

解析能力を損なわない共有疫学データのプライバシー保護[§]

Protecting Privacy of Shared Epidemiologic Data without Compromising Analysis Potential

John B Cologne Eric J Grant 中島栄二 Yun Chen 船本幸代 片山博昭

要約

目的 疫学データを外部の研究協力者と共有する場合に研究対象者のプライバシーを保証するため、データのマスクング(修正)を行うが、過度のマスクングは実用性(解析能力)を損なう可能性がある。プライバシー保護を目的とする統計データの開示制御の方法は、小規模な共同研究に携わる個々の研究者にとっては実用的とは言えない。

方法 我々は、疫学研究者が簡単に導き出すことができる、開示リスクと解析の実用性との比較尺度に基づく簡単な方法を調べた。日本人原爆被爆者集団から得られたデータを用いて、この方法を説明する。

結果 データを適度に丸めることによるマスクングでは安全性を十分に高めることはできなかったが、相対精度で数桁に丸めると、実用性が大幅に低下することなく、個人が特定されるリスクが効率的に減った。グループ化またはランダムノイズを加えると顕著なバイアスが生じた。

結論 疫学データを共有する場合、丸めによるデータのマスクングを行うことを勧める。具体的な処理については、開示リスクと解析上の必要性を考慮した上で個々の状況に合わせて個別に決定すべきである。

[§] 本報告書は *J Environ Public Health* Volume 2012, Article ID 421989, 9 pages, 2012 (doi: 10.1155/2012/421989) に掲載されたものであり、その正文は同掲載論文のテキスト(英文)である。この日本語要約は、日本の読者の便宜のために放影研が作成したが、本報告書を引用し、またはその他の方法で使用するとき、同掲載論文のテキスト(英文)によるべきである。