Statistical Sampling Plan Alternatives for Seal Strength Testing on Pre-sealed Tyvek® Bags

Félix Bruno
Master in Manufacturing Engineering Program
Iván M. Avilés Segundo, Ph.D.
Industrial Engineering Department
Polytechnic University of Puerto Rico

Abstract *— Seal strength is an important quality of sterile barrier systems used for Pharmaceutical and Medical Device Packaging operations, and subject to strict quality control industry-wide. This study evaluated the feasibility of sampling plans alternatives, specific for testing of Tyvek pouches, with focus on reducing cost with minimum impact on risks. Previous studies, data analysis, and results obtained showed minimum influence by variables such as suppliers and product characteristics, providing valuable information and the framework to support cost-effective process improvements.*

Key Terms *— pouches, seal integrity, seal strength, sterile barrier*

BACKGROUND

According to the International Organization for Standardization (ISO), a sterile barrier system (figure 1) is the minimum package that prevents ingress of microorganisms and allows aseptic presentation of the product at the point of use [1] [2]. Other terms of equal importance are seal integrity (a characteristic of the seal that minimizes the risk of ingress of microorganisms demonstrated under test conditions that consider sterilization process, handling, distribution, transport, and storage) and seal strength (mechanical capacity of the seal to withstand force) [1] [2].

Figure 1: Example of a sterile barrier system

For the Medical Device manufacturing industry, maintenance of package integrity (e.g., no holes in materials and no channels in seals) through sterilization and distribution to the point of use is the bottom line in Terminally Sterilized Medical Device Packaging [1]. To that purpose, organizations are required to establish the necessary process steps for ensuring that sterile barrier systems meet specified requirements [1] [2]. That way, companies can demonstrate conformity to requirements and ensure these packaging components withstand real-world stresses of sterilization, distribution, and product load on package seals [1].

Due to the criticality of the of the component and depending on the type of industry in which the company is operating (e.g., medical and pharmaceutical companies), organizations may have practices that follow recognized industry standards/test methods such as the ones the American Society for Testing and Materials (ASTM) has developed throughout time. Among those standards, there is ASTM F88 for seal strength [3], which is the one followed by Becton Dickinson.

Specifically, within ASTM F88, seal strength is described as a quantitative measure for use in process validation, process control, and capability [3]. ASTM F88 uses a defined approach to testing with the purpose of measuring resistance during seal separation. This test is particularly applicable to peelable medical package seals.

Different techniques were established to hold samples at various angles to the pull direction to control this force (figure 2). These techniques were the result of protocols that were designed to support the development of ASTM F88 [3]. Each protocol used different material combinations, with the purpose of identifying the effects of variations in the use of the methods, on the final measured result as

well as on repeatability (r) and reproducibility (R) [3].

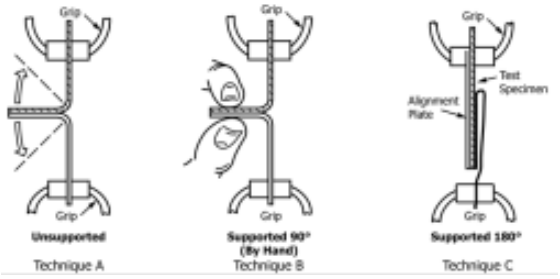

Figure 2: Examples of tail holding methods or techniques per ASTM F88 [3]

For clarity purposes on the testing techniques, when using the unsupported technique (Technique A), the angle of peel is constantly varying from 90° to greater than 90°. In the supported-tail technique (Technique C), the tail is restrained by a fixture to keep the angle of peel at 180° [3].

Techniques/methods were considered to be acceptable, but the results of the measured specimens were not equivalent [3]. Studies showed that the resulting values were higher when using Technique C in comparison to results obtained when using Technique A [3].

ASTM studies also demonstrated that even with differences in the end result when moving from one technique to the other, that comparison of sample sizes n=10 versus n=30 may not show large differences in either reduction or increase in variation [3]. The overall average of measured values differed by less than 0.1 lbs/in when comparing data series using a sample size n=10 versus a sample size n=30 [3].

