DDESB Seminar - 16-18 August 1994

A STATISTICAL ANALYSIS OF SAFETY TEST RESULTS AND IMPLICATIONS FOR INSENSITIVE MUNITIONS

by Carole Bodart (ENSIETA) and Patrick Kernen (NIMIC)

Abstract

In the last 12 months, two countries, France and the United States, have issued or updated safety policies regarding their munitions or insensitive munitions. At NATO, a Standardization Agreement (STANAG) regarding IM is in its final draft and incorporates standardized testing.

Responses to these tests will help to decide if a munition meets safety and/or IM requirements. Also cost considerations affect why a limited number of munitions may be allocated for a given test.

In a presentation at the DDESB seminar in 1992, NIMIC focused on the poor reproducibility of some of these standardized tests (qualitative aspects), and hence the necessity to couple experimental testing with modeling.

A further study by NIMIC is presented which deals with the probability of information that can be misleading as a result of interpretation of the test responses of a few items, selected from a large production lot.

Through its extended database and the commitment of some of its European points of contact, NIMIC has conducted a statistical study of data involving repetitive bullet impact test reports series on various munitions, e.g. the 155 mm M 107 artillery shell and the General Purpose MK 82 bomb. It has focused on parameters such as:

- error of first kind
- error of second kind
- the operating characteristic curve of the test

The study enabled NIMIC to propose an assessment of the degree of confidence of a test series versus the number of tests conducted. The particular case of the standardized NATO bullet impact test procedure, requiring 2 items to be tested, has been addressed and its poor level of confidence highlighted.

In this paper, some conclusions have been drawn to improve the bullet impact test procedure and its reliability. However, exchanging data on, and applying models to other subscale or

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **AUG 1994**	2. REPORT TYPE	3. DATES COVERED **00-00-1994 to 00-00-1994**

4. TITLE AND SUBTITLE **A Statistical Analysis of Safety Test Results and Implications for Insensitive Munitions**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Organization due Traite de L'Atlantique Nord,OTAN-NATO,1110 Brussels, Belgium,**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADM000767. Proceedings of the Twenty-Sixth DoD Explosives Safety Seminar Held in Miami, FL on 16-18 August 1994.

14. ABSTRACT
see report

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**29**	

similar configurations within the NIMIC countries are advocated in order to replace all- up round testing.

1. Introduction

Munitions designers and users have always been concerned with the dangers involved in manufacturing, storing, using and, for some years now, disposing of munitions.

A great effort has been made to increase the safety of personnel developing them, and adherence to very strict instructions has reduced the hazards these activities entail for the surrounding facilities. The same cannot be said of the deployment or use of such munitions. The risks of incident or accident are considerably greater in the field. As a result, the armed forces have naturally imposed increasingly strict requirements for safety in use.

In order to meet safety requirement, NATO has recently introduced the concept of less hazardous munitions, also named LOVA (Low Vulnerability Ammunition) or IM (Insensitive Munitions) or MURAT (Munitions à Risques Atténués) and a NATO Standardization Agreement should be ratified soon. (1)

2. Characterization of munitions and of Insensitive Munitions, in particular

The IM STANAG recommends evaluating the safety of a munition or pyrotechnic item through a number of standardised tests. The safety of a munition is assessed on the basis of seven to nine tests, depending on the country policy (2) (3) (4). Those tests are currently:

- fuel fire;
- slow heating;
- drop;
- bullet impact;
- light fragment impact;
- heavy fragment impact (for France);
- shaped charge jet impact;
- detonation of an adjacent munition (sympathetic detonation);
- a harsh electrical and electromagnetic environment (for France).

The IM qualification is based on compliance with the requirements of the NATO STANAG 4439 (1), which involves supportive STANAGs and national test procedures for the hazards aforementioned. The set of tests to be carried out may depend on the country/specific purpose but most of the times, the pass/fail criteria are similar. The different test program approaches, "single label" or "multiple label" or "tailored", are described extensively in an other NIMIC presentation at this seminar (5).

The safety approach is applied to the munition in its various logistic and operational configurations. Various threat hazard assessment have been conducted to analyze and characterize the hazards facing a munition throughout its life cycle, the last one being a NIMIC paper presented at this seminar (6). Some studies go so far as to assign a probability of occurrence to each situation, be it one which occurs during maintenance or operations (7).

3. Interpretation of the outcome of the standardised tests responses

During qualification approval, only a small sample is tested. Often only one or two items per test are selected from an in-service stock that may comprise from a few dozen to several thousand items. For example, if we consider the bullet impact test, the procedure advices to test the components twice and the munition twice. In the latter case, the 12.7 mm AP (armour piercing) bullets will be aimed at the most sensitive components and the pass criterion will be "no reaction more violent than burning", i.e NATO type V reaction. (8)

The possibility that a response which may be untypical of the munition in question is observed twice in succession has to be considered in view of the complexity of the phenomena involved in interaction between the stimulus and the target, as well as the problem of reproducing identical test conditions.

For example, at the last DDESB seminar NIMIC gave a presentation entitled "Think Before Testing!" (9), which included examples of misleading safety test responses on generic munitions/test vehicles. Another NIMIC presentation at the current seminar (5) points out that the current way of carrying out fuel fire and bullet impact testing on all-up rounds cannot highlight which component of the round reacted in case of detonation, and hence how the design of the round could be improved. One of the main drawbacks of the current fuel fire and bullet impact test procedures is that they are not reproducible and , in consequence, their reliability may be doubted.

Furthermore, clear cost limitations preclude the use of very large samples to be tested for the IM qualification. Hence, one cannot guarantee strictly identical test conditions from one test to another, and the physico-chemical phenomena observed in the tests are currently still difficult to understand. This is the reason why and given its involvemnt in the IM STANAG, it seemed necessary for NIMIC to try to evaluate statistically the validity of these experimental results before using them in order to establish the MURAT status/labels of an item.

