



Ministry of Defence

Owner: MOD Chief Statistician

Author: Defence Statistics

Issued: October 2016

Ministry of Defence Disclosure Control and Rounding Policy

Purpose

1. This document sets out Ministry of Defence (MOD) policy on the application of methods of disclosure control and of rounding to statistics produced for external publication.
2. In this new policy, MOD is moving to disclosure control on a case-by-case basis, enabling a larger amount of data to be released unrounded to better meet users' needs.
3. Disclosure control is necessary when presenting any statistical or numeric information to safeguard the confidentiality of individuals, to protect commercially sensitive information, or where sharing the statistics may pose a threat to security.
4. Rounding is used both as a means of disclosure control and to improve the clarity of outputs and convey appropriate levels of precision to users.

Definitions

5. For the purpose of this policy, the following definitions are used.
6. **Statistics** are those numbers generated by performing mathematical or statistical operations on a set of raw data (including counts).
7. **Personally identifiable information (PII)** is any data that could potentially identify a specific individual. Any information that can be used to distinguish one person from another and can be used for de-anonymising anonymous data can be considered PII.
8. **Disclosure** means to permit access to or the release, transfer, or other communication of personally identifiable information by any means.
9. **Disclosure risk** is a function of likelihood (related to the presentation of data), and impact of disclosure (related to the nature of the underlying data).
10. **Disclosure control** or **disclosure avoidance** refers to the efforts made to reduce the risk of disclosure, such as applying statistical methods to protect PII in aggregate data tables. These safeguards often referred to as disclosure control or avoidance methods, can take many forms (e.g., data suppression, rounding, recoding, etc.).

11. **Unsafe cells** are those parts of the statistical table that could lead to disclosure, commonly, cells containing small counts.

Scope

12. This policy applies to all MOD statistics produced for external publication.

13. Exemptions to disclosure control include:

- a. When the data do not refer or relate to an individual, business or organisation, for example animals or public buildings.
- b. When permission to release confidential information has been given by the individual, business or organisation.

Context

14. This policy has been developed in accordance with the UK Statistics Authority's Code of Practice for Official Statistics.

15. Principle 5 of the Code of Practice guarantees confidentiality to those who provide private information for the production of official statistics: "Private information about individual persons (including bodies corporate) compiled in the production of official statistics is confidential, and should be used for statistical purposes only."

16. The National Statistics Code of Practice: Protocol on Data Access and Confidentiality, which underpins this guarantee, provides guidance on how the standards for protecting confidentiality should be set: "Statistical disclosure control methods may modify the data or the design of the statistics, or a combination of both. They will be judged sufficient when the guarantee of confidentiality can be maintained, taking account of information likely to be available to third parties, either from other sources or as previously released National Statistics outputs, against the following standard: It would take a disproportionate amount of time, effort and expertise for an intruder to identify a statistical unit to others, or to reveal information about that unit not already in the public domain".

Risk assessment for disclosure control

17. Disclosure control is necessary when presenting any statistical or numeric information to safeguard the confidentiality of individuals, to protect commercially sensitive information, or where sharing the statistics may pose a threat to security.

18. To ensure that no commercially sensitive information or information that may pose a threat to security is released, producers consult subject-matter experts within MOD and apply appropriate methods of disclosure control, including the withholding of data from publication if judged necessary.