With regards to specification levels in Medical Device Packaging there has been an arbitrary seal strength minimum value of 1.0 lbs./in, for many years [4]. This specification is often perceived as a regulator requirement, and it is not [4]. There is no definitive answer on where the value came from [4]. To that point, ASTM WK57656 is intended to provide a standard guide for determining minimum seal strength for Medical Device Sterile Barrier Packaging Systems [4].

Based on available information regarding the development of guidelines on how to conduct testing [3] and how to establish minimum seal strength limits [4], the possibility of alternative sampling plan options would be of further evaluation, given the fact that a reduction in samples may not be a factor that increases risk. However, data analysis must consider that the results will vary depending on the support condition or testing technique, so test results for different testing techniques cannot be compared [5].

PROBLEM STATEMENT

This study was performed in a Becton Dickinson (BD) facility, located in Nogales, Sonora, Mexico. This manufacturing site is dedicated to the manufacture of medical devices focused on gastrointestinal care and urological drainage. Most of the products are packaged/stored inside Tyvek® sealed bags, which are then sterilized and labeled prior to final distribution. To guarantee the device's sterility, the bags must be completely sealed; otherwise, air may enter the bag and contaminate the device. Due to the wide variety of products at the site, there are 31 different types of bags that are currently supplied by three different suppliers (A, B, and C), supplier A being the preferred one. Although these bags use the same material type and grade, they differentiate in size and shape. Pouches have three pre-sealed sides, called the sealed perimeter (also known as seal width), which is common among the entire family of pouches.

Based on current procedures, each supplier's lot of pre-sealed bags is subject to quality inspection using a Zero Sampling Plans [6]. From the series of tests performed, one of the most critical is seal integrity. For this, a total of six sealed area specimens are cut and individually tested following internal test methods based on ASTM F88. Therefore, the quantity of independent samples is exponentially incremented by the number of sampling points of each bag.

The current inspection plan for seal-strength testing (destructive testing) is established as C=0, AQL = 0.65 [6]. Data from a set of 82 lots received in a four-month period showed that a sample size ≥68 is taken 90% of the time (75 out 82 lots) (table 1).

Table 1: Distribution of the samples per lot received at incoming inspection

Sample Size	Count of Lots	% of total lots	Cumulative % of total lots
53	7	8.54%	8.54%
68	17	20.73%	29.27%
77	22	26.83%	56.10%
96	24	29.27%	85.37%
119	12	14.63%	100.00%
Total	82	100.00%	100.00%

Time study performed at the site showed that the time invested in the activity is the top contributor impacting the flow of materials through the area. The length of the inspection (up to 8 hours) is mainly driven by the number of readings recorded from each bag (total = 6). Figure 3 shows an example of the sampling points subject to test for seal strength.

This whole scenario negatively affected the goals, efficiency, and resources in the area.

Due to the cost involved and the great performance of this type of materials, it was considered as a potential opportunity to improve processes at BD, mainly because of the healthy performance noted while reviewing historical data and controls in place at the receiving site.

Data from the purchasing department showed that cost for this type of pouches ranges from 5¢ to $3.40 each, with an average of 21¢ each. In fact, total inspection cost varies from $40 to $760, depending on the material and amount of samples taken. This range in cost only considers labor cost and material waste due to destructive testing.

OBJECTIVES

The initiative is aimed to evaluate alternative sampling plans for seal strength testing, taking into consideration historical product performance, cost, resources, operational efficiency, and compliance to regulatory requirements, without increasing the risk associated with the sampling process. The experiment will be focused on the following:

- Evaluating alternative sample plan seal strength testing with focus on reducing cost without impact on sampling risks.
- Increasing efficiency in the area.

Implementation of alternative sampling plans for destructive seal strength testing will reduce cost, increase capacity, and, therefore, give a much-needed boost to operational efficiency without compromising the quality of the products and compliance to regulatory requirements. The project also has the potential to be extended to any of the 80+ other facilities within BD.

The experiment could be applied to consider alternative sampling for destructive testing on other products with similar performance. Research may also provide alternatives to traditional sampling for companies working in the regulated industry, to reduce cost without impact on risk.