4. A NIMIC method of assessing the level of confidence of safety test reports

To tackle this problem, NIMIC decided to assess the reliability of the test procedures, applying a statistical analysis of the responses. This work has been carried out by a graduate student, Carole Bodart, working at NIMIC for a 3-month period in 1994 using the NIMIC safety test reports database. Given the short period allocated to this work, it was decided to limit the scope of the work to one of the main hazards of concern to the munitions community. It appeared that the most useful documents (i.e those including the best described test responses) and in the largest number in the NIMIC data base addressed bullet/fragment impact testing. Therefore, this work will address the aspects relating to the bullet impact test

STANAG and its reliability (8).
The complete calculations and the full description of the examples used are included in a NIMIC-Limited report. (10). As this avenue of study is undertaken for the first time, NIMIC welcomes comments and encourages other organizations to address the other threats.

4.1 Examination of the NIMIC database

People in charge of ordnance disposal have obviously no limits on the number of munitions to be tested whereas munition designers are constantly limited by budgets. This is the reason why the EOD scope has been taken into account in the selected database. About 660 different test reports have been selected in the NIMIC database relating exclusively to bullet/shells impact testing on all-up-rounds. The reports have been released to NIMIC by the Royal Dutch Air Force, the UK Ordnance Board, Royal Ordnance. the Australian Ordnance Council and the French Service Technique des Systèmes Navals (Technical Board for Naval Systems).

They deal with all sorts of shells, mines, bombs and missiles. The annex 1 describes in details the number of repetitive tests for 15/20 different types of bombs/shells/mines impacted by 10 different projectiles (including 5 different types of 12,7 mm bullet). Other similar comprehensive tables in the report (10) address the cases of missiles.

4.2 Examination of how the tests are performed

Despite its publication as long ago as 1988, little or no heed seems to be paid in the choice of bullet (the standard is the 12.7 mm AP), the bullet velocity (850+/- 20 m/s) or the accurate description of the level of reaction observed (NATO types), as standardised by STANAG 4241.
Thorough examination of the test examples in this database enables certain comments to be made:

- Most of the 12.7 mm bullets used fall into the API, APT, APIHC and APIHE categories, and these may cause different reactions. Their characteristics vary significantly from the 12,7 AP bullet characteristics which seems to be reserved for tests on missiles.

- The database also contains records of cases where 7.62mm, 20mm and 30mm calibre projectiles are used. These correspond to different attack scenarios and their results cannot therefore be compared directly with those of the standard test. (Nevertheless, they have been included for the records in Annex 1 and some have been used for statistical calculations, given the similarities of the variations of responses to attacks).

- The velocity at impact varies between 750 and 1200m/s depending on the type and loading of the projectile as well as the distance from the muzzle to the point of impact.

- Despite the NATO classification of reactions according to their violence (type I to type V), the levels of reaction are evaluated differently from one country to another, and even from one organization to another, making it difficult to compare the conclusions drawn from several tests.

- Large-calibre munitions and missiles are usually tested only once for the type of attack defined in STANAG 4241, instead of twice, as it should be, according to the procedure.

- The average repetition is relatively high (about 10) except for missiles;

- The test reports seem to assign the same value to the results obtained when firing single rounds or short bursts: some single targets which had to be disposed have been submitted to a series of tests before being destroyed but according to our calculations of proportions of violent reaction applied to 9 different set of munitions, the probability of violent reaction do not differ in the two statistical addressed cases (all test reports, only reports with single shots). Therefore, this could be a means of increasing the number of possible tests for a given target in order to understand the phenomena. A precaution would be to aim subsequent projectiles at non-predamaged zones.

4.3 Observations on the selected database

In conclusion, most of the bullet impact tests in our data base comply with the "philosophy" of STANAG 4241. However, they seek more to evaluate the vulnerability of a pyrotechnic item to a type of stimulus than to demonstrate its safety vis-à-vis the standard stimulus described in the STANAG.

So the basis usable for the study was far from being perfect, due particularly to the lack of harmony for the description of the test results and the variety of projectiles. Some of the statistical work has taken account of impact testing of various bullets, regardless of their type (AP, API, APT, APIHC, APIHE) on munitions such as missiles, mines, bombs, shells. An example of their repetition is given in Figure 1 for 12.7 mm bullet impact tests (note that the bottom left figure addresses the STANAG particular case of the 12.7 AP bullet):

- Column 1 indicates the total number of firings carried out on the generic type of munition indicated (shells, mines, bombs, missiles).

- Column 2 indicates the total number of munitions tested (the total number of tested individual items out of the population which the generic type constitutes).

- Column 3 indicates the number of different munitions configurations tested.

- Column 4 indicates the average number of repetitions for each munition configuration. It represents the quotient of Column 2 by Column 3. The average does not take account of firings repeated at different points on the same target.

4.4 Scattering of the responses

The responses to the bullet impact vary widely, as it is described in Fig 2 for 7 of the munitions tested. The fillers in many cases were melt cast explosives but some of them were pressed or cast cured PBXs or composite propellants. Given the wide distribution of the responses, it appears obvious for these examples that random would have made the first two tested munitions comply or not with the STANAG 4241 requirements (no reaction more severe than burning) and that a wrong conclusion might have been easily drawn from two first successful tests. These qualitative assumptions will be quantitatively proved later.

It might be argued that it is not wise to test now items filled with melt cast explosives. But there is often uncertainty regarding the sensitivity behavior of many new energetic formulations which is not yet comprehensively predicted at large scale. These new energetic materials can be considered as the melt cast compositions of yesterday, with regards to our knowledge of behavior at munitions scale.

4.5 Introduction to inspection sampling

The essential aim of inspection sampling is to formulate a diagnosis about a given population from observation of samples drawn from that same population.

The munition has the property x that it will behave in a certain way when the impact occurs (for example, x=1 if the reaction is a combustion or less severe, x=0 in the other case), and there is therefore an unknown proportion p* of munitions which have the property sought.

In our particular case, the experiment involves subjecting the munition to the bullet impact twice under the conditions required by STANAG 4241. To obtain an observed value x_1 of the result of the experiment, one firing is sufficient. The STANAG requires at least two observed values (n = 2).

