A Proposal for the Validation of Control Banding Using Bayesian Decision Analysis Techniques

Paul Hewett
Exposure Assessment Solutions, Inc.
Morgantown, WV

John Mulhausen
3M Company
St. Paul, MN

Perry Logan
3M Company
St. Paul, MN

1 ABSTRACT

COSHH Essentials (CE) is basically an exposure prediction model that results in control recommendations without actual exposure measurements. A current concern involves the external validation of its exposure prediction capabilities. Because CE fits a Bayesian decision framework, it is proposed that validation data be analyzed and summarized using methods based on Bayesian statistics. Furthermore, these methods can adapted for use by employers to efficiently determine whether or not exposures are adequately controlled. Sample datasets will be analyzed and discussed for both applications.

A project is underway to develop a methodology, based on Bayesian statistics, for industrial hygienists to use when evaluating the work environment. This methodology, or Bayesian Decision Analysis (BDA), is based upon the concepts of prior, likelihood, and posterior distributions of "decision probabilities". The prior distribution represents what is already known about an exposure scenario, and can be based upon professional judgment; company, industry or trade organization experience; historical or surrogate exposure data; or exposure modeling predictions, such as those found in CE. The likelihood distribution represents a set of decision probabilities given only the current data. The posterior distribution is calculated using both the prior and likelihood distributions, and represents the final set of decision probabilities. The calculation techniques are complex, but the outcome is a simple bar chart of decision probabilities.

With the following three assumptions, BDA can be applied to the validation of CE: (1) the upper end of each control band is statistically interpreted as an upper percentile exposure (e.g., 95th percentile exposure), (2) a process is considered "controlled" relative to the target control band whenever the true 95th percentile exposure is less than the upper limit for the band, and (3) exposures tend to be lognormally distributed. The first two assumptions are consistent with the original published literature on COSHH Essentials. Given these assumptions, exposure data for a given combination of process, exposure prediction band, and control strategy can be analyzed using BDA to determine the probability that the true 95th percentile exposure falls within each of the A, B, C, and D control bands.

A nearly identical analysis could be used by employers (and their IH professionals) to verify that actual exposures are consistent with the target control band. The initial CE assessment could be adapted to represent a prior decision distribution, which would then be combined with a number of worker or process exposure measurements to determine the posterior probability that exposures are controlled relative to the target control band. The combination of CE and actual data has the potential to permit efficient and effective decisions regarding the actual control of exposures.

In summary, it is proposed that Bayesian decision analysis techniques be applied to both the validation of the COSHH Essentials model and the verification by employers that exposures are indeed controlled appropriately. These techniques have several advantages: decision probabilities are estimated using a commonly accepted distribution model and exposure parameter, more accurately reflect the predictive ability of COSHH Essentials, and are easy to understand and visualize.

2 BACKGROUND

The COSHH Essentials scheme, developed by the UK Health and Safety Executive (HSE), is basically an exposure prediction model. By identifying three attributes of the process (or unit operation) -

- quantity used per day,
- dustiness or volatility, and
- type of existing controls,

- the user predicts which of four Control Bands - A, B, C, or D - is most likely. For particulates, the upper ends of the control bands are 10, 1, 0.1, and 0.01 mg/m^3. For vapors, the upper ends are 500, 50, 5, and 0.5 ppm.

The HSE acknowledged that COSHH Essentials (CE) was introduced and released for use without being fully validated:

> "It has proven to be surprisingly difficult to find quality data against which to compare the ... exposure predictions."
> (Maidment, 1998)

In place of rigorous validation there was a "high degree of reliance" on peer review (Maidment, 1998). Consequently, users have little to no assurance that CE accurately and reliably predicts exposures.

2.1 Validation Studies

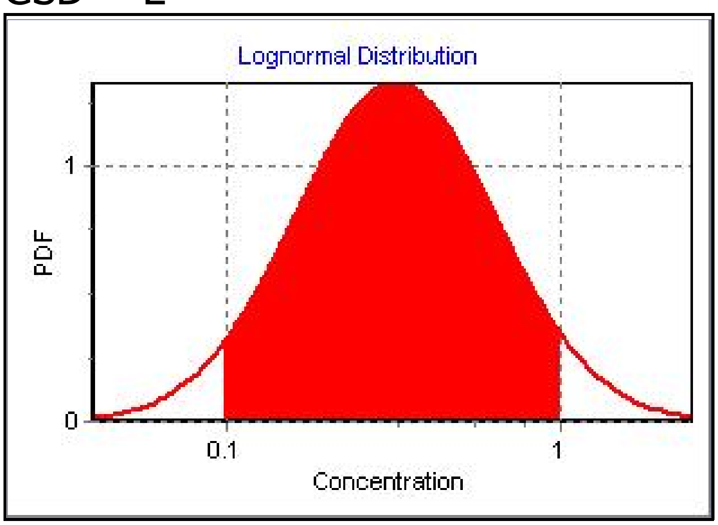
There are few studies - HSE sponsored or otherwise - attempting to validate CE. Probably the best published thus far is that by Tischer *et al*. (2003). Using exposure data collected for 18 unit operations at dozens of work sites in Germany, the authors determined the <u>fraction of measurements</u> that fall into each control band for each unit operation. We will call this the "fraction agreement" method.

