

Validation of Test Methods for Characterizing and Specifying Materials Used in the Construction of Sterilization Packaging

By Christopher William Berry^{1*} and Lorraine Harding²

¹Smithers Pira, Cleeve Road, Leatherhead, Surrey KT22 7RU, UK

²Sterile Barrier Association, St Weonards, Herefordshire UK

This article details the validation of preexisting test methods from EN 868 to enable their adoption into ISO 11607. ISO 11607-1:2006 specifies the requirements and test methods for packaging materials and barrier systems, which are intended to maintain the sterility of terminally sterilized medical devices until the point of use. To evaluate each method, we conducted a precision experiment as described in ISO 5725-2. The methods assessed are as follows: water repellency (EN 868 parts 2, 3, 6 and 7), pore size (EN 868 parts 2, 3, 6 and 7), chloride and sulfates content by hot extraction (ISO 9197 and ISO 9198). Copyright © 2012 John Wiley & Sons, Ltd.

Received 24 June 2011; Revised 26 March 2012; Accepted 13 April 2012

KEY WORDS: sterilization; packaging; medical devices; EN 868; ISO 11607

INTRODUCTION

The sterilization of medical devices before use in aseptic conditions is essential for reducing the risk of infection in hospital environments.¹ In practise, it is impossible to prove that all microorganisms have been destroyed; therefore, sterility assurance levels are used to describe the probability of viable, contaminating organisms surviving after terminal sterilization. In Europe, to enable a medical device to be labelled 'sterile', a sterility assurance level of 10^{-6} must be met. This means that there is less than or equal to one chance in a million that an item remains contaminated or nonsterile.

The packaging used for medical devices allowing them to be sterilized is known as a *sterile barrier system*. This provides a microbial barrier and maintains sterility effectively up to the point of use. A sterile barrier system is an essential part of a sterile medical device and is considered to be an accessory to the device, as defined under the amended European Medical Device Directive (Council Directive 93/42/EEC), and thus must conform to the same legislation as the device.

A sterile barrier system is defined within ISO 11607:2006 as 'the minimum packaging configuration that provides a microbial barrier and allows aseptic presentation of the product unit at the point of use'. This is distinct from protective packaging, which is defined as the 'packaging configuration designed to prevent damage to the sterile barrier system and its contents from the time of their assembly until the point of use'.

The medical device is typically sterilized after it has been sealed within its sterile barrier system, and this process is defined as *terminal sterilization*. The sterile barrier system must allow an effective sterilization of the medical device, withstand the sterilization process and maintain the microbial barrier after sterilization. Terminal sterilization for medical devices can be achieved through a variety of technologies. The main ones used in the medical device industry are as follows:

* Correspondence to: C. W. Berry, Smithers Pira, Cleeve Road, Leatherhead, Surrey KT22 7RU, UK.
E-mail: cberry@smithers.com

Heat	Steam, dry
Radiation	Electron beam, gamma
Gaseous	Ethylene oxide, formaldehyde
Low-temperature oxidative	Vaporized hydrogen peroxide, hydrogen Peroxide/gas plasma

With the exception of radiation and dry heat, all the previously mentioned processes require porous materials to allow the sterilant to permeate the sterile barrier system.

Sterilization packaging can take several forms,^{2,3} including bags, pouches and preformed packs, and is usually made from a combination of paper, plastic and nonwoven materials. Consequently, the materials used for sterile barrier systems must conform to requirements,^{4,5} which dictate a range of critical properties that include porosity, water resistance, heat sealability, acidity, and so on. However, historically, these requirements have not been defined within a single standard, which has presented a source of confusion⁶ for the packaging industry.

Over the last 10 years, there has been a continuing effort to harmonize global standards for medical packaging. The industry now has a global harmonization of standards for medical packaging, including the packaging process, in the form of ISO 11607—Packaging for Terminally Sterilised Medical Devices: Part 1—Requirements for Materials, Sterile Barrier Systems and Packaging Systems and Part 2—Validation Requirements for Forming, Sealing and Assembly Processes. Both parts of the standard were designed to meet the relevant essential requirements of the European Medical Device Directives.