In general, the interpretation of probability in terms of frequency then gives:

$$p = (x1 + x2 + ... + xn)/n$$

One can readily understand that there will be sampling fluctuations. Thus if the series of n observations is repeated on another sample under the same conditions the results will give us the values x'_1, x'_2, x'_3,...,x'_n and in general:

EQUATION

$$(x'1 + x'2 + ... + x'n)/n \neq (x1 + x2 + ... + xn)/n$$

If n becomes very large, the sampling provides so much information that the fluctuations tend to cancel out, and all the samples will lead to an estimation of the same value for p*. If one wishes to know whether the munition is safe when subjected to the standard bullet impact tests, one will understand that the statistical information provided by these two firings seems a priori inadequate, particularly when the number of manufactured munitions is very high. The experiment should be repeated a large number of times to obtain as much information as possible.

It is not always possible to carry out a large number of experiments to verify or quantify a characteristic present in a population. So one must estimate the value of that characteristic according to the observed results, always bearing in mind that these results are not systematically representative of the general case. One must then evaluate with a probabilistic model the risk of committing an error by assimilating p, resulting from the estimation, to p*, the real but unknown value.

4.6. Choice of a suitable probabilistic model

The use of a multinomial law could be introduced in order that all the types of reaction as well as other parameters should be taken into account, e.g. in order to understand the behavior of the target. But, considering the small size of each category of response for a given munition from our selected database and the fact that calculations are made much more complicated by the multidimensional character of this law, it was decided not to apply it.

The same decision has been taken as far as the Gaussian law is concerned because the size of the different samples available was considered too small to meet the so-called "normality hypothesis".

The binomial law fits perfectly the acceptance criteria of the STANAG 4241 bullet impact test and the IM/MURAT STANAG that consider two categories of munitions (i.e. two different values for the response which will be considered as the random variation parameter K of the binomial law):

- those whose responses to the test are "reaction not more severe than combustion (NATO reaction type V or no reaction or propulsion); (K = 1)
- and the other ones. (K = 0)

The other parameter of the binomial law is n, the total number of items tested.

4.7 Considerations on the applicability of the binomial law

Knowledge of the reactions of explosive substances is obviously not zero. During all the phases of their development, such substances undergo a whole range of tests, all yielding useful information. Thus we know that a given family of substances such as cast cured PBXs, which have low sensitivity, react non-violently to bullet impacts in several different configurations (mock-ups and munitions). One can therefore not consider their behaviour as random and maybe avoid AUR testing.

Yet the choice of explosive substances is sometimes complicated by the search for a trade-off between performance and sensitivity which may involve taking a risk on the safety side, e.g. by including more octogen or selecting a pressed explosive. Response to impact is then much less obvious. Very often, either the phenomenon observed behaves randomly or the information available from other studies cannot be used or is not available. In such cases the phenomenon must be considered completely unknown. This is also the case with the data we have been able to gather, <u>hence the potential interest of this analysis for munitions in development.</u>

In the absence of knowledge about the behaviour of the munitions, the pattern to which the munition corresponds is then identical to that of Bernouilli's box. The box contains two categories of balls (e.g. black balls [or munitions which react too violently] and white balls [munitions which comply with the STANAG's acceptance criterion]) in the proportion q=1-p and p. n independent draws are made, and the balls drawn are put back.

The probability of obtaining k balls in the category represented by proportion p, whatever the order of their appearance, is given by the binomial law B(n,p):

EQUATION

$$P(k)=C_n^k p^k q^{n-k} = \left[\frac{n!}{(k!(n-k)!)}\right] p^k q^{n-k}$$

A pattern in which the samples are not put back, which is the case with all destructive tests, may be considered equivalent to its counterpart in which they are put back provided the fraction sampled remains less than one tenth of the batch of munitions whose safety one is trying to prove. This condition is of course fulfilled in view of the high cost of the safety tests and the size of the sample destroyed in a particular EOD program. In particular, this condition was met for the two examples considered in paragraph 4.8.

4.8. Characteristics yielded by a binomial law

p is the proportion of items satisfying the STANAG 4241 acceptance criterion among the population constituted by all munitions of the same type. It varies from one sample to another and therefore does not avoid the need for other more complete statistical calculations.
There are different standardised methods to estimate p. (11). These are as follows:

- exact methods, based on the binomial law and using a table of that law or a table of the "fractiles" of F (law of Fischer-Snedecor);

- approaching methods, based on the approximation of the binomial law by the normal law (directly or by the formula of Molenaar) or by the law of Poisson (table of that law or fractiles of the χ_2 law).

These methods have been developed through two specific examples, the first one providing an analogy with an other projectile:

- MK 82 bombs filled with Tritonal submitted to 30 mm shells impacts;
- 155 mm M107 shells filled with TNT submitted to 12.7 mm bullet impacts .

4.8.1 Demilitarization of Tritonal containing MK82 bombs with 30 mm shells impacts

This bomb can not be considered as safe and we will determine to what extent. Figure 3 shows the responses of 18 bombs: 2 reactions acceptable and 16 non acceptable. Hence the frequence of non-violent reactions observed on this sample is 2/18 = 0.11 and can be considered as a rough particular estimate of p. In this EOD case, the lower is p, the more difficult it will be to dispose the munitions.

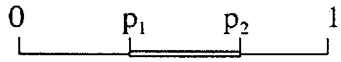
In the following items, the characteristics of the binomial law will be described on a general standpoint and exemplified as soon as possible with the specific example.

