# Using HIPE data as a research and planning tool: limitations and opportunities: A Response

The paper by O'Loughlin et al (2005)[1] raises a number of issues in regard to the operation of the Hospital Inpatient Enquiry that require clarification and/or a response. In our view, there are many potential applications for these data that are not explored in this paper; our comments here, however, are of necessity limited to those issues raised by O'Loughlin et al (2005) in an attempt to ensure that a more balanced perspective on the operation of this system may be portrayed.

*MM Wiley*

*The Economic & Social Research Institute, 4 Burlington Road, Dublin 4.*

## INTRODUCTION

While O'Loughlin et al (2005)[1] recognises in the introduction that the Hospital Inpatient Enquiry is intended to provide a "standard minimum dataset on inpatient morbidity and mortality" subsequent comment would seem to assume a broader set of objectives for this system. It is important, therefore, to reiterate at the outset that the HIPE system is intended to collect information on *discharges* treated in the *inpatient setting*. Notwithstanding the desirability of being able to track individuals through the health system, with it's establishment the HIPE system was never intended for this purpose. While the 'inpatient' setting has been broadened in recent years to include day case activity, it was never intended that the HIPE system would collect information in the outpatient, A&E or community settings. Again, while the establishment of information systems in such settings might be desirable, such an initiative demands consideration of a whole range of factors specific to the relevant areas.

In Table 1, O'Loughlin et al (2005) list what they call 'Data collected by HIPE'. This listing is incomplete and out of date. In Table 1 here, a complete and up to date listing of data collected by HIPE is presented.

### Practical Issues
### Data quality
It is extraordinary that nowhere in the paper by O'Loughlin et al (2005) is there any mention of the HIPE software (Windows HIPE and Windows HIPE Reporter). Custom designed software for data entry and reporting of HIPE data have been produced by the ESRI's HIPE & NPRS Unit. An independent review of clinical coder training programs and data quality audit procedures commissioned from the University of Sydney in 2004 went so far as to state that "the 'ace' in the Unit's pack of data quality

initiatives is the HIPE computer system".[2] Every HIPE hospital uses the same software, standardised to the same version for the entry of HIPE data. This software is updated regularly. Integrated within this software are a whole range of procedures directed at safeguarding data quality. Automated edits and validations enable hospitals to correct errors at the source and are performed on individual data fields as data are entered in the system. There are also inbuilt combination checks between two or more data fields and final cross comparisons between all data fields when data entry for the episode of care has been completed. Table 2 gives an indication of the range of data quality and consistency checks integrated within the HIPE software. Messages are displayed to prompt or guide the user in correcting any errors. Fields are also colour coded and 'red' indicates the field with an error.

Flags are inbuilt to the data entry software to reject certain codes or coded combinations. They also allow users to confirm, endorse or explain their choice of codes. Some fields allow the user to enter a textual explanation of why they selected a certain code (particularly the .9 unspecified codes). Two special flags, audit and report flags, are automatically triggered when users select certain codes. Queries are marked in logs and the logs can be viewed and accessed centrally by the HIPE & NPRS Unit for analysis.

In addition to the edits built into the HIPE computer system, approximately 140 validation checks are routinely performed on national data by the HIPE & NPRS Unit. These cover a range of coding conventions and guidelines. If any problems are found, the Unit produces query reports for hospitals to verify and correct. Routine quality checks are also applied to administrative and demographic data in addition

**Table 1**

**DATA COLLECTED BY THE HOSPITAL INPATIENT ENQUIRY***

**ADMINISTRATIVE DATA**

- Patient name (retained within hospital)**
- Case reference number and hospital number**
- Dates of admission and discharge**
- Dates of first and principal procedure
- Day case indicator**
- Admission type and admission source**
- Discharge status and discharge destination**
- General Medical Services status**
- *Medical Card Number (GMS patient number)*
- Admitting and discharge consultant (encrypted)**
- Intensive care days and private care days**
- *Public Care days (optional)*
- *Infant admission weight (for all neonates and low weight infants)*
- *Date of transfer to Pre-Discharge Unit (optional)*
- *Admission Mode*
- *Waiting List Indicator*

**CLINICAL DATA**

- Principal** and up to *19* secondary diagnoses (ICD-10-AM wef 01/01/2005)
- Principal and up to *19* secondary procedures (ICD-10-AM wef 01/01/2005)

**DEMOGRAPHIC DATA**

- Date of birth**, sex**, marital status**
- Area of residence by county**

---

\* Data elements in *italics* are in addition/different to the data listing presented in Table 1 of O'Loughlin et al (2005)[1]