Historically, the European standards EN 868 parts 2–10 have provided specific performance requirements for particular materials and preformed sterile barrier systems intended to maintain the sterility of terminally sterilized medical devices. Although not mandatory, each standard provides examples of particular requirements and test methods that can be used to demonstrate compliance with one or more of the requirements of ISO 11607 part 1. It has therefore been desirable for the test methods described in the EN 868 parts 2–10 to be transferred into ISO 11607. However, some of these test methods have no record of validation, and validated alternative methods are not available or are not appropriate for the materials or products involved. Therefore, the Sterile Barrier Association, whose members produce more than 70% of the total European medical market demand for sterile barrier systems, has commissioned the following series of interlaboratory studies (ILS), which have been performed by their members, to determine the repeatability and reproducibility of these methods with the aim of validating some of these methods.

EVALUATED TESTS

Several currently nonvalidated tests are regarded by members of the Sterile Barrier Association to be essential tools in assessing the performance of materials used in the construction of sterile barrier systems. It is therefore essential that their validity be determined. A brief description of each the currently unvalidated tests methods are given in the following sections, together with their importance for sterilization packaging.

Pore size

The pore size test found in EN 868 parts 2, 3, 6 and 7 is used to determine the mean diameter of pores on the surface of a material. The test is performed by measuring the pressure required to force air bubbles through the interstices of a porous material, wetted by a liquid and having a layer of the same liquid applied to its upper surface. The surface tension of the liquid and the pressure applied are used to estimate the size of the interstices in the material. This method is important because papers with controlled pore size distribution are used in sterile barrier systems to provide an environmental and biological barrier, while allowing sufficient airflow to be used in gaseous sterilization methods such as ethylene oxide, steam and gas plasma.

Although several studies investigating the size of leakage routes in integral packaging exist,^{7–9} they are not practical for obtaining the quick measurements of pore size typically required within a quality

control environment. As an alternative, ASTM F2638-07 provides a method of assessing the filtration efficiency of material using a defined aerosol. However, the cost of the test equipment prevents its universal adoption.

Water repellency

Water repellency (EN 868 parts 2, 3, 6 and 7) of paper used for sterile barrier systems is essential to control the penetration of inks and coatings during processing and to ensure that the microbial barrier is not compromised during sterilization and storage. The test method determines water repellency by measuring the mean penetration time of water through the material. This is achieved by sprinkling the upper surface of a sample with a dry, water-transudation ultraviolet indicator powder. The sample is then floated on purified water and observed under controlled ultraviolet lighting to determine the time required for general fluorescence to occur, indicating that water has permeated through the structure of the material.

Although the wettability of porous surfaces has been examined in detail,¹⁰ there are currently no alternative methods for evaluating the water repellency of materials used in the construction of sterile barrier systems. EN 868 part 2 does include test methods for assessing the repellency of low-surface tension liquids (Annex A) and saline repellency (Annex E), but these are typically measured over a longer period so the principle involved in determining when penetration has occurred is less critical and therefore not as complicated.

Chloride and sulfate content

The water-soluble chloride and sulfate contents of the materials used for sterile barrier systems can affect the acidity of the material; hence, it is important to control them. This could affect the coatings, the inks and the medical device itself. The EN 868 series of standards specify that ISO 9197 and ISO 9198 can be used to measure the water-soluble chloride and sulfate contents of paper. However, it specifies that a hybrid of each test performed using a hot extraction method with boiling water is used in accordance with ISO 6588-2:2005, 7.2.

Although each of these standards already contains a section relating to precision, the hybrid nature of the tests requires that the chloride and sulfate content tests be validated using the hot extraction method.

OBJECTIVE

The purpose of this study was to validate the pore size, the water repellency, the chloride content and the sulfate content tests by determining the repeatability and reproducibility of each method as defined within the EN 868 standards.

The robustness and suitability of the test methods can then be assessed by comparing the current acceptance criteria defined within the test methods, with the typical error associated with performing the tests.

METHODS

The repeatability and reproducibility standard deviations for each test method were determined using a series of precision experiments, which were performed and analysed using ISO 5725-2. All participating laboratories were instructed to perform the tests in accordance with the appropriate EN 868 methods. Each test was performed at several levels to assess the method over the range of its typical application.

Test samples

The test samples were created from a selection of consistent paper materials, which were representative of the type used in the industry and typically evaluated using each of the test methods being assessed. All test materials were preconditioned to 30 °C, 30% relative humidity, randomized, labelled and then dispatched to the participating laboratories. The laboratories were then instructed to condition the samples to 23 °C, 50% relative humidity, before testing them in accordance with the standard.

