Using multiple sources to improve and measure case ascertainment in surveillance studies: 20 years of the British Paediatric Surveillance Unit

Rachel L. Knowles[1], Alan Smith[2], Richard Lynn[3], Jugnoo S. Rahi[1] on behalf of the British Paediatric Surveillance Unit (BPSU)

[1]Centre for Paediatric Epidemiology and Biostatistics, Institute of Child Health, London WC1N 1EH, UK
[2]Health Protection Agency, Centre for Infections, Communicable Disease Surveillance Centre, London NW9 5HT, UK
[3]British Paediatric Surveillance Unit of the Royal College of Paediatrics and Child Health, London W1W 6DE, UK
Address correspondence to Rachel L. Knowles, E-mail: r.knowles@ich.ucl.ac.uk

ABSTRACT

Background The British Paediatric Surveillance Unit (BPSU) was established in 1986 to facilitate national surveillance of uncommon paediatric disorders. This study investigated the effectiveness of using multiple source reporting and capture–recapture analysis to maximize case ascertainment in studies undertaken through the BPSU.

Methods Structured review of all surveillance studies completed through the BPSU. Quantitative and qualitative analysis of the effectiveness of multiple reporting sources and capture–recapture methods was made.

Results Of 71 studies undertaken through the BPSU, 59 were included in this review and 38 used additional data sources. Established national sources were most readily adapted for use as secondary data sources, including routine health data, communicable disease and specialty-specific surveillance units, whilst the involvement of parent groups and the media was less successful. Six studies employed capture–recapture techniques to estimate the completeness of case ascertainment.

Conclusions Active surveillance through the BPSU remains a timely and reliable primary source of cases, but employing additional reporting sources is effective in enhancing case ascertainment. When the assumptions for its valid use are met, capture–recapture analysis allows the estimation of completeness of ascertainment. It is essential to define the purpose of an additional source at the outset of a study and to ensure that subsequent analysis is appropriate.

Keywords British Paediatric Surveillance Unit, BPSU, capture–recapture, case ascertainment, paediatric, surveillance

Background

Paediatric surveillance studies have an important public health function: collecting reliable and timely information about the distribution and determinants of disease in the population and facilitating effective healthcare responses to reduce morbidity and mortality and improve health.[1] The British Paediatric Surveillance Unit (BPSU) was founded in 1986 by the Health Protection Agency, Royal College of Paediatrics and Child Health, Institute of Child Health (London), Royal College of Physicians (Ireland) and Health Protection Scotland to undertake active surveillance of uncommon conditions in childhood, including infections, and to provide a mechanism to respond to acute public health events. Currently, 2400 paediatricians in the UK and Ireland participate in the BPSU monthly active reporting scheme (the Orange Card), with over 90% response rate.[2] Now in its 20th year, the BPSU has inspired similar schemes nationally and internationally and is a founder member of the International Network of Paediatric Surveillance Units (INOPSU), a worldwide affiliation of 14 national paediatric surveillance units who have studied over 180 paediatric disorders to date.

Rachel L. Knowles, MRC Research Fellow (Health of the Public) and Medical Adviser (Non-communicable Disease) to the BPSU

Alan Smith, Specialist Registrar and Medical Adviser (Communicable Disease) to the BPSU

Richard Lynn, Scientific Co-ordinator of the BPSU

Jugnoo S. Rahi, Clinical Senior Lecturer in Ophthalmic Epidemiology/Honorary Consultant Ophthalmologist and former Medical Adviser (Non-communicable disease) to the BPSU

Public health surveillance is the ongoing systematic collection, analysis, interpretation and dissemination of data regarding a health-related event.[3] In disease surveillance, it is important to maximize the ascertainment of true cases and to remove incorrect or duplicate reports to obtain accurate estimates of incidence and to avoid selection bias. Four approaches to improve ascertainment are common to all BPSU studies:

1 Careful pre-study review ensures simple reporting case definitions and questionnaires, reducing workload and improving compliance.
2 The accuracy of the reporting database is continually monitored.
3 Monthly card returns are monitored and reminders sent out.
4 Additional reporting sources are often used to complement the BPSU reporting system.

