# A rational approach to risk reduction: what can discovery screening planners learn from volcanoes and dust detection?

Andrew Chadwick
Montreux, June 2010 andrew.chadwick@tessella.com

Tessella
Technology & Consulting

# A rational approach to R&D risk reduction

- The recent problems with volcanic ash and European and Transatlantic air travel* remind us that it is difficult to strike the balance between risk of inaction (possible accidents) and risk of action (economic losses)
- Even with limited data we have to set a threshold on a prediction, or measurement for taking action
- Pharma R&D, which has to improve performance and yet has too much variety to naively apply Six Sigma principles, is faced with the same challenges – in particular, how to optimise its safety screening
- We have developed a way of helping discovery project managers handle risk management decisions, i.e. *planning under uncertainty*:
  - First identify the tradeoffs between upside and downside consequences
  - Then help people visualise the impacts of partly uncertain information (such as potential project value)
  - Demonstrate how this can help practical decisions, with examples of past industry choices (e.g. choice of cutoff in toxicity computational prediction)
  - We offer support in
    - communicating the opportunity for change
    - providing simple toolsets that support rapid learning
    - analysis of the relevant data and decisions
    - continuing knowledge management and capture of prior information on risk and reliability
    - partnering with other providers who specialise in chemoinformatics and multi-parameter optimisation

