

## 5 METHODOLOGY

### 5.1 *Organisational aspects*

Funding for the project was approved on 10 March 2005. In the original protocol, approval was anticipated for the pilot study in December 2004 and for the main study in January 2005. There was therefore a slight delay in the recruitment process and the start of the pilot survey.

A pilot point prevalence study was undertaken in three acute hospitals between May and August 2005. These hospitals included a large teaching hospital, a large district general and a small district general hospital in three different NHS board areas in three geographical locations (East, West and South of Scotland), whose MRSA bacteraemia rates, as reported through the HPS mandatory surveillance programme represented the upper, lower and average of rates reported in Scotland. These were selected in order to test the methodology in a range of different hospital settings before undertaking the national prevalence survey. The results of the pilot study were reported in September 2005 (4). As a result of the pilot study it was concluded that the plans and methodology for the main study were feasible with minor refinements to the protocol and plans.

A Project Team consisting of a project manager, project administrator, data manager and initially four data collectors was recruited by HPS to work full time on the project. The team was overseen by a Project Steering Group consisting of the project director, project consultants, SEHD, public and key stakeholders from the NHS boards, (Appendix Table 12-3 page 236).

A letter from the CNO to all Chief Executives, Medical Directors, Nursing Directors and Directors of Public Health informed all hospitals in Scotland that 'as part of the Ministerial HAI strategy in 2005, SEHD has commissioned Health Protection Scotland (HPS) to carry out a national prevalence survey of HAI' (30). In this letter a request was made to the Caldicott Guardians of each hospital for their permission for the data collectors from HPS to access medical notes. By June 2005 signed approval had been obtained from all the eligible hospitals.

A Data Collection Protocol for use by the data collectors was prepared and tested during the pilot survey. The hospitals surveyed in the pilot survey were re-visited in the main survey in order to ensure consistency in the data collection methodology for all hospitals.

Intensive training sessions were held for data collectors at which the rationale for the survey and for the methodology was discussed. Due to the importance of using consistent HAI case definitions throughout the survey, several training sessions for the data collectors used case studies (31) (provided courtesy of Petra Gastmeier of the KISS Project, Germany and the SSHAIP team at HPS) for training in the diagnosis of HAI according to the CDC definitions. Training in the use of the data collection tool was also undertaken. During the training of data collectors for the pilot study the need for more detailed understanding of microbiology reports and definitions of surgical procedures was identified. This was addressed during further training of these four and three additional data collectors before the main survey.

An Information Pack was prepared and sent to the nominated link member of the infection control team (ICT) at each hospital being surveyed. The pack included a handout for hospital staff that outlined the rationale, methods and implications of the survey and the standard HPS

patient information leaflets used to inform patients about their rights and how their personal information is protected, posters were also supplied to the hospital giving the basic methods of the survey and photographs of the data collectors. The poster contained a space where the date of the team visits could be entered. Information packs and posters were distributed by the nominated link contact to individual ward staff who were thereby prepared and informed in advance of the arrival on the ward of external data collectors.

At regular meetings of data collectors and the project manager during data collection in the pilot survey, practical problems in data collection, data entry and HAI diagnosis were reported and discussed. Two detailed and numbered Issues Logs; one relating to Data Definitions and practical (non-IT) issues in data collection and the other to the 'Data Collection Tool (IT)' were kept by the Project Manager and updated at each meeting. These lists were used as an agenda for the meetings and used to record conclusions and refinements that were included in the main study plans and protocol. Issues lists were maintained throughout the main survey.

Shortly after the initiation of the Scottish National HAI Prevalence Survey, the Department of Health (DoH) in England commissioned the Hospital Infection Society (HIS) in collaboration with Infection Control Nurse Association (ICNA) to carry out a prevalence survey in England. The Departments of Health in Wales, Northern Ireland and the Republic of Ireland also expressed an interest in collaborating, and HIS invited the Scottish Survey Project Director to join the UK HIS prevalence steering group in order to ensure that the prevalence surveys were carried out using a similar methodology as far as was possible.