\** Completion of these data fields is mandatory

**Table 2**

**SUMMARY OF DATA QUALITY AND CONSISTENCY CHECKS INTEGRATED WITHIN THE HIPE SOFTWARE**

| ITEM | ICD-9-CM | ICD-10-AM |
|---|---|---|
| Sex checks | 1862 | 1701 |
| Admission type checks | 4566 | 2454 |
| Discharge code checks | 6820 | 4000 |
| Checks on use as Principal Diagnosis/Proceedures | 3260 | 6087 |
| Checks on use as Secondary Diagnoses/Proceedures | 216 | 144 |
| Checks on Age | 1664 | 2213 |
| Checks on LOS | 17216 | 18636 |
| Rare Diagnosis | N/A | 265 |
| Complete record checks | 114 | 30 |
| **Total** | **35718** | **35530** |

to analyses of compliance with the guidelines published regularly in *Coding Notes*[3] and the HIPE instruction manual

Requests for data quality audits can originate from a number of sources, including hospitals, the Department of Health & Children, clinicians, researchers and staff of the Unit. Audit methodology is determined by the purpose of the audit. In-house auditing software has been developed which facilitates extraction of data from the national file, the analysis of data as required and standardisation of the format of reports.

In addressing the issue of coverage, O'Loughlin et al (2005)[1] express concern at "discrepancies between the Department of Health and Children's classification of an acute hospital and that used by HIPE" and list the five hospitals concerned in Table 3. It should be noted here that the two hospitals not included in the HIPE system opted out many years ago. The three hospitals in HIPE considered to be long stay or district have been retained within the system in the interests of maintaining a data flow that might prove useful for the area concerned. To put this issue in perspective, however, it should be noted that these five hospitals combined had estimated discharges of 5,240 in 2004 while the national returns to HIPE for 2004 (estimated end

May 2005) were 963,785 discharges. The presentation of HIPE data are differentiated by Acute (length of stay 0-30 days) and Extended Stay (length of stay > 30 days) discharges specifically to take account of the fact that a small number of hospitals with long stay patients were historically included within this system. While the achievement of 100% cover is the objective for the HIPE system, the returns to date for 2004 are at the 96% level.

O'Loughlin et al (2005)[1] present as a criticism that "HIPE, however, does not record any activity in the outpatients or emergency departments". As noted above, this was never an objective put forward for the HIPE system. Any decision regarding the development of national databases in these areas will have to be informed by a range of factors including the objectives to be achieved, the costs involved and the priorities for such an investment given competing alternative demands for scarce resources within the health system. Again, to gain an appreciation for the scale of such an undertaking, it is worth noting that in 2004 there were approximately 2.4 million attendances at outpatient departments and 1.2 million attendances at casualty departments. The range of information to be collected, as well as potential applications for such data, would have to be addressed before the large scale investment of resources in these areas could be justified.

In commenting on completeness of data, O'Loughlin et al (2005)[1] note that incompleteness can render some fields useless and that one solution is to make the entry of data for certain fields compulsory. They neglect, however, to report that this is exactly what happens in the HIPE data entry system. In Table 1, completion of the data elements marked with a double asterisk are treated as mandatory by the HIPE data entry system. Records cannot therefore be returned to the HIPE & NPRS Unit or included on the national file unless these data fields have been completed.

In the section on data accuracy, O'Loughlin et al (2005)[1] quote a 1995 study of one Dublin hospital and studies of three different conditions. The authors do not report on a whole range of initiatives since the mid-1990's including the advancements with computer-based edits/checks within the HIPE data entry system, improvements in training, support and guidelines for coders, or the 18 chart-based audits that have been conducted on HIPE data since

2001. The independent review[2] of clinical coder training programs and data quality audit procedures commissioned from the University of Sydney in 2004 has also not been referenced or referred to in any way by these authors. The Bramley and Reid Report[2] (2004) was funded by the Department of Health and Children and commissioned by the ESRI's HIPE & NPRS Unit. A range of objectives were addressed by this review including an evaluation of the policies and procedures being applied within HIPE for the purpose of auditing and improving the quality assurance of coded records. A wide ranging set of recommendations have been proposed by this report and these now constitute an essential input to the agenda for the future development of the HIPE system.

The improvement of data quality and the development of more effective data quality initiatives will always be a challenge for HIPE as with other data systems. This objective is, however, accorded the highest priority in all developments being considered for this system. While recognising that much remains to be done, it is also important to acknowledge improvements achieved in recent years thanks to the commitment of those working at all levels within the system and increased investment by the Department of Health and Children in the HIPE system. When compared with the quality assurance initiatives being pursued by an agency like the Canadian Institute for Health Information (CIHI), those applied within HIPE compare favourably as, for example, both HIPE and CIHI support a coding query database, an abstracting manual, education programmes, abstracting software, system edits and re-abstraction studies[4,5].