The hygroscopic nature of paper materials means that their properties are affected by their moisture content. Therefore, preconditioning and conditioning the materials before testing are important to ensure that all samples are tested at the same moisture content. Paper materials are also affected by hysteresis, which means that the moisture content is not just dependent on the current temperature and humidity but also affected by the previous conditions in which it was exposed.

Interpretation of results

After the completion of the test, the results were returned to Pira for statistical analysis. The results were critically assessed using the Cochran and the Grubbs tests to identify any stragglers or statistical outliers. Any results that were identified as stragglers or outliers were then examined to determine the cause. Where appropriate, the outliers were excluded from the determination of the general mean, repeatability and reproducibility standard deviations for each test level. The repeatability and the reproducibility of each method were then determined over the tested range as a function of the mean measurement.

Unfortunately, the limited number of laboratories who regularly perform these tests meant that the number of participants available for the ILS was considerably less than 30, recommended within ISO 5725-2. A lower number of participants for each test evaluation meant that the statistical analysis techniques are less effective in determining stragglers and outliers. However, the Cochran and the Grubbs tests are still valid tests if six or more sets of results are present, a requirement that was satisfied for each of the methods tested.

Cochran and Grubbs tests

The statistical analysis was performed in accordance with the numerical outlier technique described in ISO 5725-2. This requires that unexplained, statistically noncoherent results (outliers) should be removed to avoid unjustified overestimation of the repeatability and reproducibility. Although all unexplained, statistically suspicious results (stragglers) should be accepted as corrected but marked for identification.

The Cochran test assesses within laboratory variance at each test level by critically comparing the results of the laboratory with the highest standard deviation against those returned by the other participants.

The Grubbs test assesses between-laboratory variability by comparing the mean results returned by all laboratories at each level to identify any critically significant differences.

RESULTS

The results from the ILS performed on each test method are presented in the following sections.

Pore size

The pore size method was evaluated using nine different instruments from seven organizations; in each case, the analysis was conducted according to the method referenced earlier. All laboratories performed 10 replicate tests on four sample materials. The sample materials were labelled as levels 1–4 and consisted of three plain papers and one crepe paper. The basis weights were approximately 50, 60, 55 and 50 g/m², each with an approximate thickness of 55, 90, 75 and 130 µm. The mean values and the standard deviations of pore size are shown in Table 1.

Table 1. Test results of pore size (µm).

Instrument <i>i</i>	Levels <i>j</i>							
	1		2		3		4	
1	9.1	(1.9)	14.4	(2.5)	15.7	(2.4)	37.0	(3.9)
2	6.7	(1.3)	10.4	(1.1)	10.7	(1.3)	39.4	(5.5)
3	16.3	(3.0)	21.3	(2.3)	22.0	(1.4)	41.4	(4.7)
4	13.2	(1.9)	20.3	(3.9) ^b	24.1	(3.3)	52.4	(4.2)
5	14.4	(3.3) ^a	16.6	(2.4)	18.5	(3.8) ^b	39.5	(4.1)
6	<12.0	(0.0)	18.6	(2.1)	20.9	(2.5)	44.4	(2.7)
7	13.3	(2.6)	19.9	(2.3)	21.1	(1.8)	55.0	(12.0) ^b
8	13.3	(0.9)	16.0	(0.7)	18.4	(0.8)	39.0	(1.3)
9	2.9	(0.1)	7.4	(0.6)	8.0	(0.3)	21.8	(1.1)

Values are presented as mean (SD).

^aStraggler.

^bStatistical outlier.

The application of Cochran’s test to the standard deviations identified a straggler in level 1 and a single outlier in levels 2, 3 and 4. An investigation into the cause of the straggler and outliers revealed that each case was due to excessive variability, which was inconsistent with the other results within that level. Consequently, the outliers were excluded from the determination of the general mean, repeatability standard deviation and reproducibility standard deviation. The Grubbs test applied to the mean results identified no stragglers or outliers.

The result from instrument 6, level 1 was excluded as an absolute result was not reported, and all results from instrument 9 were also excluded because of the consistently low results relative to those from other laboratories. A summary of the pore size results for each level is shown in Table 2. By plotting the standard deviations against the general mean, as shown in Figure 1, a linear relationship for repeatability and reproducibility has been derived.

Table 2. Results summary: pore size, mean and SD (µm).

