



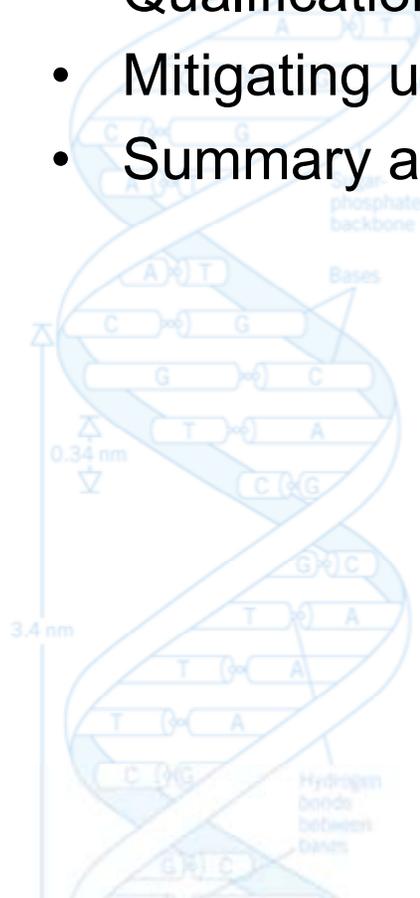
Scale-down Model Qualification and Use in Process Characterization

Nathan McKnight

CMC Strategy Forum

28 January, 2013

- Introduction and context
- Scale-down model design and key elements
- Uses and data interpretation
- Qualification
- Mitigating uncertainties from scale-down model use
- Summary and conclusions

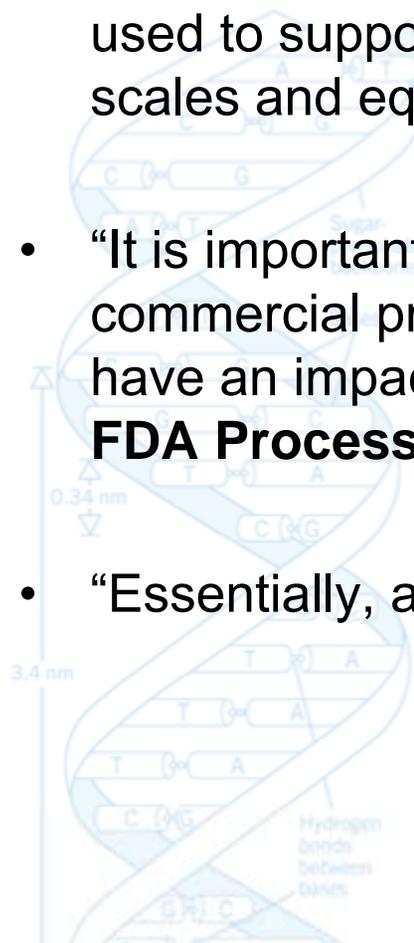


- “Small-scale models can be developed and used to support process development studies. The development of a model should account for scale effects and be representative of the proposed commercial process. A scientifically justified model can enable a prediction of quality, and can be used to support the extrapolation of operating conditions across multiple scales and equipment.” **ICH Q11 Step 4**

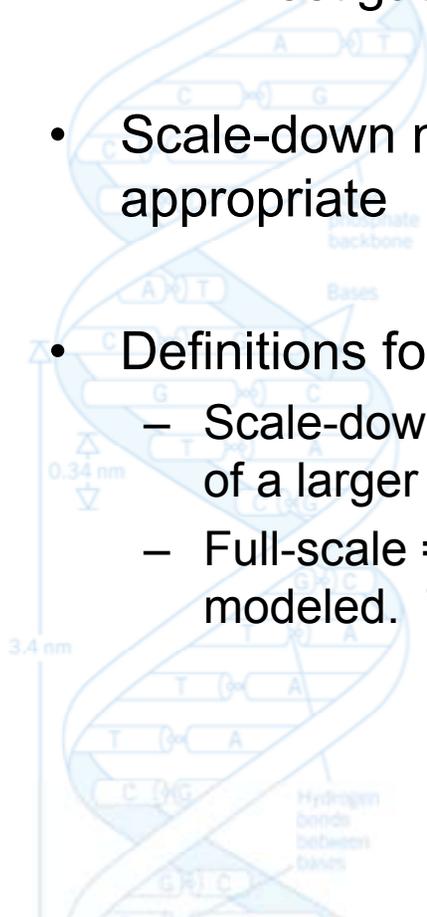
- “It is important to understand the degree to which models represent the commercial process, including any differences that might exist, as this may have an impact on the relevance of information derived from the models.”

FDA Process Validation Guidance

- “Essentially, all models are wrong, but some are useful.” **George E. P. Box**



- Scale-down models are indispensable:
 - Process optimization in development
 - Evaluating material and parameter variability for characterization
 - Investigations and improvements post-licensure
- Scale-down models need to be appropriate... and *demonstrated* to be appropriate
- Definitions for this presentation:
 - Scale-down model = small-scale model = model: a physical scale-down model of a larger system. Distinct from a statistical process model.
 - Full-scale = manufacturing-scale = commercial-scale: the system being modeled. Typically a single unit operation of a multi-unit operation process

