Stephen Bird
Department of Political Science (doctoral candidate)
Boston University School of Management

Stephen Bird designed and implemented a risk-assessment and prioritization tool that the Office of Technical Assistance (OTA) and other environmental agencies can use to prioritize outreach and enforcement to maximize their resources. OTA, a sub-agency at Massachusetts' Executive Office of Environmental Affairs, helps private and public facilities reduce chemical usage and implement pollution prevention measures across the state. The risk-assessment tool will also help agencies increase knowledge of their regional profiles, and identify non-filing facilities in their jurisdictions. The method rates a complete population of facilities that use threshold amounts of chemicals in the state, and allows for incorporation of less apparent risk factors such as a facility's credit rating. Standard risk factors derived from chemical usage and the kind of facility are also used. While still unrefined in its initial design, the development of this method represents the first time a state agency has been able to gain a comprehensive overview of facilities within its jurisdiction. Both the Environmental Protection Agency and the Massachusetts Department of Environmental Protection were briefed in late July. It is anticipated that they will be included in future policy development as the tool is refined further.

MEMORANDUM

From: Stephen Bird

Subject: OTA Risk Prioritization Tool

Date: 8/5/2004

Executive Summary

Environmental and security risk assessment for chemicals has increased in importance because many state environmental agencies are pressed to produce improved results while simultaneously losing staff and other resources in the context of tight state budgets. Secondly, in light of heightened homeland security, there is a concern that large-scale chemical use presents "low-hanging fruit" for those who would wish to conduct sabotage, theft, or terrorism.

The creation of a flexible risk assessment methodology, which can prioritize inherent risk in a wide variety of facilities using toxic chemicals, will add a new tool to the palette of regulatory options for outreach, enforcement, and improved facility knowledge. The Office of Technical Assistance (OTA), a sub-agency at Massachusetts' Executive Office of Environmental Affairs, can further attain its goal of helping facilities to reduce chemical and toxics usage, and to implement pollution prevention measures across the state. Other environmental agencies can also use this system to prioritize outreach and enforcement to maximize their resources, increase knowledge of regional facility profiles, and identify non-filing facilities in their jurisdictions.

The project is based on research conducted at the Wharton School's Risk Management and Decision Processes Center at the University of Pennsylvania. Their initial risk analysis system was based on the ratio between threshold amount for EPA reporting and the amount of different chemicals being stored. This rating system has a statistically significant correlation with accident rates. Additional research reveals other risk factors, including a significant relationship between a company's degree of debt and accident rates.

The assessment tool integrates several of these risk factors (or proxy variables) and assigns a risk measurement rating to a larger, comprehensive population of facilities "from the ground up" as a measurement of future potential risk. It uses EPA data on state facilities with high hazard chemicals and separately obtained company financial data.

Research conducted at OTA also identifies future avenues for increasing the sophistication and reach of the tool and adding different risk factors such as geographical information (using GIS methodology), transportation information, facility size, and other aspects of chemical characteristics to the analysis. While many refinements to the initial design can still be made, the development of this system represents the first time a state agency has been able to gain a comprehensive overview of facilities within its jurisdiction. Both the Environmental Protection Agency and the Massachusetts Department of Environmental Protection were briefed in July. They are expected to be included in future development as the system is improved further.

Discussion

The risk assessment system will allow OTA and other agencies to prioritize resources for facilities that are of the greatest interest for outreach (and potentially for other agencies, enforcement), increase the knowledge of the regional or jurisdictional profile, and identify non-filing facilities. However, several aspects of the tool's development can be improved. Data is sometimes incomplete, much of the data is hard to integrate, and, currently, proper weighting (multivariate regression) of different risk factors is not possible. Thus, the tool is really an "ugly duckling" of risk assessment, which, with improvements, can mature into the equivalent of a more beautiful "swan" of regulatory methodology.

Few risk assessment tools are able to analyze large populations.[1] While both state and federal regulators have created highly sophisticated risk assessment tools for specific issues, they have not done so on a broad basis. One of the advantages of the OTA tool/system is simply that it allows a broader survey of large populations (e.g. the entire state of Massachusetts). The tool is adaptable for targeting more specific sub-population issues within regions. For instance, agencies could focus on a specific industrial or manufacturing sector, a specific kind or class of chemical (e.g. reactives), or solely facilities with a certain kind of credit profile. Suggested uses for other agencies include:

- State Emergency Response Commissions (SERCs) and Local Emergency Planning Committees (LEPCs) can identify non-filing facilities in their jurisdiction.
- Dept. of Environmental Protection (DEP) can identify facilities of interest for enforcement or profiles.
- Environmental Protection Agency (EPA) can identify facilities of interest and non-filing facilities.

