Epidemiol. Infect. (2015), **143**, 2570–2579. © Cambridge University Press 2015
doi:10.1017/S0950268814003355

Using linked birth, notification, hospital and mortality data to examine false-positive meningococcal disease reporting and adjust disease incidence estimates for children in New South Wales, Australia

A. GIBSON[1]*, L. JORM[1] AND P. McINTYRE[2]

[1] *Centre for Health Research, University of Western Sydney, NSW, Australia*
[2] *National Centre for Immunisation Research and Surveillance, Kids Research Institute, at The Children's Hospital at Westmead, NSW, Australia*

Received 16 April 2014; Final revision 1 July 2014; Accepted 19 November 2014; first published online 9 January 2015

SUMMARY

Meningococcal disease is a rare, rapidly progressing condition which may be difficult to diagnose, disproportionally affects children, and has high morbidity and mortality. Accurate incidence estimates are needed to monitor the effectiveness of vaccination and treatment. We used linked notification, hospital, mortality and birth data for all children of an Australian state (2000–2007) to estimate the incidence of meningococcal disease. A total of 595 cases were notified, 684 cases had a hospital diagnosis, and 26 cases died from meningococcal disease. All deaths were notified, but only 68% (466/684) of hospitalized cases. Of non-notified hospitalized cases with more than one clinical admission, most (90%, 103/114) did not have meningococcal disease recorded as their final diagnosis, consistent with initial 'false-positive' hospital meningococcal disease diagnosis. After adjusting for false-positive rates in hospital data, capture–recapture estimation suggested that up to four cases of meningococcal disease may not have been captured in either notification or hospital records. The estimated incidence of meningococcal disease in NSW-born and -resident children aged 0–14 years was 5·1–5·4 cases/100 000 child-years at risk, comparable to international estimates using similar methods, but lower than estimates based on hospital data.

Key words: Meningococcal disease, notifiable infectious diseases, notifications, paediatrics, surveillance.

INTRODUCTION

Meningococcal disease, caused by infection with the bacterium *Neisseria meningitidis* [1], is a life-threatening disease which may be associated with a rapidly progressive haemorrhagic rash, meningitis, septicaemia and death. Rarer forms of the disease include conjunctivitis, otitis media, epiglottitis, urethritis, arthritis and pericarditis [2]. Children aged <5 years have the highest rates of notification for meningococcal disease [3] and the highest death rates (0·58 deaths/100 000 Australian population in 2005–2007 [4]).

In Australia in 2011, nearly 84% of meningococcal disease cases were serogroup B, with serogroup C accounting for only 4% [5]. Since the introduction of a vaccine for group C meningococcal disease in 2003 [6] there has been a significant and sustained reduction in cases of group C disease [5]. Greater prevention of meningococcal disease will not be achievable without a vaccine against group B disease [7]. An expensive serogroup B vaccine has been approved for use in Australia [8], but it is not yet approved for funding under the National Immunization Programme. This

* Author for correspondence: Dr A. Gibson, Centre for Big Data Research in Health, Level 1, AGSM Building, UNSW Australia, UNSW, Sydney, NSW, Australia 2052.
(Email: amy.gibson@unsw.edu.au)

same vaccine has been rejected for routine use in the UK because of low and uncertain cost-effectiveness [9], while restricted use in children at a cost-effective price is being negotiated [10].

Accurate reporting of meningococcal disease is essential in order to monitor the impact and cost-efficacy of public health policy including vaccination programmes, and provide an appropriate public health response to cases and their close contacts [11, 12]. Yet, meningococcal disease poses some unique challenges for monitoring systems. It is a rare disease; in 2012 only 65 cases were reported in New South Wales (NSW), Australia (population over 7 million) [13], therefore many person-years of data are needed to accurately estimate its incidence. Moreover, infection can progress from initial symptoms to death within hours, and many early cases of meningococcal disease present with only non-specific symptoms [14], which can lead to both misdiagnosis and over-diagnosis [15]. Many cases are referred to specialist centres for assessment, which can complicate hospital meningococcal diagnosis reporting. These challenges mean that even the most rigorous of recording systems, such as notifiable disease registers, can be inaccurate [16]. Paediatric meningococcal diagnoses might be particularly prone to higher rates of false-positive reporting [11]. Accurate estimates of disease incidence must both minimize case under-ascertainment and quantify the proportion of false-positive diagnoses [17].

In NSW, several databases record cases of meningococcal disease. These include notification data (the Notifiable Conditions Information Management System), hospitalization data (the Admitted Patient Data Collection), and mortality data (from the Australian Bureau of Statistics; ABS). The current study used linked notification, hospitalization and mortality data from 2000–2007 inclusive for all NSW children aged 0–14 years to: (i) estimate rates of false-positive meningococcal disease diagnoses in hospital records; (ii) quantify case under-ascertainment using capture–recapture estimation methods; and (iii) estimate the incidence of meningococcal disease, adjusted for false-positive reporting and case under-ascertainment.