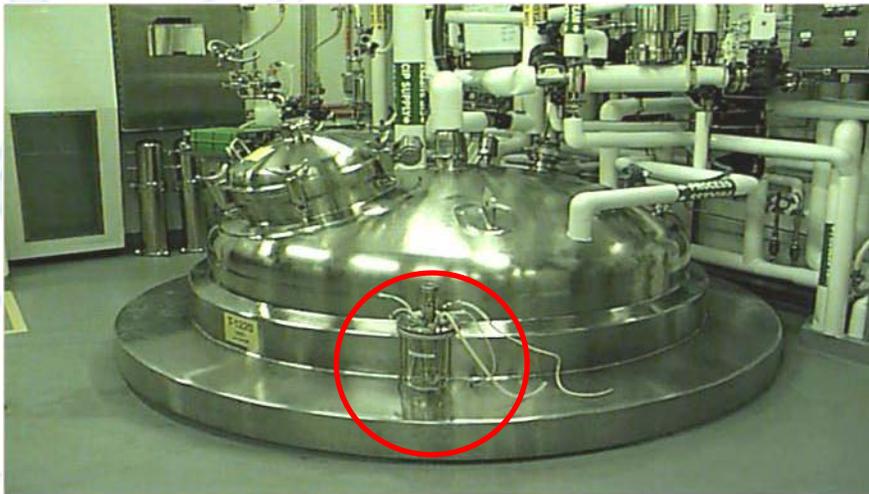


Scale-down Model Design: Design Basis Options

Slide 5

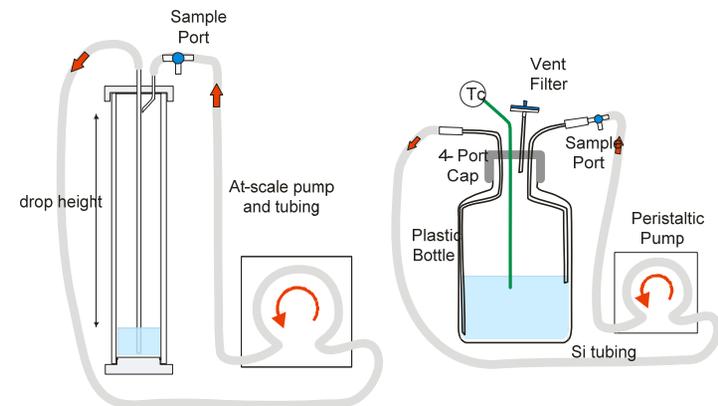
- By definition, a scale-down model is an incomplete representation of a more complicated, expensive and/or physically larger system.
- Scale-down models can be designed based on two general concepts:

Miniaturization of full-scale unit operation



Bioreactor culture

Partial, or “Worst-Case”, model of specific properties



Recirc. rate,
drop height

Thaw rate,
shear

Drug Substance thaw

Scale-down Model Design: Miniaturization

Slide 6

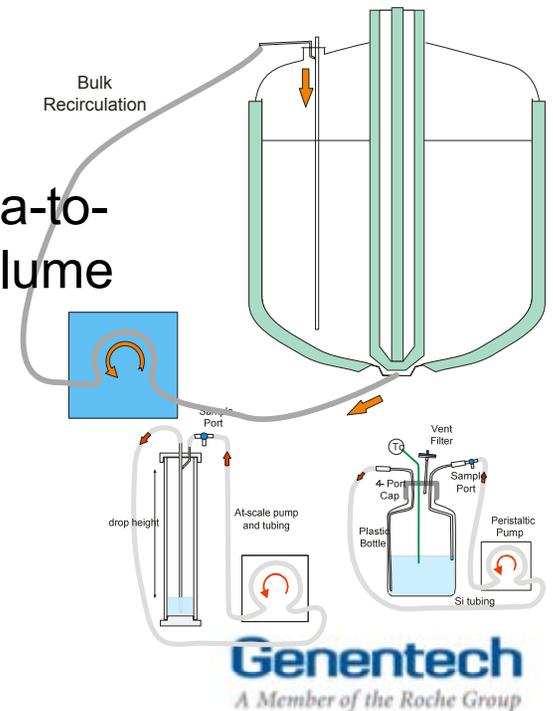
Whole unit operation model – miniaturization from full scale

- Designed to represent the physical and (bio)chemical environment of an entire unit operation
- Typically a reduced size version of the full-scale equipment
- Examines effects of process parameters and materials
- Even when system is well understood, comparison to full-scale performance is needed
- Examples: bioreactor cultures, chromatography columns



Partial, or “worst-case,” model

- Designed to represent a specific sub-set of physical and/or (bio)chemical properties of a unit operation, e.g., shear, surface area-to-volume, cell lysis.
- May use miniaturized equipment, or an apparatus imparting a desired force, property or environment.
- Typically used to test worst-case conditions of a subset of parameters.
- Examples:
 - Solution chemical stability: worst-case surface area-to-volume, gas-exchange interface, or headspace volume
 - Dilution of Drug Substance w/ formulation buffer:
 - Small-scale mixing to assess shear stress
 - Separate study for homogeneity (at-scale)
 - Harvest hold: Use of “fully lysed” feedstock

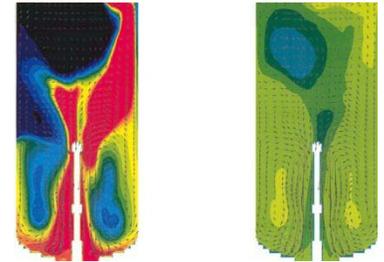


Scale-down Model Design: Key Elements

Slide 8

Key elements to consider for all scale-down models:

- Inputs: raw materials and components, feedstock/cell source, environmental conditions
- Design: selection of scaling principle(s), equipment limitations, operational procedures, parameter control concepts, on- and off-line analytical instruments
- Outputs: performance and product quality metrics, sample handling/storage, analytical methods.