<u>Estimation of p by interval</u>

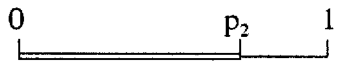
The interval is a function of the confidence sought in the result. It may be:

EQUATION

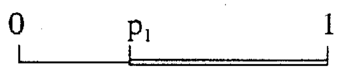
- two-sided, in the form (p_1,p_2) where $0 \leq p_1 < p_2 \leq 1$,



- one-sided on the left, in the form $(0,p_2)$ where $0<p2 \leq 1$,



- one-sided on the right, with the form $(p1,1)$ where $0 \leq p_1 < 1$,



The type of interval depends on the nature of the problem considered and not on the observed value k meaning the number of results equal to 1 (success). In our case, we shall first seek an interval likely to contain the exact value of p (two-sided test) before postulating a hypothesis and studying its validity.

<u>The confidence level and the risk α</u>

The confidence level is directly linked to the type I risk (error of first kind) called α. It expresses the probability of having the proportion p within the defined interval. For a given α, the confidence level is α. Thus a confidence level of 99% corresponds to a probability of correctly bracketing the exact value of p of 99% and a risk of incorrectly estimating it of 1%.

In general, the value of α is fixed either by a standard or by the test director. In our particular case, there is no α mentioned in STANAG 4241 (any more than there is any mention of any numerical safety target to be achieved). We shall therefore study the classic values of α, namely 1%, 2%, 5% and 10% depending on the objective pursued.

The boundaries of the confidence intervals are defined on the basis of knowledge of n, k (number of successes) and α.

Definition of p_1 and p_2

It is reminded that k is the number of non violent reactions. By convention, for k=0, p_1=0. For k=1, 2, 3,..., n, p_1 is defined by the following condition: the probability of obtaining a value of x greater than or equal to k is equal to $\alpha/2$ (two-sided test) or α (one-sided test on the right). In the case of interest for us here (k=2), p_1 is therefore the solution of the following equation:

$$P(x \geq 2) = \sum_{j=2}^{18} C_{18}^{j} p^{j}(1-p)^{18-j} = \alpha/2$$

For k=n=18, by convention p_2=1. For k varying from 0 to n-1, p_2 is the value of p which solves the equation:

$$P(x \leq 2) = \sum_{j=0}^{2} C_{18}^{j} p^{j}(1-p)^{18-j} = \alpha/2$$

Fig 4 shows the values of p1 and p2 for the binomial law applied to our case, where n =18 and k =2. The complete formulae and calculations with the 2 other methods, the fractiles of F and the approximation of the binomial law by the normal law, are given in (10); Fig 5 and Fig 6 compile their respective results. Hence, the 3 methods yield about the same results: this munition is not safe and will be easy to dispose. Given the similarities for the results, the tables of the binomial law will be used for the other calculations.

Comparison of a proportion to a given value

It may seem more interesting to know which side of a fixed limit, p0, p falls, particularly for comparison with a value imposed by regulations (a safety objective) or by common sense. In our case, an arbitrary choice is p0= 0.19 (the so-called "null hypothesis") which is valid because k<npo.

Meaning of the risk α

The risk α, also known as the producer's risk, is the probability of rejecting the null hypothesis (p<0,19) when it is true. It reflects both the confidence that can be ascribed to the calculation and the probability of excluding from acceptance, as a result of the test, a type of munition which would nonetheless have had the required qualities.

Validation of the hypothesis $p \leq p_0$ with a significance level α corresponds to the definition of a critical region for the observed value k. The boundaries of that critical region (0,c) are defined by the equation:

$$P(x \geq c) = \sum_{j=c}^{18} C_{18}^{j} p_o^{j} (1-p_o)^{18-j} \leq \alpha$$

It is an integer, rounded up owing to the discrete nature of the binomial law. α is directly linked to c, and the fact of rounding c modifies α. The real significance level α' can therefore be appreciably below the desired level for α (see figure 7).

Meaning of the risk β

The type II risk, (error of second kind) also known as the consumer's risk, is the probability of validating the null hypothesis H_0 (p<p0) when it is false, instead of another hypothesis H, known as the alternative hypothesis, which would be true. It represents the probability of accepting, following the test, a batch of munitions of unacceptable quality.

The risk β is fundamental to the guarantee offered by a safety label, yet it is often neglected in favour of α. The user - the serviceman - will not be interested in whether the producer is likely to make a mistake by wrongly stating that a batch of munitions is bad. He will, on the other hand, take care to ensure that there is no risk that the munition he chose because it was accepted is in fact unsafe. His survival depends upon it. *The credibility of IM thus depends on the risk* β .

β is defined as a function of the alternative hypothesis H: $p>p_1>p_0$, and its variations as a function of p_1 are represented on a graph, the operating characteristic curve of the test.

β and α are linked, since the determination of c depends on α. When $H \equiv$ Ho, the relationship $1-\beta=\alpha$ is verified.

Fig 8 shows the value of β for different values of p1. For example, the probability to draw a wrong conclusion (p<0.19) whereas p> 0.3 is 94% for a confidence level of 99%. In that case, the probability of making an accurate estimate of the exact proportion is impossible because β is high. The non-safety of a munition will be easier to prove than its safety.

4.8.2 Bullet impact tests on 155 M 107 shells

Within 34 shells tested, 16 successes and 18 failures have been observed. Hence the proportion p is about 0.47. Let us choose the hypothesis p= p0= 0.47. The critical values of c1 and c2 are determined by the formulae $P(x \leq c1)$ and $P(x \geq c2)$.

Figures 9, 10, 11 and 12 show the respective values of p1, p2, c1 , c2, α and β.

Figure 13 shows the operating characteristic curve of this test. In general, the increase of the confidence level $1-\alpha$ means a higher β. For example, for $\alpha=1\%$, there is a risk to make a wrong statement on the safety level of the munition as high as 98% whereas, for $\alpha=10\%$, this risk is 86%. Hence, a compromise has to be decided between the risk α and the risk β.

4.9 Efficiency of the statistical test on small samples

$$n, \beta, \alpha, p_0 \text{ and } c$$

are linked by the equations of the binomial law. Guaranteeing the safety of a munition entails achieving an acceptable compromise between these variables.

The choice of c=1 and k = 0 is imposed by the STANAG 4241 procedure and acceptance criterion.

p_0 will be chosen as from now as the proportion of violent reactions and represents a safety objective which the munition to be qualified must achieve. It cannot be circumvented if the results of a test are to be used statistically.