METHODS

Databases and data linkage

This study used four health and administrative datasets:

- The *NSW Perinatal Data Collection* holds records of all births in NSW, and was used to calculate the population at risk for incidence rate calculations. Records for live births from 1 January 1994 to 31 December 2007 were used.
- The *NSW Notifiable Conditions Information Management System* is a register of cases of infectious and other conditions specified in the NSW Public Health Act 2010, including confirmed and probable cases of meningococcal disease, that are notified to the NSW Ministry of Health. Confirmed cases of meningococcal disease require either (1) laboratory-confirmed evidence or (2) laboratory suggestive evidence in combination with clinical evidence. Probable cases require only clinical evidence, including the absence of evidence for other causes of clinical symptoms and either (*a*) clinically compatible disease or (*b*) clinically compatible disease and recent close contact with a confirmed case [4]. Cases are recorded as a particular serogroup, serogroup 'NOS' ('not otherwise specified') or serogroup 'not typable'. Records for cases for NSW-resident children (0–14 years) with estimated disease onset dates from 1 July 2000 to 31 December 2007 were used.
- The *NSW Admitted Patient Data Collection* records all inpatient separations (discharge, change of service category, death or transfer) from all public, private, psychiatric and repatriation hospitals in NSW. Meningococcal disease was identified on the basis of the International Statistical Classification of Diseases and Related Health Problems, Tenth revision (ICD-10) diagnostic codes A39·0–A39·9 inclusive in any diagnosis field. Diagnoses are coded for each new separation. A patient may have a number of separations corresponding to a single episode of illness if they are transferred between hospitals, have a change of service category within a single hospital, or are admitted several times. Separations were considered to be part of the same illness episode if no more than 24 h elapsed between discharge from one hospital separation and admission into the next. Admission dates of 1 July 2000 to 31 December 2007 were used for NSW-resident children aged 0–14 years.
- The *Australian Bureau of Statistics* collates coronial information and mortality data from the State Registry of births, deaths and marriages. Deaths from meningococcal disease were identified on the basis of ICD-10 codes A39·0–A39·9 inclusive in any of the cause- of-death fields. Records for deaths of NSW-resident children (0–14 years) occurring between 1 July 2000 and 31 December 2007 were used.

The datasets were linked together probabilistically by the NSW Centre for Health Record Linkage (CHeReL) using personal identifiers, including full name, date of birth, sex and address. The CHeReL's probabilistic linkage procedures are designed to achieve around 5/1000 incorrect links and 5/1000 missed links [18]. Following data linkage, each child (with a unique identifier) had a set of data records which could include notification/s, hospitalization/s, birth and death records. Only de-identified data were analysed by researchers.

Ethical standards

Ethics approval was received from the NSW Population and Health Services Research Ethics Committee (Ref. No. 2009/11/193) and the University of Western Sydney Human Research Ethics Committee (Ref. No. H10651). The authors assert that all procedure contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

Study population

The study population was all NSW-resident children aged between 0 and 14 years (a mean of 1 324 925 children/year) who had the potential to have notification, hospital separation or death recorded between 1 July 2000 and 31 December 2007. The Perinatal Data Collection recorded a total of 1 223 313 children born alive in NSW between 1 January 1994 and 31 December 2007.

Analyses

All analysis was performed using SAS v. 9·2 (SAS Institute Inc., USA). The proportion of false-positive reports of meningococcal disease in hospital data was estimated by examining those children with an additional hospital/mortality record after the original hospital admission. Comparisons between subjects were performed using χ^2 and two-sample t tests. Capture–recapture analysis was performed using the Chapman estimator [19] for notification and hospitalization data. Capture–recapture estimation using three datasets (notification, hospitalization, mortality) was performed with the GENMOD procedure using a log-linear Poisson regression model [20].

The remoteness of residence was defined by using the mean Accessibility/Remoteness Index of Australia (ARIA) value corresponding to a subject's statistical local area of residence [21] in NSW. Mean ARIAs <0·20 were considered to be major cities, between 0·20 and 2·40 were inner regional areas, and >2·40 were outer regional and more remote areas. A subject's sex and Aboriginal/Torres Strait Islander status were taken from the first hospital record where meningococcal disease was diagnosed.

Approximate yearly incidence rates between 2000 and 2007 (per 100 000 child population) were calculated by dividing the number of notified cases of meningococcal disease in NSW-resident children by the NSW child population (0–14 years). This incidence rate uses all available data to maximize the number of recorded cases of this rare disease. To calculate an incidence rate/100 000 child-years at risk, the dataset was limited to only those children resident and born in NSW (i.e. those with a perinatal data collection record) to give both the total child-years at risk of meningococcal exposure and the number of notified cases of meningococcal disease in this subset of children. Incidence rates were presented both as cases/100 000 child population and cases/100 000 child-years at risk to facilitate comparisons with other published works. Incidence rates were adjusted for the estimated false-positive rate in hospital data, the degree of under-ascertainment found through capture–recapture analysis, and the estimated interstate migration rate for children.

