

Special Report

Evaluating Sharps Safety Devices: Meeting OSHA's Intent

Gina Pugliese, RN, MS; Teresa P. Germanson, PhD; Judene Bartley, MS, MPH; Judith Luca, RN; Lois Lamerato, PhD; Jack Cox, MD; Janine Jagger, MPH, PhD

ABSTRACT

The Occupational Safety and Health Administration (OSHA) revised the Bloodborne Pathogen Standard and, on July 17, 2001, began enforcing the use of appropriate and effective sharps devices with engineered sharps-injury protection. OSHA requires employers to maintain a sharps-injury log that records, among other items, the type and brand of contaminated sharps device involved in each injury. Federal OSHA does not require needlestick injury rates to be calculated by brand or type of device. A sufficient sample size to show a valid comparison of

safety devices, based on injury rates, is rarely feasible in a single facility outside of a formal research trial. Thus, calculations of injury rates should not be used by employers for product evaluations to compare the effectiveness of safety devices. This article provides examples of sample-size requirements for statistically valid comparisons, ranging from 100,000 to 4.5 million of each device, depending on study design, and expected reductions in needlestick injury rates (*Infect Control Hosp Epidemiol* 2001;22:456-458).

On November 6, 2000, President Clinton signed into law the Needlestick Safety and Prevention Act, requiring healthcare employers to provide safety-engineered needles and sharp medical devices for use in their institutions.¹ The law also instructed The Occupational Safety and Health Administration (OSHA) to revise the Bloodborne Pathogen Standard, to mandate the use of these safety devices, and to require the implementation of a sharps-injury log for recording exposure incidents. The full enforcement of this law by OSHA, which began July 17, 2001, is expected to result in a dramatic increase in the use of new safety technology and ultimately, it is hoped, in a significant reduction in occupational exposures to bloodborne pathogens.¹

As part of the revised Bloodborne Pathogen Standard, OSHA requires employers to maintain a sharps-injury log that records, among other items, the type and brand of device involved in each injury. However, federal OSHA does not require needlestick-injury rates to be calculated by brand or type of device. The information in the log is intended to establish priorities and identify trends that bear further investigation rather than for rate calculation. The log provides only one source of information for guiding prevention programs and device selection. Incident reports, information gathered on walking observational

rounds, and staff interviews will be useful for supplementing this process.

Employers are required to identify, evaluate, and select appropriate and effective sharps devices with engineered sharps-injury protection. Appropriate devices are those that, based on reasonable judgment, will not jeopardize patient or employee safety or be medically contraindicated. OSHA has indicated that an "effective" safer medical device, based on reasonable judgment, will make an exposure incident involving a contaminated sharp less likely to occur.² As employers evaluate new technology, there will be an increase in the opportunities to assess the effectiveness of the new devices in preventing injuries during the identification, evaluation, and final selection process.

There are several different approaches that may be taken for product evaluation. The most common is the informal evaluation or product trial that elicits subjective feedback from users and that has no sample-size requirements. Informal product trials usually are brief and involve clinical observations of a relatively small number of devices. They can provide valuable information about user preferences and product characteristics and represent the most common method used to evaluate devices for selection and adoption. However, these relatively small product

From the Premier Safety Institute (Dr. Cox, Ms. Pugliese, Ms. Bartley, Ms. Luca, and Ms. Lamerato), Premier Inc, Oak Brook, Illinois; IdeaMetrics (Ms. Germanson), and the International Healthcare Worker Safety Center (Dr. Jagger), University of Virginia, Charlottesville, Virginia. Address reprint requests to Gina Pugliese, RN, MS, Premier Safety Institute, 700 Commerce Dr, Oak Brook, IL 60523. 00-SR-123. Pugliese G, Germanson TP, Bartley J, Luca J, Lamerato L, Cox J, Jagger J. Evaluating sharps safety devices: meeting OSHA's intent. *Infect Control Hosp Epidemiol* 2001;22:456-458.

evaluations cannot be used to draw objective conclusions about user injury rates or safety performance of needle devices.

Other factors that make it difficult to obtain reliable injury data for rate calculations include the high degree of underreporting and the variability in degree of underreporting over time. Even if sample sizes are large enough to conduct statistically valid studies, there are a number of considerations in obtaining reliable injury data and calculation of rates for device comparison. For example, there is a high degree of underreporting (up to 70% in some studies) and variability in degree of reporting over time, both of which have an impact on the accuracy of injury data. There also is no national consensus on the best denominator to use to calculate injury rates for comparisons, and the choice of the best denominator often depends on the data that are available. Examples of denominators being used include number of devices, number of full-time equivalent staff, or number of occupied beds.