The use of additional reporting sources has been encouraged as a means of identifying cases not ascertained through the BPSU, confirming true cases and permitting the estimation of the completeness of ascertainment. However, there can be both advantages and difficulties to constructing new reporting sources, and their use and interpretation should be carefully considered for individual studies. In this article, we present the findings of a structured review of all studies undertaken through the BPSU over the last 20 years, focusing on the appropriateness of different types of reporting sources and the application of capture–recapture analysis in these national paediatric surveillance studies.

Capture–recapture analysis uses surveillance data from two or more sources, covering the same population and with known overlap, to estimate the number of additional cases that are more likely to have been missed by all reporting sources, hence the 'true' number of cases (N) in the population (Fig. 1).

Capture–recapture techniques originated in wildlife studies, e.g. to estimate the number of fish in a pond. A net was used to catch a number of fish, which were counted, marked and returned to the pond. The net was used again to catch more fish that were also counted, and the number of marked fish, those caught twice, was noted. If all fish caught on the second occasion were marked, then it could be assumed that there were only these fish in the pond, whereas if no fish in the second catch were marked, then there is likely to be a large number of fish in the pond. Statistical formulae allow us to estimate the overall number of fish in the pond based on the number of fish caught on each occasion and the number caught more than once (Fig. 1).

Capture–recapture has been used in demographic,[4] epidemiological[5] and disease surveillance studies[6–9] for over 60 years, and there are good examples of its recent application to paediatric epidemiology in the UK.[10,11] In surveillance studies, different sources identify or 'catch' disease cases in the same way as different throws of the net catch fish. However, the application of capture–recapture methods relies on four assumptions being met:[12]

1 Matching pairs must be identified between sources. In surveillance studies, cases identified by one source are matched with cases identified by the second source, using unique identifiers, and the number of cases 'caught twice' is thus known. It is important that the method and the accuracy of diagnosis are the same for each source to allow good matching.
2 The population under study must be 'closed'. Just as there should be no way for fish to enter or leave the pond, thus reducing the likelihood of them being caught, in disease surveillance, there should be no migration or loss of cases from the population during the study period.
3 Each case must have the same probability of being identified by each source (known as homogeneity). Sources should not preferentially identify certain types of cases, e.g. more severe or younger cases; in other words, a net with the same size of holes must be used each time.
4 The two sources should be independent. The identification of a case by one source should not affect it being identified by the other. For example, if the net was not emptied between catches, some fish inevitably would be caught twice. If cases identified by one source are thereafter more likely to be identified by the other source, sources show 'positive dependence'. The completeness of ascertainment is overestimated, and the total population (N) is underestimated. Conversely, when cases identified by one source become less likely to be identified by the second source, then 'negative dependence' and an overestimate of the total population (N) result (Fig. 1).

Case homogeneity and the independence of sources are difficult assumptions to verify and in reality, in health care, are unlikely to be fully achievable.[13–15] Cases tend to be heterogeneous depending on age, sex, geography, social status
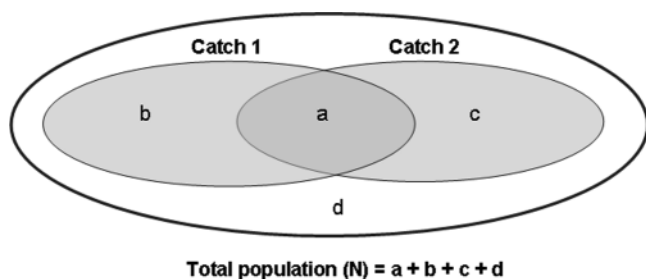

Fig. 1 Two-source capture–recapture.

Total population (N) = a + b + c + d

or severity of disease. Adjustment for heterogeneity can be made by stratifying the analysis by the variables thought to be related to 'capture', e.g. case severity.[16] Complete independence of reporting sources is also difficult to achieve,[8,9] but a more complex statistical model can be developed, based on stratification, where there are three or more sources, to account for dependence bias.[12]

The aim of this study was to review the use of multiple sources and their effect on improving case ascertainment in all paediatric surveillance studies undertaken through the BPSU between 1986 and 2006.