RESULTS

A total of 595 cases of meningococcal disease were notified between July 2000 and December 2007. No subject had more than one notification of meningococcal disease. The majority of meningococcal disease cases [343 (66%) of 522 laboratory-confirmed cases] were serogroup B. Ninety cases (17%) were classified as serogroup NOS, 66 cases (13%) were serogroup C, 17 cases (3%) were 'not typable', and six cases (1%) were one of the less common serogroups (groups W, Y or 29E).

There were 684 children who had one or more episodes of hospitalization coded as meningococcal disease between July 2000 and December 2007, of whom 466 (68%) also had a notification of meningococcal disease. ABS mortality data recorded 26 children who died from meningococcal disease. A total of 813 cases were identified as having meningococcal disease in at least one of the three datasets. Figure 1 illustrates the overlap between cases of reported meningococcal disease identified in the three datasets.

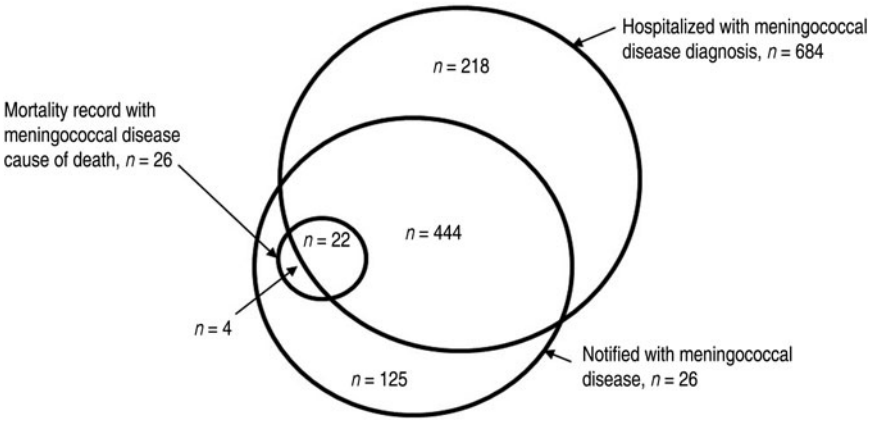

Fig. 1. All 813 cases recorded in notification, hospitalization, and/or mortality data as meningococcal disease in children, July 2000–December 2007. (Figure is not to scale.)

Of the 466 children both hospitalized and notified with meningococcal disease, 58% were male, 12% were resident in outer regional or more remote areas of NSW, and 7% were Aboriginal or Torres Strait Islander.

Fatal cases of meningococcal disease

All 26 deaths with a cause of death listed as meningococcal disease had a notification of meningococcal disease, although four of these children did not have an associated hospital admission (see Fig. 1). In two of these non-hospitalized fatal cases, the causes of death included Waterhouse–Friderichsen syndrome, or fulminant meningococcal disease, the most severe presentation of the disease [22]. It is possible that some of these children died before they could be admitted to hospital.

Estimating false-positive disease reporting by observing multiple clinical records for single subjects

Just over one-third (38%, 257/684) of cases with a hospital diagnosis of meningococcal disease had at least one further clinical record (all had hospital admission/s and 24 cases had a death record) which followed the initial meningococcal hospital admission within a period of 24 h. The 466 notified cases with either single ($n = 323$) or multiple ($n = 143$, including 22 deaths) clinical records were considered 'proven cases' of meningococcal disease (group 1). 'False-positive cases' were not notified, had multiple clinical records, and lacked a diagnosis of meningococcal disease on the final record (group 2, $n = 103$ including two deaths). These represented 90% (103/114) of all non-notified cases with multiple clinical records. Non-notified subjects with a single clinical record ($n = 104$) or with multiple clinical records where the final clinical record included a diagnosis of meningococcal disease ($n = 11$) could not have their status determined and were classified as 'unknown' (group 3, $n = 115$). Such subjects (90% of whom had a single clinical admission) may represent either proven cases who were not notified, or cases who did not have a subsequent clinical record and the opportunity to revise the diagnosis. Table 1 demonstrates significant differences between group 3 'unknown' cases and either groups 1 'proven' or 2 'false-positive' subjects. Group 3 subjects were not significantly different from either groups 1 or 2 subjects by sex, remoteness of residence, time in intensive care, or Aboriginal or Torres Strait islander status.

A single NSW hospital accounted for 40% (41/103) of all group 2 false-positive cases of meningococcal disease during the study period. All other NSW hospitals recorded no more than eight group 2 cases each.

