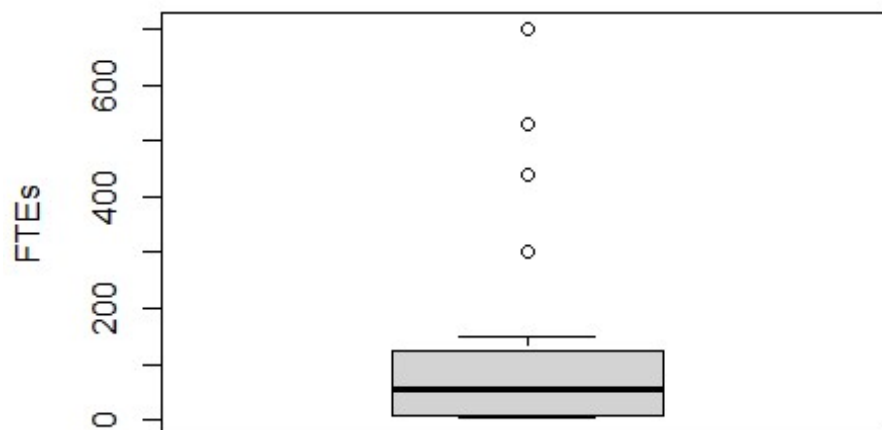
 small-sized	1
 middle-sized	9
 big	14



4. How long have you been in your current position (in years)?



5. What is the size of your statistics department in terms of FTEs?



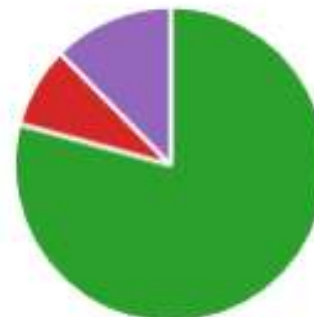
6. Is your Statistics department a global or a regional/local department?



Virtual Statistics Leaders Meeting 2020

7. What is the one core business of your Statistics department?

 Non-Clinical Statistics	0
 Statistics for Translational Me...	0
 Clinical Statistics	19
 Statistics for Medical Affairs	2
 None of the above	3



Virtual Statistics Leaders Meeting 2020

7. What is the one core business of your Statistics department?

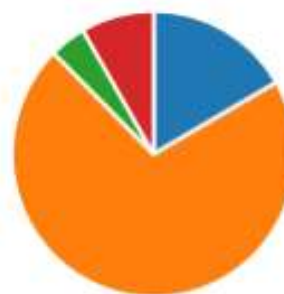
8. Please specify

- Actually we do all except for Non-Clinical
- Clinical statistics
- Statistics/Pharmacoepidemiology/Data Science/Scientific writing
- I am leading a statistical team in Germany supporting global Medical affairs and Market Access
- All of the above
- Biostatistics for the development projects. Biostatisticians are working mainly in the role of the project statistician
- Data sciences for Danone. so covering Media, toppling, sales, operations , procurement, HR, finance, R&I...etc
- "We don't have one core business of Biostats, we have the following: 3 Clinical Statistics team, Oncology, Non-Oncology, Vaccines, Non-Clinical and Translational Statistics including CMC, Research Statistics, Biomarkers, Translational Statistics, Statistical Data Science, Clinical Programming and Governance, Advance Biostatistics and Data Analytics, India Biostatistics Group, Regional Biostatistics teams in Japan and China (dotted line to central team)"
- I assume question 7 refers to a specification of question 6: Then to add here that in our Statistics department all roles mentioned in question 6 as options can be found and besides Clinical Statistics, TMCP Statistics and Medical Affairs Statistics play an equal role.
- Department covers early development and non-clinical Biostatistics (Early Clinical Biostatistics Phase I-IIa, Translational Medicine, CMC, Non-Clinical Efficacy and Safety)
- Focus on non-interventional studies. PASS.
- Focus is Clinical Statistics but we contribute to Non-Clinical Statistics, TMCP, Medical Affairs as well
- Phase 2-4 support but we cover all phases from non-clinical up to market access.

