Please note that the content of this document is not in the public domain and should not be disseminated outside participating laboratories or their local HPUs.

Comparison of hepatitis C reports from routine national surveillance with sentinel surveillance data

Introduction

As part of the national hepatitis surveillance scheme, laboratory-confirmed hepatitis C virus (HCV) infections are reported to the HPA Centre for Infections on a voluntary basis. Reporting is known to be incomplete, and varies between centres [1]. Data from sentinel surveillance (the Denominator study) can be used to investigate the extent of this under-reporting, since test results on all individuals tested for hepatitis in participating centres are extracted electronically from laboratory records, providing a complete data set for comparison against national surveillance reports [2].

In order to calculate the proportion of positive hepatitis C tests from laboratories participating in the sentinel surveillance study which were also reported to the national surveillance scheme, a process of matching between the two data sets was undertaken.

Methods

Details of all laboratory-confirmed hepatitis C virus infections reported to national surveillance between January 1996 and April 2007 were extracted from Labbase2 using the Hummingbird program. These data were imported into a Microsoft Access database and duplicates removed using the unique OPIE ID identification code.

All samples with a positive anti-HCV antibody or HCV RNA result from the sentinel surveillance dataset between 1st January 2002 and 31st December 2006 were identified. After exclusion of samples tested for quality control purposes, records were imported into a table in the existing Microsoft Access database for matching with the national surveillance data. The participating laboratories and the first and last dates of samples used in matching are shown in table 1.

Table 1. List of participating centres and the dates of first and last samples from the sentinel surveillance dataset used for matching.

Participating laboratory	Date of first sample	Date of last sample
Ashford	02/01/2002	29/12/2006
Birmingham	02/01/2002	30/12/2006
Cambridge	02/01/2002	2912/2006
Centre for Infections (CFI)	03/01/2002	29/12/2006
Chester	02/01/2002	29/12/2006
Dulwich	03/09/2004	30/12/2006
Ealing	16/11/2002	15/10/2003
Grimsby	01/04/2004	29/12/2006
Hull	01/04/2002	30/09/2003
Leeds General Infirmary	01/09/2004	30/12/2006
Leeds HPA	02/01/2002	29/07/2005
Liverpool HPA	02/01/2002	29/12/2006
Manchester	02/01/2002	30/12/2006
Newcastle	02/01/2002	28/12/2006
North Middlesex	29/07/2002	29/12/2006

Nottingham	02/09/2002	29/12/2006
Portsmouth	02/01/2002	30/12/2006
Preston	02/01/2002	30/12/2006
Royal Liverpool Hospital	02/01/2002	31/12/2006
St Bartholomew's Hospital	01/08/2004	31/12/2006
St George's Hospital	01/01/2006	31/12/2006
University College Hospital	01/09/2004	29/12/2006

Individuals were matched on laboratory number, hospital number, date of birth, soundex, sex, region, date of test, and a combination of these variables.

Individual patients from the sentinel surveillance dataset whose samples had been matched to a national surveillance report were then identified using a unique patient reference number, allowing calculation of the total number of individuals known to national surveillance.

The same process of matching was subsequently undertaken in order to assess the percentage of individuals with a national surveillance report from laboratories participating in the sentinel surveillance study who were present in this latter dataset, in order to check the matching process used.

Results

In total, 27063 (43.8%) of 61800 anti-HCV and/or HCV PCR positive individuals identified through the sentinel surveillance study between January 2002 to December 2006 could be matched to reports from routine surveillance of hepatitis C, suggesting that these individuals had been reported to this scheme.

There is considerable variation in matching between participating laboratories (table 2). The percentage matched in a previous analysis of the centres that participated in the pilot phase of the study is also shown for comparison [2].

When interpreting this data, it should be remembered that laboratories are responsible for reporting cases identified as part of primary screening within their laboratory to the national surveillance scheme. Laboratories testing samples for reference/confirmatory purposes are not responsible for reporting these cases to national surveillance; it is the responsibility of the original testing laboratory. Therefore the amount of reference work undertaken by individual laboratories may affect the level of matching in that sentinel centre.

Detailed analyses suggest that, in some centres with a high reference workload, matching of positives identified through primary screening (i.e. cases which that laboratory is responsible for reporting) is much greater than the overall figure. For example, data from Birmingham – shown in table 3 – suggests that reporting of cases from Heartlands Hospital (where the laboratory is based) and from primary care services, such as general practice and GUM clinics, is good, but that low reporting of cases referred to the laboratory from other locations skews the overall figure, due to the large volume of reference work undertaken in this laboratory.