## 5.2 Study design

### 5.2.1 Definition of acute hospitals

These were defined as per Information and Statistics Division (ISD) classification of hospital type. Hospitals in Scotland were classified as acute hospitals and non-acute hospitals. Acute hospitals were defined using the classification proposed by ISD (32). 'Acute hospitals provide a wide range of specialist care and treatment for patients. Typically, services offered in the NHS acute sector are diverse. They include: consultation with specialist clinicians (consultants, nurses, dieticians, physiotherapists and a wide range of other professionals); emergency treatment following accidents; routine, complex and life saving surgery; specialist diagnostic procedures; and close observation and short-term care of patients with worrying health symptoms' (32). A full list of acute hospitals<sup>1</sup> in Scotland is listed in Appendix Table 4-3 page 178.

---

<sup>1</sup> These hospitals are ISD main category A sub-category 1 to 3

## 5.2.2 Definition of non-acute hospitals

Non-acute hospitals are hospitals which offer long term care, for psychiatric, elderly or community patients. The majority of their inpatients are cared for within the specialties of Care of the Elderly and Psychiatry (32). A full list of non-acute hospitals<sup>1</sup> is included in Appendix Table 4-4 page 181.

## 5.2.3 Structure of the survey

The survey consisted of two parts: the 'prevalence survey' (Figure 5-1) of all patients which involved collection of a limited data set and the 'burden study' in which more detailed data were collected so that the burden of HAI in Scotland could be estimated in terms of health service utilisation and costs. A sample of 25% of inpatients was included in the burden study. Inpatients were allocated to the burden or prevalence survey in ward units.

Detailed data were collected from inpatients included in the burden study including inpatients with and without a prevalent HAI. These data included detailed information on surgeries within the last year and prevalence of invasive devices used (Figure 5-2).

All inpatients with a prevalent HAI were included in the LOS analysis and therefore more detailed data was collected for them. Discharge information was collected for inpatients within the burden study and all inpatients with HAI (Figure 5-3). These data have been used to estimate the additional burden (bed days used and cost) of HAI. For a complete listing of the data collected for the prevalence and burden parts of the survey see Appendix Table 2-1 page 162.

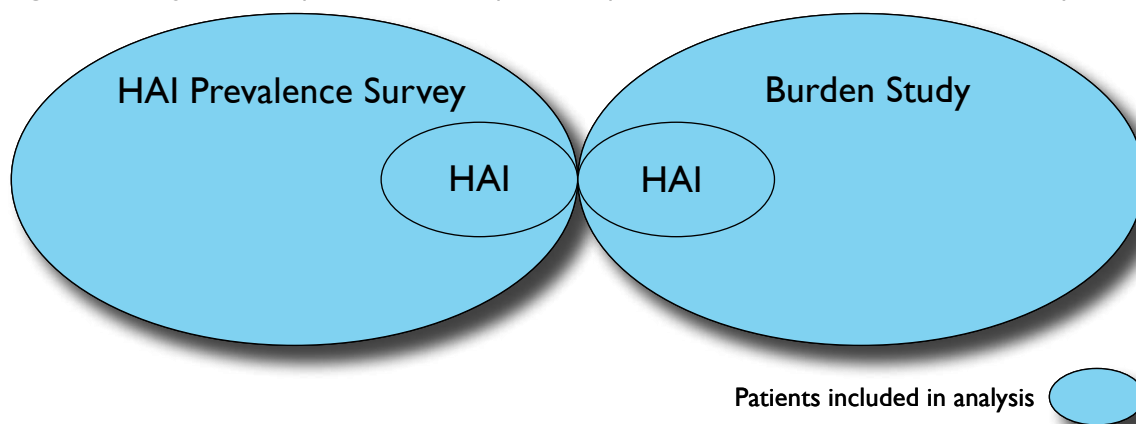
All eligible acute adult inpatient beds (a total of 11 608 patients) were surveyed in each hospital for the prevalence component of the survey, and a random sample of 25% of wards were included in the burden study. All eligible adult inpatient beds were surveyed in a sample of non-acute hospitals representative of Scottish NHS boards and hospital size (2146 patients in non-acute hospitals). The non-acute sample was included in the burden study, however the non-acute hospitals were not included in the additional LOS calculations. This decision was made due to the observed longer lengths of stay in the non-acute hospitals and it was decided that prevalence was not a sound indicator of additional LOS in non-acute hospitals.

For inpatients in the prevalence survey, detailed data were collected on inpatients with a HAI and a limited dataset on inpatients without HAI. These data were combined with those collected in the burden study to provide age/gender and specialty specific prevalence of inpatients with HAI for each hospital (Figure 5-1).

