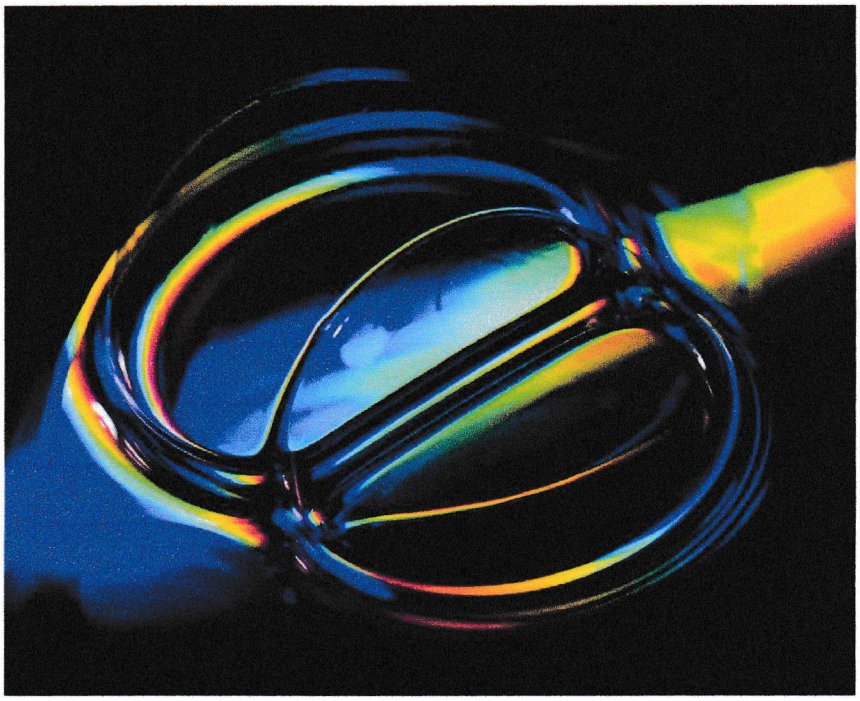


[HANS HENNING WENKI/GETTYIMAGES

Cleaning Validation for Biologics

Can alternative approaches to the permitted/acceptable daily exposure be justified?

ANDREW WALSH

Many pharmaceutical companies have adopted the permitted daily exposure (PDE), or acceptable daily exposure (ADE), to set cleaning validation limits. Some regulatory agencies now require its use, but some companies continue to argue against it, particularly biologics manufacturers. Critics of the approach argue that protein molecules are denatured or degraded by their cleaning processes, and that alternate approaches should be used. This article explores whether alternative approaches to using the PDE/ADE for determining cleaning validation limits for biologics can be justified, and, if not, what approaches could be used.

In 2010, the International Society for Pharmaceutical Engineers (ISPE) published its Risk-MaPP (Risk-based Manufacture of Pharmaceutical Products) baseline guide (1). This guide introduced ADE as the starting point for evaluating the risk to both patients and workers in the manufacture of pharmaceutical products in shared facilities.

FDA participated in the development of the Risk-MaPP guide. In particular, FDA wanted to see Risk-MaPP provide a tool that could be used to identify highly hazardous compounds and also wanted to know how this tool could be used for cleaning validation (1). The European Medicines Agency (EMA), Ministry of Health, Labour and Welfare (MHLW),

ANDREW WALSH, M.S., a certified Lean Six Sigma black belt, is principal and owner of PharmaClean Group LLC, and technical director of the Center for Pharmaceutical Cleaning Innovation in Hillsborough, NJ. He can be reached at andy@pharmacleangroup.com.

Japanese Pharmaceuticals and Medical Devices Agency (JPMDA), World Health Organization (WHO), Health Canada, Swissmedic, Brazilian Health Surveillance Agency (ANVISA), and the Pharmaceutical Inspection Convention and Pharmaceutical Inspection Co-operation Scheme (PIC/S) also reviewed Risk-MaPP and provided comments (1).

The Risk-MaPP guide defines the ADE as the most appropriate and accurate assessment of the potential hazard presented to both patient and worker from the compound being manufactured, because it is based on all the available toxicological and clinical data for that compound. In particular, the main role of the ADE is for use in informing the decision process for accepting a compound into a multi-product manufacturing operation, or designating that compound for manufacture in a dedicated facility based on the risk it presents to patients and workers.