Virtual Statistics Leaders Meeting 2020

9. Whom do you consider your main internal 'customer'?

Research	4
Medicine	17
Marketing	1
Other	2

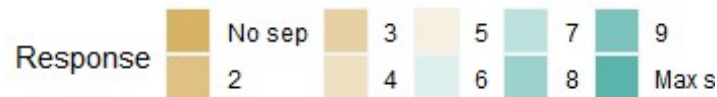
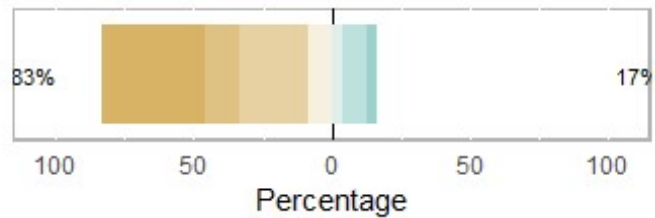
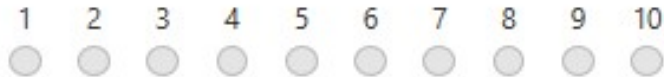


- Multiple. Research, medical, marketing all play a part but other functions also are key (eg medical writers, data managers)
- Multiple customers in R&D, Medical Affairs, Value Evidence and Outcomes including HTA, Regulatory & Safety, Portfolio Governance

Virtual Statistics Leaders Meeting 2020

11. How strictly is your Statistics organization separated into strategic versus operational part?

1 = no separation, one organization vs. 10 = max separation, two organizations



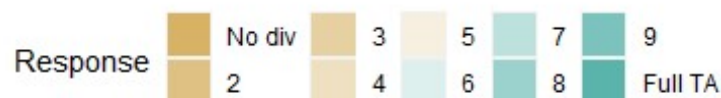
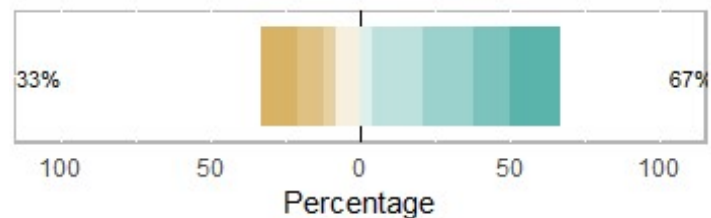
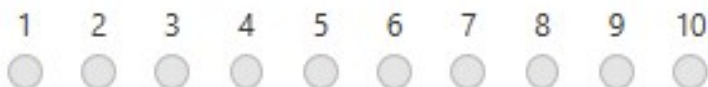
12. Textbox for additional explanation to the last question

- There is separation in what people do as they develop skills etc. but no separate track or structure, just done by assignment
- Statistical production has been extensively outsourced (platform). Less mature activities (Pepi) or activities requiring large interactions with internal partners (ex medical teams) are internalized
- We have a separate strategic group for scientific consultation and guidance but all statisticians report into the one organization and interact together.
- No study operations in-house; full contracting out model for Biometrics
- The statisticians have combined responsibility of strategic and operational roles
- We contract out most of our operational work and concentrate on strategy
- There is a separation somewhat but not strictly speaking
- No formal separation but some resources (depending on expertise, experience, and leadership skills) focus more on strategic or operational tasks
- It's one organisation. However, operational activities on study level are outsourced
- Operations highly outsourced.

Virtual Statistics Leaders Meeting 2020

13. To which extent is your Statistics organization oriented towards therapeutic areas?

1 = no division, vs. 10 = full division into TAs



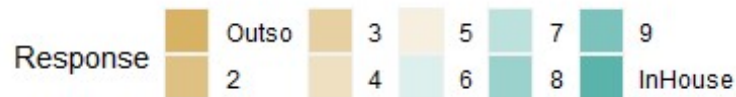
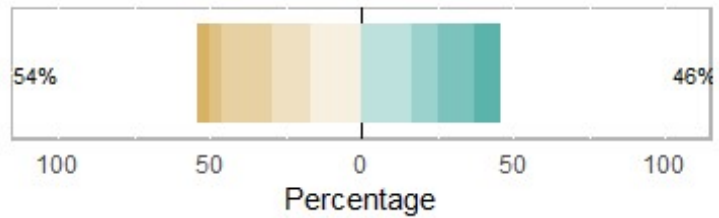
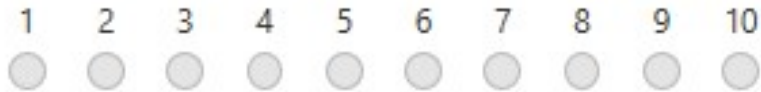
14. Textbox for additional explanation to the last question