Tischer *et al*. found "reasonably good" agreement between the measurements and the predicted control band. For 10 of the 18 operations the majority of the measurements fell within the control band predicted by the CE assessment. For the remaining 8 operations the majority of the measurements was in the next lower band. However, for 6 operations an appreciable fraction of exposures was in the next higher band.

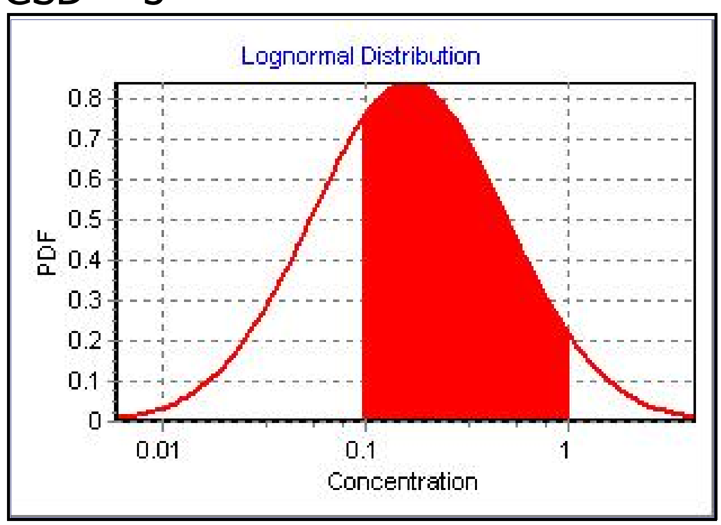
Tischer *et al*. neither calculated summary statistics nor assumed that a particular statistical distribution applied. For this presentation we will assume that whenever exposures are aggregated across several work sites the data can usually be well described by a lognormal distribution.

For example, let us set the 95th percentile of a lognormal distribution equal to 1 mg/m³, the upper end of the B band for particulates. The figures below show us that we should expect to observe a fraction of the exposure profile to be in both the higher and lower bands, and that this fraction will vary with the underlying geometric standard deviation (GSD) and location of the true 95th percentile within the target band.

GSD = 2



GSD = 3



The proper interpretation of a "fraction agreement" analysis is not obvious. For example, the majority of measurements could easily fall in the predicted band or lower, but with a true 95th percentile that is well into the next higher band. We propose an alternative analysis.

3 VALIDATION USING BAYESIAN METHODS

We submit that COSHH Essentials (CE) fits a Bayesian decision framework. The user reaches a decision as to which of four exposure bands applies to a particular process. By assuming that a statistical distribution can be applied to occupational exposures and by using Bayesian calculation techniques (Hewett, 2003), we can calculate the probability that the overall 95^{th} percentile for the sites studied fell within each control band.

To use Bayesian Decision Analysis (BDA), we first assume the following:

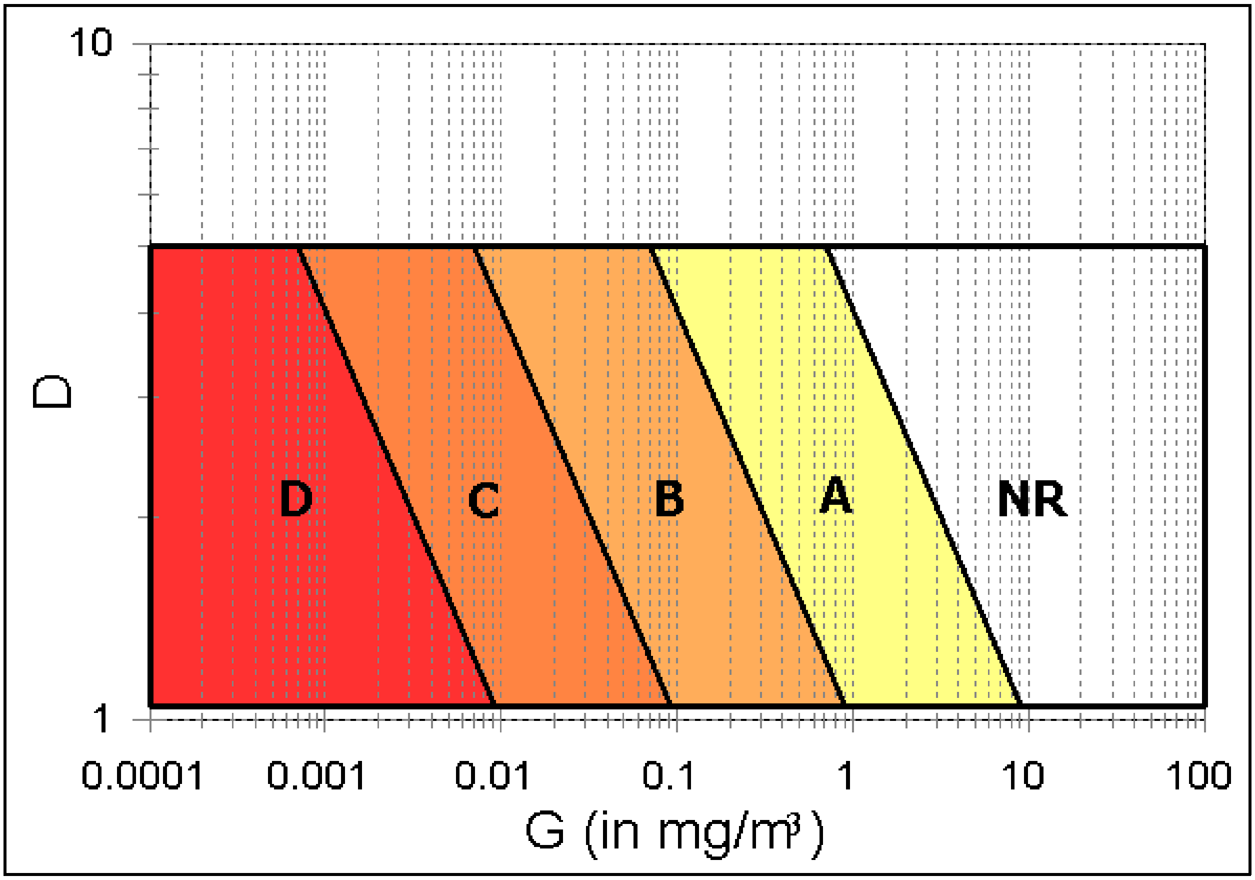
- The exposure profile for each "unit operation" is basically lognormal.