Level <i>j</i>	No. instruments <i>p_j</i>	General mean \hat{m}_j	Repeatability SD s_{rj}	Reproducibility SD s_{Rj}
1	7	12.3	2.3	3.9
2	7	16.7	2.0	4.1
3	6	18.1	1.8	4.6
4	7	41.9	4.0	6.4

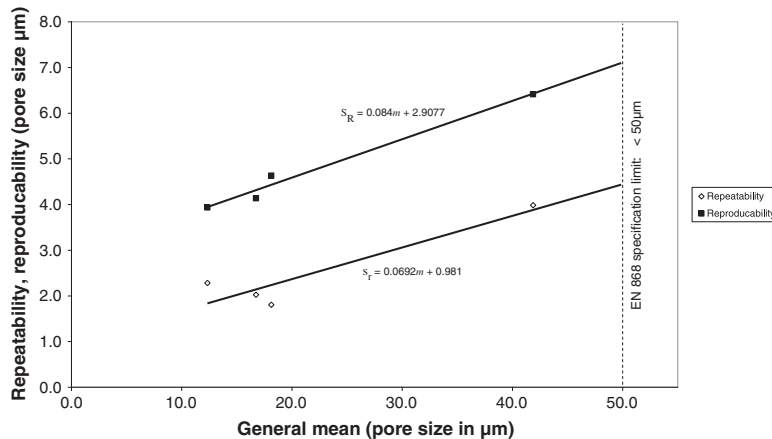


Figure 1. Pore size: plot of repeatability and reproducibility against general mean.

Water repellency

The water repellency method was assessed by 16 technicians from three organizations. Each technician performed 10 replicate tests at three test levels. Test levels 1–3 represented a fine crepe paper, a crepe paper and a plain paper. The approximate basis weights were 45, 65 and 80 g/m², and the thickness of each material was approximately 100, 160 and 110 µm. The mean water repellency and the standard deviations are shown in Table 3.

The application of Cochran's test to the standard deviations identified a straggler in level 1 for technicians 14 and 16. A single outlier was also present in level 3 for technician 12. An investigation into the cause revealed that the stragglers and outliers were caused by variability within the samples, which were inconsistent with the results from other technicians. The results of the Grubbs test applied to the mean results identified technician 14 as an outlier for level 1 and a straggler for level 2. An examination of the results revealed that technician 14 consistently reported results significantly higher than any other technicians, indicating that either the test method was not followed correctly or the apparatus was faulty. Therefore, the general mean, the repeatability standard deviation and the reproducibility standard deviation shown in Table 4 were determined, with all results from technician 14 excluded. The level 3 results of technician 12 were also excluded.

By plotting the standard deviations against the general mean as shown in Figure 2, a linear relationship for repeatability and reproducibility has been derived.

Chloride content

The chloride content method was assessed by seven technicians from three organizations. Each technician performed two replicate tests at four levels. All test materials were plain sterilization papers. The mean values and ranges are shown in Table 5.

Table 3. Test results of water repellency (s).

Technician <i>i</i>	Levels <i>j</i>					
	1		2		3	
1	16.1	(1.1)	32.0	(1.3)	43.6	(2.1)
2	16.1	(0.9)	29.3	(1.9)	38.0	(1.1)
3	15.9	(1.4)	30.5	(1.9)	40.6	(1.8)
4	17.5	(1.1)	31.5	(1.4)	44.2	(2.4)
5	16.9	(0.8)	30.3	(1.2)	41.7	(1.4)
6	12.7	(1.7)	22.8	(2.7)	31.2	(2.4)
7	15.0	(0.4)	30.3	(2.2)	46.9	(2.2)
8	17.0	(0.8)	32.3	(1.1)	52.3	(1.8)
9	16.7	(0.7)	32.3	(1.2)	46.4	(2.1)
10	12.4	(1.4)	30.9	(3.1)	43.7	(2.0)
11	16.1	(0.9)	32.5	(0.8)	44.4	(1.8)
12	14.5	(1.1)	37.8	(1.8)	43.2	(4.1) ^b
13	14.8	(1.4)	29.0	(1.8)	39.9	(1.1)
14	23.3 ^b	(1.9) ^a	44.9 ^a	(3.1)	53.8	(1.8)
15	15.5	(1.0)	33.3	(3.4)	43.4	(1.7)
16	16.6	(2.0) ^a	33.3	(2.9)	45.0	(1.8)

Values are presented as mean (SD).