- Only 2 main subteam by group of TAs (Rare Diseases vs Other)
- Basically three groups. Early development (mainly CP), oncology and everything else
- Teams working in Oncology are generally dedicated to this TA
- We have TA alignment for segments of our business but staff will work across TAs too.
- organization is also too small for separations; only separation is in terms of regions (but with connotation of flexible resourcing across)
- Statistical resources are allocated to the TAs.
- it is in functions in my case but only for a specific amount of time. it is part of SAL with business. specialized resources but for an agreed timeframe e.g 2 years so to keep resources motivated and learning
- Development Statistics is broken into 6 therapy areas, separate Oncology and Vaccines teams
- Different statisticians are assigned to certain TAs, but we are so small that we often have to work across projects
- My organization is a platform for early development for all TA but Oncology. Some People are experts in specific TAs. Late Phase organization in the Company is organized by TAs.
- No division for programming & division for statistics (general comment about the overall department)

Virtual Statistics Leaders Meeting 2020

15. What is the percentage of your in-house activity (vs. outsourced)?

1 = 0 % all outsourced, vs. 10 = 100 % in-house activity



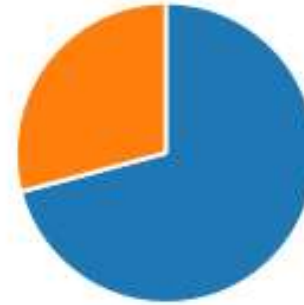
Virtual Statistics Leaders Meeting 2020

16. Textbox for additional explanation to the last question

- All operational activities outsourced, all scientific / strategic / regulatory activities internally
- Outsource much study level but not project level, so most operational outsource
- For a more drug/project-oriented collaboration, more methodological/statistical/analytical are led and carried out internally benefiting from technical platform offers. Outsourced projects are focused on lack of expertise or access to technology or data
- with respect to clinical studies operations; the design of the studies is with us (contents) and also some data utilization projects (re-use of data)
- We have a mix of FTE, Functional Service Provider and Full Service Outsourcing resource models
- We outsource production programming but do ad hoc analyses in house
- Individual Trials are outsourced
- Inhouse are strategic activities. Operational activities are mainly outsourced or off-shore

Virtual Statistics Leaders Meeting 2020

17. Do you consider your Statistics department as the custodian of the data?



Virtual Statistics Leaders Meeting 2020

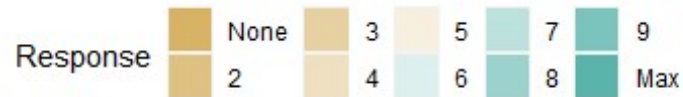
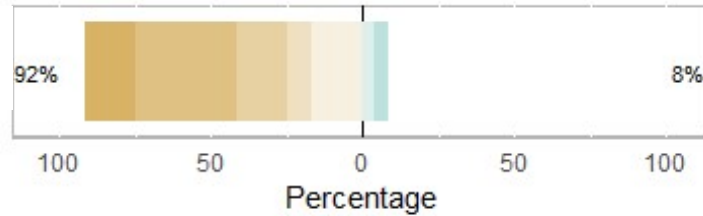
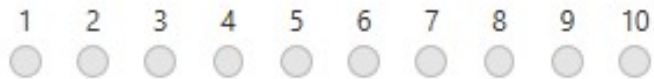
18. If "No": How is the custodian of the data / access to data organized in your department?

- We are joint custodians with our Data Management department. They ensure data integrity and store the raw data. We support with ensuring data integrity and utilize/manage the data for mapping, analysis and reporting.
- via data excellence group and IS/IT custodian. We follow the DAma structure with MDM team, Data excellence team and data sciences team.
- Data Office
- I would describe statistics as custodians of analyses of the data. The data may be accessed via browsing tools for clinical review (blinded), but not to conduct analyses.
- It depends on the data. Clinical data for Primary use are handled by Biostatistics. For secondary use data and samples an Independent process is being established.