One would expect the regulations to impose a value of p_0. Not a bit of it! STANAG 4439 on MURAT calls for the probability of accidental initiation occurring to be minimised, without further explanation. The experts we asked were unable to advise or had no opinion. We therefore established the value of p_0 ourselves. A limit of 1/1000 seemed draconian, while 1/10 appeared inadequate. We chose 1/100, which incidentally was also chosen in STANAG 2818 as the failure limit for firing systems of pyrotechnic items.

Once p_0 and c are given, n, α and β are interdependent.

We established in the previous paragraph that β is the essential parameter for ensuring safety. It would therefore seem logical that the user should first specify the risk he is prepared to run in service. The munitions manufacturer would then chose the best compromise between α and n to meet his customer's requirements.

Such an approach would already consume too many munitions: <u>for $\beta=20\%$, at least 70 munitions would have to be destructively tested, which is unthinkable.</u>

As the values of n are limited, we shall therefore evaluate the reliability of the standard tests and the confidence which can be placed in the results observed.

Reconsideration of the general procedure for a test

In this paragraph we shall seek to determine the proportion of items or munitions that have the property that they react too violently.

The value 0 is therefore assigned to any reaction less severe than burning and the value 1 to any reaction strictly more severe. The appropriate probability law is still the binomial law.

It is necessary to define a null hypothesis and its alternative H1 in relation to the population of the munition under consideration.

$$H_0 : p \leq p0 = 10^{-2} \text{ versus } H_1 : p > p_1 > p_0$$

On the basis of this hypothesis, this chapter will study the values of risks α and α in order to establish the confidence that may be placed in acceptance or rejection of H_0.

The STANAG's acceptance criterion (k=0 for munitions reacting non violently) is the discriminant function of the observations.

The choice of confidence level 1-α defines a critical region for the values of k. The lower boundary of this critical region is called the critical value c.α then represents the probability, if H_0 is true, such that the criterion is contained within this region.

The figure 14 compiles the risks β of stating that the probability of a violent response for the munition tested is less than 1% (null hypothesis) whereas it is in reality 10% or more (alternative hypothesis), for most of the samplings involving 2 to 20 items.

If α is increased, β generally decreases. But this is not possible for small samples. Given the discrete feature of the binomial law, α has a limited number of values, e.g only 2 % for n =2. Hence, the value of β will be high.

Consequence for the bullet attack STANAG

Can one assess the safety of a munition from only two tests? It all depends on what one expects from the IM label. The test as it is currently performed according to the STANAG procedure cannot guarantee an acceptable safety objective $p_{0:}$: among other cases, the particular case of n=2 is exemplified within Figure 15 and shows that the risk of stating that the probability of a violent reaction for the munition tested is less than 1% when in reality it is 10% or more is very high, of the order of 80%, which is unacceptable, in the view of the IM status.

4.10 Alternative methods/samplings

The current standardised sampling method not being satisfactory, we have therefore studied other classical statistical methods that could be applied to the bullet impact STANAG procedure:

- progressive samplings
- double and multiple samplings

These two methods are commonly used in the industry for the acceptance of a batch. The double sampling method has been applied to our most comprehensive database: 12.7 API bullet impact tests on 51 shaped charges. They resulted in the fact that there was no significant improvement, regarding the number of tests necessary for acceptance/rejection decisions, in case of severe safety objectives (1%) (10).

There is another alternative: the use of knowledge of the behavior of the munition. The following methods might be of interest in the future:

- Bayes evaluation method.

More information regarding the projectile, its trajectory in the target and the behavior of the target would be necessary for the use of this method.

- Comparison of paired observations - sign test

The development of mock-ups or models of parts of munitions makes us wonder whether it is possible to link these results to those obtained from the real munition. One of the 2 major defects which raise doubts about the application of this technique to IM testing is the need to obtain 8 pairs exhibiting a difference in the results; the other defect is the misperception of risk β.

5. Improvement of the standardized test

The bullet impact test codified by STANAG 4241 aims at proving, from 2 tests, the safety of a munition in an operational context. The acceptance criterion, though severe, cannot guarantee achievement of a safety objective consistent with the IM requirements: the probability of an interpretation error leading to acceptance of a batch of unsafe munitions is unacceptable (about 80% for fillers whose behaviour in the event of a bullet impact is hard to predict). Therefore, as the size of the samples tested cannot be drastically increased, the risk of making a bad diagnosis on the basis of the observed results will be unacceptable, whatever statistical method is used and whatever precautions are taken in the conducting of the impact test. Hence, the current STANAG procedure is useful for rejecting munitions, not for accepting munitions.

One of the great weaknesses of the pass/fail test as set out in STANAG 4241 is the lack of a measurable characteristic quantity. On the whole, statistical tests relating to qualitative characteristics are less powerful than their counterparts relating to measurable quantities, especially as the behaviour of the projectile during its penetration may assume one of an infinite number of possible configurations: the bullet swings. This random behaviour makes the results more difficult to interpret.

To demonstrate a munition's safety effectively one must be able to predict, step by step, the phenomena which occur in the munitions when subjected to impact. A correctly conducted test must, if not be able to prove directly the safety of the munition studied, at least provide better knowledge of the mechanisms controlling the behaviour of this munition. The test no longer acts as a censor (a role it cannot currently play as it is unreliable) but sets out to gather useful information.

It seems that the recording of the blast overpressure (in the event of explosion of the target munition) and the film of the reaction currently supply all the information gathered, but are not enough to explain what has been observed. The test must therefore be more completely instrumented. For example, systematic X-raying of target-munitions which have not or have scarcely reacted after impact would be interesting to characterise the projectile's behaviour in the material and the damage which results from it.