Methods

We obtained the details of all studies undertaken through the BPSU since 1986. We acquired copies of all published papers and BPSU annual reports for these studies, identifying and excluding ongoing studies and those that had never fully reported. We reviewed these studies and, using a standard proforma, collected the following information: number of studies using additional reporting sources, number of reporting sources used per study, number of cases reported by each source and by both sources and the percentage of all study cases ascertained through the BPSU. We reviewed qualitatively the contribution of different types of reporting source to increasing case ascertainment and defined good reporting sources as those which identified true cases, used active surveillance and had good national coverage. Where capture–recapture analysis was described, we verified whether the assumptions for its valid use were met. Data were entered into an Excel database before analysis.

Results

Between July 1986 and January 2006, 71 BPSU studies (24 involving communicable and 47 non-communicable diseases) were undertaken. Published reports were available for 59 studies, 10 were ongoing and no reports could be found for two.

Of the 59 included studies, 38 used one or more secondary reporting sources in addition to the BPSU (Table 1). In 47% (18/38) of these studies, only one additional source to the BPSU was employed; however, 34% (13/38) of studies used two additional sources and 18% (7/38) used three or more (Table 2). No secular trend was observed in the use of additional sources.

Few reports provided the number of cases by source, i.e. the number of cases reported through the BPSU, the additional source and both sources. It was therefore difficult to

Table 1 Reporting sources used by studies undertaken through the British Paediatric Surveillance Unit

	Number of studies	As a proportion of total studies included in the review (%)
Total studies included	59	–
Studies using additional reporting source(s)	38	64
Studies using one additional source	18	31
Studies using two additional sources	13	22
Studies using three or more additional sources	7	11

interpret reported data about the proportion of cases ascertained through the BPSU or to compare the completeness of ascertainment for different sources.

Of 38 studies with additional reporting sources, 27 actively requested notifications—16 of these contacted sources on a monthly basis, whereas 11 requested a single report at study completion. Effective additional sources included communicable disease and laboratory surveillance systems,[17,18] specialty-specific surveillance units,[19] specialty interest groups[20,21] and clinical record reviews.[16] Although improving case ascertainment overall, particular disadvantages in relation to the use of capture–recapture analysis were identified with some sources, including non-independent reporting among healthcare professionals,[22] preferential identification of severe cases by the media[23] and non-independence of diagnosis by parents.[24]

When investigators rely on the BPSU as the only reporting source, cases occurring within certain categories of children may be missed. Thus, additional sources were most often used to extend population coverage in studies where it was more likely that some children would not be seen by paediatricians who report to the BPSU. For example, older children may be treated by an adult clinician,[20] a specialist[16,21,25–28] or a chronic care team[22] rather than a paediatrician. Atypically presenting cases may be managed by a variety of adult and paediatric clinicians who should all be included in surveillance.[14,29] Microbiological laboratory tests and infectious disease surveillance systems cover both hospital and community infections.[17,30] Finally, postmortems and coroner's reports may be required as additional sources for deaths occurring before hospital admission.[31]

Additional sources have also been used for other purposes in surveillance studies. In a BPSU study of cerebral oedema in children with diabetic ketoacidosis (DKA), a survey of paediatric units provided the population denominator of

Table 2 Characteristics of studies undertaken through the British Paediatric Surveillance Unit and using multiple reporting sources