Virtual Statistics Leaders Meeting 2020

19. To which extent are 'self-service tools' used and enable researches to conduct the analyses on their own? *

1 = no self-service tools, 10 = 100% self-service tools



20. Textbox for additional explanation to the last question

- Not much but moving more in that direction
- Mainly biologists are using self service tools
- We have tools to navigate the data and dive into patient detail for safety reviews but otherwise all analyses are performed by statisticians.
- Statistical consultancy is needed in addition to the "self-service-tools", further there is need to develop these tools
- part of our strategy is to provide self service BI tools to whoever need it. it helps in maintaining a frugal data organisation and allow for data citizens in business Organisation while DS/stat can be working on more at scale use case
- Later this year we are rolling out a new platform which will enable researchers to access data and run their own exploratory analyses. We also have a data repository where researchers can request access to data and then run their own analyses.
- We are providing some data visualization tools via R for our PV group
- Not yet many self-service tools available but efforts under way
- Self-service tools are available to perform medical review of the data from individual patients and to look for trends in blinded summaries.
- Not established yet, but some activities discussions are ongoing.
- Used for medical review and safety surveillance
- Progress on going for providing more "administrative" pictures on the data in a self-service mode

Virtual Statistics Leaders Meeting 2020

21. Did you already implement agile working models and techniques (e.g. SCRUM, Kanban,...) in your Statistics department?

 Yes

6

 No

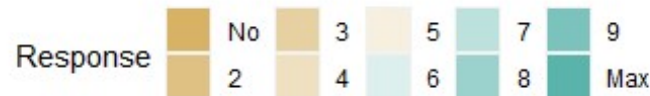
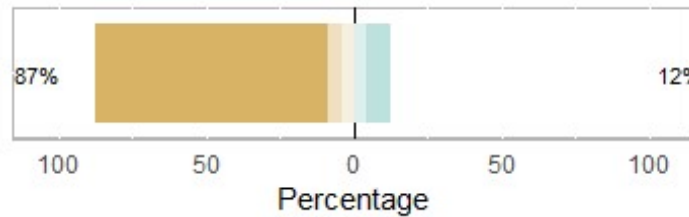
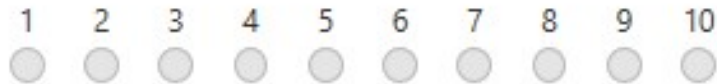
18



Virtual Statistics Leaders Meeting 2020

22. If "Yes", to which degree did you implement agile working models and techniques (e.g. SCRUM, Kanban,...) in your Statistics department?

1 = no implementation, vs. 10 = fully implemented



23. If "Yes", since when have you implemented agile working methods?

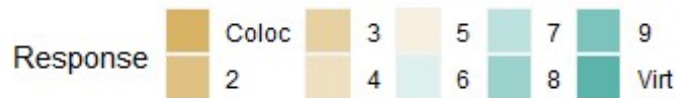
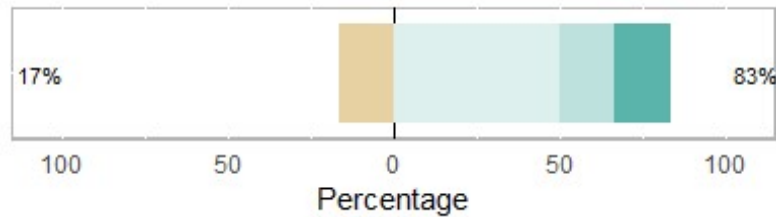
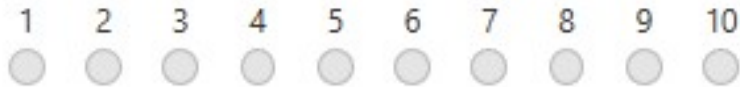
● less than 1 year	1
● 1 - 2 years	5
● 2 - 5 years	0
● > 5 years	0



Virtual Statistics Leaders Meeting 2020

24. Do those agile teams work co-located or virtually together?

1=fully co-located to 10=fully virtual



Virtual Statistics Leaders Meeting 2020

25. Is there a Data Science subgroup within your department?



Virtual Statistics Leaders Meeting 2020

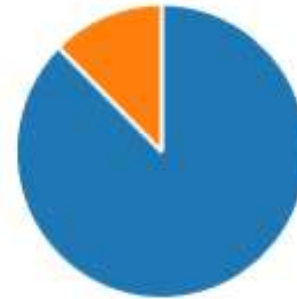
26. If "Yes", what is the composition of this Data Science group with regard to job profiles?