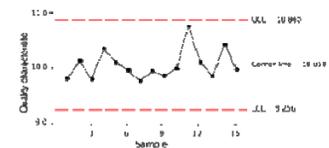
$$k_L a = K(P/V)^\alpha (v_s)^\beta$$

$k_L a$: hr⁻¹ P/V : W/m³ v_s : m/s



$$v_s = \frac{Q_G}{\pi \left(\frac{D}{2}\right)^2}$$
$$vvm = \frac{Q_G}{\pi \left(\frac{D}{2}\right)^2 H}$$

- Use of sound scientific and engineering principles for scaling
- Match full-scale as much as possible and feasible. Understand and/or control for differences between scale-down and full-scale (e.g., materials of construction, use of different assays)

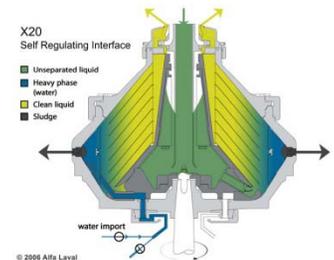
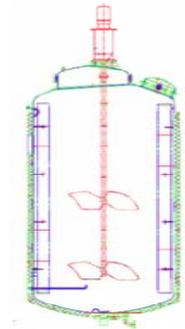
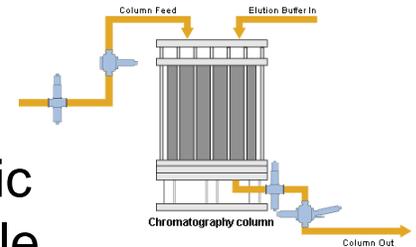


These elements should be described and justified as part of the overall qualification of a scale-down model.

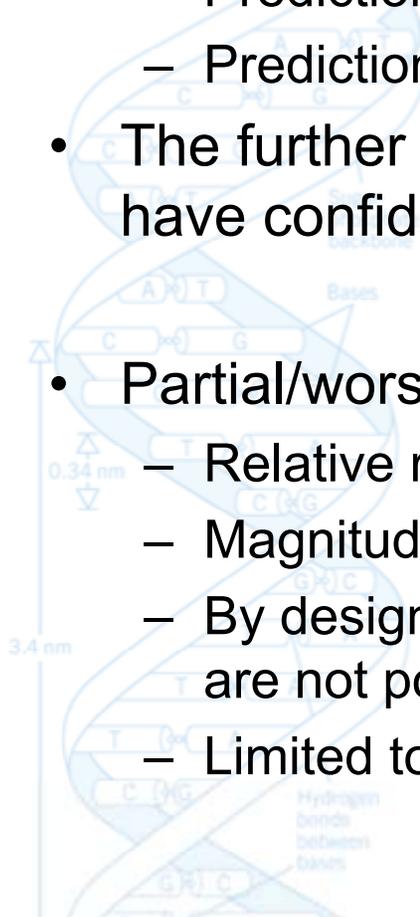
Scale-down Model Design: A spectrum of scalability

Slide 9

- Unit operations are not equally scalable...
 - Chromatography and filtration steps scale well
 - Bioreactors require trade-offs that can be system-specific for an optimal match (e.g. mix times, shear/power, bubble residence, etc.)
 - Harvest/centrifugation least scalable, best suited to at-scale characterization
- The ease of scaling an operation is based on having established and accepted scaling laws, and equipment limitations.
- The ease of scaling translates into a loose continuum of *a priori* confidence in scale-down models (though the field continues to evolve...)



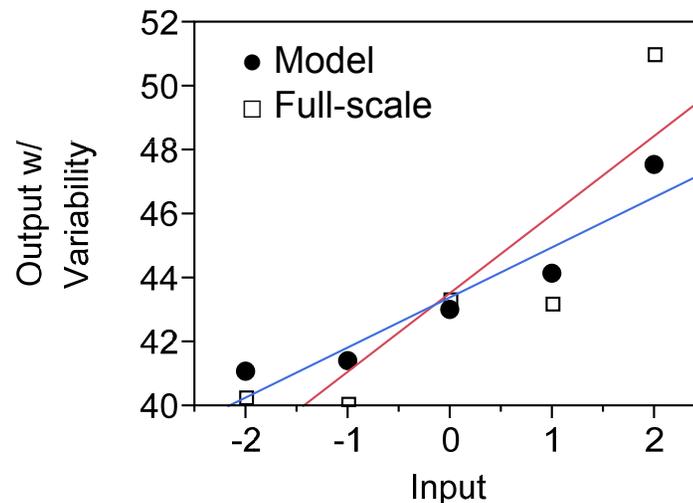
- Whole unit operation models (in order of decreasing confidence):
 - Relative ranking and directionality of factor effects
 - Magnitude of factor effect sizes
 - Prediction of a process output
 - Prediction of system variability
- The further down the list, the more rigor in verification is needed to have confidence in a conclusion
- Partial/worst-case, models:
 - Relative ranking and directionality of factor effects
 - Magnitude of factor effect sizes
 - By design (generally) direct predictions of process output or variability are not possible from partial model results.
 - Limited to effect(s) of properties model is designed to represent



Scale-down Model Qualification for Process Characterization

Slide 11

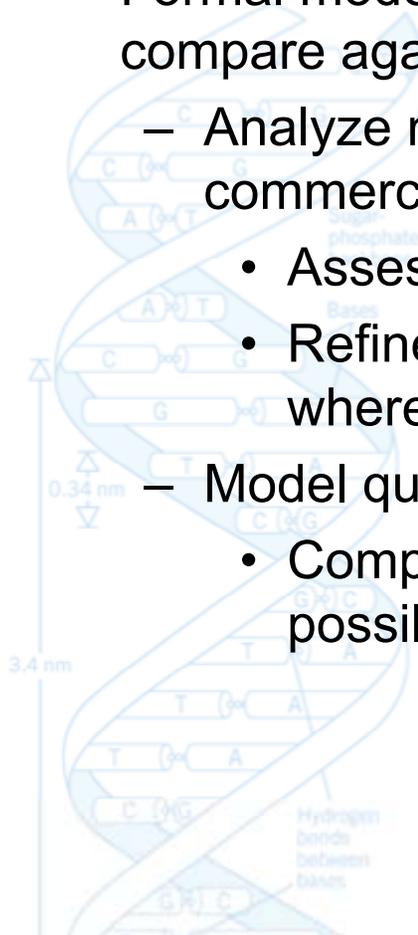
- Qualification is documenting evidence a model is suitable for evaluating the effect of input material and parameter *variation* on process performance and product quality outputs.
- What defines “suitable” depends on the type of model:
 - Partial/worst-case, model: design accurately applies the intended physical and/or (bio)chemical environment.
 - Whole unit operation model: the same *change* in inputs results in a substantially similar *change* in outputs.