Condition under study	Objectives of study	Start date	Duration of study (months)	Number of additional sources	Details of additional sources
HIV/AIDS	To undertake surveillance of paediatric HIV infection and AIDS in the UK and Ireland in children under 16 years old	June 1986	Ongoing	3	RCOG PHLS Direct reports
Subacute sclerosing panencephalitis (SSPE)	To determine the incidence of SSPE in children under 16 years old	June 1986	181	1	PHLS
Reye's syndrome	To describe the epidemiological and clinical features of Reye's syndrome, and to monitor secular trends, in children under 16 years old	June 1986	130	2	ONS (death registrations) CDSC
Haemolytic uraemic syndrome (HUS)	To determine the clinical and epidemiological features of HUS in children under 16 years old	June 1986	43	1	PHLS
Haemorrhagic shock encephalopathy	To determine the incidence of haemorrhagic shock encephalopathy syndrome (HSES) in children under 16 years old	June 1986	30	2	PHLS Pathologists
Drowning and near drowning	To determine the pattern of drowning/near drowning in children under 15 years old in Britain and to identify means of prevention	January 1988	24	3	ONS (death registrations) Royal Society for the Prevention of Accidents (ROSPA) Coroners
Galactosaemia	To determine the incidence and features of clinical presentation of galactosaemia in children under 16 years old	January 1988	45	1	Specialist laboratory reports
Insulin-dependent diabetes mellitus (IDDM) in under 15s	To ascertain the annual incidence rate of IDDM in children under 15 years old during 1988 and to compare the results with those from 1973 to 1974	January 1988	12	2	Diabetologists Diabetic nurses/health visitors
Higher-order births	To assess prospectively the number of triplet and higher multiple births born in 1989, their methods of conception; obstetric factors and the effect on neonatal medical services	January 1989	12	3	ONS (birth registrations) General Register Office (Scotland) Survey of consultant neonatologists
Congenital toxoplasmosis I	To determine the incidence of congenital toxoplasmosis in live born infants in the UK	June 1989	12	1	PHLS
MMR meningo-encephalitis	To determine the incidence of meningo encephalitis within 6 weeks of MMR vaccination	January 1990	24	1	PHLS
Chemistry set poisoning	To determine incidence and outcomes of poisoning because of toy chemistry sets in children under 16 years old	January 1991	16	1	National Poisons Information Service
Acute flaccid paralysis (polio)	To determine the incidence of acute flaccid paralysis (poliomyelitis) in children under 16 years old	July 1991	36	1	PHLS
IDDM under 5s	To establish the incidence of IDDM diagnosed in children under 5 years old	January 1992	12	2	Diabetic nurses/health visitors Regional health authorities
Long-term parenteral nutrition	To quantify the need for support services for children receiving parenteral nutrition for more than 6 weeks	February 1992	3	2	Hospital nutritionist teams Direct reports

Table 2 Continued

Condition under study	Objectives of study	Start date	Duration of study (months)	Number of additional sources	Details of additional sources
Juvenile dermatomyositis	To determine the incidence of juvenile dermatomyositis in children under 16 years old	June 1992	19	2	British Paediatric Rheumatology Group British Dermatology Association
Invasive *Haemophilus influenzae* infection	To determine the incidence of invasive *H. influenzae* infection and assess the efficacy of an accelerated vaccination schedule in infants	October 1992	36	1	Microbiologists
Biliary atresia	To determine the incidence of biliary atresia in children in the first year of life	March 1993	24	1	Gastroenterologists and surgeons (survey)
Congenital dislocation of the hip (CDH)	To estimate the incidence of operative procedures for CDH with adjustment for under-ascertainment by capture–recapture techniques	April 1993	13	2	British Orthopaedic Association Hospital case notes review
Neonatal necrotizing enterocolitis (NNEC)	To establish the incidence of NNEC and to determine whether early diet can influence onset and severity	October 1993	13	1	NICU/SCBU survey
Medium chain acyl-CoA dehydrogenase (MCAD) deficiency I	To investigate the diagnosis and outcome of MCAD deficiency in the UK	March 1994	25	1	Specialist laboratory reports
Water births	To compare perinatal morbidity and mortality for babies delivered in water with rates for babies delivered conventionally	April 1994	24	1	NHS maternity units (survey)
Congenital syphilis	To measure the incidence of syphilis in pregnancy and congenital syphilis in the UK	July 1994	36	1	Genitourinary medicine specialists
Pyridoxine-dependent seizures	To study the epidemiology of pyridoxine-dependent seizures and other forms of pyridoxine-responsive seizures in children under 16 years old	September 1995	26	2	Direct reporting Personal (investigators cases)
Congenital cataract	Use of capture–recapture analysis to determine completeness of ascertainment in a national study of congenital cataract	October 1995	12	1	Ophthalmologists (BOSU)
Cerebral oedema in children with diabetic ketoacidosis (DKA)	To determine the risk and outcome of cerebral oedema complicating DKA in children under 16 years old	October 1995	35	1	National reporting system for DKA
Neonatal meningitis	To determine the incidence of neonatal meningitis in England and Wales	July 1996	18	4	Hospital microbiologists National reference laboratory CDSC ONS
Hepatitis C	To investigate the prevalence, distribution and clinical details of paediatric hepatitis C virus infection in the UK and Ireland	April 1997	21	3	CDSC/SCIEH National blood authority National Paediatric HIV Programme
Food allergy	To discover the incidence of fatal and severe allergic reactions to food in children aged under 15 years in the UK	March 1998	24	5	Register of allergy deaths Parent support group Letters to experts Web-based information service Newspaper reports