- Statisticians or Data Scientists
- Data Scientists, Statisticians
- So far own data science resources are limited, focus on the collaboration with academics and selected vendors. Recruitment plan in place
- mixed
- in R&I a few DS profiles vs stats profiles - 50%
- Statisticians, data scientists, programmers
- Statisticians
- Mixed: statisticians and data scientists
- Statisticians, Programmers, Data Managers, Business Analysts
- Statisticians and Data Scientists

27. If "Yes", what is the main focus of this Data Science group?

- Subgroup discovery, targeted population, Images analysis, bibliographic research,
- Device Data
- Machine learning and signal detection - identification of right treatment for right patients using data from several data sources (clinical, wearables, omics etc)
- Data mining in complex clinical and non-clinical datasets
- RWE, external data sources, AI, etc
- Ai enabled formulation, image recognition, other ways of analyzing clinical data
- Data engineering and developing tools and applications to support predictive analytics
- Data analysis of Real World Evidence sources
- Clinical Development
- Predictive Analytics

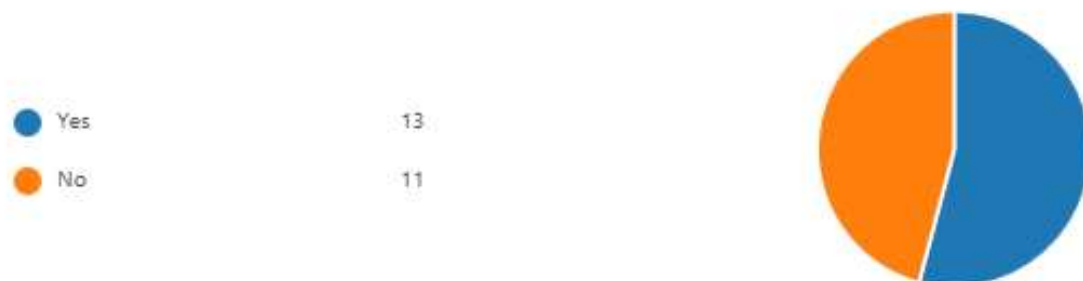
28. Are there other Data Science groups in your company?



29. If "Yes" please indicate which groups and which department they belong to

- IT, Scientific Knowledge
- AIA basically a RWD group
- We have an eData group and then also a Data Analytics group that support clinical monitoring group that both have data scientists/
- Data Science group in Strategy supporting projects in Manufacturing, Operations, and Finance
- Digital Transformation, IT, Marketing, Real World Evidence, Research
- Computer Aided Med Design and Biomarkers
- to Product and Development, Marketing Access, Consumer groups, Chief Data Officer teams and IT/IS
- specific group, by TA
- Finance
- R&D Tech, Computational Biology, Digital data analytics plus others
- We are part of the Data Science group. There are also separate Data Science groups within each TA.
- There are other Data Science groups or groups of employees working in a quantitative science. They are located for example in Departments close to Market Access but also in R&D
- Data Science groups within our IT organisation who also serve non-clinical and early research and marketing.
- Digital Data Science
- multiple
- IT, Finances/Marketing, HEOR...
- IT, Innovation Lab, Research, Marketing
- IT

30. If Data Science and Statistics are positioned in different departments, are they multidisciplinary teams to federate ML/AI projects for R&D perspective?