Subsequently, EMA issued a guideline in 2015 requiring companies to implement the use of PDEs for assessing the risk of cross contamination in shared facilities (2).

There is also a newly proposed guide for publication by ISPE on cleaning process development and validation (3). This guideline also includes the ADE as a starting point for evaluating the risk involved in cleaning processes, to inform the facility what level of cleanliness is needed, the level of validation necessary, and whether further risk reduction efforts, in addition to cleaning are necessary.

However, both the EMA guideline and the proposed ISPE guide contain language that may be seen to allow the use of alternative approaches, as long as these approaches can be justified.

Section 4.1 of the EMA guide, Calculation of a Permitted Daily Exposure (PDE) (2), states, "the use of other approaches to determine health-based exposure limits could be considered acceptable if adequately and scientifically justified."

Once a potential hazard has been identified ... manufacturers must demonstrate that the cleaning process is capable of achieving a safe level of the API.

In addition, the following passage appears in Section 5.3, Therapeutic Macromolecules and Peptides (2):

"Therapeutic macromolecules and peptides are known to degrade and denature when exposed to pH extremes and/or heat, and may become pharmacologically inactive. The cleaning of biopharmaceutical manufacturing equipment is typically performed under conditions which expose equipment surfaces to pH extremes and/or heat, which would lead to the degradation and inactivation of protein-based products. In view of this, the determination of health-based exposure limits using PDE limits of the active and intact product may not be required."

SCIENTIFICALLY JUSTIFIED ALTERNATIVES

Qualifications for acceptance of scientifically justified alternatives to ADE/PDE also appears in this passage, from the proposed ISPE Cleaning Process Development and Validation Guide (3, 4):

"Other methods may be justified for cleaning limit determination where products are demonstrated to be denatured and inactivated (4). Degradation and inactivation studies should be performed and documented. A risk analysis should demonstrate and document that:

- *Degradation has taken place and the extent of that degradation (as degradation does not always occur (5 and 6)*
- *Any by-products of cleaning processes are not hazardous (7)*

Inactivating an API via a cleaning process can be an efficient way to reduce the hazards associated with an API (8)."

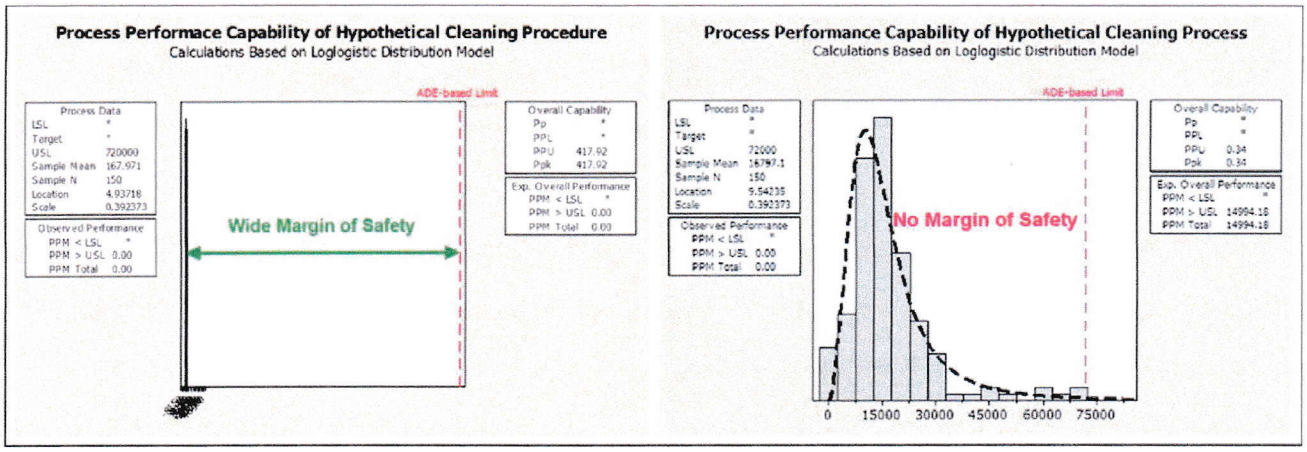
This language was added to the proposed ISPE Guide specifically for some biologics manufacturers. They have argued that they should be exempt from using the ADE to determine cleaning limits and assess the level of risk, since their cleaning processes denature their protein molecules.