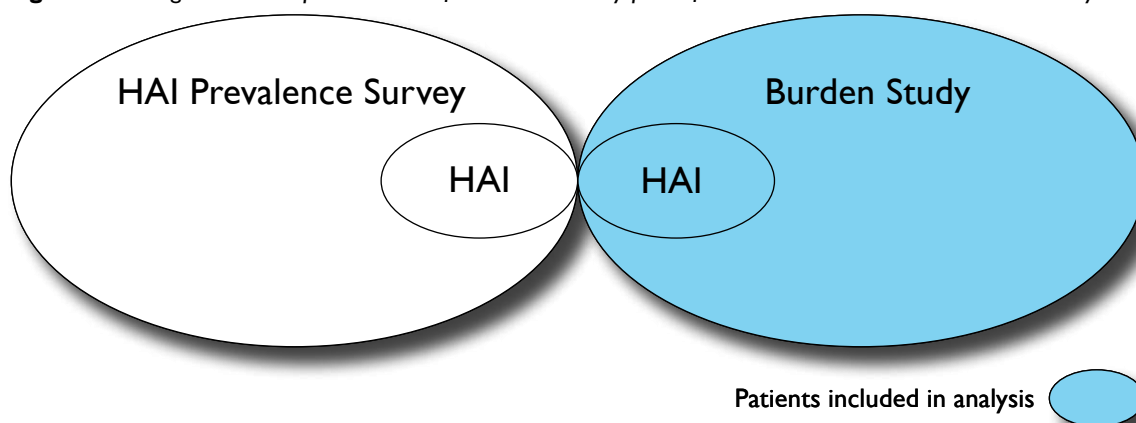
---

<sup>1</sup> These include ISD main category A sub-category 5 and main category B and C (with the exception of day care facilities)

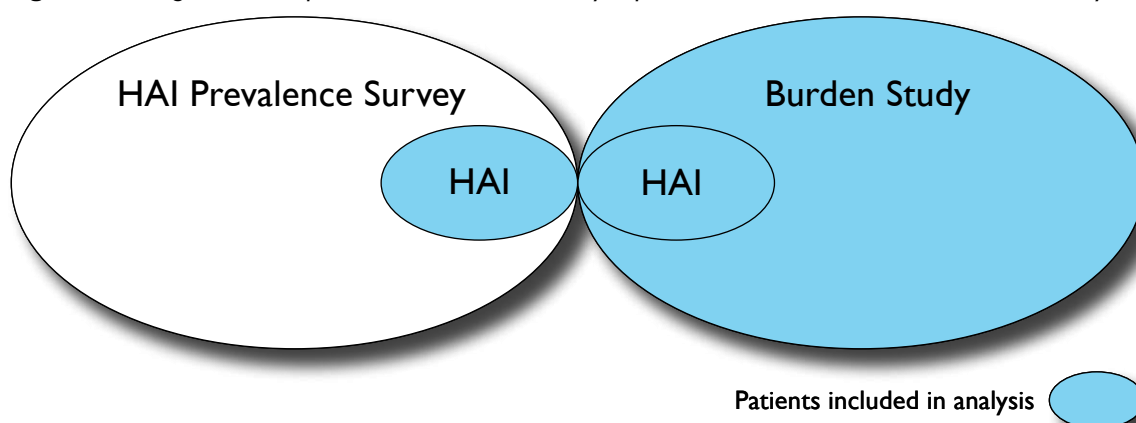
**Figure 5-1:** Diagrammatic representation of the prevalence part of the Scottish National Prevalence Survey