19. In safeguarding the confidentiality of individuals, producers of statistics weigh the needs of users against disclosure risk. To determine the proportionate application of disclosure control methods for the safeguarding of individual confidentiality, producers of statistics:
- a. Identify the main users of the statistics, why they need the figures and how they will use them in detail. This is necessary to ensure that the design of the output is relevant and the amount of disclosure control used has the least possible adverse impact on the usefulness of the statistics.
 - b. Assess how the characteristics of the data will affect any disclosure risks. In particular, in relation to personally identifiable information, risk increases as statistics become more detailed (in terms of geography and categories) and as the dimensions of tabular data grow. Risks are higher if the distribution of the counts is skewed or the data are considered sensitive.
 - c. Assess the level of disclosure risk. Assessment of the likelihood of disclosure and its impact is made by those who have a detailed understanding of the statistics and experience of the interest in the figures. This involves identifying situations where there is a likelihood of disclosure. Where a risk is identified, it is necessary to establish whether any disclosure would constitute a breach of public trust, of a legal obligation, or of a national or international policy standard for official statistics.
 - d. Apply disclosure control methods if judged necessary. If a breach is thought to be likely, i.e. where cells have been identified as 'unsafe', disclosure control methods are used to manage the risk effectively. The various methods have different advantages and disadvantages and are chosen bearing in mind users, uses and characteristics of the data. The most important consideration is maintaining confidentiality but these decisions will also accommodate the need for clear, consistent and practical solutions that can be implemented within a reasonable time and using available resource. The methods used will balance the loss of information against the likelihood of individuals' information being disclosed.

Methods of Disclosure Control

20. In many instances, the risk of disclosure will be minimal and no disclosure control methods necessary. If there are no concerns with regard to disclosure and the protection of confidentiality, it is acceptable to provide numbers in their unsuppressed and unrounded format if that is what is required by the user.
21. If risk exists and disclosure control considered necessary, one or more methods of disclosure control are applied. These methods include:
- a. Table re-design. Unsafe cells are disguised by, for example, grouping categories within a table; aggregating data to a higher level of geography or for a larger population sub-group; or, aggregating tables across a number of years/months/quarters.
 - b. Modification of cell values:
 - i. Cell suppression. Unsafe cells are suppressed and replaced by a special character, such as ~, to indicate a suppressed value.

- ii. Rounding. Rounding involves adjusting the values in all cells in a table to a specified base. This creates uncertainty about the real value for any cell while adding a small but acceptable amount of distortion in the data.
 - iii. Barnardisation. Cells within a table are adjusted by +1, 0 or -1 based on probabilities ($p/2$, $1-p$, $p/2$), where p is set by the producer.
- c. Adjustment of the data. For example, pairs of records within a micro-dataset that are partially matched are swapped to alter the variable but all other aspects remain unchanged.

Rounding for other purposes

- 22. Figures, rates and percentages may be rounded beyond the requirements for disclosure control where there is good reason to do so.
- 23. Rounding can improve the clarity of tabulated data presented in a report. In many cases, the extra detail provided by unrounded figures is not needed, and may obscure the main features of the data.
- 24. Rounding can also avoid giving a false impression of the accuracy of estimates, survey data, or where data quality issues exist.
- 25. The appropriate form of rounding will depend on the type of statistic. The form of rounding applied to each type of statistic has been agreed following discussion with subject matter experts, and taking into account disclosure, precision, specific business requirements and practical implementation issues.

Implementation

- 26. This policy should apply to all new outputs. Regular publications will be brought into compliance with its provisions as and when resources permit. Until then regular publications will remain compliant with the previous Rounding Policy, which is available to users on request from DefStrat-Stat-Enquiries-Mailbox@mod.uk.
- 27. The decision to apply methods of disclosure control or to use rounding for other purposes will be conducted on a case-by-case basis by the producer.
- 28. MOD does not mandate the application of a particular method of disclosure to statistics produced for external publication, nor does it establish a particular threshold for what constitutes sufficient disclosure avoidance, nor does it mandate the application of a particular form of rounding for other purposes.
- 29. When looking at published statistics, users should be aware that the dataset has been assessed for disclosure risk, and methods of protection may have been applied. For quality purposes, users of a dataset will be provided with an indication of the nature and

extent of any modification due to the application of disclosure control methods, or due to rounding for other purposes.

Further Information

30. The Code of Practice for Official Statistics can be accessed at:

<http://www.statisticsauthority.gov.uk/assessment/code-of-practice/index.html>

31. The National Statistics Code of Practice: Protocol on Data Access and Confidentiality can be accessed at:

<http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/about-ons/who-we-are/services/unpublished-data/access-to-ons-data-service/access-arrangements/index.html>