To draw objective conclusions about effectiveness of a new device, statistical analysis of a properly designed, large-scale safety-efficacy study must be performed. However, it should be emphasized that federal OSHA does not require efficacy studies be conducted to evaluate devices prior to adoption. In this article, we describe the sample-size requirements needed for a safety-efficacy study. These required sample sizes show that a valid comparison of safety devices based on injury rates rarely is feasible in a single facility because of the difficulty in obtaining a sample size large enough to determine if the difference in injury rates between two devices is due to more than chance alone.

SAFETY-EFFICACY STUDIES

Safety-efficacy studies are based on device-specific injury rates; in this case, the comparison of injury rates from safety-engineered devices to their conventional counterparts. When comparing needlestick-injury rates, an important consideration is achieving an adequate sample size, since needlestick injuries are statistical "rare events." According to published studies, needlesticks occur in the range of approximately 1 to 37 injuries per 100,000 devices.³⁻⁸ There is an inverse relation between rates of events and required sample size. Because occurrence rates are very low, the denominators for injury-rate comparisons (ie, number of devices required) must be very large. The large number of devices is needed to attain an acceptable statistical power of 80% or 90%. Statistical power is the ability of a test to show a statistically significant difference between groups if a true difference exists.

SAMPLE-SIZE CALCULATIONS

To determine the sample size needed to achieve a stated level of power, the researcher must establish the minimum reduction in injuries from the safety device believed to be clinically meaningful. For instance, should the device prevent at least 75% of injuries, or would it be acceptable for a device to prevent as few as 25% of injuries? Reasonable expectations for injury reductions

TABLE
SAMPLE-SIZE REQUIREMENTS FOR DETECTING REDUCTIONS IN
NEEDLESTICK RATES*

Injuries per 100,000 Conventional Needles	% Reduction With Safety Devices	No. of Devices Required per Device Type
5	25	4,600,000
	50	1,000,000
	75	375,000
10	25	2,300,000
	50	500,000
	75	186,000
15	25	1,540,000
	50	340,000
	75	125,000
20	25	1,150,000
	50	250,000
	75	94,000

* Fisher's Exact Test, alpha=0.05, power=80%, two-tailed testing.

from safety-engineered needles can be derived from the literature. Published safety-efficacy studies show reductions from safety-engineered needle devices ranging from a low of 25%⁵ to a high of 89%.⁷ It is up to the researcher to decide what is acceptable under the specific circumstances of the trial. Specification of the minimum reduction worth detecting has a profound effect on the number of devices required for the study. It is worth noting that a 100% reduction is an unrealistic expectation unless the sharp is eliminated completely from the safety device.

The Table presents several scenarios for sample-size calculations based on different injury rates with conventional devices and different expected reductions with safety devices. The number of devices needed in each group (safety and conventional) to achieve statistical power of 80%, when alpha is set to 0.05 in two-tailed testing, is shown. Sample-size calculations were based on the Fisher's Exact Test using Query Advisor (version 4.0; Statistical Solutions, Saugus, ME).⁹ Since the distribution of rare-event data may not meet the normality assumptions of Pearson's chi-square test, an exact test is the preferred analytic procedure. The Table shows the effect of different baseline injury rates of conventional needles and expected injury reductions from the use of safety needles on the number of devices required for evaluation. The selected parameters are derived from current literature and should provide realistic assumptions for determining sample sizes.

For example, if the baseline injury rate projected for conventional needles is 20 per 100,000 and the researcher hopes to demonstrate a 75% event reduction through the use of safety devices, then 94,000 devices would be required in each group to allow for statistical power of 80%. On the other hand, 375,000 needles would be needed if the baseline rate were projected to be 5 per 100,000 to

demonstrate the same 75% event reduction. If the investigator instead establishes 25% as the minimum reduction worth detecting, then the sample-size requirement would be 1,150,000 per group with a baseline rate of 20 per 100,000.

DISCUSSION

What is most notable about the calculations presented in the Table is the enormous number of devices required for testing the difference in needlestick rates of two devices, even for the smallest sample-size requirements. At one extreme, more than 4 million devices per group would be needed to achieve statistical significance if a device causing 5 injuries per 100,000 devices were compared to a device reducing injury rates by 25%. At the lower end of the spectrum, 94,000 devices per group would be needed to achieve statistical significance if a device causing 20 injuries per 100,000 devices were compared to a device that reduced injury rates by 75%.

In practical terms, the number of devices needed may exceed the annual device use of an individual hospital. In the largest hospitals, an estimated 1.5 to 3 million syringes might be used per year, while 75,000 to 300,000 intravenous catheters might be used, with similar numbers for winged steel needles and phlebotomy needles. In average-size or small hospitals, device use would be much less.