**Figure 5-2:** Diagrammatic representation of the burden study part of the Scottish National Prevalence Survey

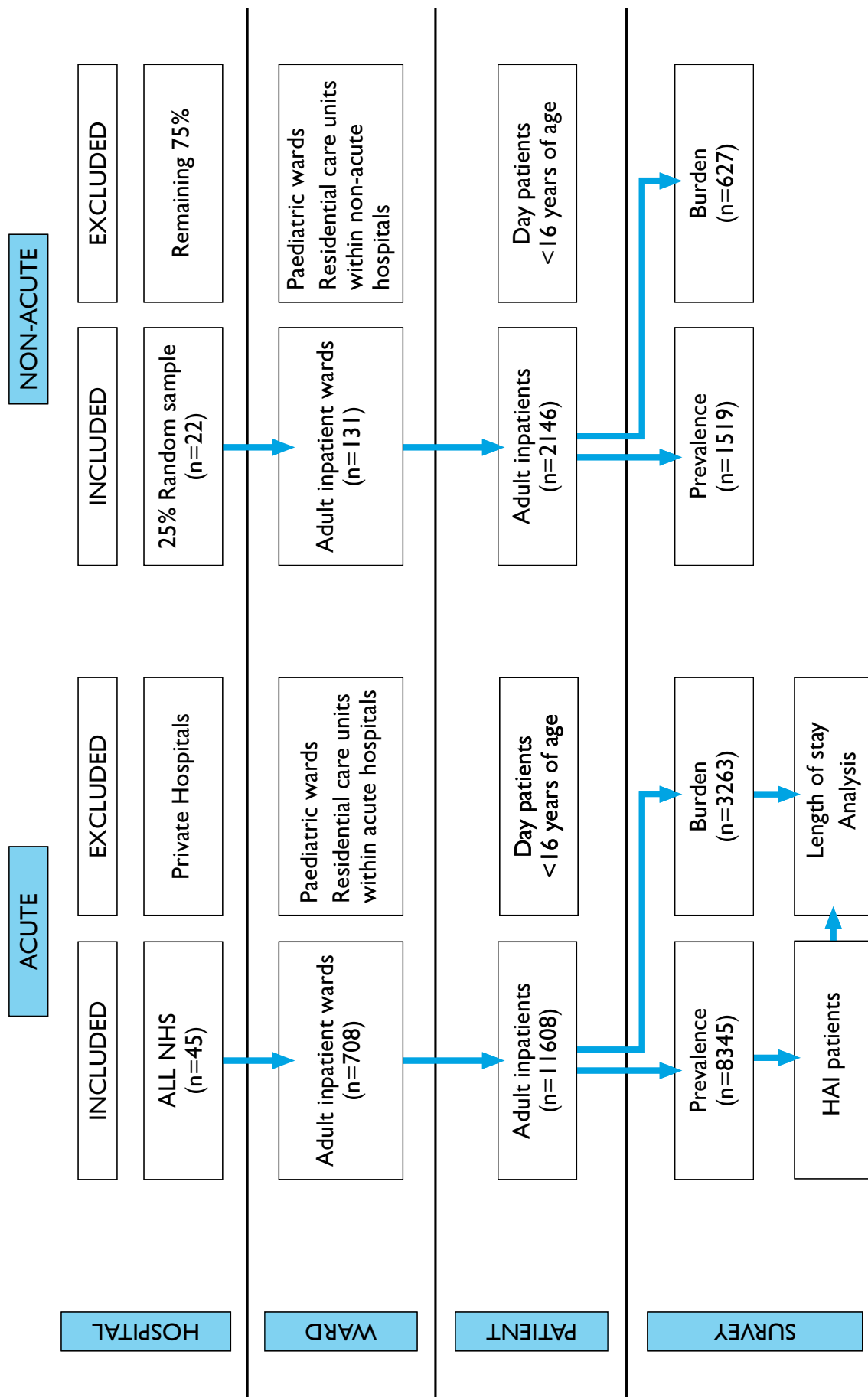


**Figure 5-3:** Diagrammatic representation of the LOS analysis part of the Scottish National Prevalence Survey



Additional work was undertaken to allow sampling of individual hospitals throughout the year of data collection. Hospitals were stratified into small, medium and large acute hospitals, obstetric and teaching hospitals. Detailed plans were made which distributed hospitals into one of four 3-month periods. Each stratum of hospital was represented equally in each three-month period (based on published bed numbers (33)). To do this, an assumption was made that hospitals of a similar size and type contain a similar specialty mix and inpatients with similar case mix. This allowed investigation of a possible seasonal effect on prevalence of HAI, an aspect of prevalence surveillance that has not been addressed previously.

**Figure 5-4:** Diagrammatic representation of Scottish national HAI prevalence survey showing inclusion and exclusion criteria for prevalence and burden parts of the survey.



