What Do/Does the Data Show?

Stacy L Christiansen

You’re editing a document when you come across this sentence: “Follow-up data from a large sample was used to estimate the incidence of carcinoma.” Sounds good. Right? Wait, should it be “Follow-up data from a large sample were…” instead? Is the word data a singular or plural noun? Yes.

There are generally 2 approaches when it comes to words such as data and bacteria, common terms that are of Latin origin and appear often in scientific documents. One is more of a “purist” approach, treating these words as the plural nouns they are in Latin (singular forms would be datum and bacterium). The other is a more contemporary-language approach, recognizing that modern English speakers would rarely if ever use the word datum in common parlance, and perhaps not even in formal scientific writing.

So which is it? Well, reference sources themselves vary in their recommendations. The Plurals chapter (chapter 9) in the AMA Manual of Style, 11th edition, notes: “A few nouns are usually used in the plural form; however, the distinction between plural and singular should be retained where appropriate” and gives the following examples: data/datum, criteria/criterion, media/medium, and phenomena/phenomenon.1 Ok makes sense: “The data are what they are.”

But wait, there is a footnote on that page:

Exception: when referring to social media, news media, or the media, use a singular verb. The same applies when referring to big data as a term for extremely large, often unstructured data sets that can be mined for business or social uses.

So if you follow AMA style, data would indeed be considered a plural in most contexts, with the caveat above. The same is true for those who follow APA style, which also recommends observing the distinction between singular and plural forms.2

In checking with Merriam-Webster’s dictionary regarding the term data, the ambivalence in usage is pronounced: both singular and plural constructions are considered standard.3 I note with slight amusement the last sentence of the dictionary’s usage guide: “The plural construction is more common in print, evidently because the house style of several publishers mandates it.”

The AP Stylebook echoes the dual approaches: data as singular for lay audiences and plural for scientific and academic writing.4 The Chicago Manual of Style, 17th edition, also acknowledges both approaches, but similar to those mentioned above, points out that in the sciences and other formal contexts, the term data is usually plural.5

A variety of other blog entries, podcasts, and articles on grammar and usage echo these sentiments; essentially, the plural construction of data is still widely used in scientific communication. The singular form is embraced for most other contexts. In this usage, data is thought of as a collective noun and, when considered as a unit rather than as the individual items of data that compose it, it takes the singular verb.

But whichever approach you take, heed the Chicago Manual’s recommendation: “[M]ake your play and be consistent—vacillating will not win the admiration of readers and listeners.” Which of course is sound advice for any editing decision.

References and Links