Medical Xpress

Anonymization remains a powerful approach to protecting the privacy of health information

December 8 2011

De-identification of health data has been crucial for all types of health research, but recent articles in medical and scientific literature have suggested that de-identification methods do not sufficiently protect the identities of individuals and can be easily reversed. A recent review conducted by researchers at CHEO entitled "A Systematic Review of Re-identification Attacks on Health Data" and published in *PLoS ONE*, did not uncover evidence to support this. "If re-identification rates were as high as some of these articles suggest, it would be worrisome," says lead author, Dr. Khaled El-Emam. "But our review did not support these claims – there is no broad empirical support for a failure of anonymization."

Such a failure would have significant policy implications. For example, it may become necessary to obtain patient consent before data is released (a time-consuming undertaking), incentive to de-identify would decline, and the likelihood of breaches would increase. For this reason, Dr. El-Emam and his team conducted a review that set out to characterize known re-identification attacks on health data and compare them to attacks on other types of data, calculate the number of records correctly identified in these attacks, and assess whether the results indicate a weakness in current de-identification methods.

After identifying 14 relevant studies and analyzing them in detail, the group was unable to find convincing evidence that existing de-identification methods are not effective. Few of these attacks involved health data which is naturally protected more strenuously. Secondly, many of the attacks were on small databases with large confidence intervals around their success rates. Most importantly, the majority of re-identified data was not de-identified according to existing standards. "Of the 24 studies we examined, only six were attacks on health data and only one of these was de-identified according to standards," Dr. El-Emam points out. "In that particular study, the proportion of correctly re-identified records was very low: about 0.013%."

In certain well-publicized re-identification attacks, adversaries were able to make use of such information as an individual's date of birth, gender, and residential zip code. Since these 3 features were not modified in any way, the database would not meet basic standards for de-identification. If anything, such a breach serves to underscore the importance of implementing existing de-identification standards.

Dr. El-Emam concludes by saying that in order to have a more accurate picture of the extent to which de-identification protects against real attacks, future research on re-identification attacks should focus on large databases that have been de-identified according to existing standards, and that success rates should be correlated with how well de-identification was performed. In the meantime, it is suggested that data custodians continue to de-identify using current best practices.

More information: Link to report: www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0028071

Provided by Children's Hospital of Eastern Ontario Research Institute

Citation: Anonymization remains a powerful approach to protecting the privacy of health

information (2011, December 8) retrieved 2 May 2024 from
https://medicalxpress.com/news/2011-12-anonymization-powerful-approach-privacy-health.html