The limited number of full scale tests that can be performed on munitions for reasons of cost and availability has led to the appearance of analytical instruments and mathematical simulation models. One may therefore consider replacing bullets, in the tests performed, by projectiles whose movement within an explosive material is empirically known. Research carried out using steel balls (12) (13) has demonstrated the reproducibility of the impact and of the ball's penetration into explosive materials. Unlike a bullet, a ball does not become destabilised and does not tumble. One may therefore hope to demonstrate in the experiments special mechanisms which had hitherto been confused with the consequences of a bullet which can not be monitored.

6. The necessity of using models and cooperating

The necessity of combining testing and modeling has already been advocated in other NIMIC papers, as well as by other organizations. Nevertheless, it is impossible to model everything. The calculations are long and costly. Tests must therefore be established to determine the parameters that influence the reaction (solidity of the confinement, loading density, porosity, projectile velocity, etc), in order to provide data that can be used in the models and to define avenues of study. Tests simulating a few conditions of use of the munition must therefore be replaced by experimental plans taking account of the data already available (reduced-scale tests or tests in similar configurations). If such methodological rigour is applied, the possibility that we may one day be able to understand and predict the response of a munition from knowledge of a few characteristics of the materials and configurations used cannot be ruled out. Once the reliability of the mathematical modelling is established, such modelling will doubtless be able to replace all the full-scale tests.

For that purpose, NIMIC could be a focal point for safety test reports on munitions/munitions components/test vehicles within the NIMIC countries. This data could be released to NIMIC by means of forms (see annex 2 for the bullet impact threat) to be filled as comprehensively as possible and made non confidential (meaning that any mention of the weapon would be deleted). NIMIC would then computerize the data. That would enable NIMIC to contribute to the enforcement of international and national IM qualification procedures, such as MILSTD 2105 B, the French national doctrine, the OB Pillar proceeding, among other national documents and the IM NATO STANAG which reflect to the idea of "demonstration" as well as standardised testing. At the occasion of the release to requestors, for a given purpose (IM design most probably), of all the information made available and regarding the studies of similar/subscale configurations, those at the origin of the data would be informed and could also make contact with the requestor, if willing.

Hence, the time of costly all-up round testing would be over!

REFERENCES

(1) NATO/AC310, STANAG 4439, "Policy for Introduction, Assessment and Testing for Insensitive Munitions (MURAT)", July 1994 (Draft)

(2) MIL-STD 2105 B, "Hazards Assessment Tests for Non Nuclear Ordnance", January 1994

(3) P. Lamy and J. Guillien, French MOD, "The French National Doctrine with Regards to Less Hazardous Munitions", Insensitive Munitions Technology Symposium, June 1994, Williamsburg, VA

(4) Ordnance Board, Pillar Proceeding 42657, "Insensitive Munitions Pillar Proceeding", September 1990

(5) M. DéFourneaux, NIMIC, "Streamlining Safety Testing for Munitions", DDESB Seminar, August 1994, Miami, FL

(6) J. deW. FitzGerald Smith, NIMIC, "A Methodology for Evaluating Potential Threat Hazards to Military Munitions", DDESB Seminar, August 1994, Miami, FL

(7) A. Victor, "Insensitive Munitions Threat Hazard Assessment: Methodology and Examples", May 1993

(8) STANAG 4241, "Bullet Attack Test for Munitions"

(9) M. DéFourneaux and P. Kernen, NIMIC, "Think before Testing!", DDESB Seminar, August 1992, Anaheim, CA

(10) C. Bodart, ENSIETA, "Analyse Statistique de la Fiabilité des Tests de Sécurité pour MURAT (A Statistical Analysis of the Reliability of IM Safety Tests)", June 1994

(11) AFNOR, NFX-06-068 for Statistical Methods

(12) H. Trumel, CEG, "Study of the Penetration of a Rigid Ball in a Non-linear Viscous Material", 1990

(13) E.N. Ferm & J.B. Ramsay, LANL, "Spherical Projectile Impact on Explosives", Detonation Symposium in Portland, 1989

ANNEX 1

AUR TESTING OF BOMBS, SHELLS AND MINES

7,62 OTAN

munitions	repetitions
BLU 86	10
MIFF	10

7,62 AP

munitions	repetitions
120 mm	7
76mm	14

12,7 mm

munitions	repetitions
HB 876	16
SG 357	2

12,7mm AP

munitions	repetitions
81 mm	5
Mk 12 1	12
Mk 12 2	45
155 mm 1	16

12,7mm API

munitions	repetitions
155mm 1	34
155mm 2	6
203mm	31
120mm	51
HB 876	2
Mk 82	14
Mk 13	4

ANNEX 1
AUR TESTING OF BOMBS, SHELLS AND MINES

ANNEX 1 (CONTINUED)
AUR TESTING OF BOMBS, SHELLS AND MINES

12,7 mm APIHE

munitions	repetitions
Mk 82	14
Mk 10	2

12,7 mm APIHC

munitions	repetitions
HB 876	25
SG 357	19
Mk 82	28
AN-M64	5
Mk 11	10
Mk 13	1
Mk 20	1
FAB 240	1
MUSA	8
155 mm 3	8
175 mm	3

12,7 mm APT

munitions	repetitions
Mk 82	8
AN-M64	1

12,7 mm APIHC

munitions	repetitions
Mk 82	5
AN-M64	3
Mk11	3

20 mm APT

munitions	repetitions
155 mm 1	8

20 mm SAPHEI

munitions	repetitions
155 mm 1	12

ANNEX 2

BULLET/FRAGMENT IMPACT SAFETY TESTING FORM FOR WORKABLE DATABASE

Projectile

- Type (bullet, fragment preformed or not, sphere, cylinder)
- Mass
- Material with mechanical properties (Young modulus, maximum stress, failure criteria...)
- Obliquity at impact (theoretical and/or measured)
- Velocity at impact (theoretical and/or measured)

Tested item/munition (excluding the energetic material)

- Type (munition, dummy, test vehicle) and name, if non confidential/limited
- Purpose of test (research, development, disposal)
- Material and its mechanical properties (same as above)/specific weight
- Dimensions (diameter, length, thickness, other as necessary...) including for end plates
- Theoretical static burst pressure
- Shock Hugoniot
- Presence of vents

Energetic material

- Type (high explosive, rocket propellant, gun propellant)
- Purpose (main charge, booster, igniter, other...)
- Formulation and family of substance (melt-cast, cast cured, pressed, composite, ...)
- Mechanical properties (same as above)/density/TMD or porosity
- Critical diameter, velocity of detonation
- Shock Hugoniot and mechanical properties (same as above)
- Popolato plot
- Card gap tests responses (various diameters/porosities)
- Reactive models
- Friability or Susan test

NATO type reaction observed

- Delays for initiation and transition to various levels of reaction
- Instrumentation used in the target and vicinity

Number of similar tests reports available

- A sheet should be drafted for every test.