## 5.2.4 Eligibility

Infection control contacts from eligible hospitals were asked to provide details of ward type, specialty and age of patients. The final decision on eligibility was made by the project team based on the ward type information supplied to the team by the link infection control nurse. Table 5-1 shows the survey inclusion and exclusion criteria. All patients who occupied beds in the selected wards at a pre-specified time of day were included in the survey and data collection. A record was made of the number of unoccupied beds and beds occupied by ineligible inpatients.

**Table 5-1:** Inclusion and exclusion criteria for National HAI prevalence survey

		Inclusion Criteria	Exclusion Criteria
Hospitals	Acute	All NHS	Independent (Private) Hospitals
	Non-acute	Selected Sample (25%). All included in possible sample	75% not selected
Wards		All wards serving adult inpatients ( $\geq 16$ years old) except those that meet the exclusion criteria	Wards serving paediatric ( $< 16$ years old) inpatients Residential care units within acute hospitals
Patients		All adult patients except those who meet the exclusion criteria	Day patients (Patients admitted for one day for treatment or for diagnostic procedures.) Inpatients ( $< 16$ years old)

## 5.2.5 HAI definitions

In this survey a HAI was an infection which arose  $\geq 48$  hours or more after admission to hospital and which was not present or incubating on admission. A prevalent HAI was considered present when the patient had signs and symptoms which met one of the CDC definitions, or had one or more signs or symptoms included in one of the CDC definitions and was being treated for the infection (with therapy). CDC's HAI case definitions (14) were adopted as these are widely used internationally. These definitions comprehensively categorise HAI according to the organ/tissue system affected.

This survey included every type of HAI which can occur and therefore examined the full spectrum of HAI in hospital inpatients which met the survey definitions. HAI are grouped into major CDC categories based on the main physiological systems and surgical interventions. (Appendix Table 3-4 Mapping of specific infection sites to high level HAI page 174 lists the specific HAI within each major category). These broad categories conceal the different types, numbers and severity of specific infections included within the major categories. Some major categories (e.g. Bloodstream Infections, Pneumonias, Surgical Site Infections) are more homogeneous than others (e.g. Eye, Ear, Nose, Throat or Mouth).

## 5.3 *Methods*

### 5.3.1 *Data collection on wards*

Data collection was undertaken on weekdays. All ward and patient data were entered onto a specially designed database held on a small portable 'tablet' personal computer (PC) while the data collectors were on the ward. All data collection on a ward was completed within one day.

Data collectors followed a standard procedure in their surveillance of a ward (See Figure 5-5) Local nominated link members of the HAI control team introduced the data collectors onto the wards. Prior to commencing the inpatient data collection, data on ward characteristics on the day of data collection (ward type, bed numbers, staff numbers and types) was collected with assistance from the nurse in charge.

The data collectors sought information on eligible inpatients from all relevant sources including case records, all results of special examinations including microbiology reports, X-ray reports, temperature charts, prescribing records, nursing notes and where necessary through discussion with clinical staff and by direct clinical observation. The design of the survey required the data collector to make an initial decision based on this information as to whether the inpatients showed signs of a specific HAI, criteria for which were included and accessible on the PC. They were required to check every sign and symptom included in the relevant CDC HAI definition which was met by a patient they had decided had an HAI. The decision as to the presence or absence of an HAI was that of the data collector. They were able to seek further help from epidemiology consultants at HPS if they had any remaining doubts about the diagnosis of an HAI according to the CDC definition.

In very rare instances eligible inpatients were omitted from the survey because both they and their clinical records were out of the ward for the entire time that the data collectors were present on the ward.

**Figure 5-5:** Standard procedure in data collectors' surveillance

### **Step 1**

- Introduction to ward staff
- Collect bed numbers occupied by inpatients/day patients/empty beds
- Collect staff numbers
- Ask about any standard antimicrobial prophylaxis regime

### **Step 2**

- Mini ward round. This gave an idea of how long the ward is likely to take to survey, look for initial proxy indicators, overall picture of the ward and case complexities etc.
- Review Inpatient Charts
- Review Drug Kardex, make notes on antimicrobials, date started
- Review temperature and wound charts
- Observe patients for invasive devices

### **Step 3**

- Detailed case note review
- Where details of care not clear ask medical or nursing staff

## **5.3.2 Data management**

Data were exported from each data collector's tablet PC on a weekly basis. The export procedure produced Microsoft Excel® files. These were subsequently imported into a Microsoft Access® database. Within the Microsoft Access® database algorithms were used to examine data consistency and validity. Algorithms were used to confirm that the criteria recorded met CDC HAI case definitions. Data quality and the performance of the data collection tool were monitored. A copy of the Data Management Standard Operating Procedure (SOP) is included in Volume 2.