Another important initiative for the HIPE system not mentioned by O'Loughlin et al (2005)[1] is the introduction of ICD-10-AM for morbidity coding beginning January 2005[6]. This development followed from the conduct of a review of morbidity coding schemes internationally and the completion of a pilot study in Ireland to determine the best available option for use within the HIPE system[7]. As the ICD-9-CM system had been in place since 1990, the introduction of ICD-10-AM involved the implementation of an extensive training programme for all clinical coders nationally, together with the introduction of such initiatives as the *ebook* for use in coding morbidity data. In addition to updating the clinical coding systems to the ICD-10 level, this development also provides opportunities for skill and

knowledge transfer between the National Centre for Classification in Health in Australia and the Irish HIPE system. Data quality tools used in Australia like the Performance Indicators of Coding Quality (PICQ) 2004 and the Australian Coding Benchmark Audit have already proved to be a useful resource for the development of data quality checks and audit software within the Irish system.[8]

### Events versus patients

There is no doubt that the inclusion of PPS Number on HIPE would greatly enhance the range of potential uses for these data. Use of PPS Number is, however, currently confined to specified agencies within the public sector and any further extension of its use requires legislative provision and consultation with the Department of Social and Family Affairs.[9] In the Health Information Strategy there is, however, a commitment to the introduction of a system for unique identification within the health sector using PPS Number.[9] Delivery on this commitment will, of course, have to ensure that the necessary safeguards are in place to protect patient privacy.

### Socio-economic group variable

When the HIPE system was originally introduced in the 1970s, a data variable for occupation was included. Prior to the ESRI taking management responsibility for HIPE, this variable was regrettably dropped in the 1980s due to the very low level of response achieved. There is an annual review of the data elements collected within HIPE and the inclusion of a variable to enable the assignment of socio-economic group has been considered on a number of occasions. When data changes to the HIPE system are considered, however, a number of factors have to be taken into account. Firstly, data can only be collected for the HIPE system if they are collected initially by hospitals. Where a number of changes to the system are being considered, priorities must be assigned to determine which changes are considered more urgent or important. Finally, there is a cost to each change to the HIPE system both in terms of the workload generated for those collecting and inputing the data and also because of software changes that have to be made locally and nationally. While the inclusion of a variable to facilitate improved measures of socio-economic status remains an objective for the HIPE system, the inclusion of information on public/private status and medical card status should facilitate an assessment of equity issues within the system as currently structured.

### Small area coding

In the interests of protecting patient and doctor confidentiality, the Department of Health and Children agreed in the mid 1990s that consultant codes would be encrypted and patient name and address would not be collected on the national database. In the absence of address, it is therefore not possible to determine a small area identifier for discharges on the HIPE system nationally. Given the information held locally on patient address, individual hospitals could, of course, choose to include a small area identifier on their Patient Administration Systems (PAS). The recent government announcement regarding the development of a post code system would, however, be expected to enable some advancement on this issue if, when available, the post code of the patient can be collected on the HIPE system.

### System issues
#### Access and Confidentiality

In the past, all health boards requesting access to HIPE data for their residents have been provided with data sets and software to facilitate analysis of these data. All health boards have now been provided with data sets for their residents and these will be updated annually. Comparable, national level data can, and have been, provided to health boards and other users on request.

Each year, a Shared Information System is developed and circulated to all hospitals involved in the national casemix programme. This data set includes data that the hospitals have agreed to share amongst those participating in the programme. Using this system, it is therefore possible for an individual hospital to benchmark performance against other peer hospitals for the same conditions.

### Reporting

To date, the resources available to the HIPE system have been concentrated on improving coverage, quality, timeliness and access to the data. Within the HIPE & NPRS Unit at the ESRI, approximately 1.5 FTEs have been allocated to data management and analysis functions. The responsibilities of these staff include processing data received from over 60 hospitals on a monthly basis and preparation of the national file for the Department of Health & Children together with report development and responding to data requests. In 2004, 119 such requests were received from a range of sources including clinicians, hospitals, health agencies, government departments,

researchers, voluntary organisations etc. Many of these requests result in the publication of HIPE data in academic journals and other sources. Given the constraints on the resources available, combined with increasing demands for access to data, it is regrettable that it has not been possible for the HIPE & NPRS Unit to produce more analytic reports on the system. Beginning in 2005, however, the Department of Health & Children have made funds available for an additional resource to work on data management and analysis so it is hoped that within the next year, further reports on the system will be available for publication.

