

# Connecting statisticians and scientists: statistical training as a key tool to ensure research integrity and reproducibility within pharmaceutical drug discovery

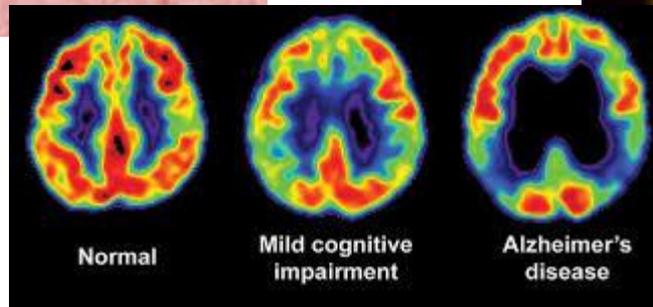
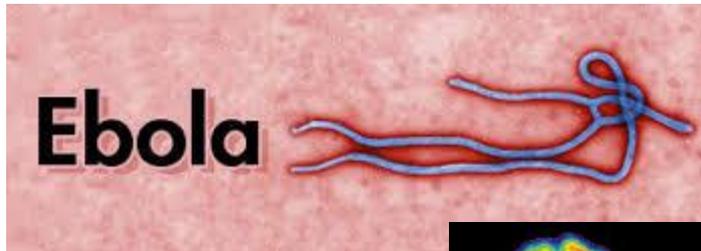
**Tom Van de Castele, Marjolein Crabbe, Bie Verbist, Kjell Johnson<sup>\$</sup>, Fetene Tekle, Luc Bijnens, ... *et al.***

**Janssen R&D, Beerse, Belgium and <sup>\$</sup>Arbor Analytics LLC, Ann Arbor, US**

[tvdcaste@its.jnj.com](mailto:tvdcaste@its.jnj.com)

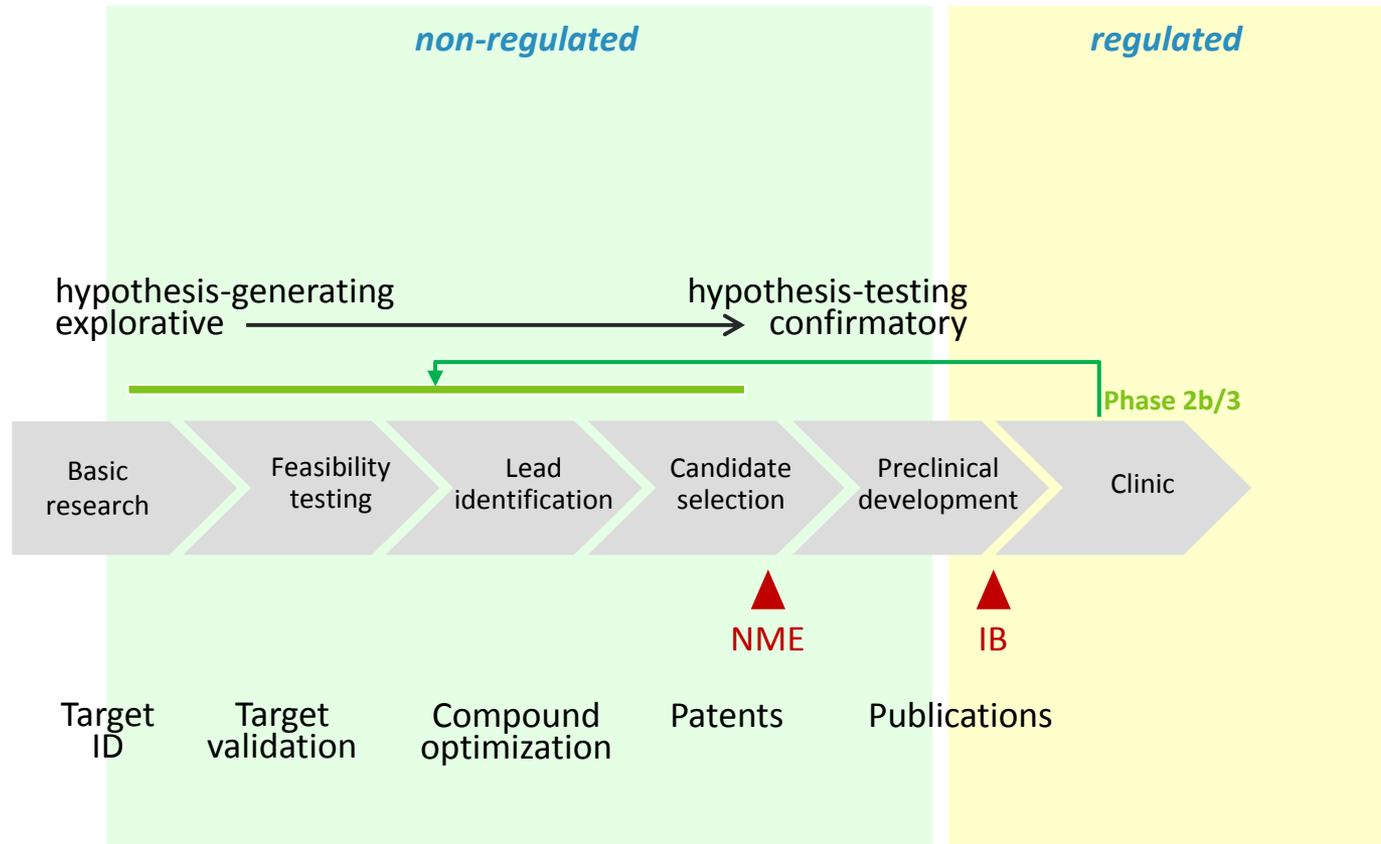
# Drug development in pharmaceutical industry