The most frequent primary diagnoses in the final clinical record for those where meningococcal disease was excluded by the final clinical record (group 2, $n = 103$) were: viral infection, unspecified ($n = 40$), unspecified viral infection with skin and mucous membrane lesions ($n = 7$), acute upper respiratory infection, unspecified ($n = 5$) and fever, unspecified ($n = 5$).

Using capture–recapture methods to estimate under-ascertainment of meningococcal disease

Capture–recapture methods assume that subjects are equally likely to be 'captured' by the individual datasets [23], which would require that the same definition

Table 1. *Significant differences between group 1 (proven), group 2 (false-positive) and group 3 (unknown) subjects*

Categorical variables*	Group 1 'proven cases' ($N = 466$) *n* (%)	Group 2 'false-positive cases' ($N = 103$) *n* (%)	Group 3 'unknown cases' ($N = 115$) *n* (%)	Significance tests
Age group				**Group 1 *vs.* 3**
0–5 months	80 (17·2)	18 (17·5)	16 (13·9)	$\chi^2 = 14{\cdot}5$, 5 D.F., $P = 0{\cdot}013$
6–11 months	70 (15·0)	7 (6·8)	9 (7·8)	**Group 2 *vs.* 3**
1–2 years	120 (25·8)	27 (26·2)	24 (20·9)	$\chi^2 = 6{\cdot}72$, 5 D.F., $P = 0{\cdot}24$
3–4 years	61 (13·1)	22 (21·4)	17 (14·8)	
5–6 years	46 (9·9)	9 (8·7)	10 (8·7)	
⩾7 years	89 (19·1)	20 (19·4)	39 (33·9)	
Presenting hospital role†				**Group 1 *vs.* 3**
Paediatric specialist	221 (47·5)	17 (16·5)	54 (47·0)	$\chi^2 = 5{\cdot}33$, 2 D.F., $P = 0{\cdot}070$
Major hospital	140 (30·1)	12 (11·7)	25 (21·7)	**Group 2 *vs.* 3**
District/other	104 (22·4)	74 (71·8)	36 (31·3)	$\chi^2 = 36{\cdot}43$, 2 D.F., $P < 0{\cdot}0001$
Presenting symptoms				**Group 1 *vs.* 3**
No meningitis/sepsis	41 (8·8)	53 (51·5)	39 (33·9)	$\chi^2 = 50{\cdot}58$, 2 D.F., $P < 0{\cdot}0001$
Meningitis only present	176 (37·8)	12 (11·7)	25 (21·7)	**Group 2 *vs.* 3**
Sepsis present	249 (53·4)	38 (36·9)	51 (44·4)	$\chi^2 = 7{\cdot}96$, 2 D.F., $P = 0{\cdot}019$
Year of admission				**Group 1 *vs.* 3**
2000	54 (11·6)	6 (5·8)	19 (16·5)	$\chi^2 = 4{\cdot}77$, 7 D.F., $P = 0{\cdot}69$
2001	71 (15·3)	8 (7·8)	15 (13·0)	**Group 2 *vs.* 3**
2002	91 (19·5)	10 (9·7)	20 (17·4)	$\chi^2 = 20{\cdot}96$, 7 D.F., $P = 0{\cdot}0038$
2003	67 (14·4)	17 (16·5)	12 (10·4)	
2004	54 (11·6)	10 (9·7)	17 (14·8)	
2005	47 (10·1)	16 (15·5)	14 (12·2)	
2006	34 (7·3)	20 (19·4)	8 (7·0)	
2007	48 (10·3)	16 (15·5)	10 (8·7)	
Continuous variable*	Mean (S.D.)	Mean (S.D.)	Mean (S.D.)	
Total length of stay	8·3 days (11·8)	3·1 days (3·5)	5·1 days (5·7)	**Group 1 *vs.* 3** $t = 4{\cdot}25$, $P < 0{\cdot}0001$ **Group 2 *vs.* 3** $t = 3{\cdot}17$, $P = 0{\cdot}0018$
Total time on mechanical ventilation	7·7 h (29·8)	1·3 h (7·6)	2·1 h (13·1)	**Group 1 *vs.* 3** $t = 3{\cdot}0$, $P = 0{\cdot}0027$ **Group 2 *vs.* 3** $t = 0{\cdot}56$, $P = 0{\cdot}58$

Group 1 'proven' cases: those notified as having meningococcal disease. May include subjects with single or multiple clinical records.
Group 2 'false-positive' cases: non-notified subjects who had been discharged without a meningococcal diagnosis recorded on their final clinical record. Only includes subjects with multiple clinical records.
Group 3 'unknown' cases: non-notified subjects who had been discharged with a meningococcal diagnosis on their final (or only) clinical record. May include subjects with either single or multiple clinical records.
* Analysis of all categorical variables is based on the first hospital separation for the illness episode. Continuous variables are summed over all hospital separations in the illness episode.
† One missing value in group 1.

of meningococcal disease in hospital and notification data. As this assumption was not met, we adjusted our capture–recapture estimates for the greater proportion of false-positive cases in hospital data.