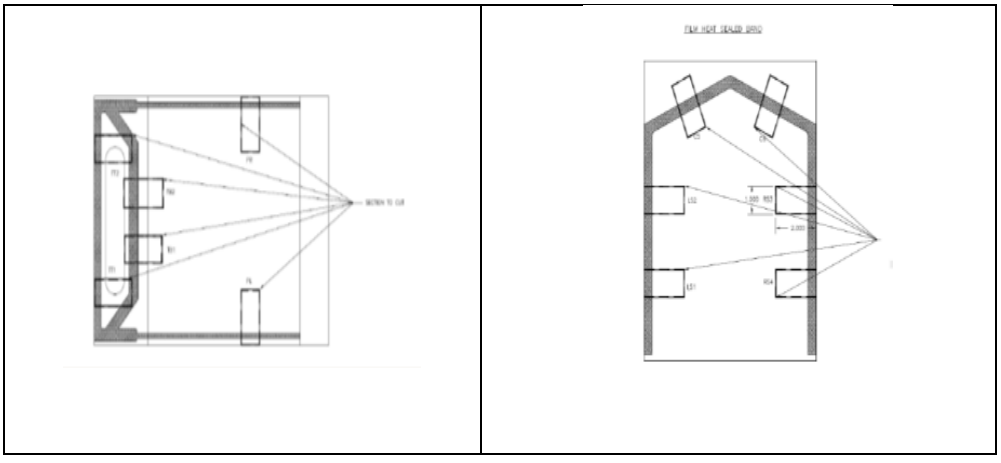

Figure 3: Examples of package seal perimeter subject to seal strength testing

METHODOLOGY

Methodology will consist of the use of DMAIC and its structured problem-solving tools. Phases in this methodology build one over the other with the goal of implementing long-term solutions to problems when risks are high [7]. Each of the phases used are described as follows:

- **Define:** This phase consists in defining the problem statement, scope, goals, and project objectives. It also focuses on the identification of customers and their requirements, and the determination of skills and areas that need representation on the project [7]. One of the main tools that could be used on this phase to describe the project opportunity may be the project charter.
- **Measure:** The objective of this phase is to focus on the identification of the process steps, and corresponding inputs and outputs. Baseline are established with trustworthy and relevant data, which is collected from the process [7].
- **Analyze:** This phase consists of identifying inputs and their relationship with the outputs as well as determining root causes impacting process performance [7]. The key components of this phase may include statistical tests and analysis of variance, among others.
- **Improve:** This is where potential solutions are identified and evaluated, and the process is optimized. The critical inputs that must be controlled to maintain performance that reliably satisfies the customer are determined. Process capability and project financials are estimated [7].
- **Control:** This phase establishes mistake proof, long-term measurement, and reaction plans. The team develops or updates procedures or documentation and establishes process capability. Project financials are updated, verified, and reported. Control plans may be established with reaction plans and ownership [7].

DATA COLLECTION AND CHARACTERIZATION

Relevant data was collected from the process with the purpose of identifying similarities among the product family or differences that may be considered potential sources of variation. Table 2 shows the data collection plan used for the experiment.

For the purposes of data analysis, a comparison of the characteristics within the family of sterile barrier pouches was conducted (table 3).

Table 2: Data collection plan

Requirements	Characteristics	Objective / Comments
Financials	Cost by part, Inspection Cost, Time	• Calculate destructive testing cost and time/resources invested
Volumes	Lot Qtys, Lot Sizes, Samples Sizes	• Set sample size baseline • Calculate destructive testing cost and time/resources invested
Risk Analysis Tools	FMEA	• Identify potential failure modes and associated risk
General	Inspection Procedures and Specifications, Suppliers	• Identify similarities and differences
Physical Requirements	Sampling Plan	• Set sample size baseline
	Dimensions (L x W)	• Identification of characteristics within the product family
	Material Types	• Identification of characteristics within the product family
	Seal Width	• Identification of characteristics within the product family
	Measuring point (6 sides)	• Identification of characteristics within the product family
Functional Requirements	Testing techniques	• Identification of characteristics within the product family

Table 3: Comparison of the characteristics within the product family

Requirements	Characteristics	Common	Different
General	Inspection Procedures	✓	
	Sampling Plan	✓	
	FMEA	✓	
Physical Requirements	Dimensions (L x W)		✓
	Material Types		✓
	Seal Width		✓
	Measuring point (6 sides)		
Functional Requirements	Testing technique	✓	✓

Common characteristics were discarded as potential sources of variation. On the other hand, differences were further investigated to understand its contribution to the end result.