*To say nothing of deepwater drilling for oil and gas …

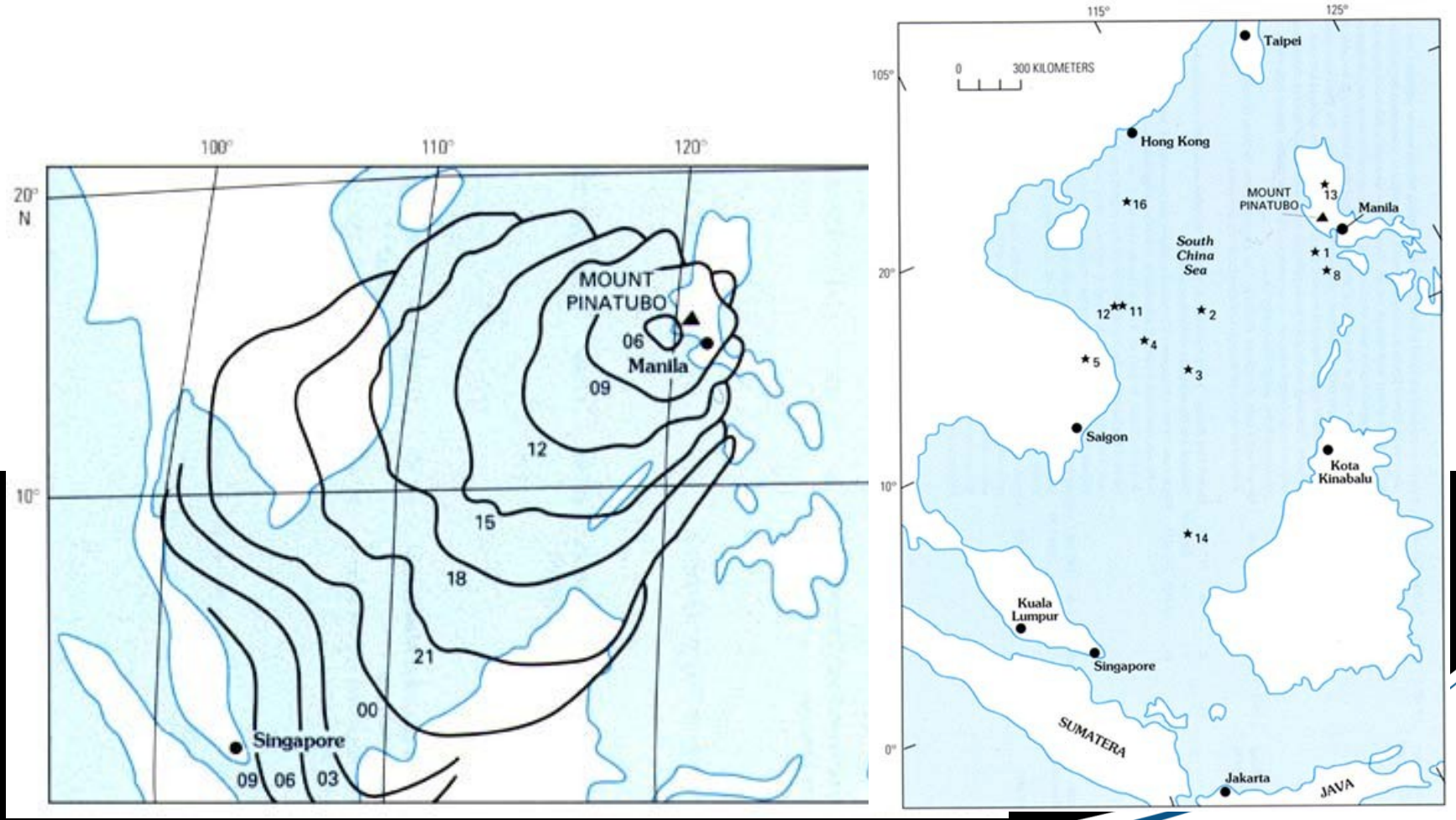
Tessella
Technology & Consulting

# A volcano is a dread threat. Toba nearly wiped out the human race about 70,000 years ago

Eruption of Mount Pinatubo in 1991

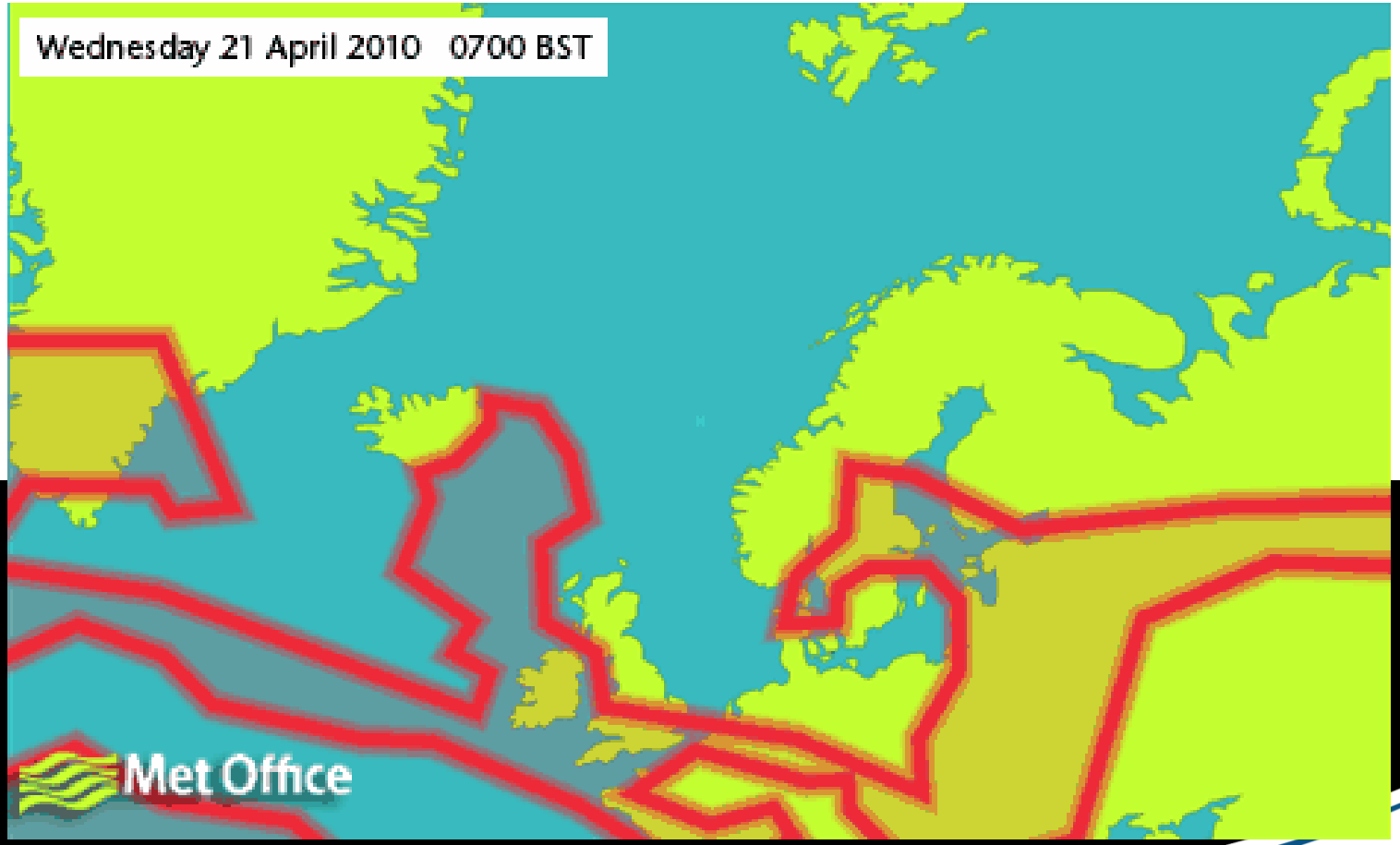
Tessella
Technology & Consulting

# Satellite (GMS) observations 1991 and reported aircraft incidents



# No smoke without fire?

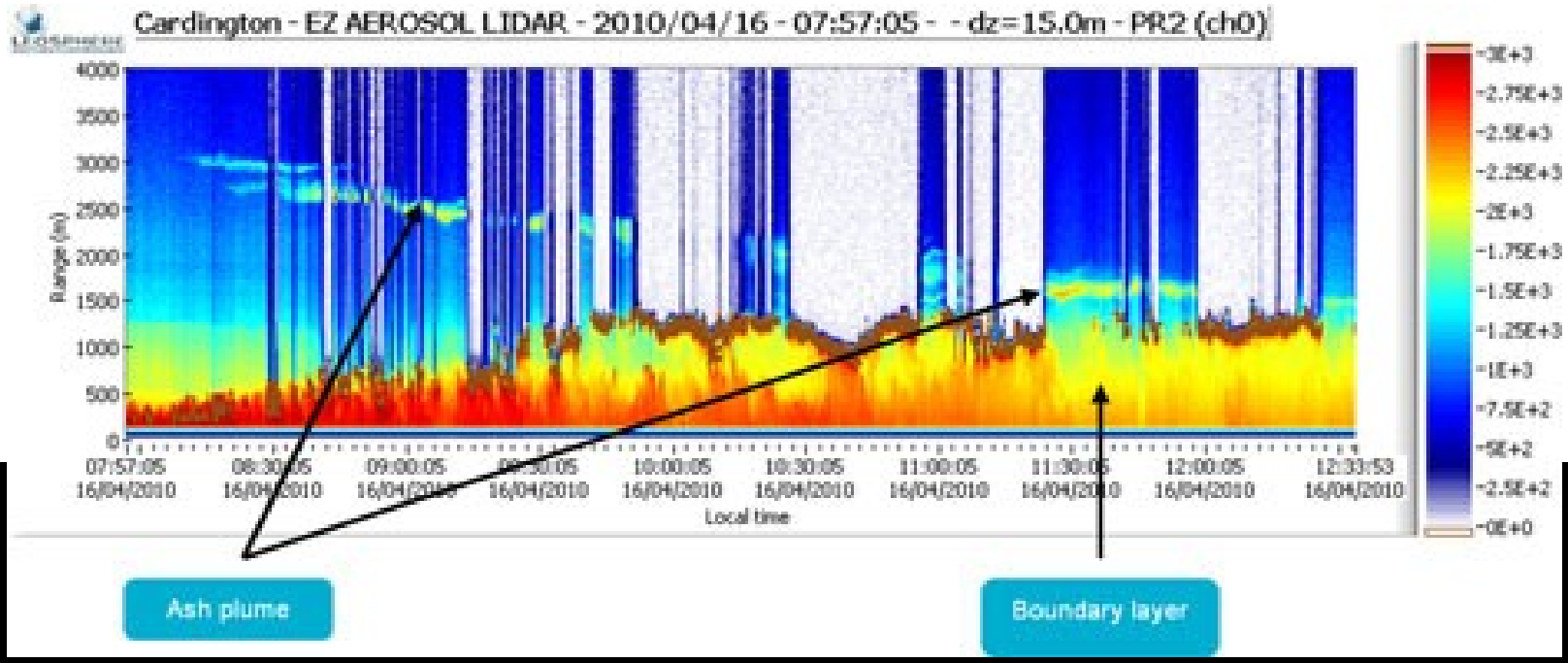
# Models make rapid but sometimes inaccurate predictions which may overstate the actual risk