After a delay of a minimum of two months, local nominated link persons at the hospitals were sent a list of selected patient identifiers and were asked to supply the discharge dates of these patients. Data from each data collector were combined into a master Microsoft Access® database file and passed to the statistician. STATA® Version 9 software was used for these analyses. Data were entered in a standard manner as developed during the pilot study (34).

## **5.3.3 Validation**

Inter-Rater Reliability (IRR) validation exercises were undertaken on two occasions during the survey to measure the consistency of data collection between data collectors. A crossover study design was adopted, requiring a sample of patients to be surveyed by the whole data collection team over the course of a single day. While the overall level of IRR was reassuringly high for the selected data items, these exercises revealed limitations to the assessment methodology in a dynamic healthcare setting. The validation recorded a 100% agreement for diagnosis of HAI type.

Repeated testing of data collectors, using a library of replica case notes, drug Kardex and lab reports, was subsequently identified as a superior method for evaluating data collection quality without the problems associated with live, time sensitive patient records. A library of case notes has been developed for use in future HAI prevalence surveys and this will be appropriate for both training and ongoing data quality assessment. This approach will allow reliability and validity to be measured between future data collection team members (see Appendix section on Validation page 228).

### **5.3.4 Invasive device data collection**

Invasive device data was collected for inpatients within the burden study sample.

### **5.3.5 Surgery data collection**

Surgical procedures undergone in the year preceding the survey were collected for all burden study and all patients with HAI. Surgical Site Infections (SSI) with implants according to CDC HAI definitions occur within one year of the surgery, while surgical site infections in the absence of implants occur within 30 days of the surgical procedure.

The decision to collect surgical procedures undergone within one year was made during the pilot survey. It was found that infections following surgery without implants were prevalent for some time, and if only surgery within the last month were recorded, a number of procedures related to infections would be missed. Therefore it was agreed that procedures for the preceding year regardless of surgery type would be recorded.

These inpatients were part of the burden study and were all surveyed while admitted to acute hospitals. The data collection protocol permitted three implant procedures and three non-implant procedures to be recorded per inpatient in one year preceding the date of survey.

### **5.3.6 Length of Stay (LOS) data collection**

The survey collected the following data items which were suitable for LOS analysis: inpatient's age; inpatient's gender; type of hospital; size of hospital; specialty for patient; time of year when admitted (season); whether patient died; HAI status. These factors make useful proxies for the complex mix of factors which affect an inpatients' LOS in hospital.

### **5.3.7 Prevalence and incidence data collection**

Data from the HAI Prevalence survey (all acute, burden study inpatients) were compared to the SSHAIP SSI Incidence Surveillance programme (all procedures between 1 October 2005 and 31 of September 2006). Within the National HAI prevalence survey, burden study surgery types were mapped to same those collected by the SSHAIP SSI surveillance programme.

## 5.4 *Statistical analysis*

The data collected in this study have a hierarchical structure. Patients are in wards which are themselves in hospitals. Multilevel models recognise that individuals are not independent of each other e.g. patients within a ward may be more alike than patients sampled randomly from within a hospital. Traditional multiple regression techniques treat the patients as independent observations. A consequence of this is that standard errors of regression coefficients are underestimated and may lead to an overstatement of statistical significance (35).

Regression analysis was carried out in STATA® using the GLLAMM procedure. The survey included 13 754 inpatients within 839 wards within a total of 67 hospitals. These analyses show that the ward level had a much greater effect on HAI prevalence than the hospital level. As a result of these analyses it was decided that all subsequent analyses of HAI prevalence should allow for clustering at ward level but not at hospital level.

### 5.4.1 *Prevalence calculations*

Prevalence was calculated as the total number of HAI patients divided by the total number of inpatients. Prevalence was calculated for both acute and non-acute hospitals, then prevalence was calculated by age category, gender, hospital type, hospital size and ward type.