Risk assessments tools (e.g., Failure Mode Effects Analysis [FMEA]) documented at the site showed an associated severity of 9 and occurrence of 1 for Failure Modes such as Blown Seal, Seal Creep, or Fractures. Detection for this type of failure modes was set as 4, in a scale of 1-10.

Although severity of this failure mode was classified as "9 - Critical Health Hazard," the occurrence is at its lowest and pouches get inspected by incoming quality as a control checkpoint.

With regards to the material type, no evidence was found in previous studies (conducted prior to the development of ASTM F88) to establish that the material causes a difference in the resulting value [3]. Something similar happened with seal width, where the specified requirement was found to be common among all part numbers (3/8"), but tolerances were different by 1/16" maximum in some of them, thus causing no impact to the study. Therefore, material type and seal width were also discarded as potential sources of variation.

Previous studies also found that resulting seal strength value may vary depending on the testing technique used [3]. This is why functional requirements at BD established an acceptance criterion of 0.75 LB/IN minimum, with a target of 1.0 LB/IN or greater per bag, for items inspected using Technique C. The maximum value accepted is 4.0 LB/IN. When using Technique A, individual results must be 1.25 LB/IN minimum, with a target of 1.5 LB/IN or greater per bag. The maximum value accepted is 5.0 LB/IN. Technique B is not used. Values over the upper specification limit are considered acceptable if film fracture or separation do not occur, since by design those sides are not intended to be opened by the end user.

Another difference noted was the ID used to identify each of the sides/specimens. In this case, those were divided into three different groups.

Experimentation on these differences was conducted to evaluation potential impact.

HYPOTHESIS, EXPERIMENT ANALYSIS & RESULTS

Statistical analysis was conducted to confirm the impact of the differences found in the characterization exercise. Data collected from the 82 supplier lots was used for the experiment. Data was traced back to receipts from a period of four months.

A two-sample T-test was performed with a significance level of 0.05, to confirm results from previous studies on the differences between testing techniques (e.g., A vs C) [3]. The test result failed to reject null hypothesis (Ho: $\mu_1 - \mu_2 = 0$), confirming that resulting values using Technique C are greater than the ones obtained when using Technique A.

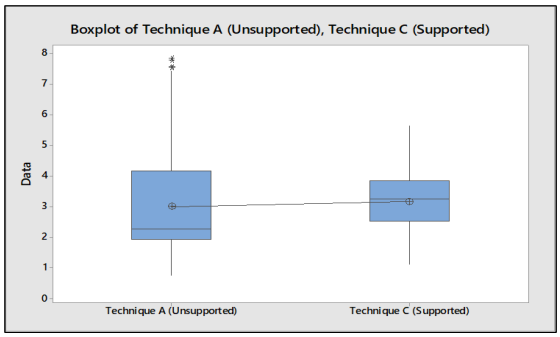

Figure 4: Box plot result from the two sample T-tests

Therefore, additional analysis considering the values corresponding to those techniques were evaluated separately from each other, as two different populations.

Additional evaluation noted differences between some of the measuring points (sides) inside both groups (Technique C – Supported and Techniques A – Unsupported). After running an ANOVA experiment, results for each group showed

that the average for some of the measuring points was found to be statistically significant. Figure 3 shows examples of the measuring points subject to testing.

Among measuring points, differences were mainly noted when observing results from the Unsupported Group (sides C5 and C6) as well as the Unsupported Group (sides TB1 and TB1), in comparison with mean value for other sides within the same group.

With regards to the supported group, it was observed that sides C5 and C6 were also distanced from the others with respect to the mean value.

Although differences in the mean value for these measuring points (sides) were considered as statistically significant, in reality the values were located at a healthy distance from the specification limits (e.g., 0.75 LB/IN minimum, 1.25 LB/IN minimum). Capability Analysis studies were conducted, and results showed Cpk values > 1.29 in all cases, suggesting results were robust enough within specification limits and above 2.0 LB/IN.

Figures 5 and 6 show that there were no differences noted on all other measuring points (LS1, LS2, RS3, RS4) or (TB1, TB2, TT1, TT2), suggesting that improvements were also possible in that area as well for reduction of samples taken.