Assuming that the 'false-positive' rate of 90% that we identified in non-notified cases with multiple hospital records was applicable to all non-notified hospitalized cases ($n = 218$), we estimated that up to 22 of these may actually be 'true' (confirmed or probable) cases of meningococcal disease. Table 2 shows capture–recapture estimates [19] for the true number of cases of meningococcal disease, after adjustment

Table 2. *Cases 'captured' and 'not captured' by notification and hospitalization data, after adjustment for estimated false-positive reporting in hospital data*

	Hospitalization data		
Notification data	Captured	Not captured	All cases
Captured	466 ($n_{N,A}$)	129	595 (n_N)
Not captured	0 to *22**	0 to *6*	0 to *28*
All cases	466 to *488* (n_A)	129 to *135*	(N) 595 to *623*

Non-italicized numbers have been observed in the data. Italicized numbers have been calculated based on assumptions described in the text. N was calculated using the Chapman estimator (here showing the upper limit):

$$N = \frac{(n_N + 1)(n_A + 1)}{(n_{N,A} + 1)} - 1;\ N = \frac{596 \times 489}{467} - 1;\ N = 623.$$

* An estimated upper limit of 22 true cases of meningococcal disease which were admitted to hospital, but not recorded in notification data

for the estimated false-positive rate in the hospitalization data.

Capture–recapture methods estimated that up to six cases of meningococcal disease in children aged between 0 and 14 years might not have been reported in either hospitalization or notification data between July 2000 and December 2007. The capture–recapture estimate for the total number of cases of meningococcal disease in this period ranged from 595 to 623.

For the subset of NSW-born and -resident children included in Table 3, capture–recapture methods estimated that up to four cases of meningococcal disease may not have been captured by either hospitalization or notification data in the study period, and the total number of cases ranged from 435 and 455.

Incidence rates

The incidence of notified meningococcal disease calculated using the total annual population of NSW-resident children aged 0–14 years averaged 5·6 (95% CI 4.4–7.0) cases/100 000 child population over the study period. Annual incidence declined from 7·7 (95% CI 6·3–9·4) cases/100 000 child population in 2001 to 4·1 (95% CI 3·1–5·3) cases/100 000 child population in 2007. Annual incidence of serogroup B also decreased from 3·4 (95% CI 2·5–4·6) cases/100 000 child population to 2·9 (95% CI 2·0–3·9) cases/100 000 child population over the same time

Table 3. *Cases 'captured' and 'not captured' by notification and hospitalization data, after adjustment for estimated false-positive reporting in hospital data in the subset of NSW-born and -resident children*

	Hospitalization data		
Notification data	Captured	Not captured	All cases
Captured	356 ($n_{N,A}$)	79	435 (n_N)
Not captured	0 to *16*	0 to *4*	0 to *20*
All cases	356 to *372* (n_A)	79 to *83*	(N) 435 to *455*

Non-italicized numbers have been observed in the data. Italicized numbers have been calculated based on assumptions described in the text. N was calculated using the Chapman estimator.

period. Throughout the years, the highest incidence of meningococcal disease occurred in children aged <5 years (Fig. 2).

In the subset of NSW-born and -resident children the linked datasets contributed 8 456 477 child-years at risk. Using the capture–recapture estimates for the number of cases in hospital and notification data, adjusted for false-positive reporting in hospital data (435–455 cases) yielded an incidence density estimate of between 5·1 and 5·4 cases of meningococcal disease/100 000 child-years at risk in NSW children aged 0–14 years over the study period.

The likely impact of interstate migration was estimated indirectly, using previously published figures. Between 2001–2002 and 2010–2011, NSW had an average net population loss of 23 349 persons/year [24], and between 5% and 8% of these people (about 1500/year) were aged between 0–14 years [25], so that about 9750 child-years were lost to interstate migration over the study period. Subtracting this figure from the total child-years at risk left the incidence density estimate unchanged. Other international population-based estimates of the incidence of meningococcal disease in children are given in Table 4.

DISCUSSION

Due to its rarity, diversity of presentations, and difficulties in diagnosis, we tested the possibility that cases of meningococcal disease were under-ascertained in existing NSW recording systems using capture–recapture analysis of linked administrative datasets. We used linked notification, hospitalization and mortality data to investigate potential under- and over-