- Unmet medical need



- Develop the most effective treatment with minimal safety risk as soon as possible and in a cost-efficient manner

# "The" drug discovery process



A stepwise scientific research process determined by a sequence of data-driven go/no go decisions

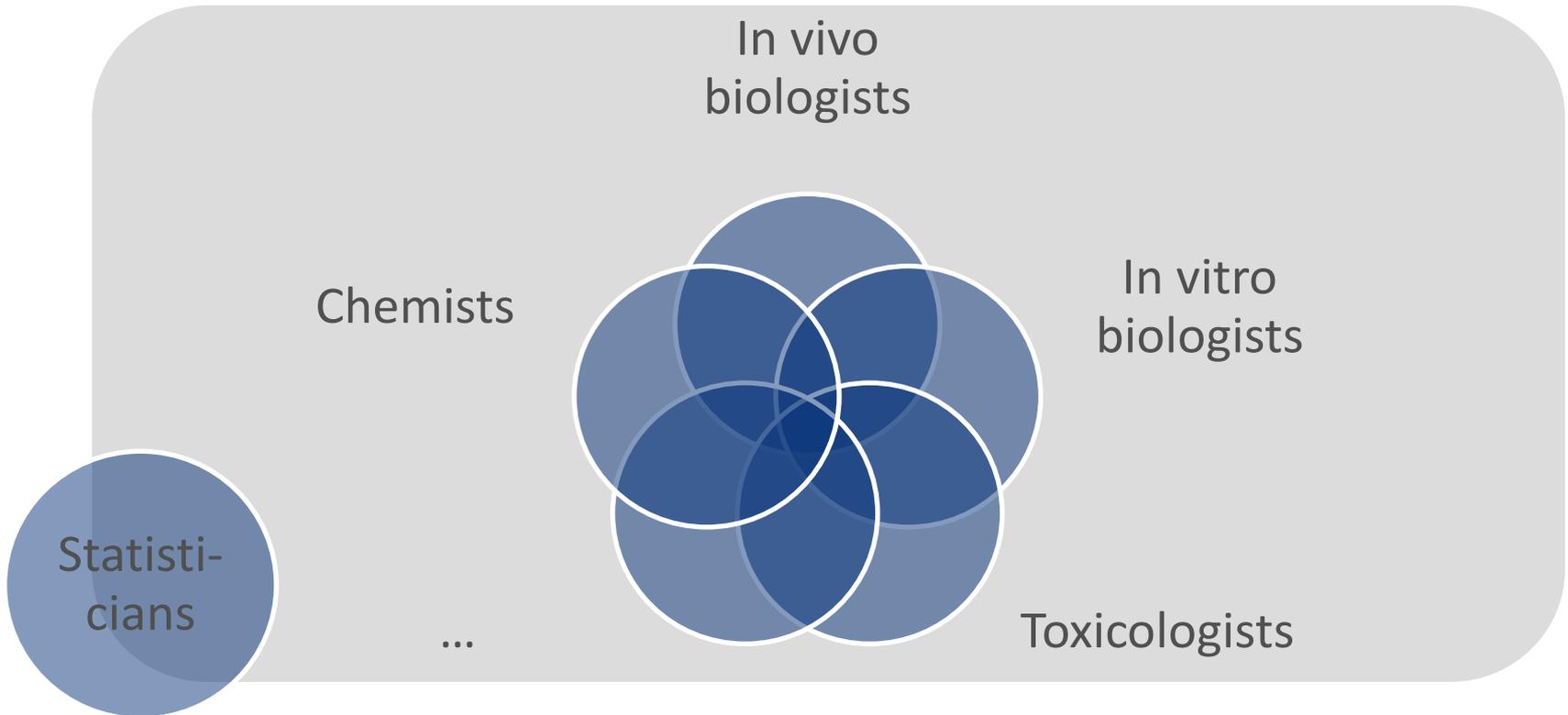
# Research objectives across drug discovery process

- Explore
- Select/filter
- Optimize
- Compare
- Estimate
- Predict

Learn/confirm, integrate, translate

Different research objectives require different statistical methods

# Drug discovery team



Statistical skills?