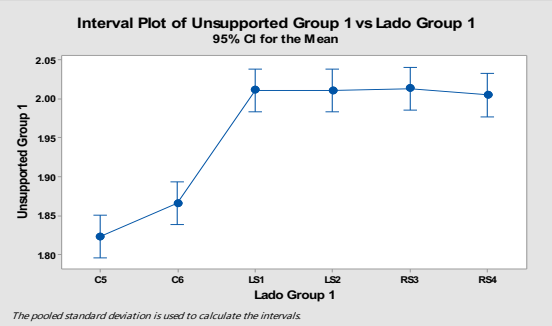

Figure 5: Examples of differences between measuring point (sides) - Group Unsupported

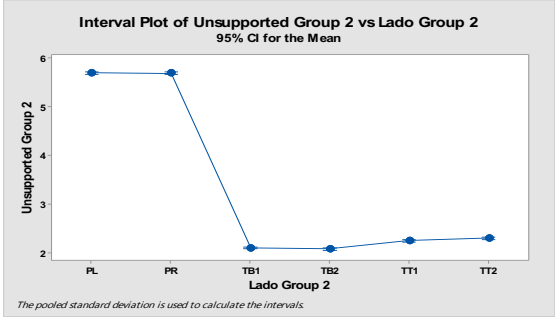

Figure 6: Examples of differences between measuring point (sides) - Group Unsupported

Figures 6a and 6b display a closer view to distinguish the behavior of measuring points/sides (PL and PR) and (TB1, TB2, TT1, TT2) is well located far from the specification limits.

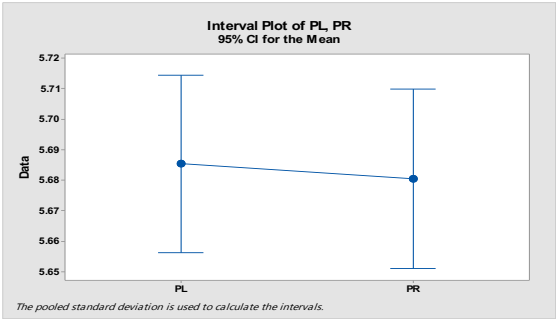

Figure 6a: Zoom in to PL and PR measuring point (sides) - Group Unsupported

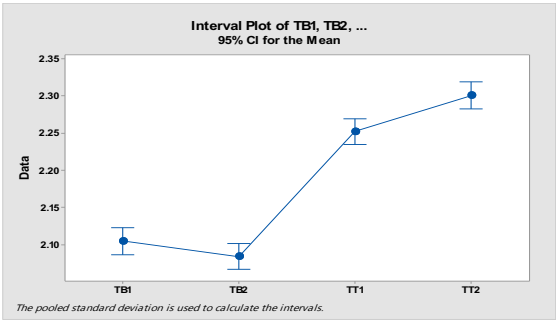

Figure 6b: Zoom in to TB1, TB2, TT1, and TT2 measuring point (sides) - Group Unsupported

Information collected through statistical analysis suggested that associated risk of experiencing expected failure modes would be minimum. Additionally, it was determined that any alternative sampling plans capable of maintaining sampling plan properties (e.g., consumer and producer risk) within acceptable limits would be acceptable to judge quality characteristics for the entire population.

IMPROVEMENTS

Based on results obtained, it was decided that there were at least three methods providing immediate and robust solutions to the pursuit of alternative sampling. Methods considered were divided as follows:

- Since data showed no statistical significance difference between measuring points in the perimeter (e.g., LS1 and LS2 sides on figure 5, PL and PR sides on figure 6), an alternative to the existing sampling would be to allow for the selection of a single measuring point from each side in the perimeter instead of taking two. As shown in figure 7, this approach allows for the selection of a single measuring point (e.g., A or B, C or D, E or F) reducing the quantity of measuring points subject to testing by 50% (from 6 to 3). This change lowers the time invested in the inspection by the same percentage (50%), having an impact on cost of at least $4.8k per year, considering that the projected volume of receipts is ~250 lots/yr. No additional risk is expected with this alternative, since sampling plan (AQL = 0.65 / C=0) remained the same.

