

PSEUDONYMISATION WORKSHOP

22nd SEPTEMBER 2011

PSEUDONYMISATION FROM THE REC AND PATIENT PERSPECTIVE

One of the functions of a Research Ethics Committee (REC) is to ensure that participants' privacy and confidentiality is protected.

I have seen very many Participant Information Sheets (PIS) that blithely state that the participants identity or personal information will not be disclosed or recorded (the word anonymised is often used) when the researchers will be using personally identifiable data (PID) in the research records. There seems to be some confusion amongst many researchers as to what constitutes anonymisation/pseudonymisation and what data is PID.

The NHS Information Centre (NHSIC) which administers the Hospital Episode Statistics (HES) database, the largest of its kind in the world, lists the following as "Sensitive and Patient-Identifiable Data":

Name	Not kept by HES
Address	Not kept by HES
Postcode	
Date of Birth	
Legal Category and Group of Patient	
NHS Number	
Local Patient Identifier (Hospital Number)	
Consultant Code	
Referring GP Code	
Person Referring Patient	
Birth Date of Baby	Maternity Data
Birth Date of Mother	Maternity Data
Detention Category	Psychiatric
Legal Group of Patient	Psychiatric
Legal Status Classification	Psychiatric

Researchers frequently state that they are using "anonymised data" when they are including NHS Number and/or Hospital Number in the patient record identifier.

Postcodes are often collected from patient records as the researcher will want to obtain deprivation scores which can only be obtained from full postcodes. There are a number of systems whereby the postcode can pass through a lookup table to obtain a deprivation score and is then deleted.

The only really satisfactory method of pseudonymisation from the patient perspective is for a "code" to be allocated to each patient and a reference list (code = patient ID) kept by someone who has the legal authority to access that particular patient's data.

Terence W. Wiseman