[Editor's Note: the author of this article was a principal writer of the proposed ISPE cleaning guide and only agreed to the inclusion of this wording with the stipulations indicated by the references (5–8)].

Although nothing formal has been proposed or justified, some of the possible alternatives discussed among biologics manufacturers have included using:

- Water for injection (WFI) specifications
- Default limits such as 10 ppm (9)
- Process Capability of the cleaning process (9)
- Thresholds of Toxicological Concern (TTC) (10).

Figure 1: Comparison of cleaning process capabilities.



The wording in these documents raises two questions:

- Should regulators allow biologics manufacturers to be exempt from using the PDE/ADE to set limits for assessing risk?
- If so, what alternative approach(es) should they be allowed to substitute?

The answer to the first question lies in the definition of risk, and the way that it is assessed.

DEFINING RISK

EMA also writes in Section 1, the Introduction to its 2015 guideline (2): "Cleaning is a risk-reducing measure...." Thus, EMA recognizes that cleaning is meant to reduce potential risk due to exposure to "pharmaceutical substances," especially APIs.

Both RiskMaPP and the proposed ISPE *Cleaning Process Development and Validation Guide* are based and structured on International Conference on Harmonization (ICH) Q9 (11). ICH Q9 provides a framework under which risk can be quantitated, so that it can be accepted, mitigated or controlled. For these purposes, ICH Q9 defines risk as:

Risk = f (Hazard, Exposure)

If the hazard that has been identified is an API, then this general equation can be further refined as:

$$\text{Risk} = f(\text{ADE}_{\text{API}}, \text{Exposure}_{\text{API}})$$

(Where the ADE_{API} provides a measure of the severity of the potential hazard that the API presents in long-term exposure).

While there may be several different potential hazards associated with cleaning, such as cleaning agent residues and microbial growth, the hazard of principal concern is the potential, unwanted carryover of API(s) of the product. So, first and foremost, the risk associated with the API must be evaluated.

RISK EVALUATION IN CLEANING VALIDATION

In the proposed ISPE's "Cleaning Process Development and Validation" guide, the principal means for evaluating the potential risk presented by an API is to measure the process capability (Cpk/Ppk) of the cleaning process. This approach provides the ability to quantify the effectiveness of a cleaning process for a particular compound so its level of risk can be evaluated. Process capability also allows the cleaning process to be compared to other cleaning processes for other compounds. In addition, process capability provides a means for predicting the potential for possible future failures of the cleaning process. **Figure 1** compares two cleaning processes.

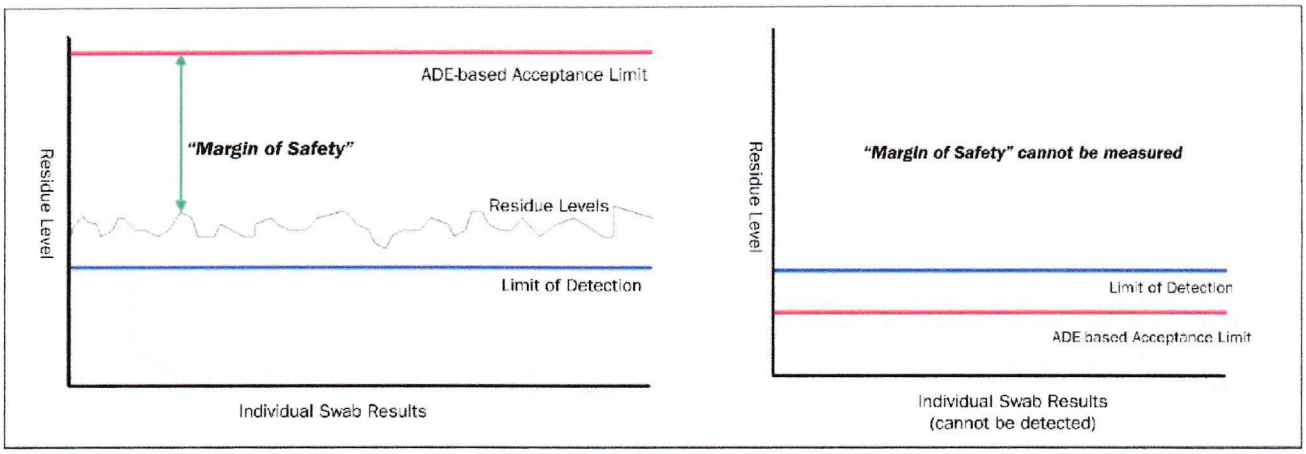
The graph on the left shows a cleaning process that is effective and capable, where the potential for failures is 0.00 out of a million, or <1 part per 100,000,000 and a Ppk of 418 (>>> 6 Sigma). The graph on the right shows a cleaning process that is not very effective and not very capable, where the potential for failures is 14,994 out of a million (1.5%) and a Ppk of 0.34 (< 1 Sigma).