Scale-down Model Qualification - When

Slide 12

- Model *development* should be continuous during clinical development as manufacturing increases in scale
- Formal model *qualification* typically requires sufficient full-scale results to compare against
 - Analyze model performance during PhIII runs (assuming at commercial scale)
 - Assess for potential offsets
 - Refine small-scale and pilot-scale procedures to remove offsets where possible and practical
 - Model qualification based on a reasonably sized data set
 - Compare to both PhIII and Qual campaign full-scale runs, if possible

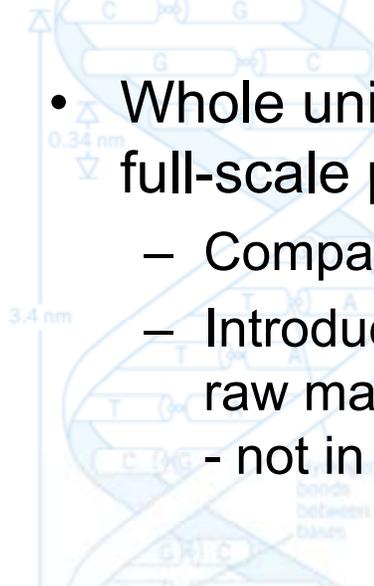


Scale-down Model Qualification - How... Depends on the type of model

Slide 13

- Partial/worst-case model: through adequate *description* and *justification* that the design provides the data it is intended to provide.
 - Design and scaling principles
 - Apparatus, materials, operational procedures
 - Feedstock source

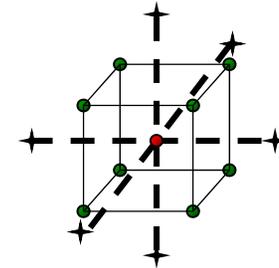
Generally, comparison to full-scale is not appropriate because the model is not intended to represent at-target performance.
- Whole unit operation model: same as above, *plus* comparison to full-scale performance.
 - Compare “at-target” performance (typically – see next slide)
 - Introduce relevant variability during small scale operations (multiple raw material lots, multiple thaws, multiple resin lots, run independently - not in replicates)



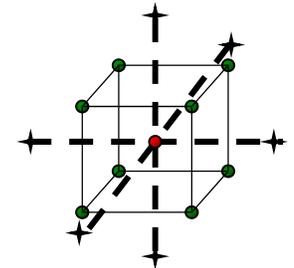
Comparing to Manufacturing-scale Performance

Slide 14

- An “Ideal Scenario”: Model is compared against full-scale at-target and off-target to verify the scale-down model is fully representative across the entire Design Space.



Design Space
in Scale-down



Full-scale PPQ?