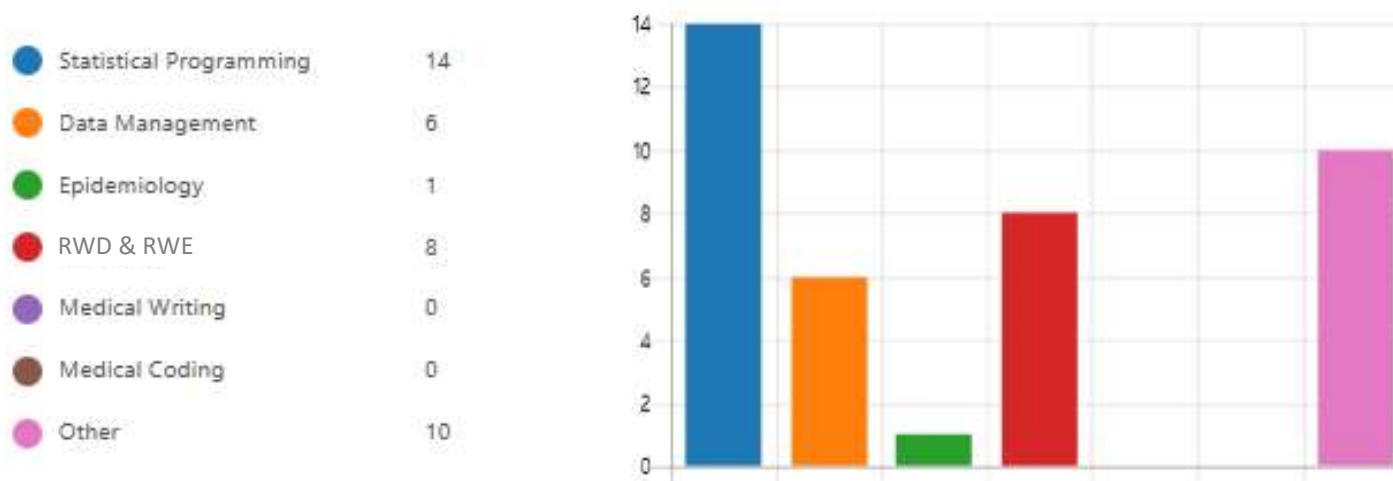


31. Do Statisticians nevertheless apply data mining and ML technics for their projects independently from any Data Scientist team?



Virtual Statistics Leaders Meeting 2020

32. Are there subgroups within your Statistics Department?



33. If "Other" please specify

- Automation
- Pharmacometrics
- within subteams (ex clinical subteam) of project/drug teams
- PK also for all other areas mentioned above we have separate groups for these (epidemiology, RED and RWE are with our late stage specialist group, medical coding sits in our data management department)
- Clinical Technology (CTMS, PRO-tool, digital initiatives in Clinical; wearables, apps)
- Digital applications
- Data science project lead
- See response to Question 7
- Statistical Innovation
- 2 programmers in the group
- There are no subgroups in the Statistics Department but certainly DM and Programming as well as other groups form a quantitative Department of its own
- ECB, Biomarker Data Science, CMC, Non-Clinical Efficacy & Safety
- Data Transparency, Independent Safety Group

34. Are there any additional teams within your department which are ensuring a smooth interface to enabling functions like IT, Sourcing, HR, Legal...?



Virtual Statistics Leaders Meeting 2020

35. If "Yes", please specify these teams

- IT, Outsourcing Department, RWE department
- programming solution sort of the gateway to IT etc.
- Biostats quality group, innovation team. All other services aligned with those mentioned above are separate departments within our company.
- IT and sourcing
- Business Operation group
- data transformation which who is taking care of portfolio management, data academy, career path, agile set up...etc
- "R&D Tech, HR, Finance, Ethics & Compliance"
- IT, Vendor Management, QM
- Digital Trials, Vendor Management, Clinical Data Systems, Clinical Data Standards, Quality, Validation & Compliance

Virtual Statistics Leaders Meeting 2020

36. If "No", how do you ensure smooth interfaces to enabling functions like IT, Sourcing, HR, Legal...?

- Systems
- Including within these subteams additional members (permanent or adhoc)
- We have many cross functional leadership groups, SMEs and initiatives. Points of contact are available for all groups.
- the "usual" way: interactions, 1-1's, project teams, etc.
- via leadership team and assignments
- portfolio manager, and team leaders
- For IT there is a link person who is focused on Data Science needs.
- With some considerable work
- This is handled with the overall statistics management responsibilities
- AA or through stats resources
- "Managed by the overall department. There a special systems group for contacts with IT is in place / chief of staff in place for interactions with planning functions"
- Some individuals are nominated in shared groups.

37. Is there a specialized Quality team within your department?



Virtual Statistics Leaders Meeting 2020

38. What are the 3 technical (hard) skills you consider will become essential to develop in your group in the next five years?

- Pharma experience, excellence in statistics and true collaboration
- modelling, machine learning, causal inference
- Statistical methodology to support decision making and Probability of Success; Data Mining, Machine Learning & AI; Statistics in the context of RWE studies
- RWD/ use of external data for design and analysis. exploratory and proper individualized medicine approaches,
- Causal inference; joint analysis of data from different sources (incl real world data and big data); reliable inference from smaller datasets
- Adaptive designs (incl external arms and complex trails). ML.
- Estimands, wider usage and understanding of R, platform trials
- programming in R (Python?, Julia?); Epidemiology (due to RWE); Missing Data / Imputation / Estimands
- - Adaptive design - Modelling and simulations -ML/AI methods
- Bayesian methodology, Utilization of RWD, Exploratory data analysis
- AI/ML, RWD/RWE and advance/complex statistical designs
- high dimensional data techniques in microbiome and MRI type of data, ML, causal relationship for non-RCTs,
- competency (complex /optimal designs), RWE experience, strategic abilities

38. What are the 3 technical (hard) skills you consider will become essential to develop in your group in the next five years?