The investigator contemplating a single-center trial must determine in advance if annual device use in his facility is adequate to meet sample-size requirements for a planned study. If not, it may be necessary to conduct a multicenter study. Another alternative is to extend the trial for the amount of time necessary to accrue an adequate sample size, if the results would still be worth knowing at the end of a prolonged trial.

The danger in undertaking a study without an adequate sample size is the increased possibility that statistical significance will fail to be achieved even if the safety device is effective. Consider the example of a study in which 150,000 needles per group were used and the true baseline injury rate of the conventional device is 5 injuries per 100,000 needles. If the true relative effectiveness of the safety device is 50%, then the probability of obtaining statistically significant results in two-tailed testing with $\alpha=0.05$ is only 27%. In other words, it is most likely that results from the study will fail to be significant.

There are many challenges facing investigators embarking on safety-efficacy trials of safety-engineered needles, but none are more important than assuring a sufficient sample size to meet statistical requirements. The challenge posed by large sample sizes can loom only larger as safety-engineered devices become the predominant technology in the marketplace. Eventually, there will be attempts to test safety devices with very low baseline injury rates against each other to find the safest of the safety devices. The sample sizes for these future studies

will reach prohibitive extremes and will require multicenter collaborations to meet statistical demands.

CONCLUSION

There are few safety-device efficacy trials reported in the literature. Although they are very valuable for making direct comparisons in the safety performance of sharps devices, the large sample-size requirements make such trials prohibitive for most hospitals to consider. Federal OSHA does not require hospitals to conduct device-efficacy trials or calculate injury rates as part of the product evaluation process put forth in its revision to the Bloodborne Pathogen Standard. Even if sample sizes are large enough to conduct statistically valid studies, there are a number of considerations in obtaining reliable injury data and calculation of rates for device comparison. These include the high degree of underreporting,⁵ the variability in the degree of reporting over time, and the denominator used (eg, number of devices, number of full-time-equivalent staff, or number of occupied beds).

Less formal, more subjective product evaluations are sufficient for meeting the revised standard. However, OSHA explicitly requires the involvement of frontline healthcare workers in the identification, evaluation, and selection of appropriate safety-engineered devices. The final selection must be guided by the clinical environment, the procedures being performed, the patient population, the cost-effectiveness of the device, and the needs and preferences of the workers. A worker-driven process, as required by OSHA, will increase the likelihood that these devices will be accepted well and used appropriately, ultimately providing maximum benefits to both healthcare workers and their patients.

REFERENCES

1. Needlestick Safety and Prevention Act of 2000. Pub. L. No. 106-430, 114 Stat. 1901 (Nov. 6, 2000).
2. Occupational Safety Health and Administration. Occupational exposure to bloodborne pathogens. Final Rule (29 CFR 1910.1030) *Federal Register*. January 18, 2001;66:5317-5324.
3. Jagger J, Hunt EH, Brand-Elnaggar J, Pearson RD. Rates of needlestick injury caused by various devices in a university hospital. *N Engl J Med* 1988;319:284-288.
4. Ippolito G, De Carli G, Puro V, Petrosillo N, Arici C, Bertucci R, et al. Device-specific risk of needlestick injury in Italian health care workers. *JAMA* 1994;272:607-610.
5. Centers for Disease Control and Prevention. Evaluation of safety devices for preventing percutaneous injuries among health-care workers during phlebotomy procedures—Minneapolis-St. Paul, New York City, and San Francisco, 1993-1995. *MMWR* 1997;46:21-25.
6. Jagger J, Bentley MB. Injuries from vascular access devices: high risk and preventable. Collaborative EPINet Surveillance Group. *J Intraven Nurs* 1997;20(6 Suppl):S33-S39.
7. Mendelson MH, Chen LBY, Finkelstein LE, Bailey E, Kogan G. Evaluation of a safety IV catheter (Insyte Autoguard, Becton-Dickinson) using the Centers for Disease Control and Prevention (CDC) National Surveillance System for Hospital Healthcare Workers database. *Infect Control Hosp Epidemiol* 2000;21:111. Abstract.
8. Chen LBY, Bailey E, Kogan G, Finkelstein LE, Mendelson MH. Prevention of needlestick injuries in healthcare workers: 27 month experience with a resheathable 'safety' winged steel needle using CDC NASH Database. *Infect Control Hosp Epidemiol* 2000;21:108. Abstract.
9. Fisher, RA. *The Design of Experiments*. Edinburgh, Scotland: Oliver and Boyd; 1935.