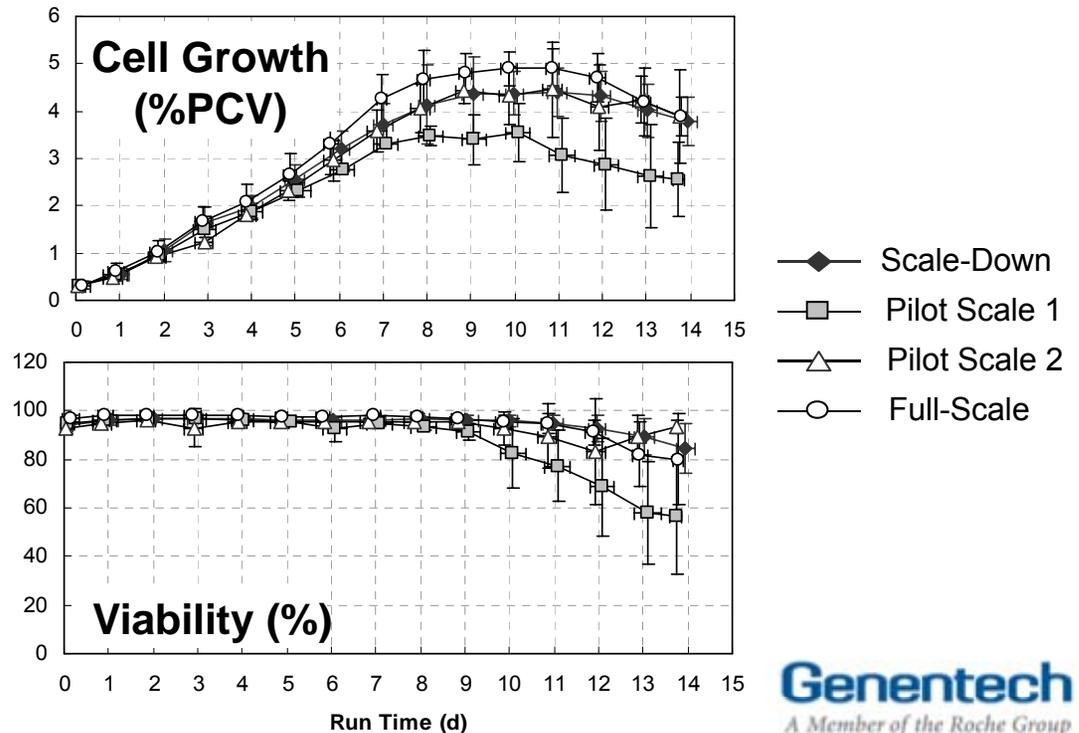
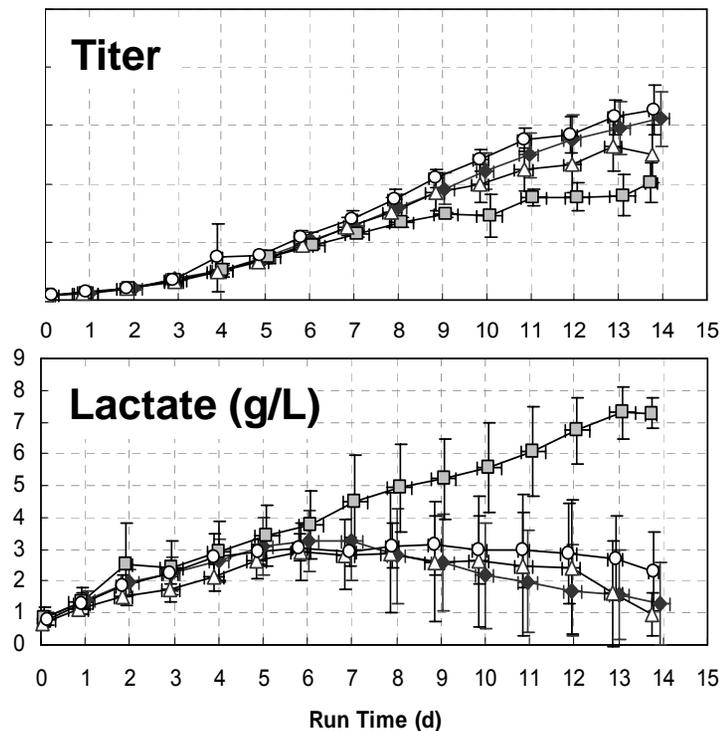
- Is this practical?
 - Short answer: No
 - Multiple additional runs, may also require sufficient replication at off-target points for statistical confidence.
 - Full scale test runs are prohibitively expensive
 - Long answer: part-way..., sometimes..., it depends...
 - Some parameters are tested: cell age, run duration, hold times
 - Select key points of the Design Space: testing process responses with an offset?
 - Testing at pilot scale instead of full-scale?
 - As part of lifecycle management of process changes in Stage 3?

Scale-down Model Qualification – Qualitative Comparisons

Slide 15

“You can see a lot just by looking” **Yogi Berra**

- Qualitative assessment of time-course trends
 - Instructive beyond typical point-value KPI comparisons
 - Similar behavior between scales supports model suitability
 - Dissimilar behavior may indicate a problem, and can be valuable for troubleshooting and model improvement

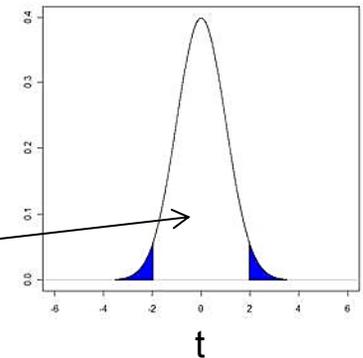


Scale-down Model Qualification – Statistical Approaches

Slide 16

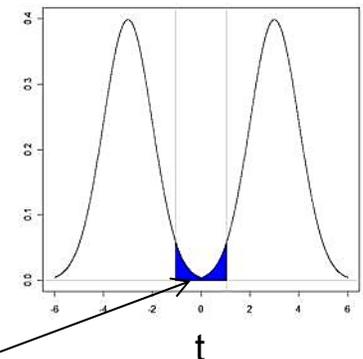
- Difference testing:

- Compute the difference in means (δ) and associated statistic comparing means (e.g., t-test and p value)
- Null Hypothesis is $\delta = 0$. Failure to achieve statistical significance (e.g., $p > 0.05$) supports “no evidence of difference” (null hypothesis not rejected)



- Equivalence testing:

- Define an interval within which a difference is not scientifically meaningful, a “practically significant difference” (PSD)
- Compute the difference in means and associated statistic testing if difference is within the PSD (e.g., two-one-sided-t-test [TOST] and p value)
- Null Hypotheses are $\delta > \text{PSD}$ or $\delta < -\text{PSD}$. Achieving statistical significance (e.g. $p < 0.05$) supports “equivalence” (both null hypotheses rejected)



Scale-down Model Qualification – Statistical Approaches

Slide 17

- Difference testing:
 - Adaptable to multivariate data analysis
 - Low replication biases outcome towards “no evidence of difference” (can mitigate with a power analysis and minimum sample size)
 - In case of a statistically significant difference, may still conclude equivalent if difference is “not practically significant”
- Equivalence testing:
 - Rewards greater data replication
 - Similar to Bioequivalence calculations
 - Supports a direct claim that model output is “not different”
- Statistical methods testing for *differences in variability* (unequal variance) require significant replication and are not generally applied. However:
 - Qualitative evaluation should still be performed
 - Means comparison should use methods which do not rely on equal variance