- A relevant exposure parameter is the overall (or "unit operation") 95th percentile exposure (or some other upper percentile).

- It is possible to define a *Universe* of potential exposure profiles, and to divide this Universe into control bands.

3.1 Exposure Universe - Particulates

The exposure Universe for particulates can be divided into the four CE control bands, and a "Not Recommended" (NR) band:



where G = geometric mean and D = geometric standard deviation.

Each band in the above graph consists of all combinations of G and D where the 95th percentile exposure is between the upper and lower cutoffs for the band. The Universe for vapors is similar, except that G is in ppm and the cutoffs between the D, C, B, A, and NR bands are 0.5, 5, 50, and 500 ppm.

3.2 Introduction to Bayesian Decision Analysis

A project is underway to develop a methodology, based on Bayesian statistics, for industrial hygienists to use when evaluating the work environment. This methodology, or Bayesian Decision Analysis (BDA), is based upon the concepts of prior, likelihood, and posterior distributions of "decision probabilities". The prior distribution represents what is already known about an exposure scenario, and can be based upon professional judgment; company, industry or trade organization experience; historical or surrogate exposure data; or exposure modeling predictions, such as those found in CE. The likelihood distribution represents a set of decision probabilities given only the current data. The posterior distribution is calculated using both the prior and likelihood distributions, and represents the final set of decision probabilities. The calculation techniques are complex, but the outcome is a simple bar chart of decision probabilities. A manuscript describing this method, with example calculations, is in preparation.

The following equation is the key to BDA. It calculates the probability of the ith control band, given the data:

$$P(CB_i|data) = \frac{\int^{G'}\int^{D'} \left[P(data|G,D)\cdot P(CB_i)\right]\, d(G)d(D)}{\int_{G_{\min}}^{G_{\max}}\int_{D_{\min}}^{D_{\max}} \left[P(data|G,D)\cdot P(CB_i)\right]\, d(G)d(D)}$$

where data = the available exposure measurements
CB_i = the ith control band
G = geometric mean
D = geometric standard deviation
G′ and D′ = values of G and D that fall within the ith band

For the validation of CE, we propose using a uniform (i.e., non-informative) prior decision distribution.

Integration of the above equation is not straightforward, as it represents a 3D surface where the integration ranges are irregular. We used an Excel spreadsheet and code written in Visual Basic for Applications® to implement a Monte Carlo Simulation approach to 3D integration.

3.3 Universe Boundaries

To integrate the above equation, we must define minimum and maximum values for the boundaries of the exposure Universe:

3.3.1 Vapors

G_{min} = 0.01 ppm G_{max} = 1000 ppm
D_{min} = 1.05 D_{max} = 5

3.3.2 Particulates

G_{min} = 0.0001 mg/m^3 G_{max} = 100 mg/m^3
D_{min} = 1.05 D_{max} = 5

For most processes, where the geometric standard deviation is normally in the 1.5 to 3 range, we suggest that D_{max} be set a 4 or 5. For several of the examples (described later) the sample D was excessive, in which case we set D_{max} at 10.

4 DEMONSTRATION

To demonstrate the validation of CE using Bayesian Decision Analysis we first intended to use randomly generated data. However, Martin Tischer and colleagues (see Tischer *et al.*, 2003) graciously allowed us to use several of their datasets. We were given eight datasets; however, we used only those where the total sample size was less than n=100. This was because our software for doing BDA is in development and at this point can only accept up to 100 measurements.

We first calculated summary statistics for each dataset, primarily to determine if our default maximum geometric standard deviation needed to be revised. Next, we used the above equation to calculate the probability that the true 95th percentile for this type of operation, under the observed conditions, fell into each of the control bands.