Table 2 Continued

Condition under study	Objectives of study	Start date	Duration of study (months)	Number of additional sources	Details of additional sources
Subdural haematoma	To determine incidence, aetiology and clinical features of subdural haematoma and effusion in infancy throughout the British Isles	April 1998	12	1	ONS (death registrations)
Congenital brachial palsy	To determine the incidence and study the causes and outcome of congenital brachial palsy	April 1998	12	2	Orthopaedic surgeons Plastic surgeons
Inflammatory bowel disease	To prospectively determine the incidence of inflammatory bowel disease in children under 16 years old	June 1998	13	2	British Society of Gastroenterology Research Unit Register of paediatric inflammatory bowel disease
Blindness or severe visual impairment	To determine annual incidence and classify causes of severe visual impairment and blindness in children under 16 years old	January 2000	12	1	BOSU
Neonatal group B streptococcal infection	To assess the current national incidence of group B streptococcal disease in infants younger than 90 days in the UK	February 2000	13	3	PHLS Microbiologists Parent support group website
Childhood cerebrovascular disease, stroke and stroke-like illness	To estimate the incidence of stroke, stroke-like illness and cerebrovascular disease in children under 16 years old. To determine patterns of presentation, referral and management	January 2001	13	2	ONS (death registrations) National reporting scheme for specialists
Suspected fatal adverse drug reactions in children	To study the frequency and nature of suspected adverse drug reactions with a fatal outcome in children below the age of 16 years	June 2002	13	1	MHRA Yellow Card scheme
Congenital toxoplasmosis II	To determine the birth prevalence of congenital toxoplasmosis	July 2002	25	2	BOSU Toxoplasma referral laboratories
Invasive fungal infection in very low-birthweight infants	To describe the epidemiology of invasive fungal infection in very low-birthweight infants (<1500 g) in the UK	February 2003	13	2	HPA/SCIEH UK Mycology Reference Laboratory

AIDS, acquired immune deficiency syndrome; BOSU, British Ophthalmological Surveillance Unit; CDSC, Communicable Disease Surveillance Centre (now part of the Health Protection Agency); HIV, human immunodeficiency virus; HPA, Health Protection Agency; MHRA, Medicines and Healthcare Products Regulatory Agency; MMR, measles mumps and rubella vaccination; ONS, Office for National Statistics (formerly OPCS, Office of Population Census and Surveys); PHLS, Public Health Laboratory Service (now part of the Health Protection Agency); RCOG, Royal College of Obstetricians and Gynaecologists; SCIEH, Scottish Centre for Infection and Environmental Health (now HPS, Health Protection Scotland).

children with DKA, and the incidence of cerebral oedema in such children was reported through the BPSU.[32] Furthermore, the study of childhood blindness was undertaken primarily through the British Ophthalmological Surveillance Unit (BOSU) with the BPSU as an additional reporting source, ensuring that children with multiple disabilities, who were more likely to be managed principally by general paediatricians rather than ophthalmologists, were adequately ascertained.[26]

Capture–recapture analysis was attempted in six studies: congenital hip dislocation,[16] insulin-dependent diabetes in under 5-year-olds,[22] congenital cataract,[27] childhood blindness,[26] neonatal group B streptococcus infection[24] and human immunodeficiency virus/acquired immune deficiency

syndrome (HIV/AIDS),[33] although the last report was insufficiently detailed to permit review.

All studies (congenital cataract, congenital hip dislocation, diabetes and blindness) demonstrated matching of pairs between sources and defined a closed population. The group B streptococcal study used a novel reporting source, electronic reporting by parents through the website of a parent support group, but the published report does not discuss the accuracy of self-reported diagnoses.