## DISCUSSION

Most of the issues requiring clarification that are summarised in this section, in O'Loughlin et al (2005)[1], have already been addressed. Just two outstanding points remain to be considered. Firstly, private hospitals cannot currently be compelled to participate in the HIPE system. Private hospitals have been invited to be involved in the HIPE system and two have agreed to be so involved on a voluntary basis. In the interests of completeness for hospital activity data at a national level, the inclusion of private hospitals in the system is an objective to be supported and actively pursued. It is worth noting, however, that the largest area of activity in the private hospital sector is the provision of maternity services and data on all births nationally are collected by the National Perinatal Reporting Scheme, also managed by the ESRI's HIPE & NPRS Unit[10].

Finally, it seems extraordinary that the only mention of the resource requirements for implementing the recommendations put forward by O'Loughlin et al (2005)[1] is in the last sentence of the paper where it is noted that "While a significant investment will be required to implement the recommendations arising from this paper, the improved quality and quantity of research are likely to make this investment worthwhile." Given competing demands for resources across the health system, any request for increased investment needs to be quantified and the likely returns estimated if a case for greater investment of public funds in health information systems is to have any chance of success.

Greater use of HIPE data for research and planning is fully supported by all involved in the operation and development of this system. It is therefore all the more regrettable that, while aspiring to support these objectives, O'Loughlin et al (2005)[1] present a

paper that portrays an out-of-date view of the HIPE system, criticises HIPE for limitations on functionality that are outside the scope of the system objectives and puts forward a number of recommendations that could not be implemented in the current health information environment.

There is no denying that there are deficiencies with many aspects of the operation of the HIPE system. Recognising that the system needs further development and improvement need not, however, be inconsistent with an appreciation for the many improvements introduced within HIPE in recent years. A number of these improvements noted here, including significant developments in quality assurance, audit, coder training programmes, coverage, completeness, access etc have enhanced the value of this system for the many potential applications for the data collected. As the only source of data on the work undertaken in the most expensive part of the health system, it is essential that the HIPE system is developed in a stable and coherent manner with continued support for all involved in all aspects of the system. While increased resources will achieve improvements like the collection of an expanded range of data and improved tools for monitoring data quality, optimum data quality standards can only be achieved where all involved in data reporting, coding and collection are facilitated in taking 'ownership' of the data they return to the HIPE system.

## REFERENCES

1 O'Loughlin, R., S Allwright, J Barry, A Kelly, C Teljeur *Using HIPE data as a research and planning tool: limitations and opportunities*, Irish J Med Sci, 2005; 174: p.41-46.

2 Bramley M., and B. Reid *Towards Best Practice in the Coding of Morbidity Data*. A review of clinical coder training programs and data quality audit procedures within the Hospital In-Patient Enquiry Unit, ESRI. The University of Sydney 2004.

3 *Coding Notes* is a quarterly newsletter produced by the ESRI's HIPE & NPRS Unit for coders and others involved in the HIPE system.

4 Canadian Institute for Health Information, *Quality Assurance Processes Applied to the Discharge Abstract and Hospital Morbidity Databases*. Canadian Institute for Health Information, Ottawa, 2002.

5 Canadian Institute for Health Information, *Discharge Abstract Database Data Quality Re-abstraction Study, Combined Findings for Fiscal Years 1999/2000 and 2000/2001*. Canadian Institute for Health Information, Ottawa, 2002.

6   National Centre for Classification in Health, The
    International Statistical Classification of Diseases and
    Related Health Problems, Tenth Revision, Australian
    Modification, (4th ed), Sydney, NCCH, 2004.

7   Murphy, D., M.M. Wiley, A. Clifton, D. McDonagh
    *Updating Clinical Coding in Ireland: Options and
    Opportunities,* The Economic and Social Research
    Institute, Dublin, September 2004.

8   National Centre for Classification in Health, *The
    Australian Coding Benchmark Audit (ACBA), Second
    Edition.* National Centre for Classification in Health,
    University of Sydney, 2001.

9   Department of Health and Children Health
    Information: A National Strategy. Dublin: The Stationery
    Office, 2004.

10  The Economic and Social Research Institute, *Report on
    Perinatal Statistics for 2001.* ESRI, Dublin, 2005.

*Correspondence to: Professor Miriam Wiley,
The Economic and Social Research Institute,
4 Burlington Road, Dublin 4. Tel: +353 1 667 1525,
Fax:+353 1 668 6231, email: miriam.wiley@esri.ie*