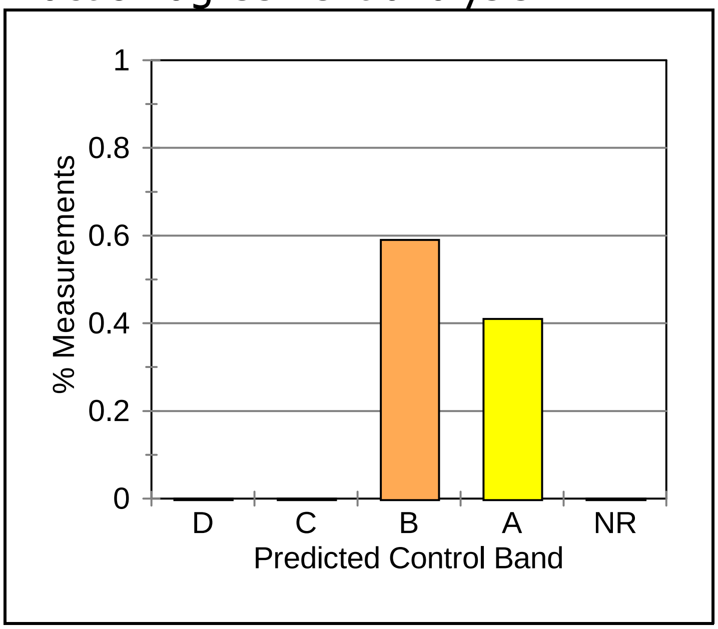
5 RESULTS - VAPORS

The "predicted exposure band" was determined by the application of CE (see Tischer *et al.*, 2003). The question that we addressed was "are the measurements consistent with the exposure band predicted by the COSHH Essentials scheme?" In the results to follow we contrasted the "percent agreement" analysis of Tischer *et al*. (2003) with our Bayesian analysis.

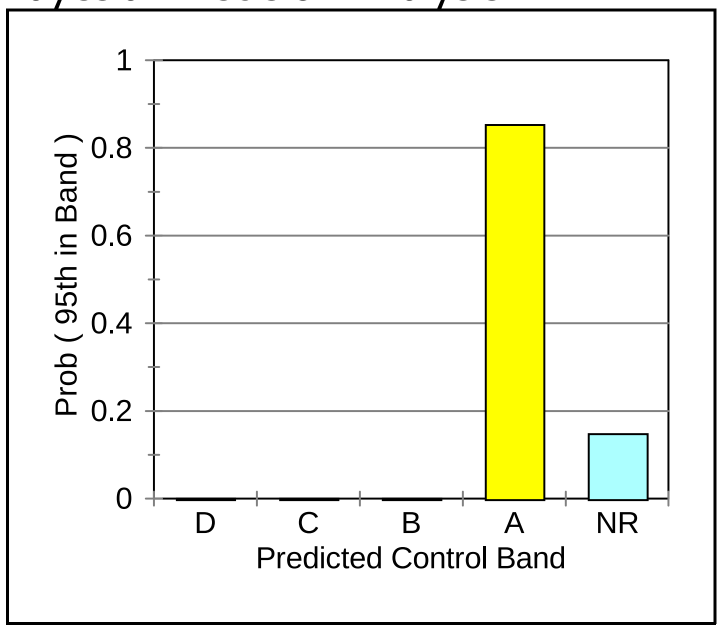
5.1 Tischer et al. Scenario #2 - Textile Industry (8 sites)

Predicted exposure band = A

Fraction agreement analysis:



Bayesian Decision Analysis:



Statistics: $n = 49$
$gm = 23.13$ ppm
$gsd = 4.75$
$X_{0.95} = 300.24$ ppm

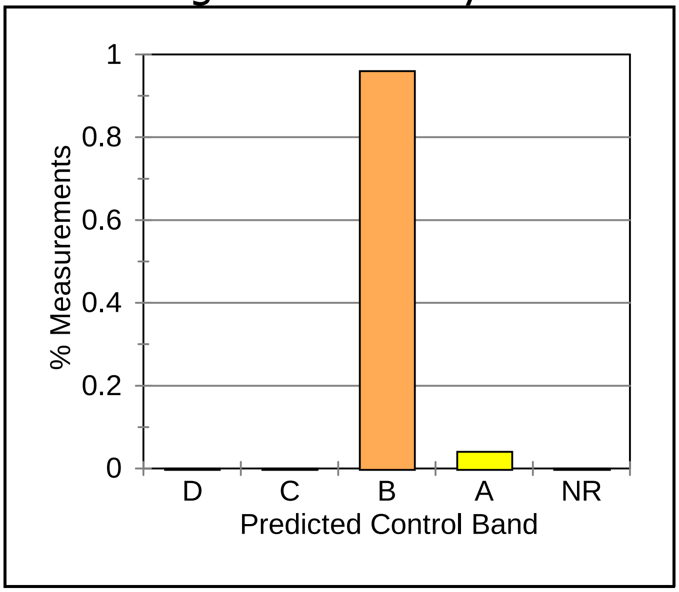
Comments: We set $D_{max}=10$ due to the large sample GSD.

No measurements were observed greater than 220 ppm, yet BDA suggests that there is a 15% probability that the a process 95th percentile exposure will be in the NR band. This is because the measurements in this dataset varied greatly, as indicated by the large gsd, making the extreme 95th percentiles in the NR band mathematically possible. We suggest that a better analysis strategy would be to analyze the data from each site separately, rather than as one large database, and simply average the probabilities for each control bands for the n sites studied. (See our 8.1 recommendation.)