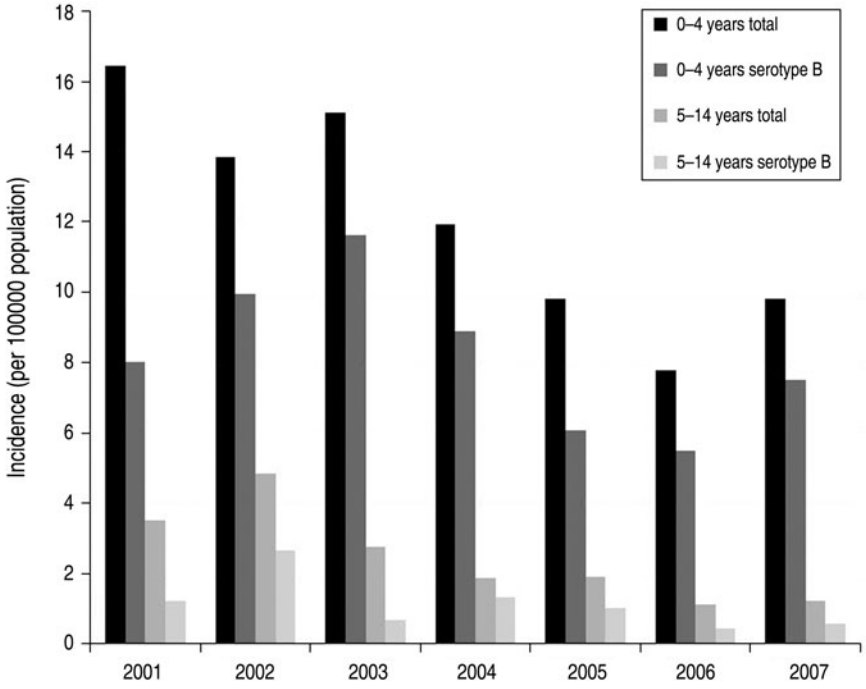

Fig. 2. Crude incidence of meningococcal disease in NSW children by year.

Table 4. *International estimates for the incidence of meningococcal disease in selected populations*

Setting	Population	Year	Incidence
Australia [4]	0–4 years	2006–2007	8·4/100 000 person-years
	5–14 years		1·1/100 000 person- years
Western Australia [28]	0–13 years	1990–1995	5·2 to 10·5/100 000 person-years
Five Australian states [30]	0–4 years, non-indigenous	2003–2006	9·2/100 000 person-years
	5–14 years, non-indigenous		1·7/100 000 person-years
Denmark [29]	All ages	1980–1999	4·3/100 000 person-years at risk
Barcelona, Spain [22]	<1 year	1987–1992	98·7/100 000 person-years
	1–4 years		65·0/100 000 person-years
	5–9 years		22·0/100 000 person-years
	10–14 years		5·6/100 000 person-years
An English electoral ward [31]	0–4 years, boys	1991–1999	41·0/100 000 person-years
	0–4 years, girls		33·7/100 000 person-years

ascertainment of cases of meningococcal disease in these recording systems, and to produce population-based estimates of the incidence of meningococcal disease. We identified high levels of capture in notification data of fatal cases (100%), and around 15% of children hospitalized with an initial diagnosis of meningococcal disease were 'false-positive' cases in whom the diagnosis was not subsequently confirmed. Incidence rates calculated using adjusted capture–recapture estimates and person-time denominators derived from linked data did not differ appreciably from those calculated using notification data only and ABS population estimates.

Completeness of notification data

Our capture–recapture methods, using hospital and notification data and correcting for estimated false-positive reporting in hospital data, demonstrated that most confirmed and probable cases of paediatric meningococcal disease in NSW are reported to the Notifiable Conditions Information Management

System. An estimated maximum of four cases of meningococcal disease in NSW-born and -resident children may have gone unreported in hospital or notification data between July 2000 and December 2007, less than one case/year.

This is an improvement on meningococcal disease surveillance conducted between 1990 and 1995 in Victoria. That study estimated seven cases unidentified in children aged <2 years, but complete ascertainment for children aged 10–14 years [26]. During this period, no culture was available for 38% of cases [26] compared to our study in the following decade, where more accurate laboratory diagnosis methods using PCR were in use, and many notifications were coming directly from the laboratories themselves.

False-positive reporting in hospital data

One of the assumptions of capture–recapture techniques is that databases have the same disease reporting criteria (which implies similar false-positive reporting rates) [23]. Hospital and notification data do not meet this assumption: hospital data records 'possible' cases in addition to 'confirmed' and 'probable' cases of meningococcal disease, so hospital data has a greater proportion of false-positive cases than notification data. While other studies have used clinical review to identify false-positive diagnoses [11], this is not possible using de-identified data linkage files. Instead, we examined the subset of non-notified patients with multiple clinical records and estimated that at least 90% were false-positive cases of meningococcal disease. This is a novel approach to estimating the degree of false-positive reporting in linked data.

To examine our assumption that group 3 'unknown' cases more closely resembles group 2 'false-positive' cases, and was less similar to group 1 'proven' cases of meningococcal disease, we refer to Table 1. In the clinical variables of both age group and time on mechanical ventilation, there was a significant difference between groups 1 and 3 (but not between groups 2 and 3), which would seem to imply that group 3 unknown cases were not cases of meningococcal disease. The significant difference in ages between groups 1 and 3 could demonstrate the greater incidence of meningococcal disease in young children (as per Fig. 2) in proven (group 1), but not unknown (group 3) cases.