Two studies appeared to have difficulty in fulfilling the criterion of homogeneity: congenital hip dislocation and group B streptococcal infection. The former compared reporting by orthopaedic surgeons with paediatric hip surgery data in case notes and Hospital Episode Statistics (HES) and found that, whilst HES data recorded cases across the whole range of severity, surgeons were more likely to report severe cases. An appropriate adjustment was made for heterogeneity before capture–recapture analysis was undertaken. In the group B streptococcal study, reporting bias related to social, economic and language factors is more likely with cases reported through the parent support group, as these relied on access to the support group and Internet.

Two studies reported dependence between reporting sources. A third reporting source (regional health data) was used to demonstrate negative dependence between paediatricians and diabetic nurses in diabetes surveillance: if a child was reported by one source, there was a lower probability of the same case being reported by the other source, suggesting that doctors and nurses working together were aware of each others reporting and adjusted their own reporting accordingly. A survey of clinicians contributing to the childhood blindness study suggested negative dependence between ophthalmologists and paediatricians, and although capture–recapture analysis was simulated, it was felt to be invalidated by source dependence. However, as the number of cases ascertained was significantly higher than predicted, collaboration and joint reporting, whilst precluding the use of capture–recapture analysis, appeared to improve the knowledge of the study and overall case reporting.

The congenital cataract study fulfilled all assumptions for two-source capture–recapture method[12] and estimated that 8% of true cases had been missed altogether, permitting the adjustment of total population estimates.

Discussion

Main findings of this study

The use of multiple sources to improve case ascertainment is an important feature of studies through the BPSU, occurring in 64% of reported studies. In approximately half of these studies, two or more additional sources were employed.

Three reasons for using secondary reporting sources in addition to a national paediatric surveillance system were identified:

1 to identify cases who would not normally be seen by clinicians reporting to the BPSU, i.e. to enhance ascertainment;
2 to define the denominator for the study and
3 to estimate the level of ascertainment of cases, by using two sources which would be expected to identify all cases, to adjust incidence accordingly.

The choice of additional source and method of data collection will be informed by the purpose; however, only health-related sources have thus far been used in studies through the BPSU. Data sources already well established to undertake passive or active surveillance were most readily used as secondary data sources and appeared to perform well. These included communicable disease surveillance and specialty-specific surveillance units. Data sources set up for specific studies, particularly involving specialty groups and case note reviews, frequently identified additional cases, whereas the involvement of other healthcare professionals, parent groups or the media proved less successful in identifying additional true cases. Specialist clinicians have also been added to the BPSU active monthly reporting system for the duration of some studies rather than being treated as an independent source.

What is already known on this topic

Maximizing case ascertainment is an important goal in surveillance studies, and capture–recapture methods have often been advocated as a means of improving and estimating completeness of ascertainment,[12] although Tilling also urges caution when applying them to epidemiological studies.[34] The use of three reporting sources has been previously reported as optimal for capture–recapture methods.[35] Capture–recapture analysis may therefore have been possible in some BPSU studies in which it was not undertaken; however, other studies appropriately avoided using capture–recapture, as there was doubt that the assumptions of independence and homogeneity held true.[26] Although these issues can sometimes be addressed at the analysis stage, some authors have cautioned against constructing very complex models for analysis to circumvent deficiencies in study design.[13,34] Capture–recapture is unlikely to be suited to disorders in which the diagnosis may be uncertain or defined differently by sources,[15] and this should be considered when parents or the media are the additional reporting source, as diagnosis either depends on contact with a health professional or remains 'unconfirmed'. Furthermore, if concerns about patient confidentiality or data protection limit the

identifiers permitted in a study,[15,36] inaccuracies in matching duplicate reports and unreliable estimates of population prevalence are more likely. Tilling has therefore suggested that single well-designed sources may be optimal for unbiased estimates of incidence, particularly as capture–recapture analysis is often difficult to apply in epidemiological studies because the necessary assumptions do not hold true.[34]

Using complementary dependent sources has been demonstrated to improve population coverage and case ascertainment in epidemiological studies,[37] and this supports the approach taken by the BPSU to include specialist clinicians in the monthly reporting system for the duration of a specific study.