Beginning in 2000, research began at the Risk Management and Decision Processes Center of the University of Pennsylvania's Wharton School.[2] This work focused on the Risk Management Plan (RMP) data kept by the EPA for the entire United States. Kleindorfer, Elliot, Lowe et al's (K, E, & L) initial research examined correlations between accident data for RMP facilities and characteristics of the entire population of facilities (just over 15,000 in total). Several papers were produced that focused on the characteristics of facilities having accidents and the accidents themselves.

The risk assessment tool for OTA was developed using many of the initial results derived from the K, E, & L research, and was conceptually inspired by their approach. However, there are some significant differences between the two projects. The K, E, & L research is "top-down" accident analysis whereas the OTA system is a "bottom-up" working tool which derives the assumptions behind its ratings from the K, E, & L research. Further differences will be described later in this document.

The K, E, & L research describes

> the data collected on one of the most ambitious government-sponsored environmental data acquisition projects of all time, the Risk Management Plan (RMP) data collected under 112(r) of the Clean Air Act Amendments of 1990. This RMP Rule 112(r) was triggered by the Bhopal accident in 1984 and led to the requirement that each qualifying facility develop and file with the U. S. Environmental Protection Agency a Risk Management Plan (RMP) as well as accident history data for the five-year period preceding the filing of the RMP. This data was collected in 1999-2000 on more than 15,000 facilities in the U.S. that store or use listed toxic or flammable chemicals believed to be a hazard to the environment or to human health of facility employees or off-site residents of host communities. The resulting database, RMP*Info, has become a key resource for regulators and researchers concerned with the frequency and severity of accidents, and the underlying facility-specific factors that are statistically associated with accident and injury rates.[3]

One of the first challenges in developing the OTA tool was that using the RMP database was inappropriate for analysis of a complete toxic-chemical-using population in the state. First, the RMP database has substantial detail concerning worst case scenarios for chemical accidents and has become a security liability in a post 9/11 world. Access to the RMP database now requires extensive security clearances and doesn't allow for adequate transparency or easy circulation. Secondly, the RMP database has very high threshold requirements for required participation. This means that its reach and population is not nearly as far-reaching as one would prefer for a tool that analyzes as many facilities as possible.

The use of RMP data for the K, E, & L research was indispensable because it provides accident data for statistical analysis and correlation that is simply unavailable in other US environmental regulatory regimes. The key interrelation between the two projects is using the statistical accident analyses done by K, E, & L (only available through RMP) and applying them to a broader population of data (Tier II) for use by OTA.

The OTA risk tool uses Tier II data from the EPA. Tier II is a national regulatory regime derived from Section 312 of the Emergency Planning and Community Right to Know Act (EPCRA). Facilities are required to file this data if they surpass storage thresholds for any of more than 600 chemicals. The vast majority of the data is now filed electronically with the EPA. It is then distributed to SERCs, LEPCs, and state environmental agencies if needed. 2295 facilities filed Tier II data in Massachusetts for 2002, however far fewer filed RMP reports. According to informal EPA estimates, this represents only 25-40% of facilities actually required to file such information. Nonetheless, the Tier II data represents the largest population of facilities storing chemicals in any state (a much larger population than RMP data) and offers the largest potential population for a risk rating system.

This initial version of the risk assessment tool uses the basic "threshold" chemical risk rating plus four additional risk factors incorporated into the numerical ratings:

1. whether a chemical is an extremely hazardous substance (EHS)
2. an additional chemical risk factor derived from the K, E, & L analysis
3. an industrial classification risk factor (Standard Industrial Classification or North American Industrial Classification System – SIC or NAICS) derived from the K, E, & L analysis
4. an additional facility risk factor derived from the facility credit rating if available

The basic chemical risk rating is derived similarly to the methodology used by K, E, & L. The records in the Tier II data refer to EPA chemical threshold amounts. Each chemical receives a numerical rating derived from the log2 of the ratio of the actual chemical stored to the threshold amount. If a facility has multiple chemicals, the ratings for each chemical are simply added. The logic behind this method is that the threshold amounts are set in accordance to the risk potential of each chemical. Additionally, if a chemical has been designated as an EHS by the EPA (a separate designation from the threshold amount), the risk rating for that particular chemical is increased by one.