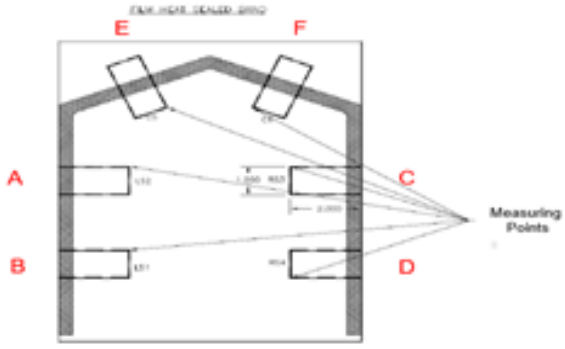

Figure 7: Examples of measuring point (sides)

- The second alternative allows for a change in current sampling plan (AQL = 0.65 / C=0) to a sampling of 50 units. As shown in figure 8, a comparison of OC Curves using current scenarios (68, 77, or 96 samples) versus the new sampling plan (50 samples) was made. It was found that, when compared to the actual sampling plans, the proposed sampling plan increases the probability of acceptance for the supplier (Producer Risk) by 8% while Lot Tolerance Percent Defective increases by only 1% (from 3.4 to 4.4), remaining below the 10% Lot Tolerance Percent Defective mark. The change in risk is shown in figure 9.

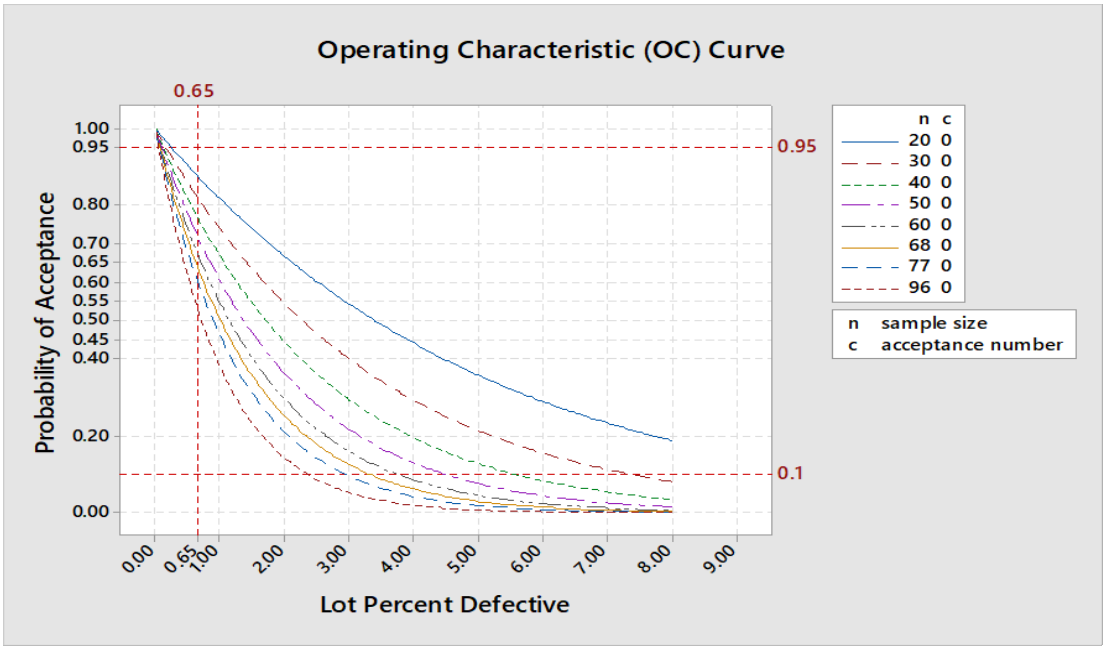

Figure 8: Sampling plan comparison using OC curves

Compare User Defined Plan(s)

Sample Size(n)	Acceptance Number(c)	Percent Defective	Probability Accepting	Probability Rejecting	AOQ	ATI
20	0	0.65	0.878	0.122	0.570	4297.3
30	0	0.65	0.822	0.178	0.534	6243.8
40	0	0.65	0.770	0.230	0.500	8066.9
50	0	0.65	0.722	0.278	0.468	9774.4
60	0	0.65	0.676	0.324	0.439	11373.7
68	0	0.65	0.642	0.358	0.416	12579.8
77	0	0.65	0.605	0.395	0.393	13863.3
96	0	0.65	0.535	0.465	0.347	16336.6

Figure 9: Alternatives using OC Curves

This approach reduces the quantity of samples taken between 6% - 58%, depending on lot size. The new plan will lower the time invested in the inspection between 6-151 minutes (5% min, 58% max), having an impact on cost (impaction time + destructive testing only) of at least $6.4k per year, considering that the projected volume of receipts is ~250 lots/yr.