Limitations of this study

There are marked variations in the reporting of methodology in the studies included in this review, which limited interpretation of the benefits and weaknesses of the multiple source approach. In particular, it was not possible to quantify the relative contribution of different sources because case ascertainment by source was not usually reported. It is also difficult to evaluate the added benefit that an additional source might provide, in terms of improving completeness of ascertainment, when the total population number is not known. Furthermore, the number and characteristics of sources varied by study, which undoubtedly contributed to their impact. Detailed reporting of completed studies and a better understanding of methodology could improve future surveillance studies.

What this study adds

This review of the experience of a national surveillance system demonstrates that multiple sources can provide a valuable means for improving case ascertainment in the surveillance of uncommon paediatric disorders. However, the choice of additional reporting sources and study design should be suited to the specific requirements of each study and an intention to apply capture–recapture analysis determined at the outset. Our review suggests that supplementing the BPSU reporting database may be a more feasible approach to improve overall case ascertainment than attempting to define an independent source suited to capture–recapture analysis. Importantly, our review also suggests that active surveillance through a monthly national reporting system involving general paediatricians is timely and achieves high ascertainment such that it remains a reliable primary source of cases of uncommon disorders of childhood. The BPSU provides a simple and unified mechanism for the conduct of high-quality paediatric surveillance studies with important implications for public health.

References

1 Nicoll A, Lynn R, Rahi J *et al*. Public health outputs from the British Paediatric Surveillance Unit and similar clinician-based systems. *J R Soc Med* 2000;**93**:580–5.

2 Hall SM, Nicoll A. The British Paediatric Surveillance Unit – a pioneering method for investigating the less common disorders of childhood. Report of a seminar held in June 1995. *Child Care Health Dev* 1998;**24**:129–43.

3 Teutsch SM, Thacker SB. Planning a public health surveillance system. *Epidemiol Bull* 1995;**16**:1–6.

4 Sekar CC, Deming WE. On a method of estimating birth and death rates and the extent of registration. *J Am Stat Assoc* 1949;**44**:101–15.

5 Wittes JT, Colton T, Sidel VW. Capture–recapture methods for assessing the completeness of case ascertainment when using multiple information sources. *J Chronic Dis* 1974;**27**:25–36.

6 Tilling K, Sterne JAC, Wolfe CDA. Estimation of the incidence of stroke using a capture–recapture model including covariates. *Int J Epidemiol* 2001;**30**:1351–9.

7 Bernillon P, Lievre L, Pillonel J *et al*. Record linkage between two anonymous databases for a capture–recapture estimation of underreporting of AIDS cases: France 1990–1993. *Int J Epidemiol* 2000;**29**:168–74.

8 Capture–recapture and multiple-record systems estimation I: History and theoretical development. International Working Group for Disease Monitoring and Forecasting. *Am J Epidemiol* 1995;**142**:1047–58.

9 Capture–recapture and multiple-record systems estimation II: Applications in human diseases. International Working Group for Disease Monitoring and Forecasting. *Am J Epidemiol* 1995;**142**:1059–68.

10 Crowcroft NS, Andrews N, Rooney C *et al*. Deaths from pertussis are underestimated in England. *Arch Dis Child* 2002;**86**:336–8.

11 Fortnum HM, Summerfield AQ, Marshall DH *et al*. Prevalence of permanent childhood hearing impairment in the United Kingdom and implications for universal neonatal hearing screening: questionnaire based ascertainment study. *BMJ* 2001;**323**:536–40.

12 Hook EB, Regal RR. Capture–recapture methods in epidemiology: methods and limitations. *Epidemiol Rev* 1995;**17**:243–64.

13 Cormack RM, Chang YF, Smith GS. Estimating deaths from industrial injury by capture–recapture: a cautionary tale. *Int J Epidemiol* 2000;**29**:1053–9.

14 Jarvis SN, Lowe PJ, Avery A *et al*. Children are not goldfish – mark/recapture techniques and their application to injury data. *Inj Prev* 2000;**6**:46–50.