The day that UK airlines started to fly again

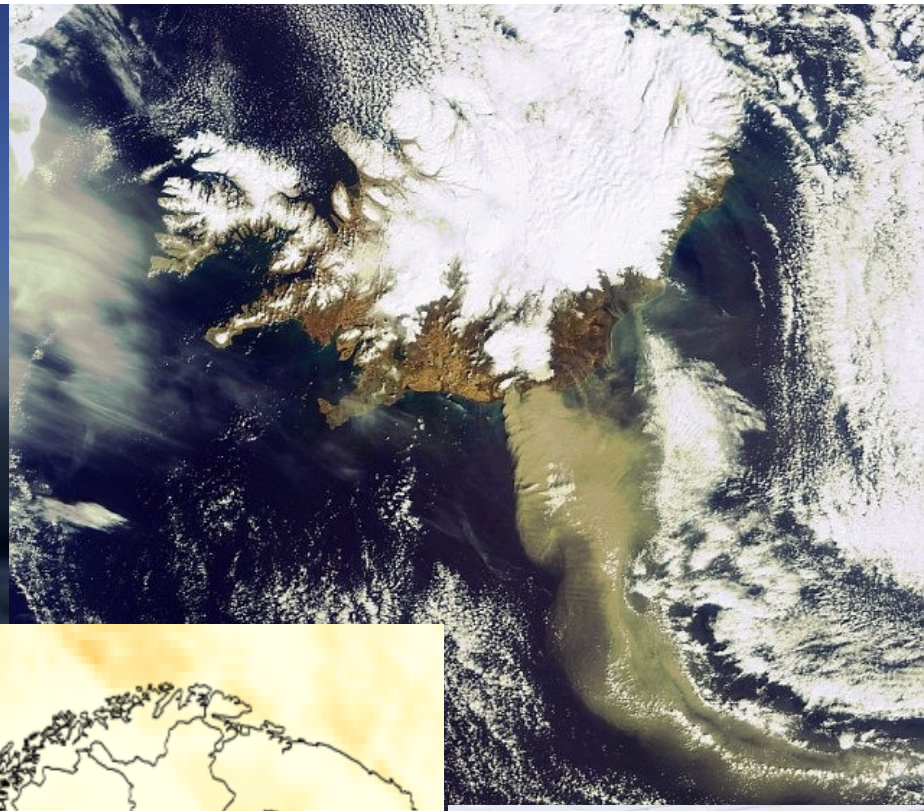
Tessella
Technology & Consulting

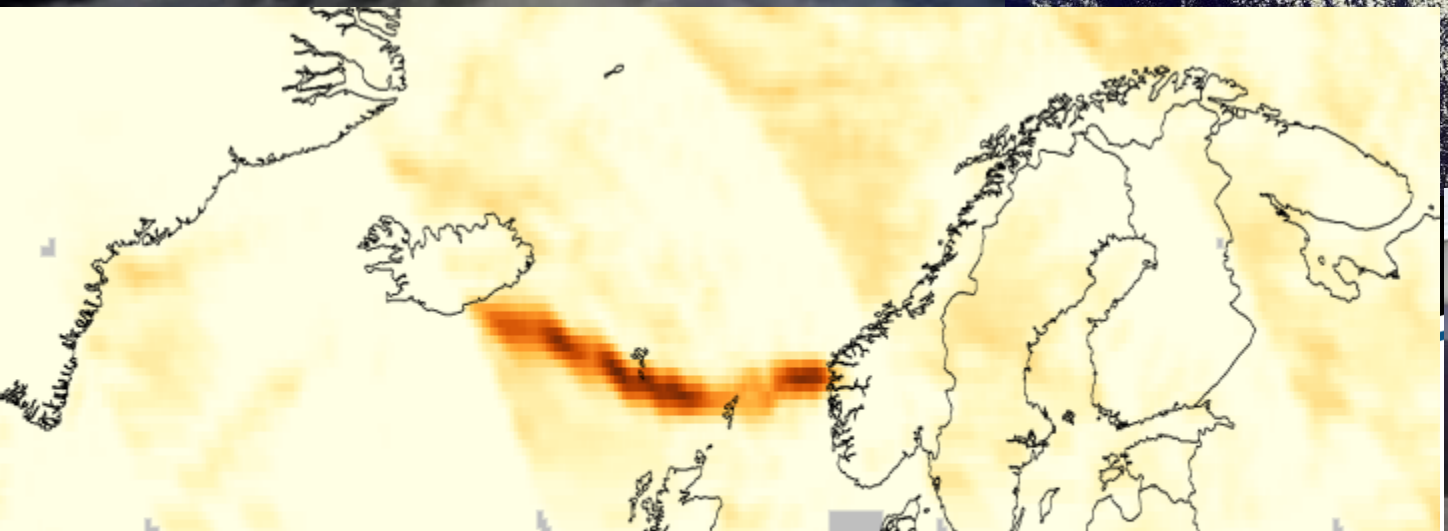
# There are an increasing number of new measurements that increase alarm and concern




Tessella
Technology & Consulting

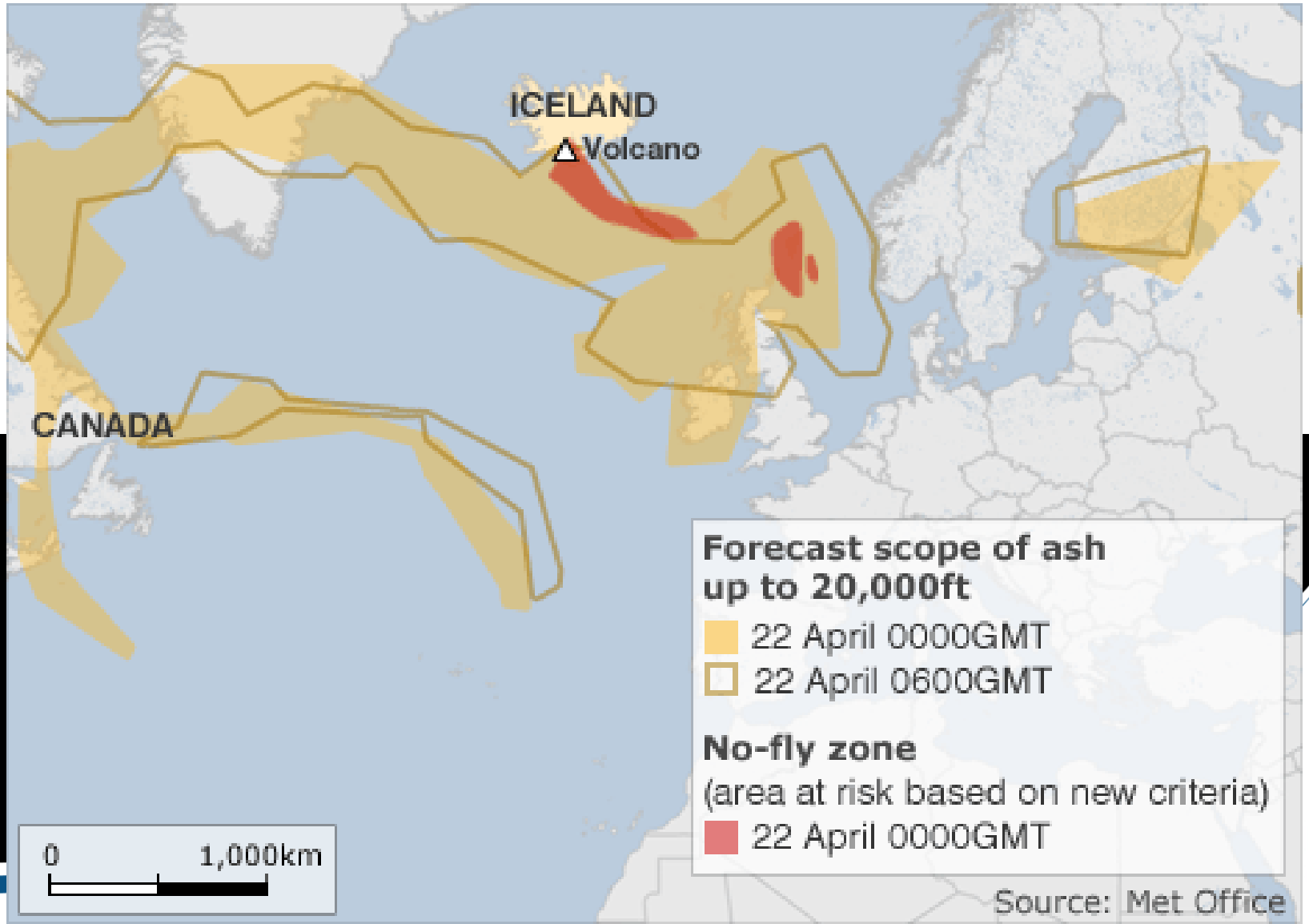
# And yet we cannot ignore that sometimes, somewhere, the hazard will really be present





# Tuning risk thresholds balances upside and downside risks:

Increasing the ash threshold by a factor of ten, to 2 mg/ $m^3$, greatly reduced the area of no-fly space. A further doubling was allowed later.



# Are pharma discovery groups also becoming too risk-averse and carrying out excessive screening?

Moncef Slaoui's view:

Financial Times, 12 December 2007

*"At least a fifth of the scientific questions currently asked were unnecessary… In every single project we look at we could have reached the critical decision with 50-60 percent fewer experiments. In a bureaucracy, if you ask more questions, no one will blame you for asking them. But we just can't afford it. In the modern world we generate lots of information we don't know what to do with."*

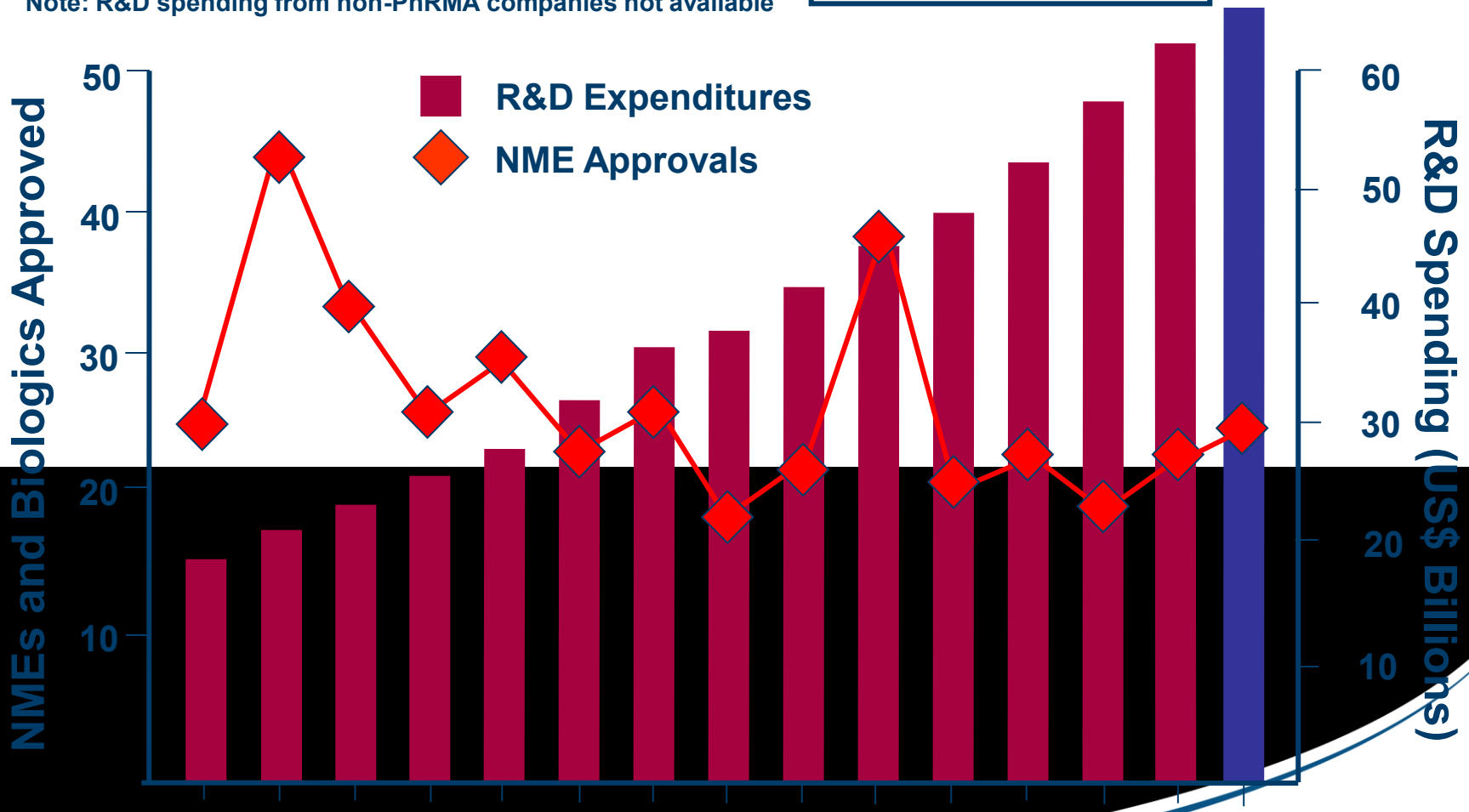
**A test is only worth performing if it could influence a decision and add value worth its cost and time**