- Alternative #3 combines the proposed approach in alternatives #1 and #2. As previously described, the risk associated to Alternative #1 is minimum, and the risk associated to Alternative #2 increases the probability of acceptance for the supplier (Producer Risk) by only 8%, while Lot Tolerance Percent Defective increases by only 1% (from 3.4 to 4.4), still remaining below the 10% Lot Tolerance Percent Defective mark.

  Based on historical data from the last two years, 100% of supplier lots/batches have been accepted. Therefore, an increase of 8% in the probability of acceptance for the supplier (Producer Risk) by 8% introduces minimum risk into the process. This approach combines the benefits of both alternatives for an overall impact on cost and savings of at least $11.2k per year, considering that the projected volume of receipts is ~250 lots/yr.

CONTROLS

Based on the results obtained, it was decided that the solutions for the scrap reduction will be divided into four approaches that, combined, will provide a robust solution to the problem. The implemented actions are:

- **Inspection Procedure (IP) changes:** The inspection procedure was updated to include details regarding the new sampling plan for inspection of the seal strength characteristic. This procedure change ensures consistent application of the plan across all items in the product family.
- **Test Method changes:** The Test Method was updated to clarify the instructions regarding the quantity of samples to be taken from the sealed perimeter that would be subject to testing. This procedure change ensures consistent application of the alternative across all items in the product family.

CONCLUSIONS

It can be stated that the objectives and main purpose of the project were achieved. Alternative sampling plans for seal strength testing were proposed, taking into consideration product performance, cost, resources, operational efficiency, and compliance requirements.

Controls will ensure the achievement of the goal of improving efficiencies, leading to an estimated annual savings of $11.2k, including materials and labor as expected. Additionally, it will improve yield in the area and may also have an impact on capacity, depending on the usage of allocated resources.

The implementation of the project provides additional opportunities to extend actions to other products in the manufacturing plant or even in other facilities. Further experimentation can be made for evaluation of the following:

- Implementation of sampling plans by variables using ANSI/ASQ Z1.9 [8]
- Evaluation of material types as a factor for variability

Management is looking forward to this since it can provide additional savings for higher-cost products.

REFERENCES

[1] *Medical devices - Quality management systems - Requirements for regulatory purposes*, ISO Standard 13485, 2016 [Online]. Available: https://www.iso.org/standard/59752.html

[2] *Packaging for terminally sterilized medical devices — Part 1: Requirements for materials, sterile barrier systems and packaging systems,* ISO Standard 11607-1, 2019 [Online]. Available: https://www.iso.org/standard/70799.html

[3] *Standard test method for seal strength of flexible barrier materials*, ASTM Standard F88 / F88M-15, 2021 [Online]. Available: https://www.astm.org/f0088_f0088m-15.html

[4] G. Pavey, "Seal strength testing: Is 1 lb. still the go-to metric? ASTM WK57656 Update," Oliver, March 13, 2020 [Online]. Available: https://www.oliverhcp.com/news-and-resources/packtalk/seal-strength-testing-is-1-lb.-still-the-go-to-metric-astm-wk57656-update

[5] Packaging Digest, "Strength and integrity, part one: The basics of medical package testing," January 4, 2016 [Online]. Available: https://www.packagingdigest.com/testing/strength-and-integrity-part-one-basics-medical-package-testing

[6] N. L. Squeglia, *Zero Acceptance Sampling Plans*, 5th ed., Milwaukee: ASQ, 2008.

[7] C. F. Berardinelli, "Best of back to basics: To DMAIC or nor to DMAIC?" *Quality Progress*, vol. 49, no. 1, p. 36, Jan 2016 [Online]. Available: https://asq.org/quality-progress/articles/best-of-back-to-basics-to-dmaic-or-not-to-dmaic?id=5c9c7d86cfd9418090852faa7d0538b7

[8] *Sampling Procedures and Tables for Inspection by Variables for Percent Nonconforming*, ANSI/ASQ ANSI/ASQ Standard Z1.9–2003 (R2013), 2018.