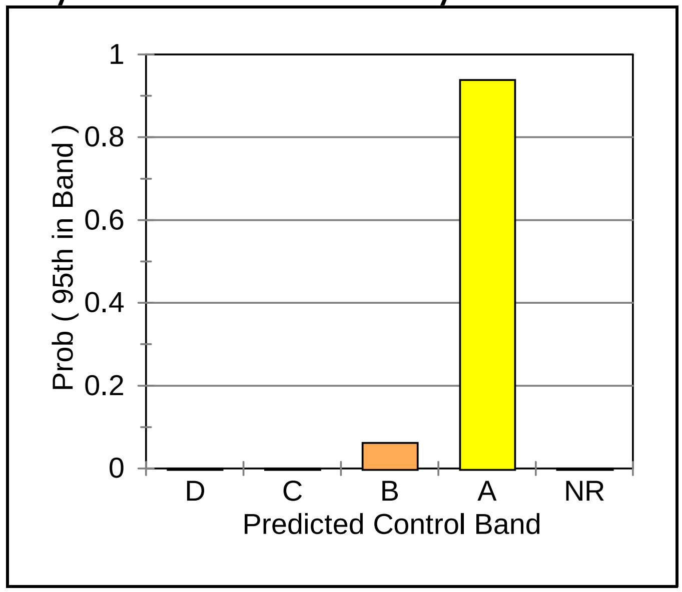
5.2 Tischer et al. Scenario #4 - Screen Printing (11 sites)

Predicted exposure band = A

Fraction agreement analysis:



Bayesian Decision Analysis:



Statistics:

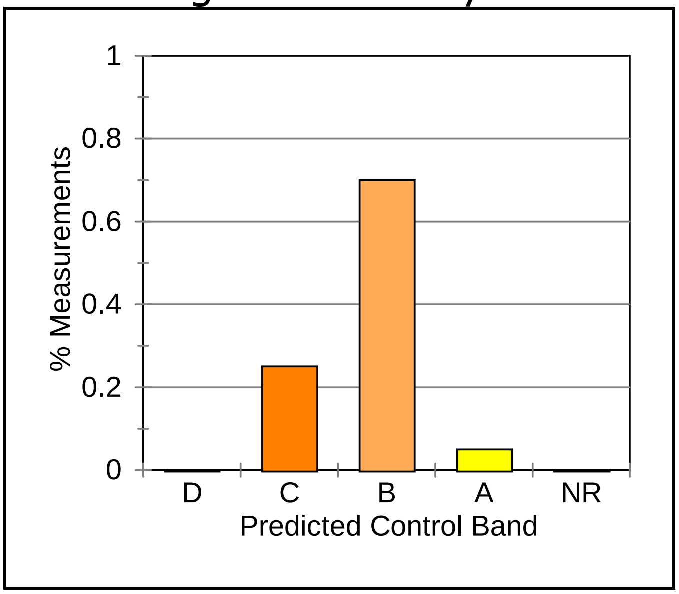
n = 87
gm = 17.09 ppm
gsd = 2.28
$X_{0.95}$ = 66.45 ppm

Comments: BDA suggests that there is a 94% probability that a process 95th percentile exposure will be in the predicted control band.

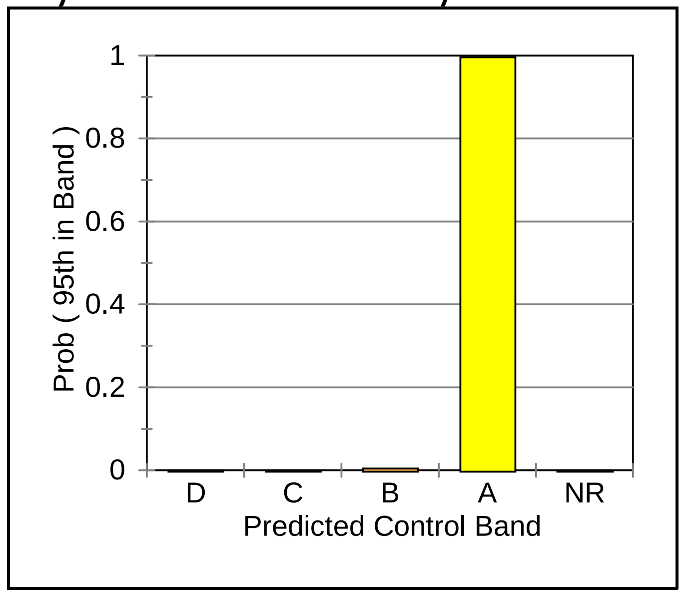
5.3 Tischer et al. Scenario #9 - Optician Workshop (14 sites)

Predicted exposure band = B

Fraction agreement analysis:



Bayesian Decision Analysis:



Statistics: n = 78
gm = 11.45 ppm
gsd = 3.471
$X_{0.95}$ = 88.68 ppm

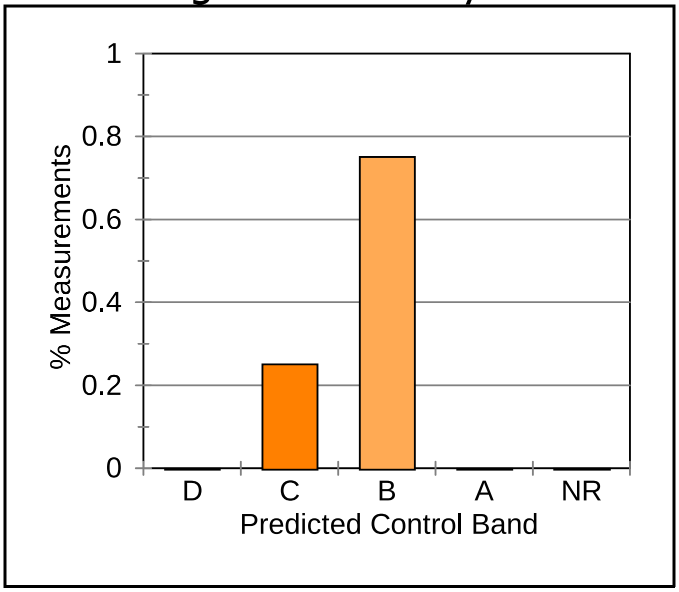
Comments: We set D_{max}=10 due to the large sample GSD.

BDA suggests that there is a >99% probability that a process 95th percentile exposure will be in the next higher control band.

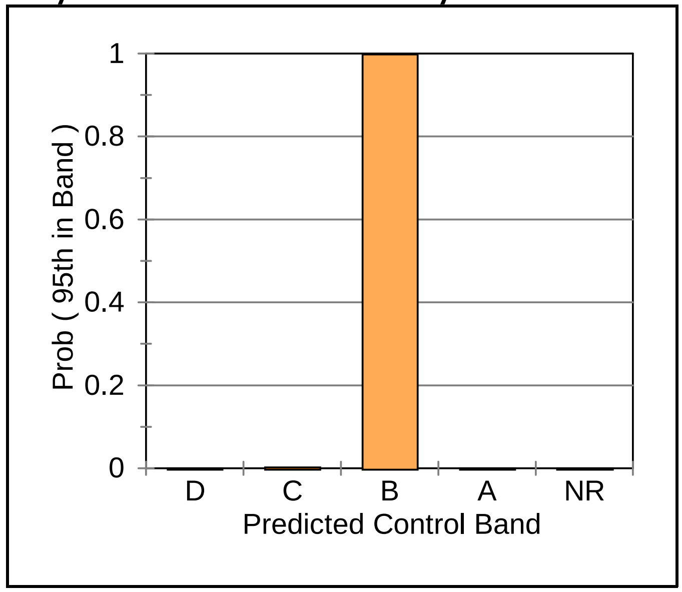
5.4 Tischer et al. Scenario #11 - Screen Printing (10 sites)

Predicted exposure band = B

Fraction agreement analysis:



Bayesian Decision Analysis:



Statistics: n = 68

gm = 1.49 ppm

gsd = 2.93

$X_{0.95}$ = 8.72 ppm

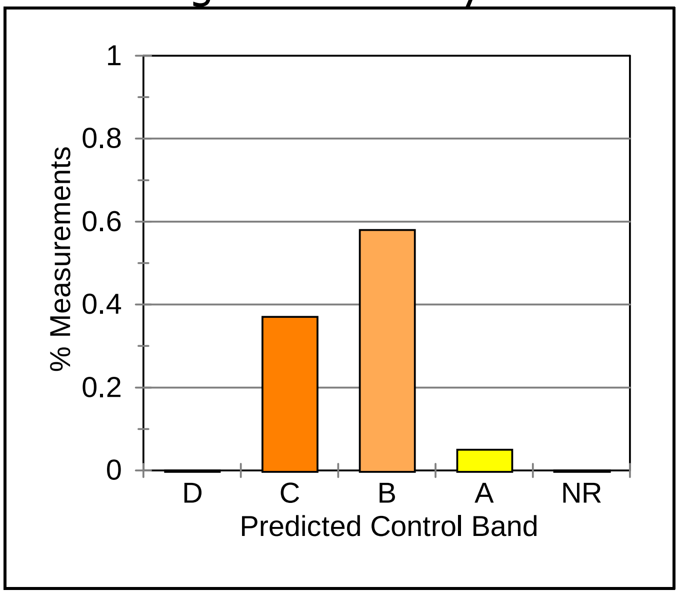
Comments: BDA suggests that there is a >99% probability that a process 95th percentile exposure will be in the predicted control band.