For the cleaning process on the left, the level of potential risk is extremely low and requires no further action. The cleaning process depicted on the right, however, requires improvement. If it could not be improved, continuous monitoring would probably be required to ensure that safe residue levels were continuously achieved, or the compound would have to be manufactured in dedicated equipment.

Figure 2 shows two graphs of the relationship of three cleaning

FIGURES COURTESY OF THE AUTHOR.

Figure 2: Cleaning processes with limits above and below detection limits.



process parameters (acceptance limits, residue data, and limits of detection) for two different cleaning situations. The box at the left depicts a cleaning scenario in which the residue data is well below the PDE/ADE-based acceptance limit, but above the analytical limits of detection so it can be measured, and demonstrates a fair "margin of safety." The box on the right shows a cleaning scenario in which the PDE/ADE-based acceptance limit is below the analytical limits of detection. In this case, the margin of safety cannot be demonstrated.

Improving the analytical method associated with these approaches should be the first course of action. But in some cases, for example, with highly hazardous compounds with very low limits, it may not be possible to make such improvements.

However, does the situation shown on the right in **Figure 2** provide no value? Upon careful consideration, it should be realized that results less than the limits of detection may still be able to demonstrate that the cleaning process is capable of achieving perhaps 70%, or 80% or 90% of the PDE/ADE-based acceptance limit. This information may not cross the finish line from a classical cleaning validation standpoint, but it should not be dismissed entirely.

The scenario on the right can occur with some (but not all) biological products. As stated previously, biologics manufacturers have argued that the PDE/ADE should not apply to them, because the cleaning processes denature or destroy the molecule and it is no longer active or present at all. Is such an argument valid? Consider that a similar argument could be made by small molecule manufacturers, who could argue that, since their cleaning processes remove compounds, or their products are freely soluble and easily washed away, then the PDE/ADE should not apply to them either. Obviously this would not be an acceptable argument for small molecule manufacturers and it shouldn't be for large-molecule manufacturers either. Small-molecule manufacturers are required to demonstrate that residues of their APIs have been removed to acceptable levels and the same should be required of large-molecule manufacturers.

Once a hazard has been identified by the manufacturer, such as the API (and biological molecules are also APIs), it is incumbent upon the manufacturer to demonstrate that the risk of patient exposure to residues of that API are at a safe level. This responsibility cannot be argued away by claiming exemption and invoking other criteria such as WFI specifications, or focusing only on showing that total organic carbon (TOC) data are in statistical control from batch to batch. Instead, manufacturers must demonstrate that the cleaning process is capable of achieving a safe level of the API.

Up until now, the condition *sine qua non* for cleaning validation has been three consecutive runs of passing swab/rinse data (including dirty and clean hold times too). This may have been sufficient evidence in the past, but in the 21st Century, and in a science- and risk-based world, is this really enough to make the claim of cleaning validation?

HAZARD IDENTIFICATION AND RISK ANALYSIS

The proposed ISPE *Cleaning Process Development and Validation* guide also provides specific guidance regarding how to go about identifying hazards and analyzing the risks associated with them (1). It should no longer be acceptable to

Table I: Comparing aseptic process validation to cleaning process validation.