Conversely, analysis of results from other centres suggests that their level of matching can be accounted for by reporting by other, referring laboratories and that matching at the sentinel laboratory is relatively poor. One example of this is the Newcastle centre, data from which is shown in table 4. At this centre, the overall percentage of matched individuals (26.6%) is largely accounted for by reporting from

referring laboratories and prison services, masking a very low level of matching from the hospitals served by this laboratory.

Matching is also low for individuals tested at the Centre for Infections, since the large majority of work undertaken by this laboratory is confirmatory or additional testing of samples referred from other laboratories which it is not CfI's responsibility to report.

Table 2. Number and percentage of individuals reported as hepatitis C positive in sentinel surveillance matched to a report from routine surveillance, by participating laboratory.

Participating laboratory	Number of individuals matched	Total number of individuals	% matched (this analysis)	% matched (pilot study)
Ashford	662	1144	57.9	-
Birmingham	3952	7215	54.8	80.2
Cambridge	2135	5287	40.4	-
Centre for Infections (CFI)	1930	10581	18.2	
Chester	730	1202	60.7	-
Dulwich	289	788	36.7	-
Ealing	1	87	1.1	0.0
Grimsby	669	1001	66.8	-
Hull	103	314	32.8	21.4
Leeds General Infirmary	1775	3313	53.6	-
Leeds HPA	1015	3412	29.7	0.0
Liverpool HPA	2202	3248	67.8	-
Manchester	5441	7834	69.5	75.5
Newcastle	1200	3173	37.8	23.7
North Middlesex	210	578	36.3	19.4
Nottingham	697	2927	23.8	2.4
Portsmouth	413	970	42.6	-
Preston	2169	2896	74.9	-
Royal Liverpool Hospital	264	1252	21.1	-
St Bartholomew's Hospital	148	2656	5.6	-
St George's Hospital	324	407	79.6	-
University College Hospital	734	1515	48.4	-
Total, all centres	27063	61800	43.8	**53.5**

Table 3. Number and percentage of individuals from sentinel surveillance matched to a report from routine surveillance from the Birmingham centre, by location.

Location	Number of individuals matched	Total number of individuals	% matched
Heartlands Hospital	251	327	77.0
Birmingham Children's Hospital	140	165	84.8
GP	89	101	88.1
GUM clinics	45	51	88.2
Drug services	74	219	33.8
Reference/hosprefall	3352	6351	52.8
Unknown	1	1	100.0
Total	**3952**	**7215**	**54.8**

Table 4. Number and percentage of individuals from sentinel surveillance matched to a report from routine surveillance from the Newcastle centre, by location.

Locations	**Number of individuals matched**	**Total number of individuals**	**% matched**
Freeman Hospital	24	61	39.3
Newcastle General Hospital	13	174	7.5
Royal Victoria Infirmary	8	146	5.5
GP	51	278	18.3
Drug services	4	13	30.8
Prison	156	247	63.2
Other hospitals/reference	944	2254	41.9
Total, all locations	**1200**	**3173**	**37.8**

Table 5 shows the results of matching of sentinel surveillance data from 2006 only.

Table 5. Number and percentage of individuals reported as hepatitis C positive in sentinel surveillance during 2006 matched to a report from routine surveillance, by participating laboratory.

Participating laboratory	**Number of individuals matched**	**Total number of individuals**	**% matched**
Ashford	105	193	54.4
Birmingham	733	1540	47.6
Cambridge	363	830	43.7
Centre for Infections (CFI)	489	2458	19.9
Chester	69	195	35.4
Dulwich	151	402	37.6
Grimsby	275	376	73.1
Leeds General Infirmary	1140	1693	67.3
Liverpool HPA	340	559	60.8
Manchester	850	1543	55.1
Newcastle	101	235	43.0
North Middlesex	100	143	69.9
Nottingham	95	667	14.2
Portsmouth	70	157	44.6
Preston	479	656	73.0
Royal Liverpool Hospital	92	383	24.0
St Bartholomew’s Hospital	70	944	7.4
St George’s Hospital	324	407	79.6
University College Hospital	306	618	49.5
Total, all centres	**6152**	**13999**	**43.9**

The level of matching at several laboratories has shown definite trends over the years since sentinel surveillance began, suggesting changes in reporting to routine surveillance. Laboratories at Leeds, Newcastle and Portsmouth have all shown an overall increase in matching since 2002. There has also been an increase in matching at Grimsby since data collection began there in 2004.

Matching of data from a number of laboratories – such as St Bartholomew's (7.4% individuals matched in 2006 and 5.6% overall) and Royal Liverpool (24% individuals matched in 2006 and 21.1% overall) – remains low. However, the level of matching may reflect the quality of patient data supplied to routine and/or sentinel surveillance rather than the level of reporting. This possibility is discussed further in the 'Limitations' section below.