- Coding in industrialized manner (ML eng), developing new also to gain competitive advantage
- Quantitative decision making (predictive inference), Predictive analytics, Bayesian statistics
- Causal Inference, analysis of recurrent events and analysis techniques relevant to large scale datasets created by the increased use of sensors.
- Development of estimands, statistical skills to support CMC, use of comparator data from external sources
- Quantitative Methodology beyond traditional statistical approaches, Visualization knowledge and how to tell the data story to people outside of statistics, Knowledge of programming in R, Python, R-Shiny
- R programming
- Skills to make use of ML/AI approaches and big data
- I wouldn't focus on a special skillset, it's at the end the whole spectrum on the methodology and practical side (e.g. graphical principles how to visualize)
- R programming. Dynamic presentations. Machine Learning.
- Causal Inference, RWD/RWE Integration in Clinical Development, Adaptive Trial Designs
- R programming, Bayesian Statistics, Data Science/machine learning

39. What are the 3 leadership (soft) skills you consider will become essential to develop in your group in the next five years?

- Influence, leadership, partnership
- training; being able to engage in meaningful dialog with non-statisticians; ability to reach across statistical disciplines
- Holistic approach of drug development (biomedicine, competitors, external data sources, master protocols, ScBRA...). Convincing skills. Risk management
- Managing in a remote environment, coaching/mentoring in a remote environment, scientific consultation skills
- intrapreneurship; agility; competitiveness/speed
- - Connect clinical questions to statistical questions - Ability to explain complex concepts in simple terms - Have the statisticians being more strategic and thinking out of the box
- Leading Teams, Influence, Agility
- Change management, motivation, feedback
- project leadership skills, explaining role of statistics and quality of data to N+2 (SrVP level), dare to take (calculated) risks
- negotiation skills, presentation skills,
- data translator = communication skills, assertivity, capabilities to take a step back
- Communication, Influencing, Leadership
- Team working between data scientists and statisticians. Communication and influencing skills (in a virtual environment). Stakeholder management skills.
- Negotiation, regulatory interactions, good communication (writing)
- Leading by intent, delegate accountability & exert smart control such that outcomes are acceptable, increase efficiency & effectiveness of work force in times of flat head counts and accelerating development speed
- None
- Consultancy skills, Negotiation skills, Influencing skills
- listening, networking, and communication
- Strategic thinking. Vision. Clear and simple thinking.
- Storytelling, Decision making under uncertainty, Coaching
- Influencing skills, presentation skills, negotiation skills

Virtual Statistics Leaders Meeting 2020

40. What are the technical (hard) skills you consider your group should train on your internal partners (biologists, clinicians, ...) in order to benefit for data access democratization?

- Basic statistical science
- FAIR principle, statistical thinking
- Statistical principles, Statistical/Regulatory guidelines, Estimands, Probability of Success
- bias, understanding of inference at the right level
- Consider the question you wish to answer.
- Data sourcing. Problem definition. Reliability/Robustness of knowledge generation
- How to effectively reviewing data dashboards, how to focus only on what is important to reduce costs and time
- Data Visualization (Power BI, Dashboards of Clinical Data)
- - Data Integration and Statistical Programming concepts
- evidence generation, reproducibility, bias avoidance
- Omics data, data collected from digital tools (wearables etc), RWD
- protecting false positive findings- avoid overinterpretation, understand structure of data, e.g data within the same patient are not necessarily independent

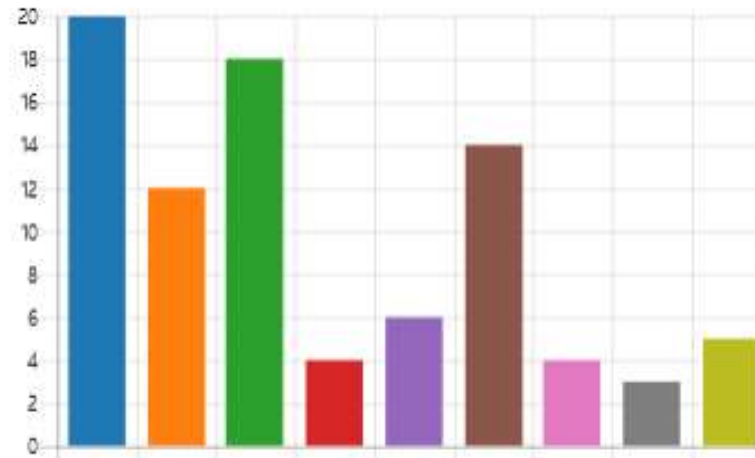
Virtual Statistics Leaders Meeting 2020

40. What are the technical (hard) skills you consider your group should train on your internal partners (biologists, clinicians, ...) in order to benefit for data access democratization?