ANNEX 2

BULLET/FRAGMENT IMPACT SAFETY TESTING FORM FOR WORKABLE DATABASE

Figure 1

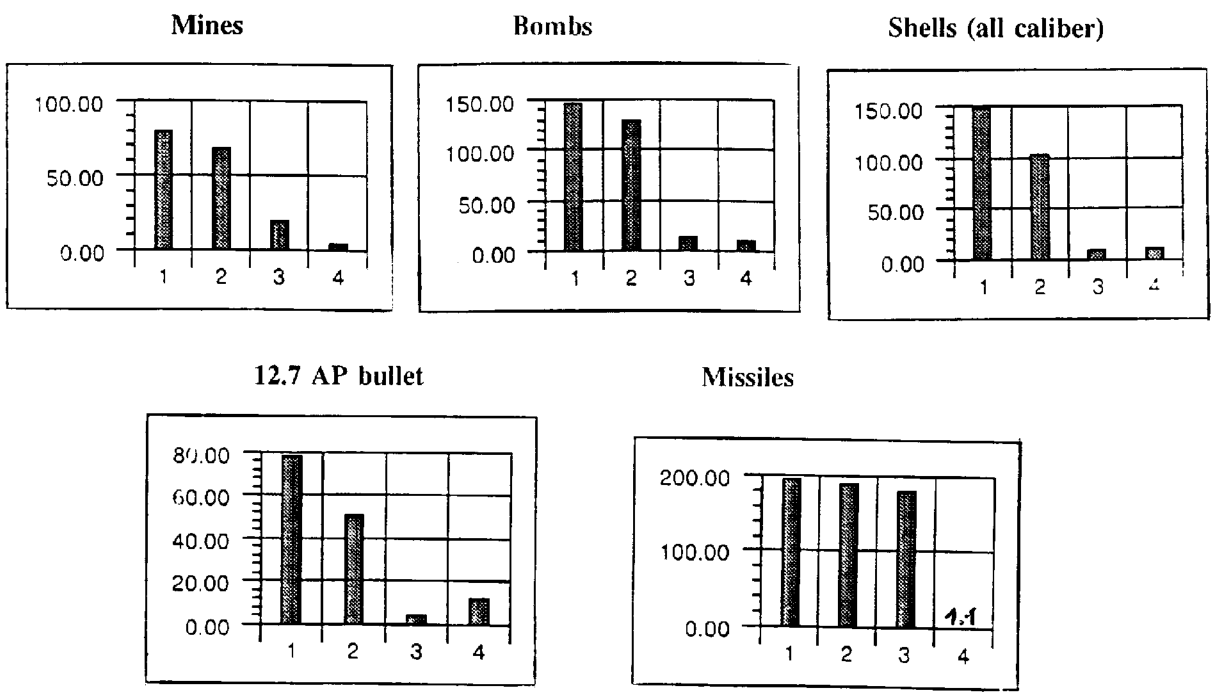
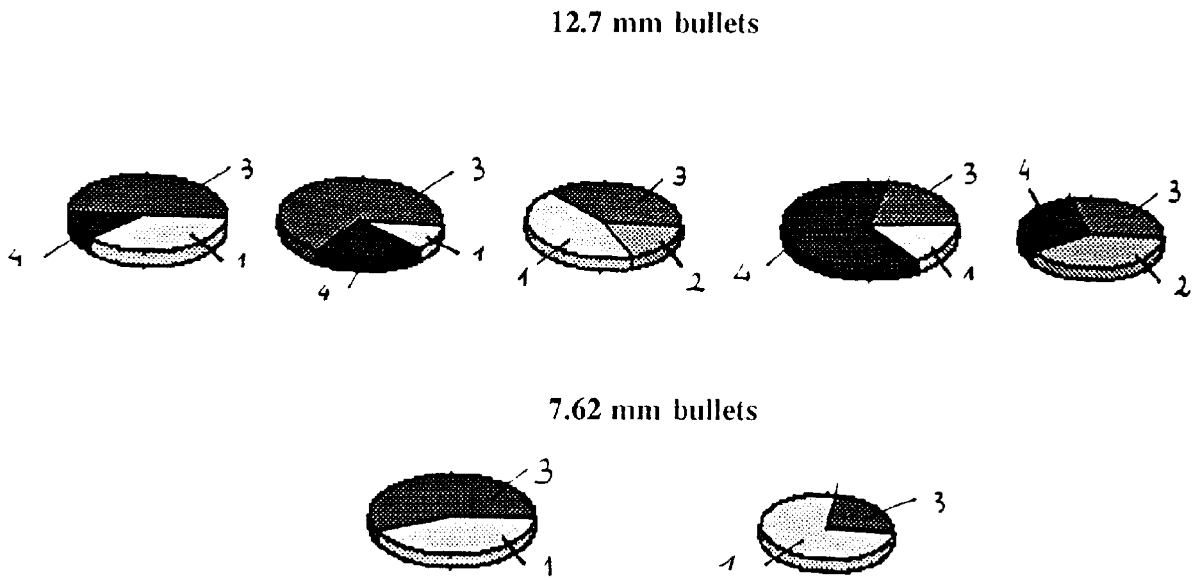