Aseptic process validation	Cleaning process validation
Sterility testing	Residue testing (swab, rinse, etc.)
Sterility assurance level determination	***Cleaning assurance level determination***
N/a	Denaturation/degradation/destruction studies
N/a	Rinseability studies
Facility and equipment design	Equipment cleanability design review
Evidence of formal written procedures	Formal review of cleaning procedures
Overall manufacturing operation (flow, process controls, etc.)	Overall manufacturing operation (separation of dirty and clean equipment, process controls, control strategy)
Operator training	Operator training
Environmental monitoring	Periodic monitoring/control charting
Periodic verification (through simulated runs)	Periodic verification
Monitoring of critical process parameters and quality attributes during routine operation	Monitoring of critical process parameters and quality attributes during routine operation

simply designate satisfactory API, cleaning agent, and micro/endotoxin swab or rinse results as the criteria for declaring a cleaning process "validated."

The proposed *Cleaning Process Development and Validation Guide* directs the user to examine all possible sources of hazards, including the following, and take steps to eliminate or mitigate them (3):

- Chemical
- Microbiological
- Design (identify potetial issues with equipment cleanability)
- Procedural (identify issues with standard operating procedure [SOP] failures).

Clearly now, cleaning validation should become a holistic approach, addressing all the sources of potential hazards. Its scope should include measures to reduce or mitigate associated risks, to achieve the highest degree of safety for the patient.

Likewise, FDA's process validation guidance defines process validation as "the collection and evaluation of data, from the process design stage through commercial production, which establishes scientific evidence that a process is capable of consistently delivering quality product" (12).

Consider also that the scenario illustrated on the right in **Figure 2** is not very different from the challenges found with sterility testing. It has long been recognized that traditional sterility testing has never been able to demonstrate sterility to any significant degree. Sterility testing has never come close to crossing the finish line for validation. It is only through the aggregation of successful sterility runs with other supporting sources of data (such as environmental monitoring, media fills, room pressure differentials, gowning practices, operator training, etc.) that the manufacturer of aseptic products can convince a regulator that the level of risk is low enough to be acceptable. FDA, in its guidance on validating sterilization (13) considers even aseptic processes to be validated when:

"Data derived from experiments and control procedures allow conclusions to be drawn about the probability of nonsterile product units (sterility assurance level). Based on the scientific validity of the protocols and methods, as well as on the scientific validity of the results and conclusions, the agency concludes that the efficacy of the sterilization process is validated. Whether a drug product is sterilized by a terminal sterilization process or by an aseptic filling process, the efficacy of the sterilization process may be validated without the manufacture of three production batches."

Perhaps in the circumstances described above, an approach should be taken similar to that used in aseptic processing: the sterility assurance level (SAL)—and start to implement a cleaning assurance level (CAL) approach, where more than just swab and rinse data are part of the validation (See **Table I**).

DETERMINING A CLEANING ASSURANCE LEVEL

How would CAL be determined? A general definition of SAL for sterilization processes is 10^{-6} reflecting the probability of a single unit out of 1,000,000 being non-sterile after it has been subjected to sterilization. For cleaning processes, a general definition for CAL could be 10^{-6} reflecting the probability of one single cm^2 out of 1,000,000 being ≥ the maximum safe surface residue (MSSR) after it has been subjected to cleaning.

CAL could also be extrapolated to the probability of not more

than one dosage unit exceeding the health-based limit (as applied to a single unit) out of one million units processed on shared equipment surface. Interestingly, the idea for CAL already exists in the medical devices industry (14).

Experimental data from the Center for Pharmaceutical Cleaning Innovation have shown a clear inversely proportional relationship between residue levels and time during a cleaning process similar to that of the model shown in **Figure 3**.