- Awareness of bias and sampling schemes challenges
- Data language: MDM, data governance principles, product owner and data ownership
- Fundamentals of Statistics - design & analysis, Quantitative decision making, Bayesian methods, Prior elicitation
- Statistical theory to develop a better understanding of how to do high quality research with a solid well defined clinical question. Utilize the Estimand framework as a way of doing this. Give examples of how not to do meta-analysis to emphasize the importance of prespecification.
- Impact of estimands. Use of data visualization tools
- Collaborate effectively, understand needs and background of one another, Risk literacy (e.g. when it comes to quantitatively based go/no go decisions)
- None
- Basic statistical concepts for decision making
- Type of data and summarization / intro. into statistical analysis; how to handle missing data and multiple comparison
- Data sharing.
- FAIR
- Interpretation of data

Virtual Statistics Leaders Meeting 2020

41. Please indicate the critical roles you have or plan to have in your organization of the future?

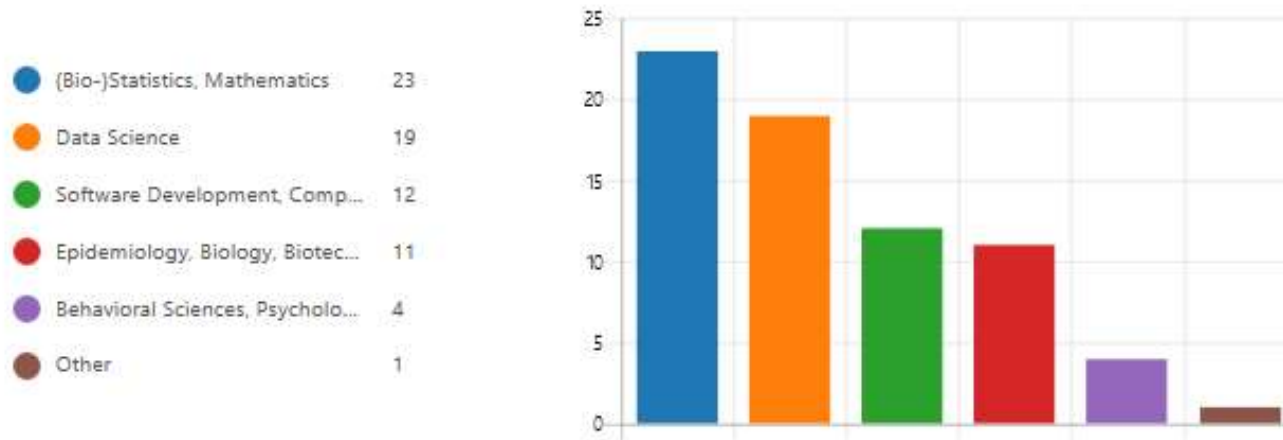


Virtual Statistics Leaders Meeting 2020

42. If you have answered "Others" please mention profile here

- Pharmacometrics
- Coordinator of multi-disciplinary/expertise teams
- Pharmacokineticists
- Epidemiologist(s); Behavioural Specialist
- portfolio manager
- Advanced methodology experts, Computer scientists
- Methodologists

43. Please indicate the critical capabilities in terms of academic backgrounds you plan to have in your organization of the future?



Virtual Statistics Leaders Meeting 2020

44. If "Other" please specify

- Results publication (medical writing) capabilities allows to cover the whole chain from data to communication
- I am not certain yet the kind of impact Machine Learning/AI will have in area of Clinical Development. It will for sure in terms of diagnostics, but not sure about treatment pathways and even part of regulatory interactions.

45. Do you plan to implement agile working models and techniques (e.g. SCRUM, Kanban,...) in your Statistics department?



Virtual Statistics Leaders Meeting 2020

46. If answered with "Yes" please specify

- Changing POC etc, to exploration approach and very small number of planned analysis.
- Separate process development project is ongoing
- For fast and relative small improvement projects, like digital tools
- Already introduced but need to further embed across the department
- Already started