Tessella
Technology & Consulting

# Productivity of Pharmaceutical R&D has been falling

**Sources: FDA/CDER, PhRMA, PricewaterhouseCoopers**
**Note: R&D spending from non-PhRMA companies not available**

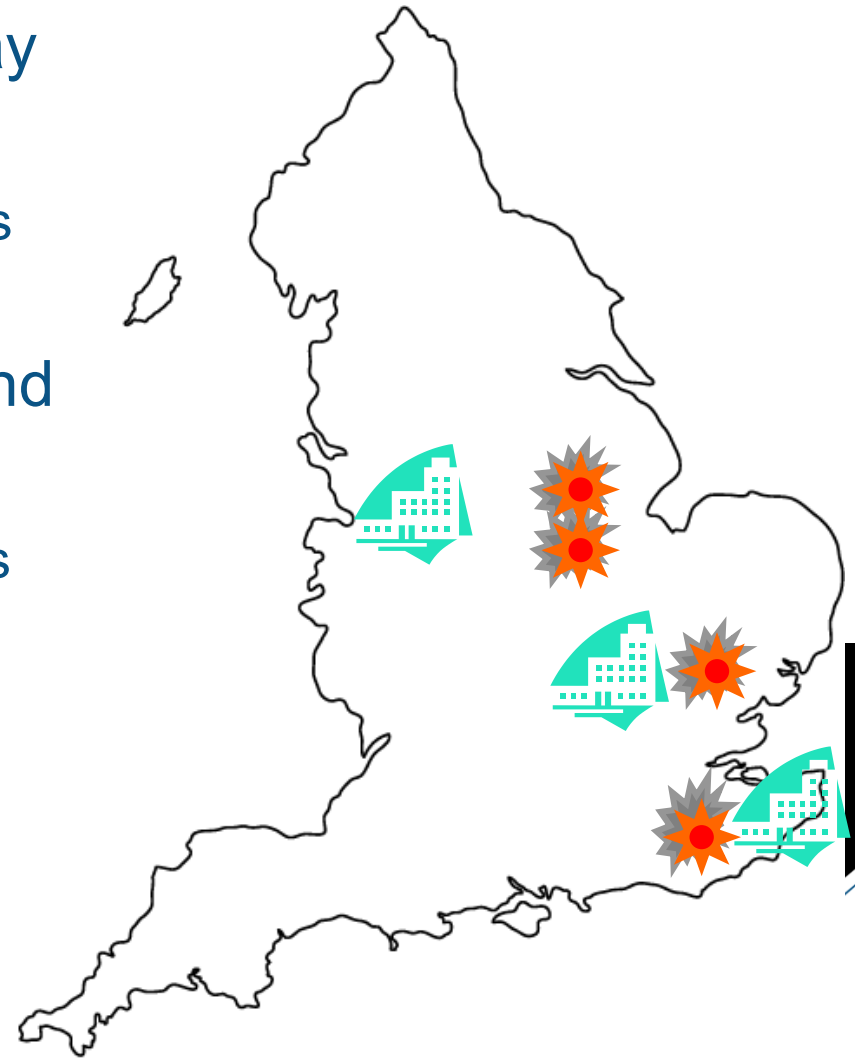
Compilation: Ted Torphy, J&J




Tessella
Technology & Consulting

# Solutions currently being pursued by big pharma

- Cost-cutting – the move away from ‘bricks and mortar’
  - especially in high-cost countries such as the US & UK
- Focusing effort on smaller and fewer disease areas
  - GSK disease performance units
- Reorganisation to bring research closer to the clinic
  - AZ innovative medicines units
  - Roche internal competition
- Process re-engineering and continuous improvement



Tessella
Technology & Consulting

# Applying ‘Six Sigma’ literally to the varied objectives of research projects breaks its first rule: reduce variability

Six Sigma (in manufacturing) means that a tolerance of six[1] standard deviations from the mean specification is required, before an unacceptable defect in the product results

| Sigma level | tolerable defects | sample size required for 95% confidence of an estimate of this population defect rate (±10%) |
|---|---|---|
| 3 | 6.7% | 5350 |
| 4 | 0.62% | 61577 |
| 5 | 0.023% | 1.7 million |
| 6 | 0.00034% | 113 million |

Largest pharmaceutical screening collection ≈ **5 million**

Number of concurrent projects in a large pharmaceutical discovery group ≈ **250**

Number of successful projects (NME’s) across the entire industry per year ≈ **25**

- ***Learning is good, but trial-and-error will not give the rate of improvement needed***
- ***Exact copying from the previous project is unwise or impossible***
- ***‘Smarter decisions’ and plans are needed***

[1] 4.5 s.d allowing for long-term process drift

Tessella
Technology & Consulting

# Project managers in research have to work out the best sequence for testing compounds e.g for safety

- Ideal screening strategies:
  - remove all hazard (‘fast failure’)
  - retain all good options (‘pipeline sparing’)
  - at zero time and cost (‘fast, cheap’)
- Real screening strategies
  - have errors
    - false alerts
    - missed alerts
  - have costs
  - are difficult to choose

Tessella
Technology & Consulting

# Making good plans involves balancing the downstream impacts of errors, but this isn't easy

- Real scientists
  - are people
  - have human biases
    - *confirmation bias* (overconfidence, self-justification, difficulty in challenging a hypothesis)
    - *calibration bias* (under-estimating uncertainty and error)
    - *availability bias* (over-attention to the recent and vivid)
    - excessive attention to small probabilities
- Research teams
  - don't have a good feel for risk
  - either don't know how reliable their predictive methods are, or tend to overestimate this
  - don't trust financial value estimates from product marketeers
- How can we help them explore the complex influences between risk, cost, project value, methods reliability, and the economic impact of their testing choices?

Tessella
Technology & Consulting

# Simplified Example: screening against a single risk

Two types of tests: *in silico*, modelling tests and results on a computer, or *in vitro*, testing compounds on biological samples

Progress project?