### 5.4.2 *95% Confidence intervals*

Statistical analyses were carried out using STATA® software, specifically the SVY: MEAN (36) procedure with ward as the primary sampling unit. This produces the slightly wider confidence intervals of HAI prevalence needed to allow for the clustering at ward level.

### 5.4.3 *Box plots*

Box plots were used to display values for LOS by specialty. The vertical line in the centre of the box represents the median value and the outer edges of the box refer to the quartiles. The dots outside of the box represent unusually large values for LOS.

### 5.4.4 *Funnel plots*

Adjusted prevalence values for HAI are displayed as funnel plots (37) (Figure 6-10 to Figure 6-21). These values are based on the output from the multivariate logistic regression analyses for acute and non-acute hospitals (Table 6-29 and Table 6-31). These analyses provide estimates of the probability of an HAI for each individual inpatient, which are then summed over all relevant inpatients to give the expected number of HAIs (E) for each hospital specialty.

The adjusted rate is calculated by the formula

$$\text{Adj}(P) = P*(O/E)$$

where O is the observed number of HAIs in each hospital specialty; E is the expected number of HAIs in each hospital specialty based on age, gender and time of year; P is the overall HAI prevalence rate for that specialty. The resulting adjusted HAI values take into account the effect of age, gender and time of year on HAI prevalence.

The results of the logistic regression indicated that separate funnel plots should be produced for each specialty for both acute and non-acute hospitals (Table 6-29 and Table 6-31). The plots show the adjusted prevalence of HAI for hospital specialties plotted against the number of patients on which the rate is based. The two funnels (one depicted by the dashed line and one by the solid line) on each plot indicate the 95% and 99% confidence limits (CL), calculated from confidence intervals throughout the range of values. Funnel plots have been produced for each specialty where patient numbers and HAI prevalence permits.

#### 5.4.5 Prevalence logistic regression analyses

The logistic analyses (both univariate and multivariate) were carried out in STATA® using the SVY: LOGISTIC procedure with ward as the primary sampling unit (36).

Acute and non-acute hospitals were analysed separately. The dependent variable was HAI status (yes/no). Several explanatory variables (and possible interactions) were investigated including: age category, gender, hospital size (small, medium, large), type of admission (planned or unplanned), hospital type (teaching, general, obstetric), calendar quarter and specialty of the consultant caring for the inpatient.

Choices between competing models were made on the basis of likelihood ratio tests for nested models or Akaike Information Criterion (AIC) for non-nested models (38). The Akaike Information Criterion (AIC) is a measure of the goodness of fit of a model and is an operational way of trading off the complexity of an estimated model against how well the model fits the data. The best model would normally have the lowest AIC value.

Likelihood is the probability that the observations could have occurred given that particular set of parameters. It is often expressed on the log scale (39).

Degrees of freedom is the number of independent units of information relevant to the estimation of the parameters in the model (39).

### 5.4.6 LOS regression analyses

LOS was ascertained for all eligible inpatients. The additional LOS due to HAI was estimated using a modelling approach taking age, gender, specialty and admission type (planned or unplanned) into account.

Date of discharge for patients who were discharged to 'Another Hospital' or 'Home/Care Home' or were 'Still in Hospital' or had 'Died' were used to calculate LOS as follows.

*Equation 1: Length of stay calculation*

LOS = Date of Discharge/Death\* - Date of Admission\*

\*from/to the hospital where the survey was carried out

Patients who were 'Still in Hospital' or for whom discharge status was 'Not known' were allocated a proxy date of discharge. Patients who were 'Still in Hospital' were given the last date of discharge known for patients from that hospital and patients whose discharge status was 'Not known' were given the census date as the date of discharge. Using these proxy dates, LOS was calculated as shown in Equation 1. The lengths of stay for these patients are unknown but are at least as long as the LOS calculated using the proxy dates. These patients are considered 'censored'.

These analyses were carried out on a subset of patients, those in the burden study and those patients with a HAI. The analyses were carried out in STATA® using the STREG procedure with ward as the primary sampling unit. A lognormal regression model was chosen (40). The lognormal distribution occurs when the log of x is normally distributed and is a good choice when analysing skewed data such as LOS. This method is suitable for censored observations (39).