^aStraggler.

^bStatistical outlier.

Table 4. Results summary: water repellency, mean and SD (s).

Level <i>j</i>	No. technicians <i>p_j</i>	General mean \hat{m}_j	Repeatability SD <i>s_{Rj}</i>	Reproducibility SD <i>s_{Rj}</i>
1	15	15.6	1.2	1.9
2	15	31.2	2.1	3.7
3	14	43.0	1.9	5.1

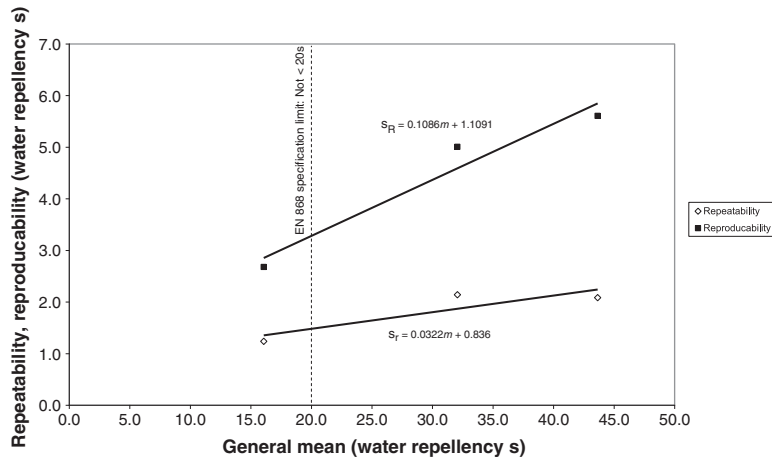


Figure 2. Water repellency: plot of repeatability and reproducibility against general mean.

Table 5. Test results of chloride content (%).

Technician <i>i</i> (laboratory)	Levels <i>j</i>							
	1		2		3		4	
1 (A)	0.0300	(0.0028)	0.0430	(0.0014)	0.0230	(0.0000)	0.0500	(0.0014)
2 (A)	0.0325	(0.0007)	0.0410	(0.0028)	0.0195	(0.0021)	0.0495	(0.0007)
3 (A)	0.0295	(0.0007)	0.0415	(0.0021)	0.0190	(0.0000)	0.0460	(0.0014)
4 (B)	0.0180	(0.0057)	0.0310	(0.0000)	0.0130	(0.0014)	0.0450	(0.0000)
5 (B)	0.0150	(0.0021)	0.0240	(0.0007)	0.0080	(0.0007)	0.0400	(0.0000)
6 (C)	0.0149	(0.0000)	0.0178	(0.0000)	0.0089	(0.0004)	0.0256	(0.0006)
7 (C)	0.0152	(0.0004)	0.0180	(0.0004)	0.0092	(0.0005)	0.0256	(0.0000)

Values are presented as mean (range).

The application of the Cochran and the Grubbs tests identified no stragglers or outliers. A summary of the general mean, repeatability standard deviation and reproducibility standard deviation is shown in Table 6. By plotting the standard deviations against the general mean, linear relationships for repeatability and reproducibility have been established as shown in Figure 3.

Sulfate content

The sulfate content method was also assessed by seven technicians from three organizations. Each technician performed two replicate tests at four levels. All test materials were plain sterilization papers. The mean values and ranges are shown in Table 7.

The application of the Cochran and the Grubbs tests identified no stragglers or outliers. A summary of the general mean, repeatability standard deviation and reproducibility standard deviation is shown in Table 8. By plotting the standard deviations against the general mean, linear relationships for repeatability and reproducibility have been established as shown in Figure 4.

Table 6. Results summary: chloride content, mean and SD (%).