Screening options

**Abandon**

*Reason:* There is no adequate screening strategy to mitigate a severe risk

| Do A? | Result A | Do B? | Result B | Action on compound | **Strategy** |
|---|---|---|---|---|---|
| No | | No | | Progress | **No screen** |
| Yes | Alert | No | | Withdraw | ***In silico* only** |
| | No alert | | | Progress | |
| No | | Yes | Alert | Withdraw | ***In vitro* only** |
| | | | No alert | Progress | |
| Yes | Alert | No | | Withdraw | **Double filter** progresses only compounds that pass both tests |
| | No alert | Yes | Alert | Withdraw | |
| | | | No alert | Progress | |
| Yes | Alert | Yes | Alert | Withdraw | **Sentinel method** withdraws only compounds that fail both tests |
| | | | No alert | Progress | |
| | No alert | No | | Progress | |

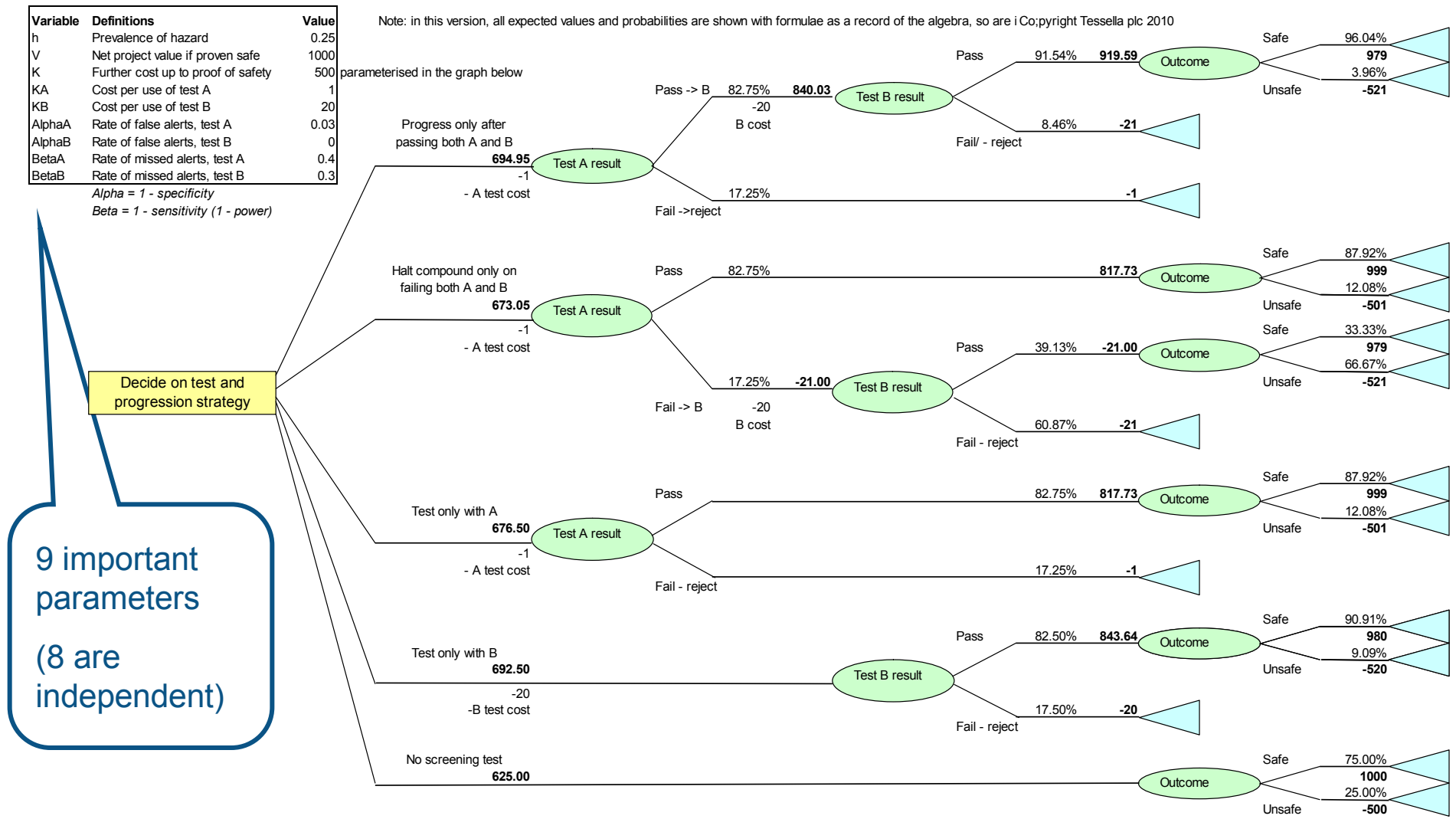
Choice of strategy depends on risk and cost of failure, value of success, risk of false alerts, and cost of different test methods

Tessella
Technology & Consulting

# A decision tree can find the best strategy, but there are ~$20^8$ (over a billion) different inputs to look at

| Variable | Definitions | Value |
|---|---|---|
| h | Prevalence of hazard | 0.25 |
| V | Net project value if proven safe | 1000 |
| K | Further cost up to proof of safety | 500 |
| KA | Cost per use of test A | 1 |
| KB | Cost per use of test B | 20 |
| AlphaA | Rate of false alerts, test A | 0.03 |
| AlphaB | Rate of false alerts, test B | 0 |
| BetaA | Rate of missed alerts, test A | 0.4 |
| BetaB | Rate of missed alerts, test B | 0.3 |

*Alpha = 1 - specificity*
*Beta = 1 - sensitivity (1 - power)*



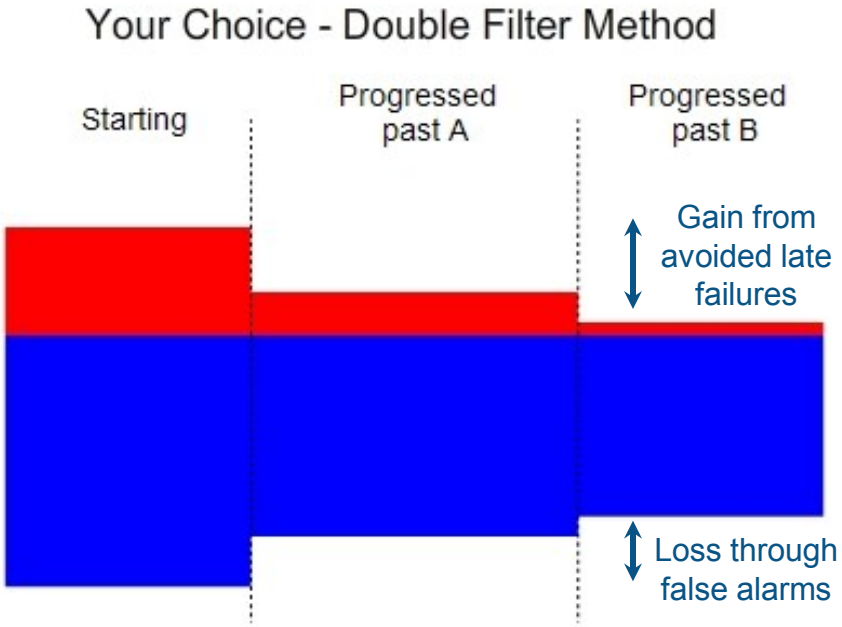

9 important parameters

(8 are independent)

# There are two important economic tradeoffs impacting the best strategy:

- Cost of early testing versus cost of late failure
- Cost of late failure versus opportunity cost
  - of projects abandoned due to 'false alarms'
  - or where the compound progressed is not actually the best possible