6 RESULTS - PARTICULATES

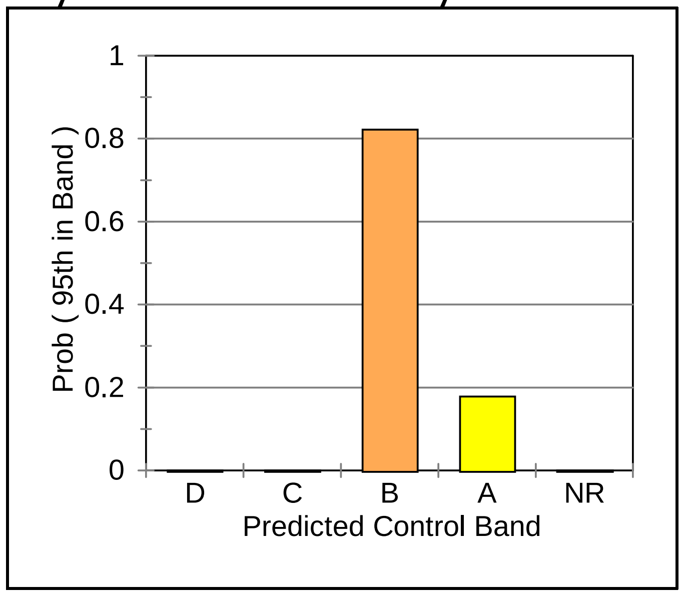
6.1 Tischer et al. Scenario #16 - Chemical Industry (3 sites)

Predicted exposure band = B

Fraction agreement analysis:



Bayesian Decision Analysis:



Statistics:

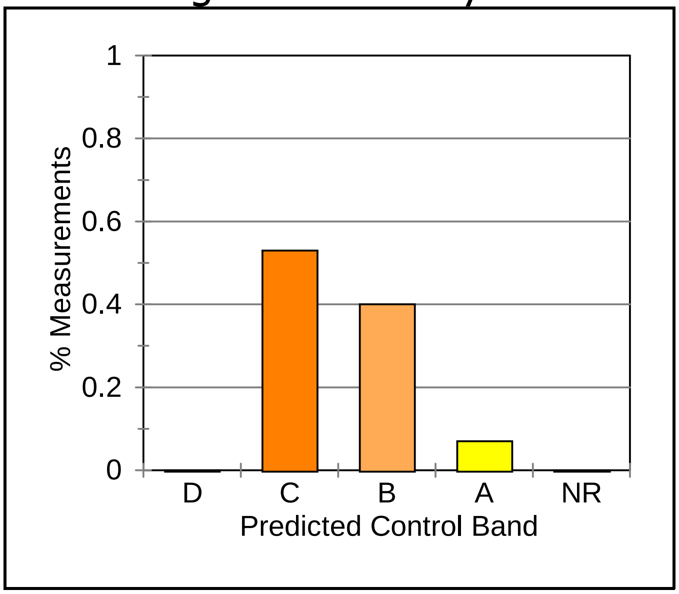
n = 52
gm = 0.15 mg/m³
gsd = 2.56
$X_{0.95}$ = 0.72 mg/m³

Comments: BDA suggests that there is a 18% probability that the a process 95th percentile exposure will be in the next higher control band.

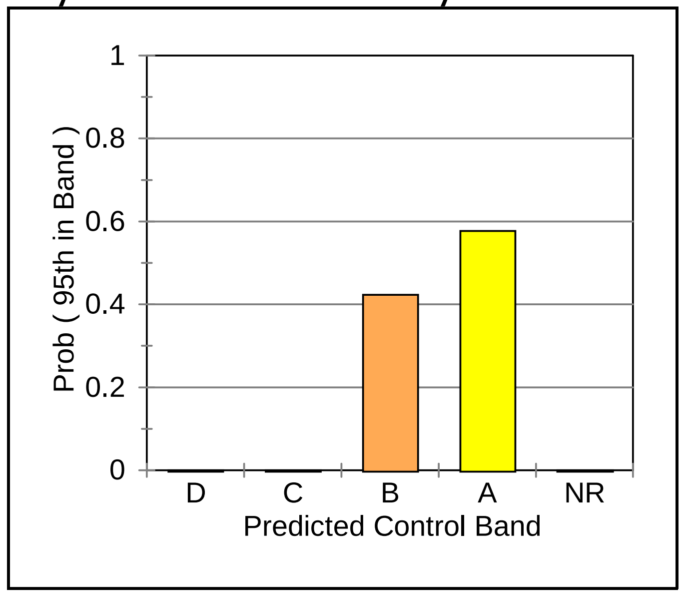
6.2 Tischer et al. Scenario #17 - Chemical Industry TRK (5 sites)

Predicted exposure band = B

Fraction agreement analysis:



Bayesian Decision Analysis:



Statistics:

n = 66
gm = 0.06 mg/m³
gsd = 5.435
$X_{0.95}$ = 1.03 mg/m³

Comments:

We set D_{max}=10 due to the large sample GSD.

BDA suggests that there is a 58% probability that a process 95th percentile exposure will be in the next higher control band.

7 DISCUSSION

The above comparisons between the "percent agreement" analysis and the Bayesian analysis suggests that, in general, a Bayesian analysis will lead to conclusions that the process exposure profiles tend to be in the higher control bands. In other words, a seemingly favorable "percent agreement" analysis may not be so favorable when examined using Bayesian techniques.

The basic problem with the "percent agreement" agreement approach is that there is no clear definition of success. Is success defined as 100% of the measurements in the predicted band? If not, then how does one conclude that the model is a success at accurately predicting exposures, or a failure due to excessive over- or under-prediction?

In contrast, when using the BDA method success can be objectively defined. For example:

> The CE model is considered a success *for a particular type of process and CE scenario* - i.e., quantity, conditions of use, and control - if the probability that a process 95^{th} percentile is within the predicted control band is 70% or greater, with no more than a 10% probability that it is in the next higher band or more than a 30% probability that it is in the next lower band.