15 Nanan DJ, White F. Capture–recapture: reconnaissance of a demographic technique in epidemiology. *Chronic Dis Can* 1997;**18**:144–8.

16 Godward S, Dezateux C. Surgery for congenital dislocation of the hip in the UK as a measure of outcome of screening. MRC Working Party on Congenital Dislocation of the Hip. Medical Research Council. *Lancet* 1998;**351**:1149–52.

17 Tookey PA, Peckham CS. Surveillance of congenital rubella in Great Britain, 1971–96. *BMJ* 1999;**318**:769–70.

18 Mucklow ES. Chemistry set poisoning. *Int J Clin Pract* 1997;**51**: 321–3.

19 Foot B, Stanford M, Rahi J *et al*. The British Ophthalmological Surveillance Unit: an evaluation of the first 3 years. *Eye* 2003;**17**:9–15.

20 Sawczenko A, Sandhu BK, Logan RF *et al.* Prospective survey of childhood inflammatory bowel disease in the British Isles. *Lancet* 2001;**357**:1093–4.

21 Symmons DP, Sills JA, Davis SM. The incidence of juvenile dermato-myositis: results from a nation-wide study. *Br J Rheumatol* 1995;**34**:732–6.

22 Wadsworth E, Shield J, Hunt L *et al.* Insulin dependent diabetes in children under 5: incidence and ascertainment validation for 1992. *BMJ* 1995;**310**:700–3.

23 Macdougall CF, Cant AJ, Colver AF. How dangerous is food allergy in childhood? The incidence of severe and fatal allergic reactions across the UK and Ireland. *Arch Dis Child* 2002;**86**:236–9.

24 Heath PT, Balfour G, Weisner AM *et al.* Group B streptococcal disease in UK and Irish infants younger than 90 days. *Lancet* 2004;**363**:292–4.

25 Rahi JS, Dezateux C. National cross sectional study of detection of congenital and infantile cataract in the United Kingdom: role of childhood screening and surveillance. The British Congenital Cataract Interest Group. *BMJ* 1999;**318**:362–5.

26 Rahi JS, Cable N. Severe visual impairment and blindness in children in the UK. *Lancet* 2003;**362**:1359–65.

27 Rahi JS, Dezateux C. Capture–recapture analysis of ascertainment by active surveillance in the British Congenital Cataract Study. *Invest Ophthalmol Vis Sci* 1999;**40**:236–9.

28 Godward S, Dezateux C. Validation of the reporting bases of the ortho-paedic and paediatric surveillance schemes. *Arch Dis Child* 1996;**75**:232–6.

29 Verity CM, Nicoll A, Will RG *et al.* Variant Creutzfeldt-Jakob disease in UK children: a national surveillance study. *Lancet* 2000;**356**:1224–7.

30 Booy R, Heath PT, Slack MP *et al.* Vaccine failures after primary immunisation with *Haemophilus influenzae* type-b conjugate vaccine without booster. *Lancet* 1997;**349**:1197–202.

31 Kemp A, Sibert JR. Drowning and near drowning in children in the United Kingdom: lessons for prevention. *BMJ* 1992;**304**:1143–6.

32 Edge JA, Hawkins MM, Winter DL *et al.* The risk and outcome of cerebral oedema developing during diabetic ketoacidosis. *Arch Dis Child* 2001;**85**:16–22.

33 Ades AE, Davison CF, Holland FJ *et al.* Vertically transmitted HIV infection in the British Isles. *BMJ* 1993;**306**:1296–9.

34 Tilling K. Capture–recapture methods – useful or misleading? *Int J Epidemiol* 2001;**30**:12–4.

35 Ismail AA, Beeching NJ, Bellis MA. How many data sources are needed to determine diabetes prevalence by capture–recapture? *Int J Epidemiol* 2000;**29**:536–41.

36 Nicoll A. Protecting health and patient confidentiality, ethics and surveillance. *Curr Paediatr* 2005;**15**:581–9.

37 Gill GV, Ismail AA, Beeching NJ *et al.* Hidden diabetes in the UK: use of capture–recapture methods to estimate total prevalence of diabetes mellitus in an urban population. *J R Soc Med* 2003;**96**:328–32.