**Impacts of a screening strategy: pipeline view and value drill-down**

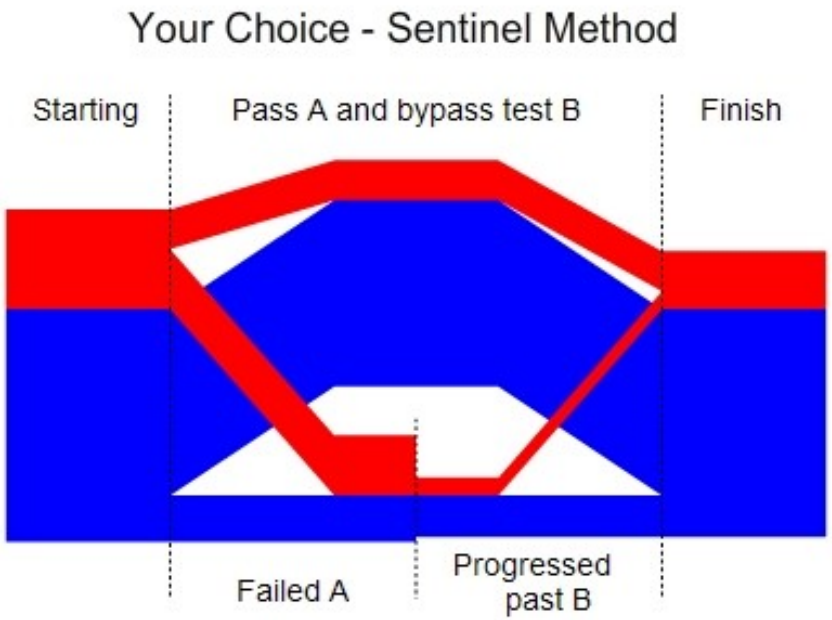



| PERFORMANCE | Per Compound | Pipeline Fraction | Overall |
|---|---|---|---|
| Maximum Possible Value (success value x (1 - risk)) | 1000 | 0.7 | 700 |
| Late Failure Cost (Missed Alerts) | 500 | 0.036 | 18 |
| Value Loss Through False Alerts | 1000 | 0.196 | 196 |
| In-silico Test Cost (A) | 1 | 1 | 1 |
| In-vitro Test Cost (B) | 200 | 0.680 | 136 |
| Total Value Lost | | | 351 |
| Net Strategy Value | | | 349 |
| Net Strategy Value / Maximum Possible Value | | | 0.50 |

# There are two important economic tradeoffs impacting the best strategy:

- Cost of early testing versus cost of late failure
- Cost of late failure versus opportunity cost
  - of projects abandoned due to ‘false alarms’
  - or where the compound progressed is not actually the best possible

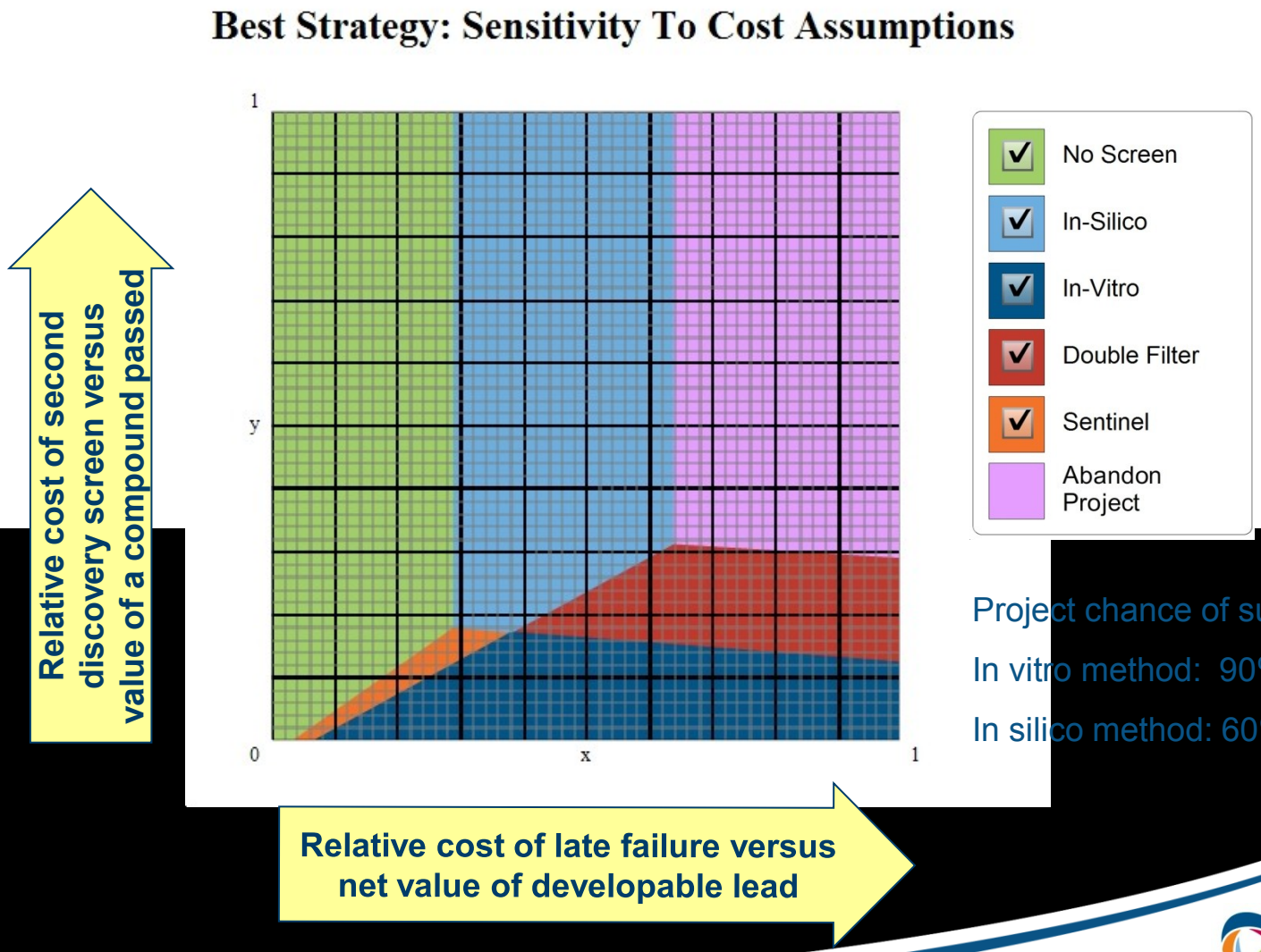
**The sentinel strategy wins over the double filter since it gives compounds ‘two chances to win through’**




| PERFORMANCE | Per Compound | Pipeline Fraction | Overall |
|---|---|---|---|
| Maximum Possible Value (success value x (1 - risk)) | 1000 | 0.7 | 700 |
| Late Failure Cost (Missed Alerts) | 500 | 0.174 | 87 |
| Value Loss Through False Alerts | 1000 | 0.014 | 14 |
| In-silico Test Cost (A) | 1 | 1 | 1 |
| In-vitro Test Cost (B) | 200 | 0.320 | 64 |
| Total Value Lost | | | 166 |
| Net Strategy Value | | | 534 |
| Net Strategy Value / Maximum Possible Value | | | 0.76 |

Toxic Safe

# Cost and value estimates are important but often tricky, so we need to see how any reasonable combination impacts the best strategy