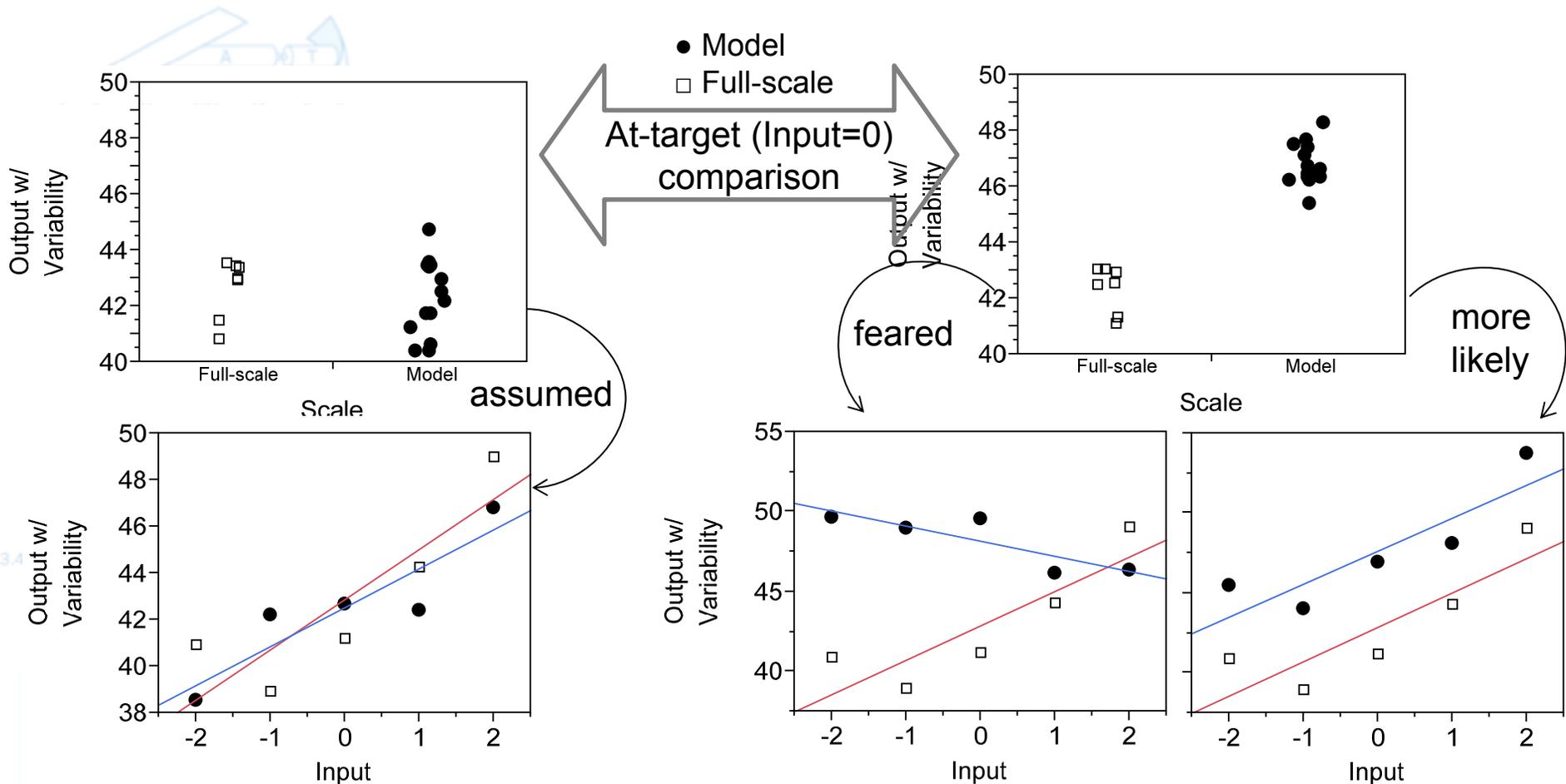
Statistical approaches in an context of overall model qualification

Slide 18

- Practically Significant Difference:
 - A difference of sufficient magnitude that it should be considered when using data from a scale-down model to predict full-scale results.
 - Should be based on a scientific/engineering considerations
 - Does not necessarily imply the scale-down data are unrepresentative of the full-scale (though it may in the case of large or unstable offsets)
- Some outputs are more important than others
 - Product quality attributes
 - Key performance indicators (e.g., yield)
 - Other characteristics (e.g, metabolic measures)
- A model can be “equivalent” for some outputs, but not all, and still be a representative model – and even still be representative of those outputs that are not statistically equivalent!

Dealing with Offsets

- The statistical evaluation of at-target performance is really an evaluation of risk, where offsets suggest higher risk
- The risk: an offset may indicate the effect of changing a given parameter will be different in the model than at full-scale

