ORTHO MEDICO
Industry News Topics

〒112-0002
東京都文京区小石川1丁目4番1号
住友不動産後楽園ビル2階

2020年10月

データの質について【第29回生物統計学】

1　概要

データは記録し、蓄積していくだけでなく、他のデータと組み合わせることや加工することでその質を向上させ、利用方法などを増やすことができます。しかし時間の経過や記録、加工、利用など元データに手が加わるほど入力ミスや重複などのエラーが起こる確率も高まるため、データの質を低下させることも起こりえます。

2　品質基準となる項目

評価項目	概要
正確性（Accuracy）	データが構文的、意味的に正しいかの度合い データに誤字、脱字がないか等
信頼性（Reliability）	データを利用する際の、そのデータの信用度 再現可能か、信頼できるツールで生成されたものか等
妥当性（Validity）	データ間に矛盾がなく、全体で整合がとれているかの度合い
完全性（Completeness）	利用目的のために収集した全てのレコードを保持しているかの度合い
網羅性 （Comprehensiveness and coverage）	意図している集団を100%カバーしているか
匿名性（Anonymity）	承認された条件下でのみ利用可能であることの保証度
結合可能性（Linkability）	ほかのデータと連結可能か
適時性（Timeliness）	収集期間と分析実施期間が重なっていないか
利用可能性（Usability）	適切に整理され、アクセス可能であり、使い易いフォーマットであるか
異時点間の一貫性 （Temporal consistency）	長期間にわたる分析に耐えうるように標準化がなされているか

3　品質の評価

データマネジメントについての定義やガイドラインをまとめる国際データマネジメント協会が策定したデータマネジメント知識体系ガイドによると品質の高いデータとは、「ビジネスの目的のためにデータが利用に適した状態」であると定義されています。

ORTHO MEDICO

お問い合わせ先　株式会社オルトメディコ　企画部
〒112-0002 東京都文京区小石川1丁目4番1号　住友不動産後楽園ビル2階
TEL: 03-3818-0610　FAX: 03-3818-0617　Mail: planning-department@orthomedico.jp

〒112-0002
東京都文京区小石川1丁目4番1号
住友不動産後楽園ビル2階

データの質を向上させるには正確性や信頼性などの項目全てを高い水準にすることが望ましいですが全て等しく高水準とすることは困難となります。またデータは相対的に評価されるものなので、同じデータにおいてもどのような目的でそのデータを利用するかによってデータの評価は変わります。
そのため目的に沿った一定以上の品質は必要になりますが、目的にそぐわない項目で過剰な品質を追及すると品質維持のためのコストの増大など長期間の継続利用が困難になるため、データの質の評価は目的によって異なりどの項目を重視するのかも変わることになります。

お問い合わせ先　株式会社オルトメディコ　企画部
〒112-0002 東京都文京区小石川1丁目4番1号　住友不動産後楽園ビル2階
TEL: 03-3818-0610　FAX: 03-3818-0617　Mail: planning-department@orthomedico.jp