In the clinical variables of presenting symptoms and length of stay, a gradient of severity was observed, but group 3 was significantly different to both groups 1 and 2, providing no further evidence to reveal the identity of group 3 subjects. For both of these variables, group 1 (proven) cases were the most severe, group 3 (unknown) cases had intermediate severity, and group 2 (false-positive) cases were less severe.

Two of the significant differences between groups 2 and 3 (the role of the presenting hospital and the year of presentation) were not clinical in nature. These differences could reflect changing policies on the referral of, and the lower capacity of smaller district hospitals to deal with cases of suspected meningococcal disease. Hospital-level variation between groups 2 and 3 can been seen clearly in the proportion of false-positive group 2 cases by hospital. A single NSW district hospital accounted for 40% of all non-notified, false-positive cases of meningococcal disease in NSW. The emergency department of this particular district hospital saw over 10 000 paediatric patients per year, yet had no paediatric unit/ward at the time and all paediatric admissions were transferred to a larger nearby hospital for assessment and treatment [27]. These final non-meningococcal diagnoses were predominately unspecified viral infections, which is consistent with false-positive meningococcal disease reporting.

Dependence of hospital and notification data

Capture–recapture methods also assume that the relevant datasets are independent, so that subjects 'captured' in one dataset are not more or less likely to appear in another dataset [23]. In our sample, there is a positive dependence between hospitalization and notification data, in that hospitalized cases are more likely to be notified. Nearly half (47%) of first notifications for meningococcal disease came from a hospital, followed by 30% from laboratories, 7% from a doctor, and 16% from other sources. This positive dependence is likely to bias the capture–recapture estimate towards underestimation [23]. We did attempt to adjust for this dependence between the databases by using a third (mortality) database and a log-linear capture–recapture model [20]. However, as all fatal cases of meningococcal disease were captured by notification data, mortality data contributed no additional information to the estimate.

Incidence of meningococcal disease

The average incidence of notified meningococcal disease in NSW children was 5·6 (95% CI 4·4–7·0) cases/100 000 child population over the study period, and this incidence decreased over the study period.

In 2007, there were 4·1 cases (95% CI 3·1–5·3) of meningococcal disease and 2·9 cases (95% CI 2·0–3·9) of meningococcal group B/100 000 child population. Using our capture–recapture estimate and the subset of NSW-born and -resident children, there was an adjusted incidence density of between 5·1 and 5·4 cases of meningococcal disease/100 000 child-years at risk (July 2000–December 2007), which was very similar to our crude estimate from notified cases only. While not directly measured in this study, levels of interstate child migration did not have a significant impact on these estimates. These incidence rates are comparable to others from Australia and Denmark where meningococcal disease notification systems and/or record linkage were used [4, 28–30], but much lower than incidence rates in the UK calculated using only hospital diagnoses [31] or Spanish rates which included all meningococcal diagnoses, whether identified from hospital or laboratory data [22] (see Table 4).

CONCLUSIONS

Obtaining accurate measures of disease incidence, including rare diseases with a difficult diagnosis such as meningococcal disease, is important so that we can best assess the impact and cost efficacy of interventions such as immunization. Our investigation shows that linked data can be used to estimate the proportion of false-positive disease reporting in hospital data where clinical review is not possible. Capture–recapture estimates show that paediatric meningococcal disease was well recorded in NSW notification data, and we obtained an internationally comparable estimate of the incidence of meningococcal disease in NSW children.

ACKNOWLEDGEMENTS

The authors acknowledge Brett Archer for comments on an earlier version of this manuscript. This project was funded by the National Health and Medical Research Council (grant no. 573 122).

DECLARATION OF INTEREST

None.