Secondly, the K, E, & L analysis also allows for an incorporation of an additional chemical risk factor. They listed the 79 chemicals involved in RMP accidents from 1995-2000. A log2 function was used to reduce risk ratings from 1-752 accidents to a 1-10.6 potential additional risk rating for each chemical. There is a statistical/analytical concern for this data because the accident data was not computed as a percentage of the facilities that store those chemicals. Until data that is more accurate is available, a subjective decision was made to use the information but also to modify its potential effect by using the log function.

Similarly, a risk factor was derived from the industrial classification (using SIC or NAICS codes) of facilities having accidents. Again, a log2 function was used to convert accident numbers that ran from 2-182 accidents to a 1-7.5 additional potential risk rating.[4] The industrial classification code weighting was set at 10% of the total facility risk rating. Again, this risk factor has less statistical validity because it was not computed as a percentage of total facilities with this characteristic. For this reason the factor was used but its potential effect was purposively lessened by use of the log function, and its small percentage weight in the formula.

Lastly, credit data was incorporated into the tool/system. In forthcoming research not yet published, K, E, & L determined that each 1 point increase in the debt to equity ratio corresponded to a 12.2% increase in the risk of injury. Similarly, a 1% increase in a facility's return on equity corresponded to a 0.8% decrease in the risk of accident. Subsequent research by K, E, & L has determined that an approximate, informal weighting for the credit risk portion of the model should account for 20% of the total risk.[5] Actual debt-to-equity ratio was not used because many of the facilities filing Tier II data are not publicly traded (and thus the information was not available). As a proxy, a percentile credit rating was obtained from InfoUSA, a company that obtains non-voluntary, high coverage, multiple factor credit assessments.

Although K, E, & L determined there was a correlation between the number of accidents and size of facility, this information was not used in the tool/system. First, it was not controlled for by using "accident per employee" data. An observant reader may note that two other risk factors (industrial classification and added chemical risk) also don't account properly for their statistical control population yet they were still used. However, as several OTA analysts reasoned, larger facilities had certain characteristics that made them less susceptible to accidents or risk. First, there is a concern that facility size is actually a spurious correlation that is already captured by the chemical threshold-ratio ratings. In addition, larger facilities have higher degrees of underlying capital and expertise to bring to risk mitigation, greater concerns for losses and bad public relations, and increased economies of scale.

Obviously, many of the assumptions and derivations used in the system/tool need to be corrected and/or improved in the future. Particularly important is the need for a proper regression analysis of several integrated risk factors. While it still represents a great improvement over the absence of any large population risk analysis methods for facilities, this is clearly a rough estimation tool for which many improvements are possible. Other problems in developing the database include the following:

- Credit data only covers 60% of the facilities listed
- Only 25-40% of facilities that should be filing Tier II data are actually estimated to be complying.[6]
- Integration of EPA, credit, chemical, and industrial classification data is extremely difficult and potentially susceptible to some error.

Future Integration of Other Risk Identifiers

Several additional risk identifiers could be added to the system/tool. Additional chemical characteristics could be incorporated (e.g., reactives, flammables, etc.). Initial analysis by K, E, & L examined these factors but they were difficult to include because a weighting formula for them was not evident even as an informally estimated factor.

Additional facility characteristics can also be included. These could include an improved consideration of size, a change in sales (particularly a reduction per the falling angel issue discussed in footnote 5) and several location/geography oriented factors. These include overall population density, population hotspots (e.g. schools, office, high day/work density), environmental justice demographics[7] which incorporate socio-economic status using zip code or census tract, or chemical transportation information using U.S. Department of Transportation chemical manifests.[8]

Policy Proposals

The following policy suggestions address the functions of the system/tool, and additional suggestions for improving its development, accuracy, and ease of use:

Tool/system Use:

1. Use database as a one of a palette of regulatory tools for comparative identification and prioritization in both outreach and enforcement applications.
2. Conduct analysis on a regional basis.
3. Disseminate database to LEPCs and SERCs for identification of non-filers and higher risk facilities.

Improvements to tool/system:

1. Encourage federal and state agencies to work towards integrated (common) identification systems that work across databases (e.g. taxpayer ID #, Dun and Bradstreet, or others).
2. Begin work to increase sophistication and finesse of broad-based risk assessment tools. Expand K, E, & L style accident analysis.
3. Increase population reach and filing compliance for improved data analysis. problematic
4. Add and incorporate additional risk factors.
5. Use multiple regression analysis of all risk factors for comprehensive weighting of causal factors.

