Hospital Estates and Facilities Statistics 2009-10

Data Quality Statement

Background

Hospital estates and facilities statistics in England are compiled from data supplied by around 400 NHS organisations. The data is collected by the NHS Information Centre for Health and Social Care (NHS IC) under a service level agreement with the Department of Health. NHS IC liaises with NHS organisations and their staff to encourage complete data submission, and to minimise inaccuracies and the consequential impact of missing and invalid data.

A secure online system is used to collect the data, the Estates Returns Information Collection (ERIC) and is the only central source of such data in England. The ERIC system has seen improvements to its core IT functionality since its launch in 2001. These changes have presented opportunities to reduce the collection burden on NHS organisations, and improve the quality, timeliness and completeness of estates and facilities data.

The NHS IC and the Department of Health work in partnership to try to minimise inaccuracies and the effect of missing and invalid data. However, responsibility for data accuracy ultimately lies with the organisations providing the data. It is not changed in any way by the NHS IC or the Department of Health. Data handling processes and methodologies are continually being considered through meetings and update reports, and implemented to improve data quality.

Methods used to improve accuracy

The data is subjected to a number of validation processes to improve both quality and accuracy.

1. The online collection system, ERIC, has a number of embedded automated validation routines to process the data at the time it is entered by the providing organisation. These routines include using minimum and maximum values to create acceptable ranges within which each data value and performance indicator is expected to fall. If a value is entered which falls outside the acceptable range the data supplier is given an instantaneous on-screen prompt to

highlight the issue and must either confirm or change the value. Similarly, significant differences (an increase or decrease of 20%) between the previous year's data are also highlighted. Each trust is required to 'commit' their return, to confirm the dataset is complete, accepting all validation messages as part of the commit/validation process.

2. Once all data returns have been 'committed', a further validation process is run by the NHS IC to identify where values are still out of range or data is missing. A validation report is provided to each trust with questionable data highlighted, and their return is automatically 'uncommitted' by the system to allow changes to be made by the data supplier. The data supplier must then either change or confirm the data by 're-committing' the return.

3. The final validation routine undertaken is when the NHS IC and the Department of Health work in partnership to identify and challenge data suppliers where, despite the routines described in 1 and 2 above, the data still appears to be erroneous in comparison to historical patterns, peer organisations or other data sources. This process takes the form of a random manual check of the dataset to identify what would appear to be significant discrepancies. Where such discrepancies are identified, data suppliers are contacted individually and given the opportunity to either change or confirm their data.

Data quality

The responsibility for data accuracy ultimately lies with the organisations providing the data. The validation routines outlined above are designed to cleanse the data as much as possible within the resources available and without knowledge of local circumstances. When the data is published, or is used to provide information, for example, when answering or contributing to a Parliamentary Question, a clear statement is made that the data has not been amended centrally and its accuracy is the responsibility of the NHS organisations.

For the 2009/10 return, ERIC collected just over 110,000 separate data items from around 400 NHS organisations. Given the volume of data collected by the ERIC system for an estate which is the largest in Europe and worth £40 billion, and the level of detail which is captured, it is inevitable that anomalies or suspected anomalies can be discovered after

publication. These issues often come to light as the information is used, often at low levels of granularity, often over an elapsed time of months and years since publication. As a general rule, once the data has been published it will not be possible to change data entries on the online system, and data suppliers are urged to correct any erroneous data in their next year's return.

Data relevance

The relevance of Hospital Estates and Facilities Statistics is maintained by reference to appropriate stakeholders and Policy groups who oversee data definitions and standards.

The new NHS landscape as set out in the recent NHS White Paper *Equality and Excellence: Liberating the NHS* will require appropriate information on estates and facilities to support it. This will bring about a change to the dataset that is currently collected and the organisations who supply it. These changes will be necessary to re-orientate data collections around the needs of the new landscape.

The ERIC data collection system is licensed each year by ROCR (Review of Central Returns) who are concerned with implementing the Government's policy in reducing the burden of data collections from the NHS.