One can argue with the specifics of this example definition, but one cannot argue that success is not defined. Of the two methods, we submit that the Bayesian approach is preferable.

7.1 Advantages to using BDA for validation

- BDA is based on a recognized distributional model.

The lognormal distribution is commonly used to characterize occupational exposure data, and is built into BDA. However, BDA does not require that the overall dataset be unimodal lognormal. Each site will have a different G, D, and 95th percentile. In most cases, this does not adversely affect the calculations. BDA simply calculates of probability that the overall 95th percentile comes from each control band, even though the site specific 95th percentiles are different.

However, as we observed in Scenario #2, extremely variable data can lead to higher than expected probabilities for the higher control bands. As discussed in the next section, we suggest that there should be both an analysis of both the aggregate data and the data from each site.

• The probability per band is calculated based upon the location of an upper percentile.

We chose to use the 95th percentile as this is consistent with the recommendations in the U.S. (see Mulhausen and Damiano, 1998), as well as with HSE studies of CE. Other upper percentiles, such as the 90th (Tischer, 2001) or 99th could easily be substituted.

• Probabilities per band are more easily interpreted.

Once we agree on the upper percentile - e.g., the 90th, 95th, or 99th - the interpretation of the probabilities per band is fairly straightforward. The probability of the primary control band will tend to become larger as additional data are added. This will not necessarily happen with a "fraction agreement" analysis.

8 RECOMMENDATIONS

8.1 Validation of COSHH Essentials

• Bayesian Decision Analysis could be used to examine the predictive abilities of COSHH Essentials.

Use of Bayesian Decision Analysis (BDA) can lead to a different conclusion than a "percent agreement" analysis. Because BDA is based on a distributional model of exposures, a reasonable and established statistical interpretation of exposure control relative to an exposure limit, a BDA-based assessment may be a more appropriate method for validating control banding. Absent any superior method being proposed, we recommend using BDA to validate COSHH Essentials.

• Analysis should be directed at both the overall dataset and the site specific datasets.

Here we examined the overall dataset, as did Tischer *et al*. (2003). It is possible that significant differences between plants may be overlooked whenever a validation analysis is focused on aggregate data. The fact that 3 of the 6 dataset we examined had very large sample geometric standard deviations suggests that exposures varied considerably within, between, or both within and between the sites studied.

We recommend that any analysis be directed at both the overall dataset for each unit operation and each site specific dataset. Analysis of the distributions of control band probabilities across all sites might lead to new and potentially valuable insights. In addition, if additional information is available regarding the actual quantity of use and the quality of the controls, it might be possible to do a components-of-variance analysis.

8.2 Employer-based Verification of Control

A nearly identical analysis could be used by employers (and their IH professionals) to *verify* that actual exposures at a specific site are consistent with the target control band. The initial CE assessment could be adapted to represent a prior decision distribution, which would then be combined with a number of worker or process exposure measurements to determine the posterior probability that exposures are controlled relative to the target control band.

Consider the following suggested steps for verifying controls:

8.2.1 Step 1 - Use COSHH Essentials to determine Initial Control Band Rating

8.2.2 Step 2 - Prioritize Unit Operations for Verification

8.2.3 Step 3 - Collect exposure measurements from high priority operations

8.2.4 Step 4 - Apply BDA to determine if the process upper percentile exposure is highly likely to be in the appropriate hazard band.

The combination of CE and actual data has the potential to permit efficient and effective decisions regarding the actual control of exposures. Work is currently underway to refine and validate this type of approach in a chemical manufacturing environment.

9 CONCLUSIONS

In summary, it is proposed that Bayesian Decision Analysis be applied to both the validation of the COSHH Essentials model and the verification by employers that exposures are indeed controlled appropriately.

Bayesian analysis has several advantages:

- decision probabilities are estimated using a commonly accepted distribution model and exposure parameter,

- more accurately reflect the predictive ability of COSHH Essentials, and

- are easy to understand and visualize.

10 ACKNOWLEDGMENTS

We are grateful to Mark Tischer and colleagues for providing the datasets for us to demonstrate Bayesian Decision Analysis.

11 REFERENCES

Hewett, P.: Improving the Effectiveness and Efficiency of IH Decision Making Process. Abstract 278, American Industrial Hygiene Conference and Exposition (www.aiha.org) (2003).

Maidment, S.C.: Occupational Hygiene Considerations in the Development of a Structured Approach to Select Chemical Control Strategies. *Annals of Occupational Hygiene* 42:391-400 (1998).

Mulhausen, J. and Damiano, J.: (editors): *A Strategy for Assessing and Managing Occupational Exposures, Second Edition*. Fairfax, VA: American Industrial Hygiene Association (1998).

Tischer, M.: What Does "Low Exposure" Mean? Exposure Considerations in the Testing of Notified New Substances. *Applied Occupational and Environmental Hygiene* 16:228-232 (2001).

Tischer, M., Bredendiek-Kämper, S., and Poppek, U.: Evaluation of the HSE COSHH Essentials Exposure Predictive Model on the Basis of BAuA Field Studies and Existing Substances Exposure Data. *Annals of Occupational Hygiene* 47:557-569 (2003).