Project chance of success: 30%

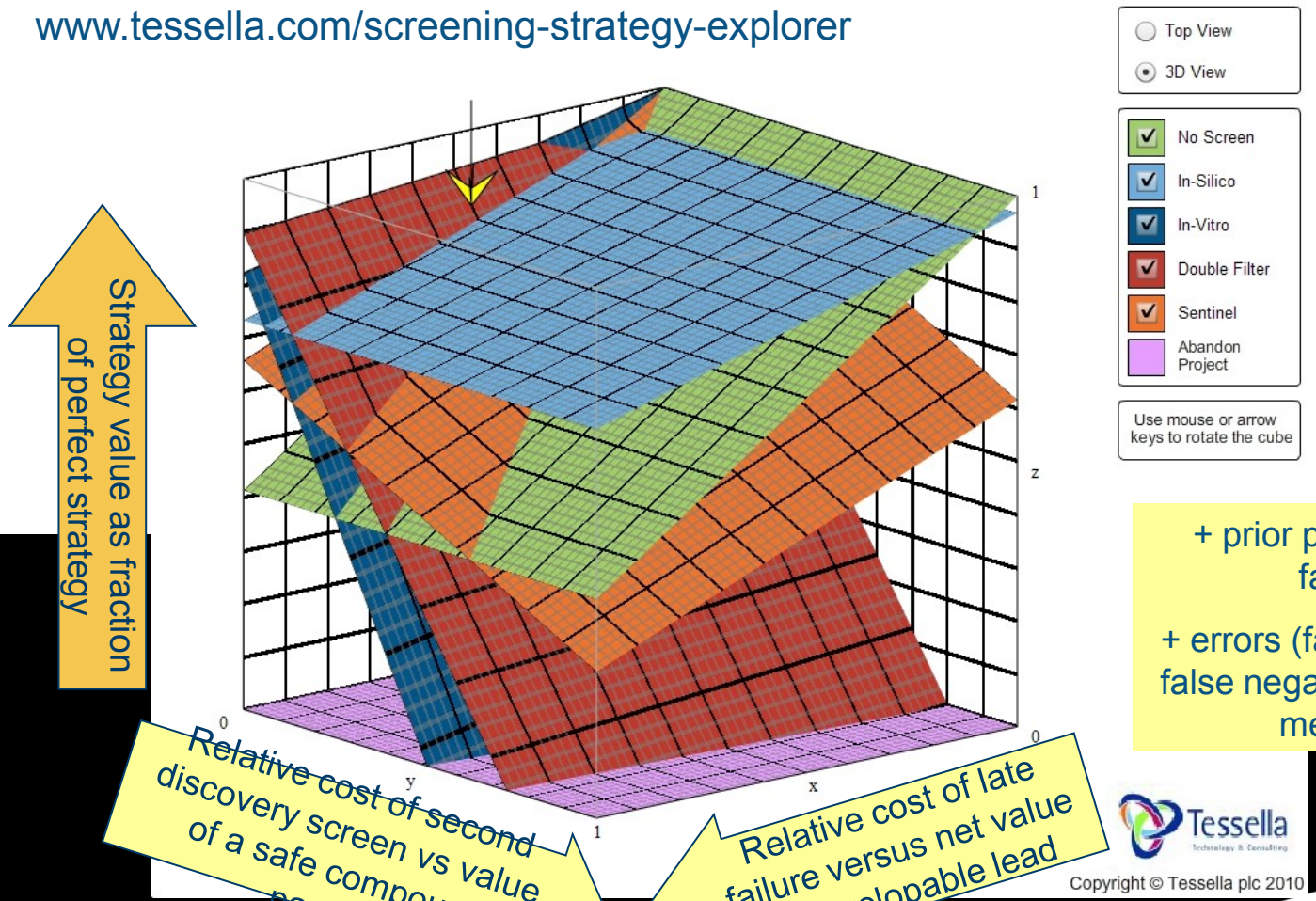
In vitro method: 90% predictive

In silico method: 60% predictive

Tessella
Technology & Consulting

# Tessella strategy cube visualizes all the eight factors that influence choice of screening strategy

www.tessella.com/screening-strategy-explorer




+ prior probability of failure

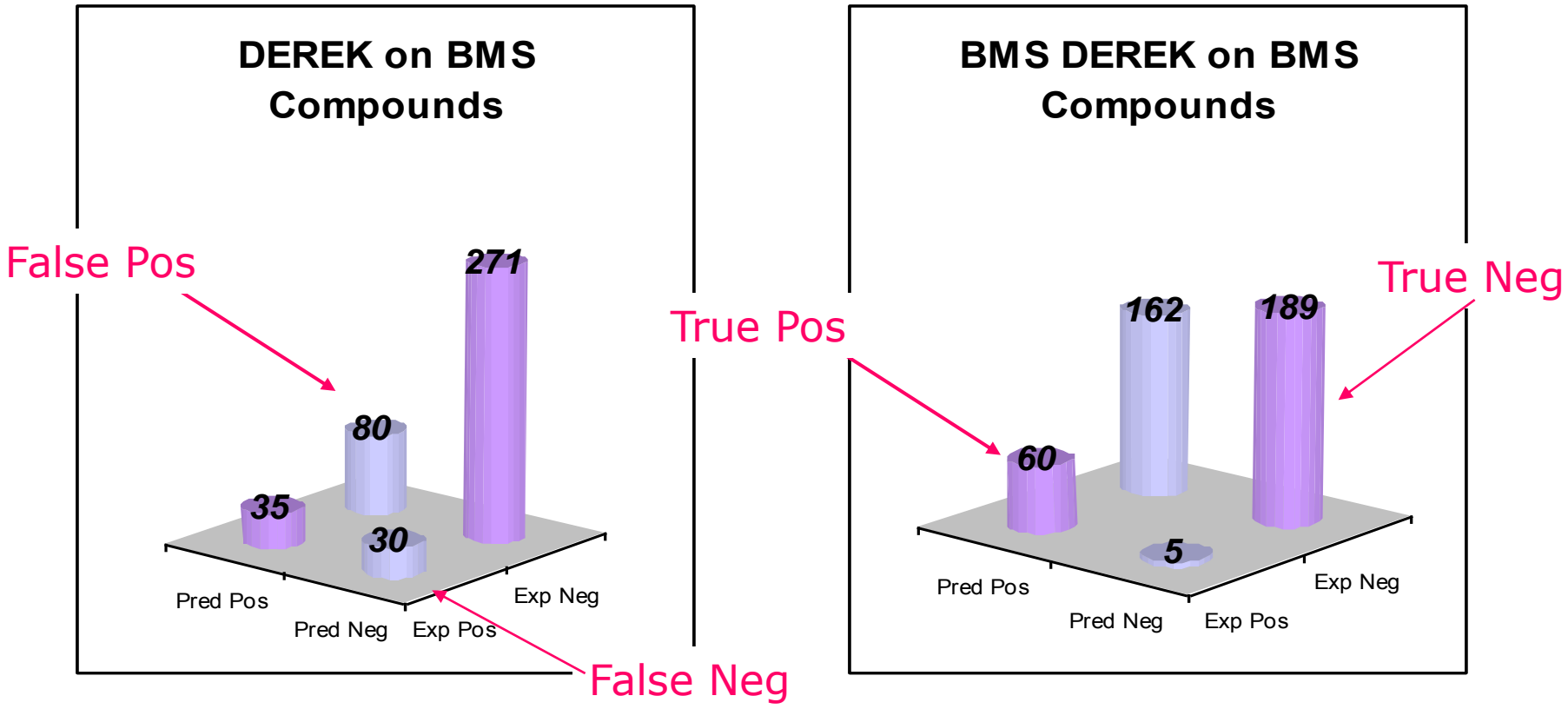
+ errors (false positives, false negatives) for both methods