**Genesis 11: 1-9**

“If as one people speaking the same language they have begun to do this, then nothing they plan to do will be impossible for them. Come, let us go down and confuse their language so they will not understand each other.” So the Lord scattered them from there over all the earth, and they stopped building the city.

# An operational challenge

- appr. **50** parallel discovery programs (post Ti/Tv, pre-NME)

**500+**

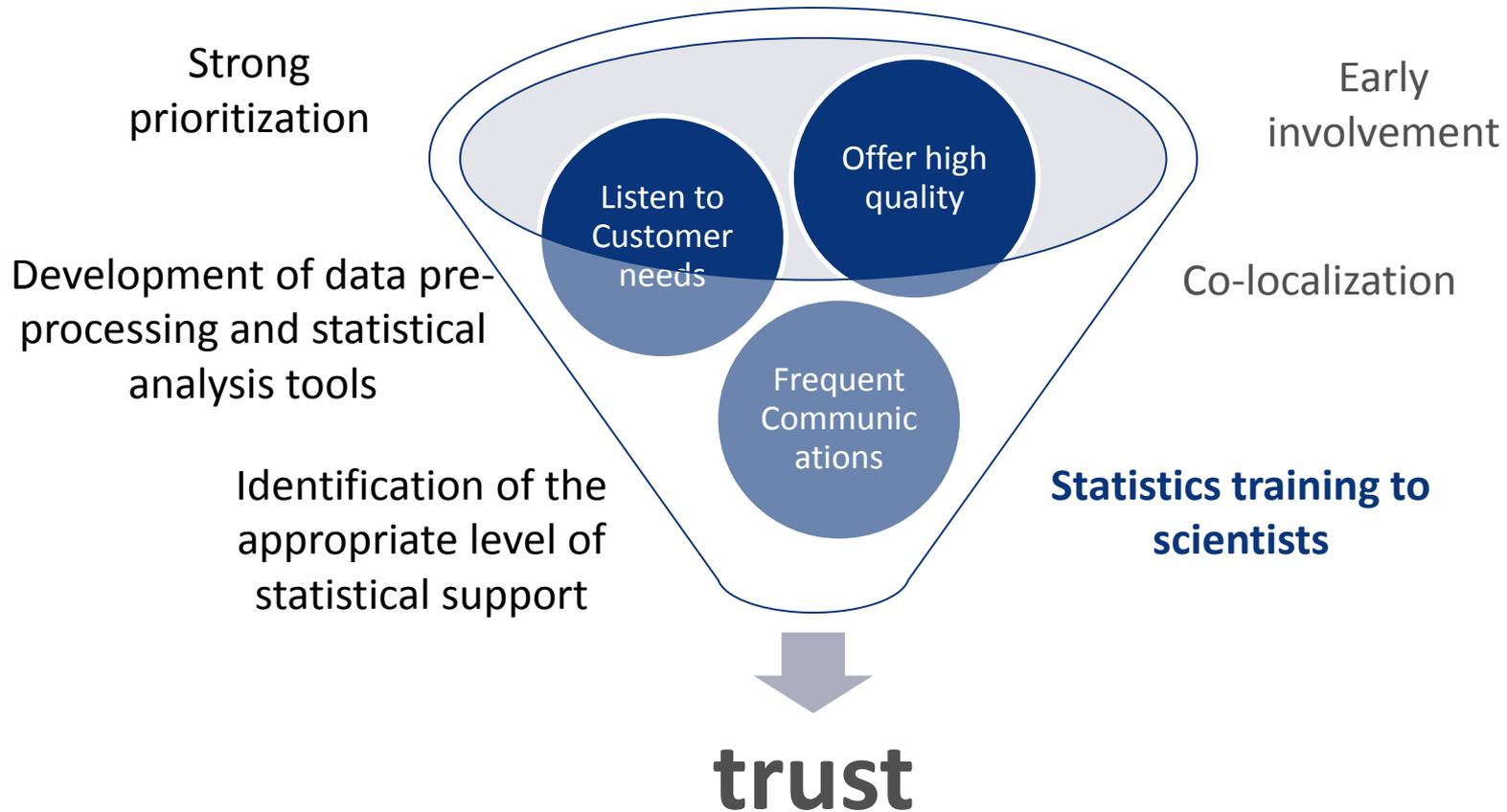
discovery  
scientists

**15**

discovery  
statisticians

- Limited guidelines and standard operating procedures to streamline the drug discovery process