APPENDIX I: RISK ASSESSMENT RATING BREAKDOWN

	Chemical A	Log2 (storage amount x threshold amount)	EHS? Add 1 to rating	Kleindorfer add'l risk factor Add Log2 (Kleindorfer accident analysis #)	
Facility X	Chemical B	Same	Same	Same	Total chemical risk factor (A + B + C) etc.
	Chemical C	Same	Same	Same	

	Total facility multiple chemical risk rating	NAICS risk rating	Credit risk rating	Total risk assessment
	70%	10%	20%	100%
Four Models Possible	80%	n/a	20%	100%
	90%	10%	n/a	100%
	100%	n/a	n/a	100%

1. I am aware of only one other attempt to assess risk on a large scale. In the mid-nineties, the Region 1 EPA developed a rating system for geographic areas based on four characteristics of each census tract. They developed numerical ratings for each tract based on socio-economic factors, population density, etc. The risk system discussed here is the only attempt to assess risk for facilities rather than census tracts.
2. Belke, James, Michael Elliott, Harold Feldman, Paul Kleindorfer, Kiwan Lee, and Robert Lowe. 2004. *Accident History and Offsite Consequence Data from RMP*Info (02-04-PK).* Risk Management and Decision Processes Center, Wharton School, University of Pennsylvania, 2002 [cited May 2004]. Available from http://grace.wharton.upenn.edu/risk/downloads/02-04-PK.pdf.

   Elliott, Michael, Paul Kleindorfer, and Robert Lowe. 2003. "The Role of Hazardousness and Regulatory practice in the Accidental Release of Chemicals at U.S. Facilities." *Risk Analysis* 23 (5):883-896.

   Elliott, Michael, Paul Kleindorfer, Robert Lowe, and Yanlin Wang. 2004. *Drivers of Accident Preparedness and Safety: Evidence from the RMP Rule (04-17-PK).* Risk Management and Decision Processes Center, Wharton School, University of Pennsylvania, 2004a [cited July 2004]. Available from http://grace.wharton.upenn.edu/risk/downloads/04-17-PK.pdf.

   Elliott, Michael, Paul Kleindorfer, Robert Lowe, and Yanlin Wang. 2004b. "Environmental Justice: Frequency and Severity of US Chemical Industry Accidents and the Socioeconomic Status of Surrounding Communities." *Journal of Epidemiology and Community Health* 58:24-30.
3. Belke, et al., Accident History and Offsite Consequence Data from RMP*Info. (02-04-PK) (Risk Management and Decision Processes Center, Wharton School, University of Pennsylvania, 2002 [cited May 2004]), 1; available from http://grace.wharton.upenn.edu/risk/downloads/02-04-PK.pdf
4. Initial chemical/threshold hazard ratings for facilities run from 0-1833, although 95% of the ratings are below 75. These ratings were converted to percentiles. If applicable, the industrial classification risk factor accounts for 10% of total risk rating. This factor was not available for some facilities. Appendix I details the weighting formulas used when data was unavailable for certain risk factors.
5. Personal communication with Elliott and Kleindorfer, 6/04. New research also indicates the importance of "falling angels." Companies which had good credit assessments but whose credit has recently and quickly deteriorated have significantly higher potential for accidents. Publication of their formal statistical analysis is forthcoming. K, E, & L also determined that risk factors for credit were mostly derived from "worst offenders." Generally, only companies with severe credit concerns had increased risk associated with credit problems. This factor was not incorporated into the initial assessment tool because the detailed research and data is not yet available.
6. Informal estimate by Region I EPA staff. Compliance is expected to increase in the future as the EPA increases use of digital information.
7. Elliott, Michael, Paul Kleindorfer, Robert Lowe, and Yanlin Wang. 2004b. "Environmental Justice: Frequency and Severity of US Chemical Industry Accidents and the Socioeconomic Status of Surrounding Communities." *Journal of Epidemiology and Community Health* 58:24-30.
8. An unanticipated result has occurred with regulatory threshold triggers for chemical storage. Some companies have increased transportation frequency of chemicals to keep chemical storage amounts below regulatory thresholds. A facility may transport 5,000 gallons of a chemical ten times a year, rather than making a single 50,000 gallon delivery once a year. The increased transportation risk may actually be greater than the risk of storing up to 50,000 gallons of the chemical on-site.