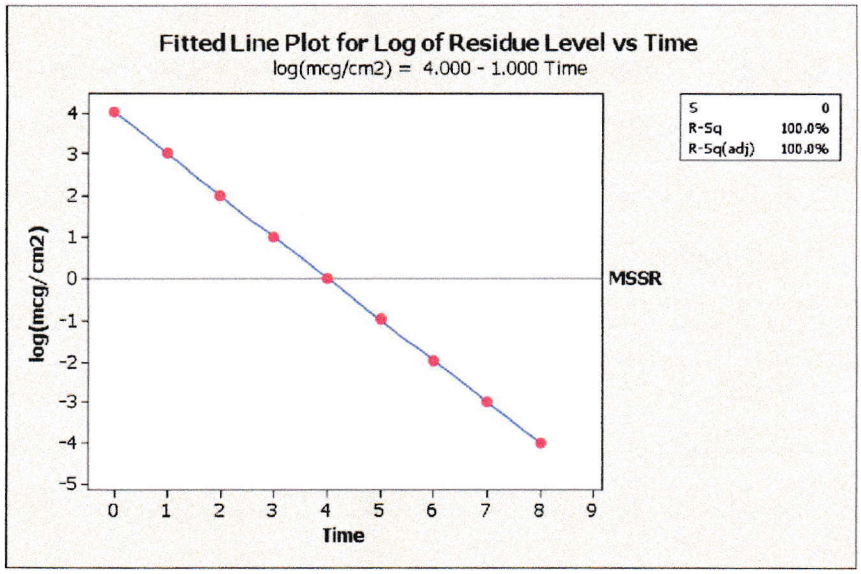
Figure 3 shows a graph of a hypothetical cleaning process that exhibits a log-10 reduction in residue level over time (minutes). (Note: this graph is strictly for illustration purposes only. The behavior of each product during cleaning must be determined individually because every product will behave uniquely.)

The graph shows a hypothetical cleaning process that starts at time zero with 10,000 µg/cm² of residue and, after four minutes, achieves a hypothetical MSSR level of 1 µg/cm² of residue. After eight minutes, the cleaning process has achieved a 4-log reduction below the MSSR. This behavior is quite similar to what is seen in sterilization processes.

One important aspect of such a graph is that, with such data in hand, it would allow manufacturers to predict when they will achieve their safe limits, enabling them to design efficient cleaning cycle times.

But more importantly, for the purposes of this discussion, this information would also allow manufacturers to demonstrate that they have a sufficient safety margin in cleaning when analytical methods cannot detect low enough to demonstrate removal of residues at the MSSR and can provide a methodology for establishing a CAL. Some readers may comment that studies such as shown in **Figure 3** are too variable, or do not reflect real-world cleaning situations; however, the very same can be said of the sterilization studies that are in current use (15).

Figure 3: Hypothetical model of cleaning assurance level determination.



Taken a step further, this information can be combined with the partial information obtained from the analytical results from the right side graph in **Figure 2** to arrive at a new probability of the level of residues. How can we do that? Bayes Theorem can be used to compute this new probability.

APPLYING BAYESIAN STATISTICS TO EVALUATE RISK IN CLEANING VALIDATION

ICH Q9 states that "risk evaluation compares the identified and analyzed risk against given risk criteria" (11). The same approach is taken in the proposed ISPE Cleaning Guide.

In the case of cleaning for APIs, the given risk criteria are their PDE/ADE-based acceptance criteria. The analyzed risk, in this case, is based on the data from the analytical method combined with the data from the CAL determination. Both provide a certain level of assurance that the cleaning process has reduced residues to a safe level. What we would like to know is how to combine these data to arrive at a new level of probability that the cleaning process has reduced residues to a safe level. Bayes Theorem allows us to do this.

Bayes Theorem starts with prior knowledge and combines it with new information to arrive at posterior knowledge. Bayes' theorem can be stated mathematically as the following equation (16):

$$P(\theta|data) = P(data\ |\theta) \times P(\theta) / P(data)$$

Where, θ represents the parameter evaluated (e.g., average surface residue level, cleaning process capability, etc.) $P(\theta)$ is "prior knowledge" (what is known about θ) before data collection, $P(data\ |\theta)$ is the likelihood function, which assesses the probability of the observed data ("new

information") given a value of θ, and P(θ|data) is "posterior distribution," which provides information on the distribution of θ given the data and P(data) is the probability distribution for observed data.

If the analytical data are substituted for the prior knowledge, and the CAL for the new information, in Bayes' equation, we can calculate a new probability (posterior knowledge) that the cleaning process has reduced residues to a safe level. If the analytical method can demonstrate that 90% of the residues have been removed down to the PDE/ADE-based acceptance limit and the CAL determination has shown that the cleaning process can reduce residues after 8 minutes by 10^{-1} then a new probability of meeting the PDE/ADE-based acceptance criteria can be calculated.

This analysis can be extended further, with additional data such as can be provided by protein denaturation studies or protein destruction studies (e.g., using compounds such as sodium hypochlorite or hydrogen peroxide). Such new information can be used to increase understanding of the probability that a given cleaning process has reduced residues to a safe level.