# Multidisciplinary collaboration built on trust



# A diverse portfolio of statistics trainings

**Statistical awareness  
lectures**

**Hands-on training  
GraphPad Prism**

**Workshop  
“Statistical thinking”**

**Mentoring/coaching  
program  
“Statistics champions”**

# Statistical awareness lectures

## Aims:

(Re-)introduce basic concepts, eliminate misconceptions, understand variability, emphasize the importance of design

## Format: 1 hour lectures during lunch

- Interactive sessions, appealing examples, avoiding mathematical formula's as much as possible

## Target audience:

Anyone interested

Topics: inspired by the “Points of significance” papers in Nature Methods (<http://mkweb.bcgsc.ca/pointsofsignificance/bibliography.mhtml>)

Basic: Population, sampling, variability,..., Error bars, Significance, P-values and t-tests, Comparing samples, Dose-response analysis and outliers; Advanced: Sources of variation, replication, and design of experiments, Power and sample size

# Hands-on training GraphPad Prism

## Aims:

Make scientists self-reliant to conduct (simple) statistical analyses themselves

## Format:

2 hour class sessions

## Target audience:

Anyone interested

## Topics:

Summarizing data and comparing two groups, Comparing multiple groups, Analysis of dose-response data, Outliers

# Workshop “Statistical thinking”

## Aims:

Make scientists think about statistical aspects of experimental designs

## Format:

1 full day offsite, class session, team work

## Target audience:

Scientists who are responsible for the development of protocols and who are designing in vivo studies

## Topics:

Solve a fictitious case study of a research problem that has nothing to do with their own work without collecting data and statistical analysis - see Vandebroek et al. (2006) J. Biopharm. Stat. 16: 61-75

# Mentoring/coaching program “Statistics champions”

## Aims:

Develop expertise to support frequently executed protocols with a clear scientific question (hypothesis) in mind:

To translate a project-related scientific question into statistical model

Implementation of statistical concepts and analysis software

Accurately interpret and report results

After the training, champion will be able to facilitate other team members in supporting good statistical practices and identify critical aspects to study design

## Format:

Class training and individual guidance by a statistician

## Target audience:

Scientists who are responsible for the development of protocols and who are designing in vivo/pivotal studies; at most one scientist of each **biological team**

**Topics:** to be discussed with scientists and biological lead

# Scientists' attendance

- Statistical awareness lectures: +/- 50 scientists
- Hands-on training Graphpad Prism: +/- 80 scientists
- Workshop “Statistical thinking”: +/- 25
- Mentoring/coaching program “Statistical Champions”: +/- 10 registered

# Feedback from scientists

1. Overall very positive feedback on training initiatives, and particularly on hands-on sessions
2. Positive feedback on covered content
3. Clear interest in training on experimental design aspects (power and sample size)
4. Clear relationship between interest and time to be invested (duration and time during the day): format of trainings is important
5. Some reluctance to engage in Statistical Champion program, due to anticipated workload on top of research activities - but need to await experiences from first participants

# Conclusions

- Statistics training within drug discovery teams can:
  - create better visibility for the statistics department within the drug discovery organization,
  - stimulate statisticians and scientists to better understand each other's challenges,
  - facilitate identification of statistical problems with different levels of complexity that require different solutions,
  - pave the way for the establishment of common research quality standards and the successful implementation of a translational drug discovery research strategy
  - be a key tool to achieve better connectivity between scientists and statisticians
- Development of statistics training material is a time investment worth the effort
- Scientists' preferences for training topics should be taken into account to acquire strongest impact

# Thanks to

- Colleague statisticians:
  - Goeyvaerts, N. , Janssen R&D, Beerse, Belgium
  - De Smedt, H. , Janssen R&D, Beerse, Belgium
  - Geys, H. , Janssen R&D, Beerse, Belgium
  - Barron, A., Janssen R&D, Raritan, US
  - Cherkas, Y., Janssen R&D, Spring House, US
  - Lubomirski, M., Janssen R&D, Spring House, US
  
- Non-regulatory QA:
  - Gilis, A. , Janssen R&D, Beerse, Belgium
  - Lavrijssen, T. , Janssen R&D, Beerse, Belgium
  - Steckler, T., Janssen R&D, Beerse, Belgium
  
- Discovery Sciences: Gallacher, D., Janssen R&D, Beerse, Belgium
  
- Janssen R&D therapeutic area heads and their teams