Many explanatory variables (and possible interactions) were investigated including: age category, gender, hospital size, type of admission, calendar quarter, HAI type and specialty. Choices between competing models were made on the basis of likelihood ratio tests for nested models or AIC for non-nested models.

### 5.4.7 Kaplan Meier analysis

Kaplan Meier analyses were used to derive curves representing the estimated proportion of inpatients with and without HAI remaining in hospital as LOS increases. The method is suitable for censored observations.

### 5.4.8 Economic analysis

The additional cost of care of inpatients with HAI was estimated by applying a cost per additional day from the additional LOS of inpatients with HAI. These costs were based on local Scottish healthcare costs (41) and assumptions about relative components of cost as reported in the study by Plowman (3). Statistical analysis determined the additional LOS attributable to HAI.

These data were used to attach a monetary value to that resource use. The term 'value' is more appropriate than 'cost' because the nature of hospital costs is that they are largely fixed irrespective of patient numbers, at least in the short-term. For example, a fully staffed 24-bed ward might have 20 occupied beds and the staffing requirements are estimated accordingly. If a HAI is prevented and a patient goes home early, there will not be fewer staff required as a result of there only being 19 patients. Another patient might be admitted to fill the vacant place – if they are more ill than the patient who went home the total amount of work might even have gone up but staff numbers are still likely to be unaffected. Similar arguments apply to numbers of medical staff, laboratories, porters, laundry staff, catering costs, and so on. Therefore the words 'cost' and 'savings' in this context can be quite misleading – the true value of preventing a HAI (aside from the health of the patient) is to allow someone else to be admitted who otherwise might not have been. This is valuable and we attach a figure to recognise that, but it is not akin to a financial cost that can be saved. The values used reflect the current costs of care from Scottish Health Service Costs (41).

Two analyses were undertaken. Firstly, the overall analysis considered the additional LOS for all acute patients. Secondly the specialty analysis considered the clinical specialties that had been sampled and where there was a statistically significant difference in LOS as a result of an infection. All the data in this section on discharges, LOS and costs were taken from Scottish Health Service Costs for the year ending 31 March 2006, accessed on the ISD Scotland website (41).

#### *5.4.9 Use of prescription of antimicrobials as a proxy indicator of HAI*

The possibility that the time of prescription of antimicrobials could be used as a proxy indicator for HAI (42) was investigated using antimicrobial data collected for all patients in the prevalence survey.

To consider this question two groups were compared; group 1 and group 2. Group 1 being the group of inpatients who were prescribed an antimicrobial 48 hours or more after admission to hospital and Group 2 being everyone else. Group 2 includes both those who never had an antimicrobial and those who had all their antimicrobials prescribed within 48 hours of admission.

Prevalence of HAI in group 1 is represented by  $p_1$  and the prevalence of HAI in group 2 is  $p_2$ .

The null hypothesis ( $H_0$ ) is that the prevalence in the two groups is the same and the alternative hypothesis ( $H_1$ ) is that they are different.

Formally this is written as:  $H_0: p_1 = p_2 \vee H_1: p_1 \neq p_2$

Subjectively this hypothesis appears logical since if patients were admitted with a community acquired infection it would be expected that they would be given a therapeutic treatment within 48 hours of their admission to hospital.

Statistical analyses were carried out in STATA® using the DIAGT procedure (36). This procedure calculates sensitivity, specificity, and predictive values; together with their 95% CI

### 5.4.10 Prevalence to incidence calculation

The Rhame and Sudderth (43) model approach was tested, as the other approaches in the published literature assume the duration of infection is known and therefore do not fit with the information collected in this study.

**Equation 2:** Formula for calculating cumulative incidence from prevalence data

$$I = P \left[ \frac{LA}{LN-INT} \right] \quad (43)$$

I=Incidence rate

P=Prevalence rate

LA=mean LOS for all inpatients

LN=mean LOS for inpatients who acquire one or more HAI

INT=mean interval between admission and onset of first HAI for those inpatients who acquire one or more HAI

Incidence data were collected from hospitals participating in the SSHAIP incidence surveillance programme, which utilises staff trained in identifying surgical site HAI using CDC criteria. Information was collated from nursing and medical documentation, including temperature and prescription charts. Only data for the time period of the prevalence survey was used from the incidence data. All non-acute hospitals were excluded from the prevalence and SSHAIP data.