Level <i>j</i>	No. technicians <i>p_j</i>	General mean \hat{m}_j	Repeatability SD s_{rj}	Reproducibility SD s_{Rj}
1	7	0.022	0.00181	0.0082
2	7	0.031	0.00103	0.0112
3	7	0.014	0.00072	0.0061
4	7	0.040	0.00058	0.0105

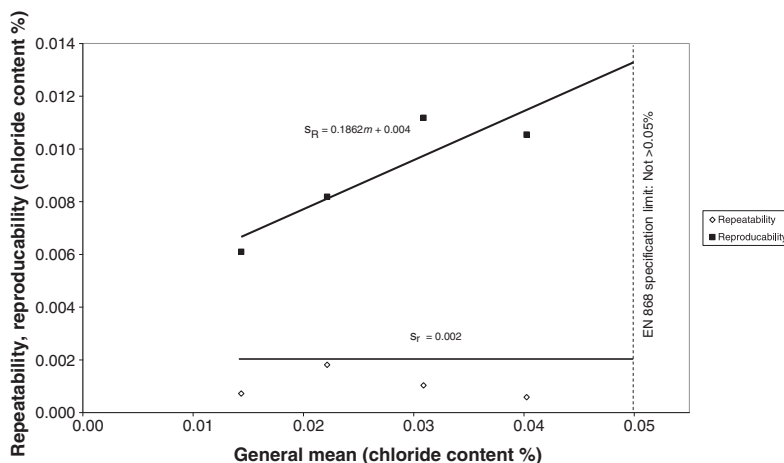


Figure 3. Chloride content: plot of repeatability and reproducibility against general mean.

Table 7. Test results of sulfate content (%).

Technician <i>i</i> (laboratory)	Levels <i>j</i>							
	1		2		3		4	
1 (A)	0.0740	(0.0127)	0.0455	(0.0064)	0.0195	(0.0021)	0.0645	(0.0106)
2 (A)	0.0955	(0.0021)	0.0425	(0.0235)	0.0235	(0.0106)	0.0450	(0.0028)
3 (A)	0.0800	(0.0071)	0.0360	(0.0042)	0.0240	(0.0042)	0.0360	(0.0000)
4 (B)	0.0390	(0.0184)	0.0060	(0.0021)	0.0070	(0.0000)	0.0110	(0.0000)
5 (B)	0.0430	(0.0078)	0.0040	(0.0000)	0.0040	(0.0000)	0.0110	(0.0000)
6 (C)	0.0538	(0.0016)	0.0080	(0.0006)	0.0097	(0.0019)	0.0119	(0.0016)
7 (C)	0.0518	(0.0006)	0.0084	(0.0003)	0.0088	(0.0002)	0.0137	(0.0004)

Values are presented as mean (range).

Table 8. Results summary: sulfate content, mean and SD (%).

Level <i>j</i>	No. technicians p_j	General mean \hat{m}_j	Repeatability SD s_{rj}	Reproducibility SD s_{Rj}
1	7	0.055	0.0062	0.0298
2	7	0.019	0.0062	0.0195
3	7	0.012	0.0029	0.0093
4	7	0.024	0.0028	0.0221

ANALYSIS AND DISCUSSION

Table 9 summarizes the repeatability and reproducibility standard deviations for each of the methods analysed. Repeatability and reproducibility are expressed as a function of measurement value and also as a percentage of the EN 868 specification value. The results show that repeatability and reproducibility are very significant when referenced to the measured value at the specification limit. Therefore, it may be appropriate to allow for this when specifying the material requirements of each test within ISO 11607.

Test materials

The test materials were requested to be evenly distributed over the measurement range of each test; in practise, the availability of commercially produced paper meant that this was not always achieved.

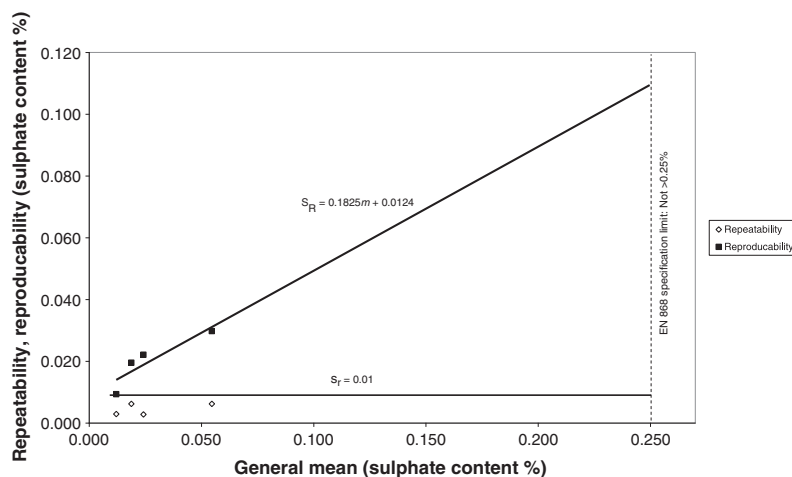


Figure 4. Sulfate content: plot of repeatability and reproducibility against general mean.