Tessella
Technology & Consulting

# Performance Assessment/ Improvement: Has DEREK been Improved?

*Source*: Kreatsoulas, BMS, 2003

**416 BMS Compounds: 351 Ames (-), 65 Ames (+)  Random Conc.: 74%**

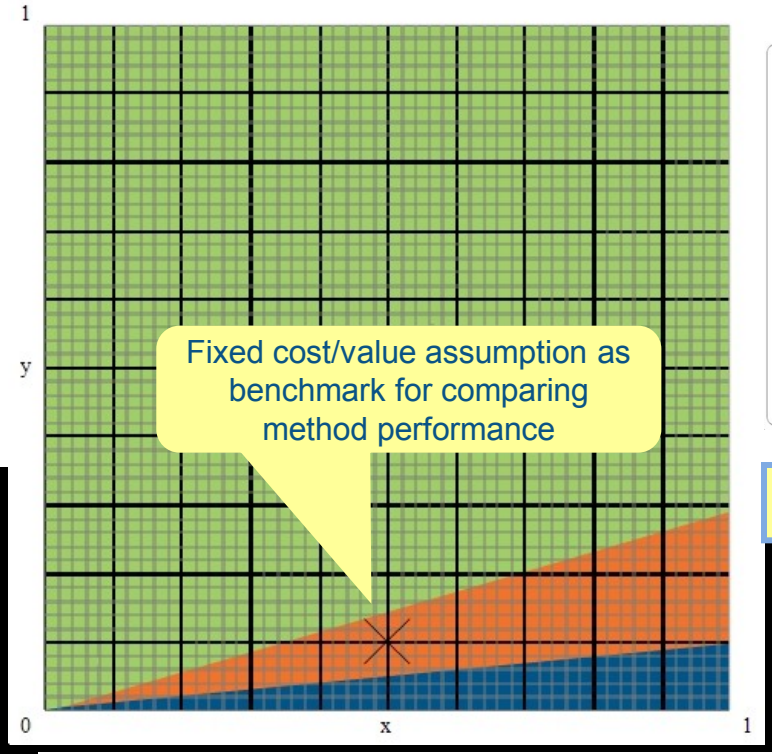



**Concordance: 74%**
**Sensitivity: 54%**
**Specificity: 77%**

**Concordance: 60%**
**Sensitivity: 92%**
**Specificity: 54%**

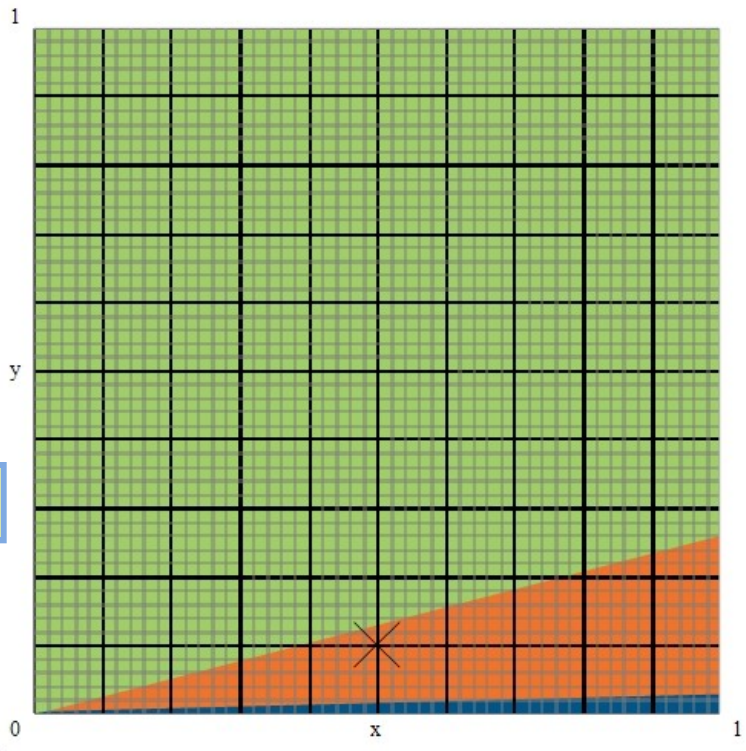
# The best strategy is likely to be the sentinel one for either standard or 'improved' DEREK

- Standard DEREK
- 'Improved' BMS DEREK





Risk = 15%



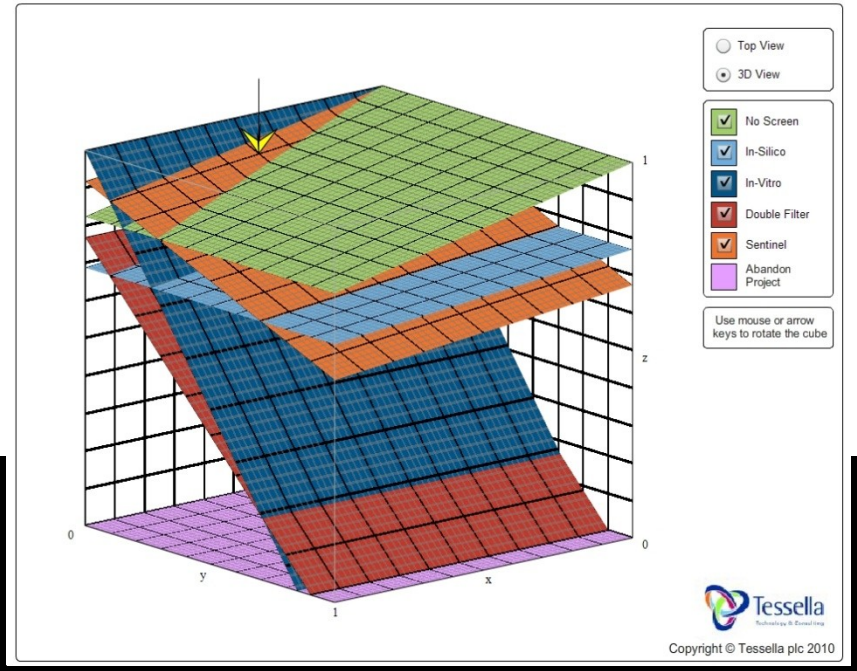

In either case, 93% of maximum possible pipeline value is reached at point shown:

x = 50% (downstream failure cost/ success value)

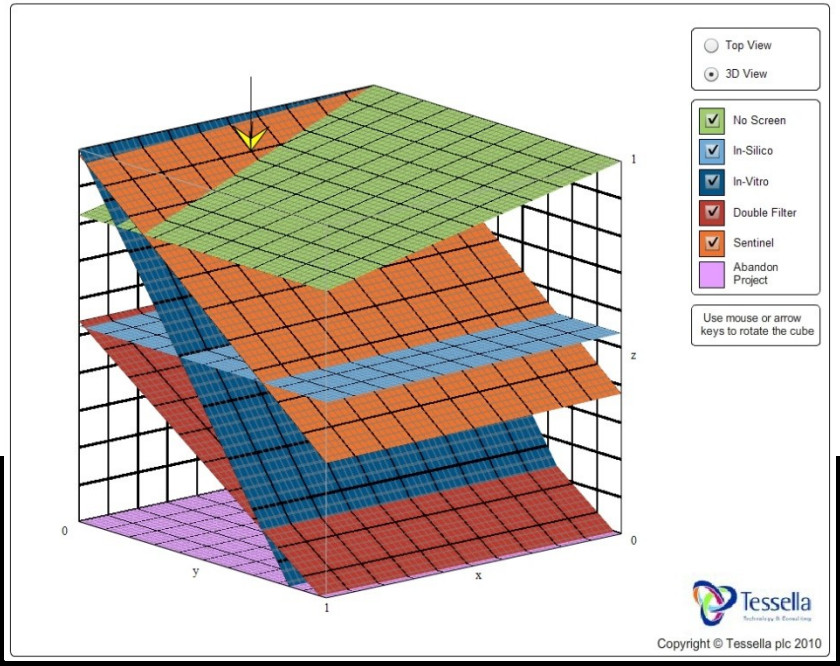
y = 10% (in vitro test cost/ value of compound passing screening sequence)

# The strategy cube shows clearly that the ‘improved’ DEREK is more fragile i.e. costly if *in vitro* costs are higher than expected and/or if project is less valuable

- Standard DEREK




- ‘Improved’ BMS DEREK




The pointer shows the benchmark location on the cube planes:

Risk = 15%

x = 50% (downstream failure cost/ success value)

y = 10% (in vitro test cost/ value of compound passing screening sequence)

If y is less than this, BMS DEREK is an improvement; if y is greater then it destroys value

Daily work decisions are also influenced by per-project or cross-project plans and guidelines. Errors in any of these can be expensive.

| | |
|---|---|
| Daily/ routine | Which compound(s) to progress?<br>Which new compounds to synthesize? |
| Project planning | Project decision: which screening cutoffs and sequence of experiments to apply to compound selection? |
| Site, research area or company standards and guidelines | More far-reaching decisions that determine the planning and technology choices available to project teams:<br>• Which methods to make available, and in which combinations?<br>• How much flexibility to give project teams in their planning and screening/ LO approaches?<br>• How to help project teams make those more flexible choices as well as possible?<br>• Stance on polypharmacology and combinations<br>• How much effort to invest in cross-project 'learning' e.g. method performance calibration? |

# Design and Optimization of Experiments

For the science / business question we trying to answer:

What tests or predictions should we do, in what order, to help answer it?

Which samples/ compounds might we evaluate?

From all the feasible experiments (tests x samples):
which set maximises the information needed for decision?

Progression criteria (cutoff, weighting) to optimise value vs risk and cost?

How can we operate the lab to run
many tests on many samples across projects
in a lean yet flexible way?

**Complex, Linked Optimisation Problems**

Tessella
Technology & Consulting

# Tessella Value Model

www.tessella.com

andrew.chadwick@tessella.com


Tessella
Technology & Consulting