REFERENCES

1. **Granerod J, *et al.*** Investigating the aetiology of an evaluating the impact of the Men C vaccination programme on probable meningococcal disease in England and Wales. *Epidemiology and Infection* 2006; **134**: 1037–1046.
2. **Rosenstein N, *et al.*** Meningococcal disease. *New England Journal of Medicine* 2001; **344**: 1378–1388.
3. **Ladhani S, *et al.*** invasive meningococcal disease in England and Wales: implications for the introduction of new vaccines. *Vaccine*, 2012; **30**: 3710–3716.
4. **Chiu C, *et al.*** Vaccine preventable diseases in Australia, 2005 to 2007. *Communicable Diseases Intelligence*, 2010; **34**.
5. **Lahra M, Enriquez R.** Annual report of the Australian Meningococcal Surveillance Programme, 2011. *Communicable Diseases Intelligence* 2012; **36**: E251–E262.
6. **Cohen N.** Introduction of the National Meningococcal C Vaccination Program. *Communicable Diseases Intelligence* 2003; **27**: 161–162.
7. **Girard M, *et al.*** A review of vaccine research and development: meningococcal disease. *Vaccine* 2006; **24**: 4692–4700.
8. **Novartis International.** Novartis vaccine Bexsero approved in Australia to help protect against MenB disease, a deadly form of bacterial meningitis. 2013 (http://www.novartis.com/newsroom/media-releases/en/2013/1723151.shtml). Accessed December 2013.
9. **Moxon R, Snape M.** The price of prevention: what now for immunisation against meningococcus B? *Lancet* 2013; **382**: 369–370.
10. **Pollard AJ, Riordan A, Ramsay M.** Group B meningococcal vaccine: recommendations for UK use. *Lancet* 2014; **383**: 1103–1104.
11. **Breen E, *et al.*** How complete and accurate is meningococcal disease notification? *Communicable Disease and Public Health* 2004; **7**: 334–338.
12. **de Greeff S, *et al.*** Underreporting of meningococcal disease incidence in the Netherlands: Results from a capture-recapture analysis bases on three registration sources with correction for false positive diagnoses. *European Journal of Epidemiology* 2006; **21**: 315–321.
13. **NSW Health Notifiable Conditions Information Management System (HOIST).** Infectious disease data – meningococcal disease. Communicable Diseases Branch and Centre for Epidemiology and Evidence, 2013 (http://www0.health.nsw.gov.au/data/diseases/meningococcal.asp). Accessed December 2013.
14. **Thompson M, *et al.*** Clinical recognition of meningococcal disease in children and adolescents. *Lancet* 2006; **367**: 397–403.
15. **Nadel S, *et al.*** Avoidable deficiencies in the delivery of health care to children with meningococcal disease. *Journal of Accident and Emergency Medicine* 1998; **15**: 298–303.
16. **Stephen C.** Capture-recapture methods in epidemiological studies. *Infection Control and Hospital Epidemiology* 1996; **17**: 262–266.
17. **Orton H, Rickard R, Miller L.** Using active medical record review and capture-recapture methods to investigate the prevalence of Down Syndrome among live-born infants in Colorado. *Teratology* 2001; **64**: S14–S19.
18. **Centre for Health Record Linkage.** *Quality Assurance*. 2011 (http://www.cherel.org.au/quality-assurance). Accessed March 2014.

19. **Chapman D.** Some properties of the hypergeometric distribution with applications to zoological sample censuses. University of California publications in statistics, 1951. Berkeley: University of California Press, pp. 131–160.
20. **Orton H, Rickard R, Gabella B.** Capture-recapture estimation using statistical software. *Epidemiology* 1999; **10**: 563–564.
21. **Australian Institute of Health and Welfare.** Rural, regional and remote health: a guide to remoteness classifications. *Rural Health Series No. 42004*, Australian Institute of Health and Welfare: Canberra.
22. **Barquet N, *et al.*** Meningococcal disease in a large urban population (Barcelona, 1987–1992). Predictors of dismal prognosis. *Archives of Internal Medicine* 1999; **159**: 2329–2340.
23. **Hook E, Regal R,** Capture-recapture methods in epidemiology: methods and limitations. *Epidemiologic Reviews* 1995; **17**: 243–264.
24. **Australian Bureau of Statistics.** Migration, Australia 2010–11 December 2013 (http://www.abs.gov.au). Accessed January 2014.
25. **State Government Victoria.** Internal migration in Victoria. 2013 (http://www.dpcd.vic.gov.au/home/publications-and-research/urban-and-regional-research/census-2011/internal-migration-in-victoria). Accessed January 2014.
26. **Robinson P, *et al.*** Laboratory enhanced surveillance for meningococcal disease in Victoria. *Journal of Paediatric Child Health* 2001; **37**: S7–S12.
27. **Clinical Excellence Commission.** Report of the review of administrative and system issues arising out of two patient deaths attributed to meningococcal disease, 2005, Sydney Hospital & Sydney Eye Hospital: Sydney.
28. **Olesch C, Knight G.** Invasive meningococcal infection in Western Australia. *Journal of Paediatric Child Health* 1999; **35**: 42–48.
29. **Snitker Jensen E, *et al.*** Neisseria meningitidis phenotypic markers and septicaemia, disease progress and case-fatality rate of meningococcal disease: a 20-year population-based historical follow-up study in a Danish county. *Journal of Medical Microbiology* 2003; **52**: 173–179.
30. **Menzies R, *et al.*** Vaccine preventable diseases and vaccination coverage in Aboriginal and Torres Strait Islander people, Australia 2003 to 2006. *Communicable Diseases Intelligence* 2008; **32**.
31. **Heyderman R, *et al.*** The incidence and mortality for meningococcal disease associated with area deprivation: an ecological study of hospital episode statistics. *Archives of Disease in Childhood* 2004; **89**: 1064–1068.