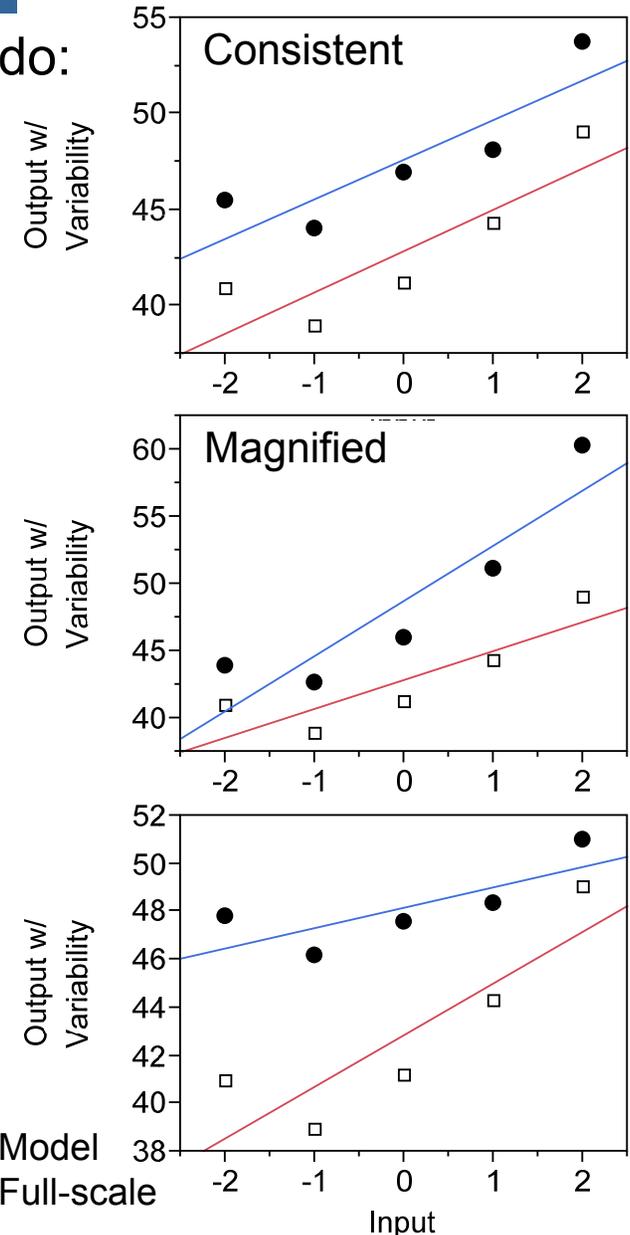


Dealing with Offsets

Slide 20

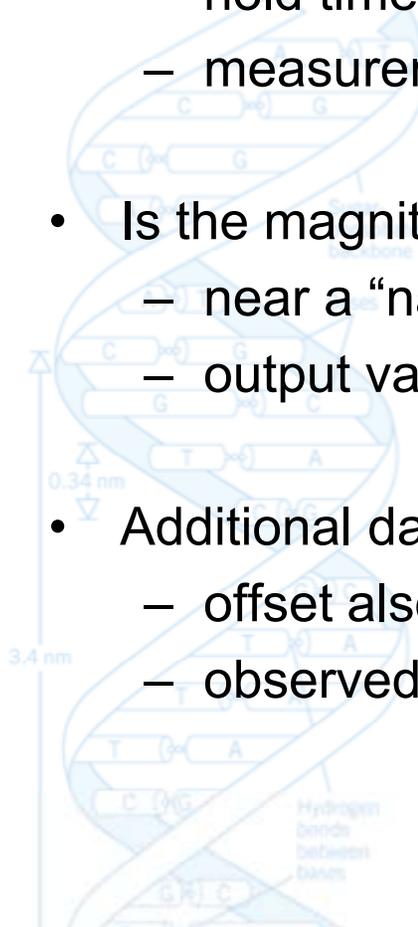
When is an offset acceptable, when not, and what to do:

- Offset consistent across parameter range - account for offset in data interpretation, need sufficient data supporting magnitude of offset used.
- Magnified parameter effect in model
 - Factor effect directionality and ranking still valid, direct prediction difficult
 - Robust interpretation possible by comparison to scale-down controls.
- Attenuated parameter effect in model
 - Same as magnified effect, but higher risk since effect sizes may be falsely interpreted as not significant.
- Output not representative (i.e., unrelated or opposite to effect at full-scale) – model not qualified for that output



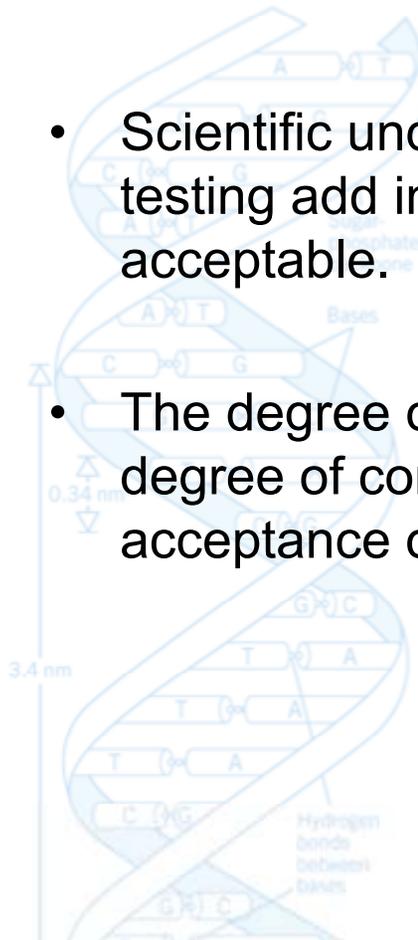
Evaluating the acceptability of an observed offset

- Is the mechanism understood and/or specific source known
 - light exposure
 - hold time differences
 - measurement/assay offset from high through-put method
- Is the magnitude of the offset likely to change parameter effect size
 - near a “natural limit”, e.g., % Monomer near 100%
 - output variation across parameter range same or bigger than offset
- Additional data
 - offset also observed off-target at full- or pilot-scale
 - observed across multiple processes (platform knowledge)



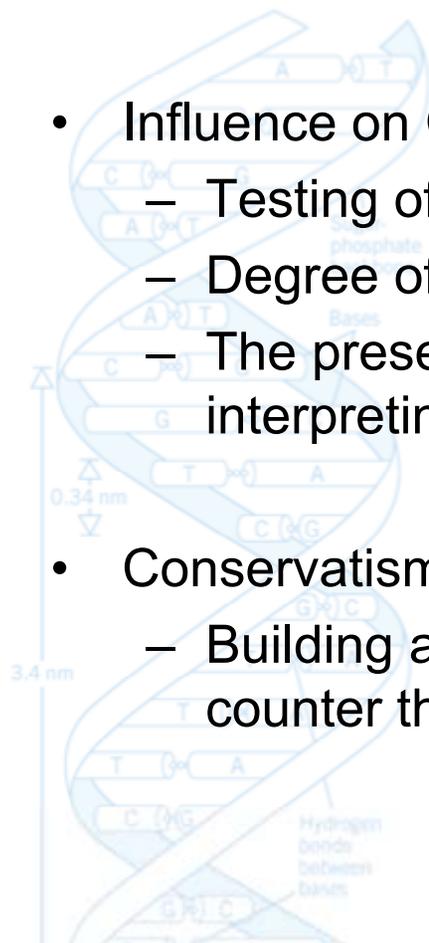
A question of confidence...