Table 9. Significance of results.

Test	Units	EN 868 specification	Repeatability s_r		Reproducibility s_R	
			Relative to mean result	Expressed as % specification limit	Relative to mean result	Expressed as % specification value
Pore size	μm	<50	$0.069m + 0.98$	8.9	$0.084m + 2.91$	14.2
Water repellency	s	Not < 20	$0.032m + 0.84$	7.4	$0.120m - 0.01$	11.9
Chloride content	%	Not > 0.05	0.002	4.0	$0.186m + 0.00$	26.6
Sulfate content	%	Not > 0.25	0.01	4.0	$0.183m + 0.01$	23.2

Consequently, the test samples for sulfate content did not test the method at or near to the specification limit of 0.25%. This means that the extrapolation of the relationships for repeatability and reproducibility determined between 0.01% and 0.05% is potentially inaccurate at 0.25%, as shown in Figure 4.

Precision

The low number of participants available to take part in the ILS meant that the statistical methods used to determine the repeatability and reproducibility standard deviations were less effective than they may have been using a larger data set.

Laboratory bias

The inspection of the original test results for the chloride content and sulfate content tests shows clear instances of laboratory bias within the results. These consistent differences in results based on the laboratory are likely due to the test equipment, procedures or environment associated with each laboratory.

CONCLUSIONS AND RECOMMENDATIONS

Sterilization packaging performs a vital role in minimizing the risk of microbial contamination of medical devices up to the point of use.

This study has shown that some of the test methods used for evaluating the performance of medical papers used in the construction of sterile packaging are neither precise nor accurate within the limits usually associated with these terms. Although the stochastic nature of paper materials may account for some of the loss in repeatability, the larger reproducibility is most likely due to differences in test

conditions, apparatus or procedures. These differences may be accounted for by multiple interpretations or implementations of the standards.

It should be emphasized that the sterile barrier materials typically used in the industry and tested in this study have proved to be fit for purpose over many years of use in the healthcare industry despite the poor reproducibility of the tests considered in this paper. However, it is always desirable to improve the consistency of the test methods used up and down the supply chain and across the industry.

It is therefore recommended that the critical sources of uncertainty associated with each of the tests should be determined so that modifications can be made at the next standard review. This should go some way towards limiting or removing the sources of error that are currently present.

The ideal solution would be for the sterilization packaging community to agree to the development of universally adopted instrumentation and methods that address the specific requirements of the tests within the context of ISO 11607. These would then ensure that the measurement of pore size, water repellency, chloride content and sulfate content could be performed consistently across the industry.

REFERENCES

1. Rutala W, Weber D, the Healthcare Infection Control Practices Advisory Committee. Guideline for Disinfection and Sterilization in Healthcare Facilities, 2008. Department of Health and Human Services USA, 2008.
2. Czerniawski B. Flexible packaging of disposable medical devices—A review. *Packaging Technology and Science* 1990; **3**(4): 203–217.
3. Czerniawski B. Requirements for single-use sterilization papers and paper packaging used for medical devices. *Packaging Technology and Science* 1991; **4**(4): 213–223.
4. EN 868-1:1997. Packaging materials and systems for medical devices which are to be sterilised.
5. ISO 11607:1997 Packaging for terminally sterilized medical devices.
6. <http://www.mddionline.com/article/iso-11607-new-standard-clears-packaging-confusion>
7. Hurme E, Wirtanen G, Axelson-Larsson L, Mattila-Sandholm T, Ahvenainen R. Reliability of Destructive Leakage Detection Methods for Semirigid Retort Packages. *Packaging Technology and Science* 1996; **9**(4): 203–213.
8. Chung D, Papadakis SE, Yam KL. Simple models for evaluating effects of small leaks on the gas barrier properties of food packages. *Packaging Technology and Science* 2003; **16**(2): 77–86.
9. Bix L, Kassarjian O, Cardoso F, Lockhart H. Using scanning electron, confocal and optical microscopes to measure microscopic holes in trays. *Packaging Technology and Science* 2005; **18**(6): 311–320.
10. Cassie BD, Baxter S. Wettability of porous surfaces. *Transactions of the Faraday Society* 1944; **40**(5): 546–551.