Estimating the Costs of a Data Breach: An Exercise at the New Hampshire State Cancer Registry

Bruce Riddle¹, Steve Nyman², Judy Rees¹
¹New Hampshire State Cancer Registry, ²Chief Information Security Officer, Dartmouth College

Background

Following a risk assessment undertaken at Dartmouth College, NHSCR performed a planning exercise to estimate what a data breach might cost our supporting institution.

In information security, risk is defined as “the potential that a given threat will exploit vulnerabilities of an asset or group of assets and thereby cause harm to the organization.” [1] Exposure is an estimate of the potential losses. If you have a $20 bill on the sidewalk, the risk is high that it will be stolen but the exposure is low. The potential risk for a central cancer registry is unauthorized access to personal and medical records due to theft, loss or electronic access to a data server or field laptop, and loss of documents from the office.

Many states require notification of breaches to a central registry [2] which lists all breaches and makes documentation of the response available for public scrutiny. In 2003, a breach was discovered at the Carolina Mammography Registry, affecting up to 180,000 patients [3]. Each patient was sent a long explanatory document with detailed information about the breach. The financial costs of that particular breach are unclear, but the intangible costs at other mammography registries and other registries with potential budgetary support are discussed later.

Estimating the Costs of a Data Breach

The Ponemon Institute, sponsored this year by the computer security company, Symantec, publishes an annual report on the cost of a data breach [4]. The most recent survey estimates the total cost of data breaches in a commercial setting as follows:

| Table 1. Potential costs associated with a central registry data breach |
|-----------------------------|-----------------------------|
| Intangible Costs | Tangible Costs |
| • Breach of trust between citizens and institution | • Costs of IT staff to evaluate and contain breach and initiate breach response |
| • Inconvenience for citizens (risk of identity theft and counter-measures) | • Certified letters to all persons or person’s families in registry |
| • Damage to the reputation of the institution as a data custodian and steward of public health information | • Credit monitoring of all living persons |
| • Damage to other activities at institution, (e.g., registries, research studies, clinical trials) | • Legal fees to attorneys supervisors breach follow-up and defending institution |
| • Adverse publicity for public health and national agencies, (e.g., NCI, NAACCR, SEER) | • Civil and criminal fines |
| • Reduced political support for registries with potential budgetary consequences | • Fees for external auditors |
| • Costs of public relations efforts | • Costs of public relations efforts |
| • Termination of business contract | • Terminations of business contract |

Of note, institutions face higher costs the first time a breach happens. Another interesting finding was the increased cost among faster responders, which was attributed to early, unnecessary notification of individuals whose records were thought to have been breached. Subsequent information showed that notification in many cases had been unnecessary.

Some of the costs in Ponemon’s model do not apply directly to the registry world, such as loss of existing or future customers. Economies of scale also need to be considered when interpreting these data.

The most recent survey estimates the total cost of data breaches as follows:

| Table 2. Estimated total cost of data breaches |
|-----------------------------|-----------------------------|
| Overall mean total costs | $214,049 |
| Institutions with a first data breach, total costs | $326,000 |
| Institutions who responded quickly (<30 days), total costs | $268,000 |
| Institutions who responded slowly (>30 days), total costs | $174,000 |
| Direct costs only | $73,000 |

Analyses

New Hampshire has a population of 1.2 million with approximately 7000 new cancer patients diagnosed annually [5]. Data have been collected since 1987. The following exercises assume loss of unencrypted personal health data. The mitigating effects of encryption and other security measures that we use in NH are discussed later.

Exercise 1: Direct Costs

**Assumptions:**

Unique individuals’ records in the database

Individuals presumed to be alive

Direct costs of a breach, per individual [4] | $73

Using Ponemon’s data, the estimated direct costs of a breach is $73 per compromised record. This gives estimated direct costs of a breach as follows:

Exercise 2: Fines

Like many cancer registries we collect cases on out of state residents, which in New Hampshire comprise about 14% of unique individuals in our database. A few states have implemented computer breach and notification laws that may come into play in the event of a breach. For example, Massachusetts can impose fines of $5000 per case [6].

**Assumptions:**

MA records in our database

Fine per breached record | $5,000

Under these assumptions, a breach might cost New Hampshire $8,025,000 in fines for Massachusetts alone, which would be added to the costs estimated in Exercise 1 for a total of over $15.6 million.

Exercise 3: Risk Modeling

Data breach liability calculators developed by interested parties are available on the internet, for example, by Symantec in conjunction with the Ponemon Institute [8], and Tech/404 [9]. Again using 155,693 unique persons, the latter calculator gives some rough estimates of potential loss. A subtotal removes regulatory investigation defense and state/federal fines or fees.

We note the potential for over-estimation by modeling software products developed by computer security or liability insurance companies, as well as difficulties applying models developed for the commercial setting in a health care setting. These factors should be considered in the interpretation.

**Table 3. Tech/404 cost estimator for a data breach [9]**

Exercise 4: Estimated total costs of data breaches

Exercise 5: Discussion

If commercial models can be applied to the registry setting, direct costs of over $7 million might result from a data breach with loss of unencrypted personal health data at a small central registry. Larger registries that have been operating for longer would have substantially larger potential exposures. These estimates should be interpreted with caution as they were based on commercial modeling and do not allow for economies of scale in estimating costs of a response protocol.

In practice, the true cost of a breach will be lower if better security measures are in place. Put another way, improvements in security may change the definition of a true breach. For example, loss of an encrypted laptop with two-factor authentication, such as those used in New Hampshire, would pose a very small risk of access to personal data, and notification of individuals may not be necessary. In contrast, unencrypted data on a lost laptop protected by a simple password is essentially unprotected. Tangible and intangible costs could be massively reduced if only encrypted, inaccessible data were lost.

We believe, given all the procedures and practices we have implemented and our ongoing scrutiny to optimize security, our potential risk is relatively low. We are discussing ways to optimize data security in New Hampshire even further.

A data breach at any cancer registry carries a large potential exposure for intangible costs both locally and nationally and would have serious repercussions for other registry and public health activities across the country.

**References**

3. [3] Ponemon Institute
5. [5] Personal health information, Sals Cherolia, MD, NH Department of Health and Human Services

This project was supported in part by the Centers for Disease Control and Prevention’s Program of Cancer Registries, cooperative agreement US/DFP00798 awarded to the New Hampshire Department of Health and Human Services, Division of Public Health Services, Bureau of Public Health Statistics and Informatics, Office of Health Statistics and Data Management. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the Centers for Disease Control and Prevention or New Hampshire Department of Health and Human Services.