In addition, evidence must be provided that the degrading agent, that has been shown to be effective in principle, actually does reach all product-contact parts of the equipment (see equipment cleanability review, **Table I**).

Some readers may wonder how probabilities from such studies can be determined, but anything can be measured with a quantifiable degree of confidence; many within the biopharmaceutical industry may be unfamiliar with these approaches (17).

Readers should also note that regulators already accept Bayesian Analysis for clinical trial studies (18). Indeed, FDA states, "the posterior distribution that has been obtained today may serve as a prior distribution when more data are gathered" (18). The same Bayesian approach can be used with the sources of additional data described above.

In short, some biologics manufacturers have proposed the use of alternatives including WFI specifications, 10ppm limits, Thresholds of Toxicological Concern (TTC), or, simply, improved process control as a way around using PDE/ADE-based acceptance limits for cleaning validation when limits are below the limits of detection of the associated analytical method. None of these alternative approaches have been adequately justified, and, more importantly, they all fail to address the question posed by the hazard identification of the API, which is "What is the risk to the patient from residues of this API?"

CONCLUSION

The ADE/PDE-based Acceptance Limits cannot be avoided and must be addressed, whether for small or large-molecule APIs and drug products.

ADE/PDE-based acceptance limits can be evaluated using existing data, through Bayesian statistical analysis, and use of the concept of cleaning assurance level.

ACKNOWLEDGEMENTS

The author wishes to thank Alfredo Canhoto, PhD, Dr. Michel Crevoisier, PhD, Andreas Flueckiger, PhD, Will McHale, Ovais Mohammad, Mariann Neverovitch, Osamu Shirokizawa, and Joel Young for their insightful comments and helpful suggestions.

REFERENCES

1. International Society for Pharmaceutical Engineering (ISPE), *ISPE Baseline Guide: Risk-Based Manufacture of Pharmaceutical Products* (Risk-MaPP), ISPE, September 2010.
2. EMA, "Guideline on Setting Health-Based Exposure Limits for Use in Risk Identification in the Manufacture of Different Medicinal Products in Shared Facilities" www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2014/11/WC500177735.pdf.
3. ISPE, "Proposed ISPE Guide: Cleaning Process Development and Validation," (United States Copyright Office Number TXu001950765, submitted to ISPE for potential publication)
4. D.G. Dolan et al., *Regulatory Toxicology and Pharmacology* 43 (1) 1–9, 2005.
5. K. Kendrik et al., *Journal of Validation Technology*, 15 (3), 69, 2009.
6. R. Sharnez et al, *Journal of Validation Technology*, 18 (4), 42-45, 2012.
7. L. Wigman et al, *American Pharmaceutical Review*, 17 (1), 2014.
8. I Valvis, et al, *Organic Process Research & Development*, 3(2), 44–52, 1999.
9. R. Sharnez et al., *Journal of Validation Technology* 17 (3), 17-25, 2012.
10. Sharnez, R et al., *Journal of Validation Technology*, 17(3)24-28, 2011.
11. ICH, Q9 *Quality Risk Management* (ICH, June 2006).
12. FDA, *Guidance for Industry, Process Validation: General Principles and Guidance* (CDER, January 2011). www.fda.gov/downloads/Drugs/.../Guidances/UCM070336.pdf
13. FDA, *Guidance for Industry for the Submission Documentation for Sterilization Process Validation in Applications for Human and Veterinary Drug Products*, www.fda.gov/downloads/drugs/guidancecompliance regulatory information/guidances/ucm072171.pdf
14. S. Richter, "Reuseable Medical Device Cleaning Validation and Testing Requirements: An Industry Perspective," www.microtestlabs.com/documents/wp_Reusable-Medical-Device.pdf
15. The limits of sterility assurance, *GMS Krankenhaushygiene Interdisziplinär*,Vol. 3(3), 2008, ISSN 1863-5245
16. W. Bolstad, *Introduction to Bayesian Statistics*, Second Edition (John Wiley and Sons, 2007).
17. D. Hubbard, *How to Measure Anything: Finding the Value of Intangibles in Business*, Third Edition (John Wiley and Sons, 2014).
18. FDA, *Guidance for Industry on the Use of Bayesian Statistics in Medical Device Clinical Trials*, www.fda.gov/RegulatoryInformation/Guidances/ucm071072.htm **BP**