Tables 6 and 7 show the number and percentage of individuals matched by age and sex, respectively. The lower percentage of matched individuals of unknown age and sex is likely to reflect a lack of patient details (in this case, DOB and sex) available for use in matching. It is interesting to note the low percentage of matched individuals aged below 1 year of age: it is possible that reporting of patients in this age group is less common due to the difficulties of interpreting a positive anti-HCV result in the presence of maternal antibody.

Table 6. Number and percentage of individuals from sentinel surveillance matched to a report from routine surveillance, by age group.

Age group	Number of individuals matched	Total number of individuals	% matched
Under 1	123	723	17.0
1-14	156	330	47.3
15-24	2284	4872	46.9
25-34	8503	17744	47.9
35-44	8709	19325	45.1
45-54	4551	10650	42.7
55-64	1410	3420	41.2
65 plus	1147	2721	42.2
Unknown	180	2015	8.9
Total	**27063**	**61800**	**43.8**

Table 7. Number and percentage of individuals from sentinel surveillance matched to a report from routine surveillance, by sex.

Sex	Number of individuals matched	Total number of individuals	% matched
Female	8740	19872	44.0
Male	17306	38697	44.7
Unknown	1017	3231	31.5
Total	**27063**	**61800**	**43.8**

Table 8 shows the number of individuals matched by IDU status as reported in the sentinel surveillance data. Questionnaire data from the pilot phase of this study suggests that reporting of IDU status is poor in sentinel surveillance, and a comparison among the matched patients indicates that IDU is better reported in routine surveillance than in the sentinel surveillance study (which uses a freetext clinical details field). The data presented here is likely to under-estimate the number of IDUs by a considerable amount.

Table 8. Number and percentage of individuals from sentinel surveillance matched to a report from routine surveillance, by IDU status as reported in sentinel surveillance dataset

Identified as IDU in sentinel surveillance	Number of individuals matched	Total number of individuals	% matched
IDU	3559	7383	48.2
Not identified as IDU	23504	54417	43.2
Total	**27063**	**61800**	**43.8**

If we assume that matching is accurate and complete, the analyses suggest that reporting to routine surveillance is not affected by the age or sex of a positive individual, although there may be some difference by IDU status. Additional work is needed to investigate this further.

Completeness of sentinel surveillance data

An analogous process of matching was undertaken to assess the proportion of individuals with a routine national surveillance report from a sentinel laboratory who could be matched to patients in the sentinel surveillance dataset. Since the sentinel surveillance data collection process involves the electronic extraction of laboratory data on all individuals being tested for hepatitis at participating centres, it would be expected that data on all patients reported to the national surveillance scheme by a sentinel centre would be available from this dataset. The results are shown in Table 9. As expected, matching is almost complete for the majority of centres, suggesting that the method used in the first matching exercise (discussed above) is an effective one. The small proportion not matched in this analysis may reflect individuals for whom personal data used in matching (such as soundex, DOB and sex) is missing or incorrectly entered.

St Bartholomew's Hospital and Royal Liverpool University Hospital are not included in this table since no hepatitis C reports from these laboratories could be found in the routine surveillance dataset. This suggests that these laboratories are not reporting hepatitis C cases and that the small proportion of routine reports which were matched to individuals in the sentinel surveillance data during the earlier analysis (Table 2) represent samples referred from other laboratories which had already been reported.

Centre for Infections is also omitted from this table: since it primarily undertakes reference testing of samples for other laboratories, it does not report cases to routine surveillance.

Table 9. Number and percentage of individuals reported to routine national surveillance who can be matched to individuals in the sentinel surveillance dataset.

Participating laboratory	Period for which sentinel surveillance data available		Number of individuals matched	Total number of individuals	% matched
	From	To			
Ashford	01/01/2002	31/12/2006	511	519	98.5
Birmingham	01/01/2002	31/12/2006	1318	1339	98.4
Cambridge	01/01/2002	31/12/2006	247	255	96.9
Chester HPA	01/01/2002	31/12/2006	722	750	96.3
Dulwich	01/11/2004	31/12/2006	231	237	97.5
Grimsby	01/04/2004	31/12/2006	163	164	99.4
Hull	01/04/2002	30/09/2003	37	40	92.5

