

Leveraging the Value in Health Data

In order to share health data to improve health outcomes, conduct surveillance, and evaluate health costs data owners often have to de-identify the data—whether it is claims and pharmacy data; clinical data in the form of EHRs; images; clinical trials; or data from registries. Because most de-identification processes remove useful information from the data it is a challenge to de-identify data and retain analytic utility, while still being compliant with HIPAA and other legislative requirements.



Common De-identification Challenges

Lack of utility. Commonly used de-identification methods such as HIPAA's Safe Harbor can eliminate too much valuable data that have analytic utility

Time Consuming and Error Prone. Manual processes involved in disclosure review of output data are time consuming, labor intensive, and typically lack rigorous and comprehensive quality control

Abundance of Individual Data. Exponential growth of available data on individuals makes frequently employed de-identification methods subject to re-identification disclosure risks

Shortage of Disclosure Experts. There are not enough experienced statistical disclosure limitation experts to meet demand

Disclosure Risk. Can lead to an inability to fully leverage the value of the data (monetary, research value), brand damage, and financial penalties

NORC at the University of Chicago is an independent research organization headquartered in downtown Chicago with additional offices on the University of Chicago's campus, the D.C. Metro area, Atlanta, Boston, and San Francisco. NORC also supports a nationwide field staff as well as international research operations. With clients throughout the world, NORC collaborates with government agencies, foundations, educational institutions, nonprofit organizations, and businesses to provide data and analysis that support informed decision making in key areas including health, education, economics, crime, justice, energy, security, and the environment. NORC's decades of leadership and experience in data collection, analysis, and dissemination—coupled with deep subject matter expertise—provides the foundation for effective solutions.

Contact

Jeffrey Hackett | Director, Business Development
(312) 759- 4266 | hackett-jeffrey@norc.org

Learn more about NORC and X-ID at www.norc.org



The NORC X-ID System is a revolutionary data de-identification process that allows clients to optimize data utility while explicitly specifying the level of disclosure risk.

The Solution: NORC X-ID produces data output that retains a high degree of analytic utility and is compliant with HIPAA and other legislative requirements.

NORC has **developed a methodology**, using proprietary statistical tools, to support an **organization's initiatives to de-identify** its protected health information. X-ID tools give data owners the **ability to define their desired risk and analytic utility thresholds**. The NORC X-ID system **enables data owners to customize the type of data they manage and deliver** to meet specific business and research needs, which makes for **higher quality** research and more valuable outcomes.



EXPERT knowledge

Unlike other systems, the X-ID process relies on NORC's renowned team of statistical experts, with proven success in disclosure limitation and expert determination. Using innovative approaches, X-ID incorporates NORC's statistical expertise into de-identification solutions that allow for the optimal balance of analytic utility and disclosure risk.

Through its work on large and complex projects for clients such as the **Centers for Medicare & Medicaid Services**, the **Centers for Disease Control and Prevention**, the **Federal Reserve Board**, the **Bureau of Labor Statistics**, and others, NORC has built a team of professional disclosure experts who use scientific principles and statistical methods to assign suitable measures of disclosure risk and data utility. The NORC X-ID experts' evaluations will meet the requirements of your organization's Compliance Office.

EXTREME utility

Compared to approaches with limited simultaneous control on risk and utility, NORC's X-ID de-identifies data via two different proprietary advanced statistical methods that yield highly usable and valuable data in a compressed timeframe. Working within the parameters dictated by the client's disclosure needs and limitations, both methods maintain important dependencies within the data and keep the data's properties as intact as possible, resulting in data sets with exceedingly high utility.

With X-ID, data users will be able to reliably estimate many quantities on the de-identified data including:

- Means, totals, and proportions
- Variances and co-variances
- Correlations and odds ratios
- Regression models
- Quality measures

EXCEPTIONAL confidence

NORC's X-ID gives you granular control over appropriate levels of risk for the dataset in question, ensuring that you meet your data goals and are also in compliance with any target regulations like HIPAA and other key guidelines.

Input De-identification

NORC X-ID uses micro-group level transformation of data to eliminate the risk of re-identification associated with using unit-level data. The X-ID system retains data value and controls disclosure risk by leveraging sophisticated survey sampling methods.

Output De-identification

Using pre-specified analysis domain level restrictions, NORC X-ID automatically de-identifies the output of queries or analysis applied to sensitive data residing behind a secure firewall and protects output from differencing attacks.