Figure 1

Figure 2



1, 2, 3 and 4 stand for no reaction, combustion, deflagration, detonation

Figure 2

Figure 3

test	1	2	3	4	5	6	7	8	9
0/1	0	0	0	1	0	0	0	0	0

test	10	11	12	13	14	15	16	17	18
0/1	0	0	0	0	0	1	0	0	0

Figure 3

Figure 4

Confidence	99%	98%	95%	90%
p_1	0,0058	0,0084	0,0137	0,0201
p_2	0,4216	0,3912	0,3471	0,3102

Figure 4

Figure 5

Confidence	99%	98%	95%	90%
p_1	0,0058	0,0084	0,0137	0,0201
p_2	0,4216	0,3912	0,3471	0,3102

Figure 5

Figure 6

Confidence	99%	98%	95%	90%
p_1	0,003	0,0023	0,0108	0,0187
p_2	0,4026	0,3470	0,3330	0,2989

Figure 6

Figure 7

α	1%	2%	5%	10%
c	9	8	7	7
α'	0,3%	1,2%	4%	4%

Figure 7

Figure 8

α	5 ou 10%	2%	1%
c	7	8	9
$\beta'_{0,9}$	0	0	0
$\beta'_{0,5}$	12%	24%	41%
$\beta'_{0,3}$	73%	86%	94%

Figure 8

Figure 9

α	1%	2%	5%	10%
p_1	0,253	0,270	0,297	0,350
p_2	0,696	0,677	0,648	0,593

Figure 9

Figure 10

α	1%	2%	5%	10%
c_1	8	8	9	11
c_2	24	23	23	21

Figure 10

Figure 11

$\alpha' = P(x \leq c_1) + P(x \geq c_2)$ where $p = p_0$

α	1%	2%	5%	10%
α'	0,9%	0,9%	2,4%	6,1%

Figure 11

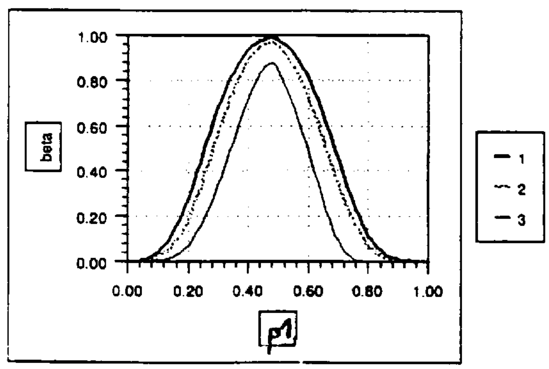
Figure 12

$\beta' = P(c_1<x<c_2)$ where $p \neq p_0$

α	1% et 2%	5%	10%
$\beta'_{0,2}$	22,7%	12,5	2,7
$\beta'_{0,5}$	98,6	96,7	85,6
$\beta'_{0,9}$	0,1	0	0

Figure 12

Figure 13



Plots 1,2 & 3 stand for α = 1, 5 & 10% respectively

Figure 13

Figure 14

n	2	3	4	5	6	7		8	
α'	1,9%	2,9%	3,9%	4,9%	6%	6,8%	0,2%	7,7%	0,3%
c'	1	1	1	1	1	1	2	1	2
β'0,1	81%	74%	66%	59%	53%	48%	85%	43%	81%

n	9		10		15		20		
α'	8,6%	0,3%	9,6%	0,4%	14%	1%	18%	1,7%	0,1%
c'	1	2	1	2	1	2	1	2	3
β'0,1	39%	78%	35%	76%	21%	55%	13%	40%	60%

Figure 14

Figure 15

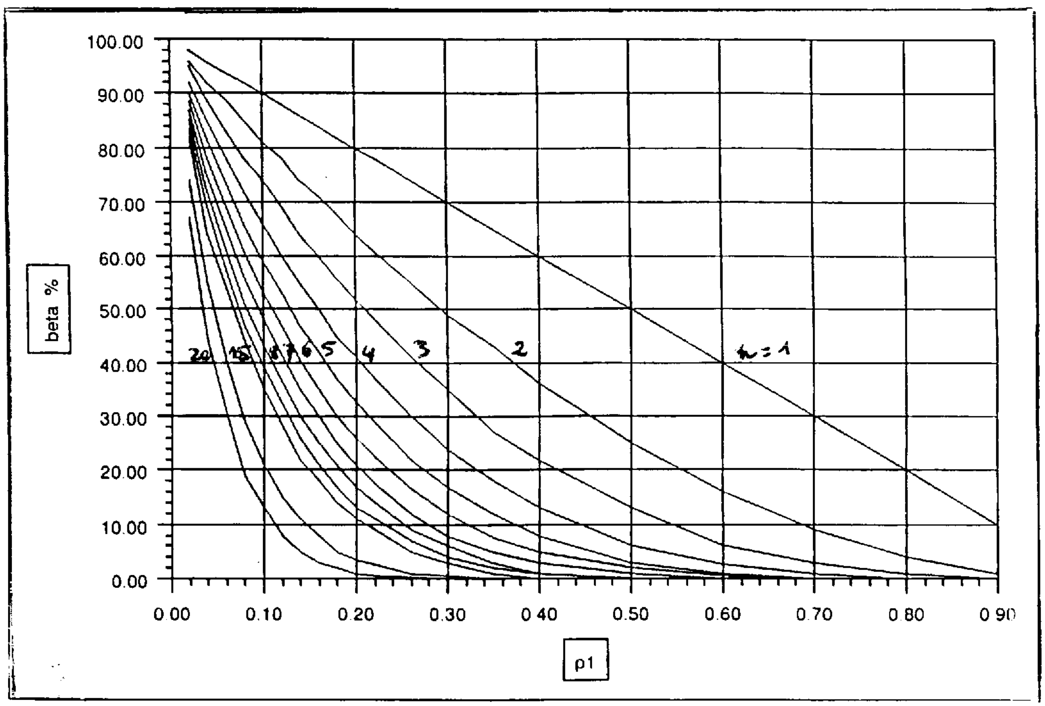


Figure 15