Leeds HPA/ Leeds Gen Infirmary	01/01/2002	31/12/2006	1372	1481	92.6
Liverpool HPA	01/01/2002	31/12/2006	1891	1959	96.5
Manchester	01/01/2002	31/12/2006	5068	5294	95.7
Newcastle	01/01/2002	31/12/2006	55	80	68.8
North Middlesex	29/07/2002	31/12/2006	189	244	77.5
Nottingham	02/09/2002	31/12/2006	422	453	93.2
Portsmouth	01/01/2002	31/12/2006	327	356	91.9
Preston	01/01/2002	31/12/2006	2005	2108	95.1
St George's Hospital	01/01/2006	31/12/2006	247	253	97.6
University College Hospital	01/09/2004	31/12/2006	447	510	87.6
Total, all centres	-	-	**15,252**	**16,042**	**95.1**

Limitations and issues for further consideration

To some extent the level of matching possible is a reflection of the quality of information provided by the participating laboratory. Absence of details such as soundex and date of birth or mistakes in data entry may result in an under-estimate of the number of samples reported to national surveillance. The quality of data from individual laboratories may also change over time, which would affect the level of matching and act as a potential confounder of an apparent increase or decrease in reporting.

Table 10 shows the percentage of individuals from the sentinel surveillance dataset used in matching who were missing key variables required to link patients to reports in routine surveillance. There does not seem to be a clear relationship between completeness of these variables and the levels of matching for each centre described above, suggesting that data quality may not be a key factor (although this table does not provide any information about mistakes in DOB, sex or soundex which could compromise matching).

Table 10. Completeness of key matching variables in sentinel surveillance data, 2002-2006

Participating centre	Individuals with unknown DOB (%)	Individuals with unknown sex (%)	Individuals with unknown soundex* (%)	Total number of individuals in sentinel surveillance dataset
Ashford	0.3	1.2	0.5	1144
Birmingham	1.2	7.0	10.1	7215
Cambridge	1.6	3.2	4.4	5287
CFI	12.4	5.4	16.1	10581
Chester	0.2	0.9	39.8	1202
Dulwich	0.8	1.8	12.9	788
Ealing	17.2	17.2	0.0	87
Grimsby	0.2	48.1	9.2	1001
Hull	0.0	0.3	0.0	314
Leeds General Infirmary	0.7	2.4	5.5	3313
Leeds HPA	2.2	4.2	6.9	3412
Liverpool HPA	1.0	1.5	18.5	3248
Manchester	1.9	1.9	14.9	7834
Newcastle	1.2	2.4	10.7	3173
North Middlesex	0.9	10.7	60.6	578
Nottingham	1.9	20.7	3.7	2927
Portsmouth	0.3	3.0	0.4	970

Preston	2.0	1.3	13.8	2896
Royal Liverpool Hospital	0.3	0.6	14.1	1252
St Bartholomew's Hospital	1.9	4.1	14.9	2656
St George's Hospital	0.0	0.5	18.4	407
University College Hospital	0.5	3.8	0.0	1515
Total, all centres	3.3	5.2	11.9	61800

* Unknown soundex was defined as missing soundex or a generic "unknown" code such as XX00.

It is possible that some samples shown here as unmatched may be from individuals who were diagnosed with hepatitis C infection and reported to national surveillance prior to the start of our Labbase dataset in 1996; however, this number is likely to be small.

It is acknowledged that there may be local factors influencing reporting patterns that this analysis does not take into account.

Although information on hepatitis C cases identified through sentinel surveillance could be used to supplement routine surveillance, this would bypass CoSurv and potentially cause problems with reporting and surveillance at a regional level. Further consideration and consultation on these issues is necessary and we would welcome any feedback you may be able to provide.

The method used on this occasion was primarily designed to identify individuals with any hepatitis C positive sample in the sentinel surveillance dataset which matched a report to the national surveillance scheme. It was therefore not able to identify individuals who matched to more than one report from the national surveillance scheme; that is, where more than one sample from the same patient had been repeatedly notified to the national surveillance scheme as separate cases.

Conclusions

This analysis suggests that under-reporting to routine surveillance occurs and that less than 50% of positive cases known to sentinel surveillance may have been reported to the national surveillance scheme. Matching to routine surveillance varies greatly by testing laboratory, but does not appear to differ by age and sex.

Emily Tweed
5th December

Reference List

(1) Gungabissoon U, Balogun MA, Ramsay ME. Hepatitis C virus: Laboratory surveillance in England and Wales 1992-2004. Epidemiol Infect 2007; 135(4):541-8.

(2) Brant L, Hurrelle M, Balogun MA, Klapper P, Ahmed F, Boxall E et al. Sentinel laboratory surveillance of hepatitis C antibody testing in England: understanding the epidemiology of HCV infection. Epidemiol Infect 2007; 135(3):417-426.