- Unlikely to have sufficient replication of on- *and* off-target conditions at full-scale for a statistically robust comparison of factor effect sizes between scales.
- Scientific understanding, offset consistency and off-target full-scale testing add incrementally to the totality of evidence that an offset is acceptable.
- The degree of reliance on data with an offset should be proportional to the degree of confidence in its accuracy - i.e., not an all-or-nothing acceptance or rejection of model-derived information.



Data from scale-down models is only one component of a process validation program. Complementary elements of an overall process control strategy, and specific actions, can mitigate uncertainties inherent to use of scale-down models:

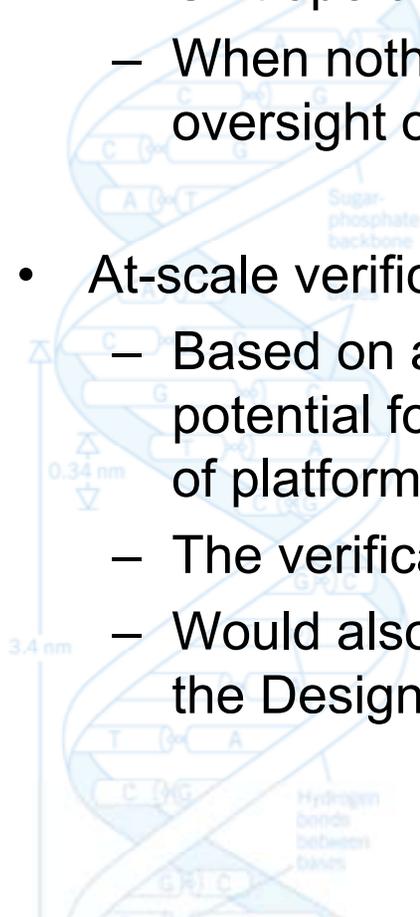
- Influence on Control System:
 - Testing of CQAs may still be required based on process capability.
 - Degree of scale-down model fidelity considered in defining CoA tests.
 - The presence of CQA testing should also be considered when interpreting data from scale-down models (i.e., risk is reduced)
- Conservatism in identification of CPPs:
 - Building a conservative bias into a system for identifying CPPs can counter the *potential* bias from a scale-down model.



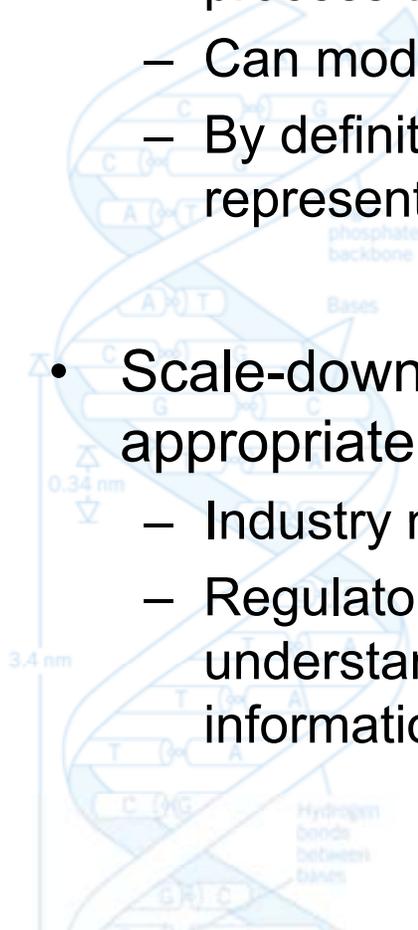
Mitigating uncertainties from scale-down model use, cont.

Slide 24

- A clearly defined and comprehensive Design Space:
 - All parameters (CPPs and non-CPPs) and raw materials
 - Unit operation descriptions and order of execution
 - When nothing is “excluded” from the Design Space, Health Authority oversight of process changes is maximized.
- At-scale verification of a process change before routine implementation:
 - Based on a risk-based assessment on the nature of a change (e.g., potential for product quality impact), fidelity of scale-down model, extent of platform or scientific understanding of change/future-state.
 - The verification could also include additional testing.
 - Would also serve as full-scale verification of an additional point within the Design Space.



- Scale-down models are a tool (one among many) for developing and characterizing “the process”
 - Enables evaluation of input material and parameter variability on a process to an extent that is simply not feasible at manufacturing scale
 - Can model a whole unit operation, or only certain aspect(s)
 - By definition of a “model”, even the best is an incomplete representation, but can still provide useful and accurate information.
- Scale-down models must be designed and demonstrated as appropriate representations of the manufacturing process.
 - Industry must demonstrate a model is appropriate and applicable
 - Regulators must recognize models cannot be absolutely perfect, but understand their value and permit industry to utilize them for the information they can appropriately provide.



Acknowledgments

Slide 26

- Tony Cano
- Dan Coleman
- Mary Cromwell
- Christian Hakemeyer
- Brian Kelley
- Bob Kiss
- Lynne Krummen
- Andy Kosky
- Fred Lim
- Steven Meier
- Melody Schmidt
- John Peattie
- Ron Taticek
- Vassia Tegoulia
- Frank Zettl

FDA and EMA reviewers for feedback and discussion

