RESEARCH ARTICLE　　　　　　　　　　　　　　　　　　　　　　　**Open Access**

Measuring underreporting and under-ascertainment in infectious disease datasets: a comparison of methods

Cheryl L Gibbons[1*], Marie-Josée J Mangen[2], Dietrich Plass[3], Arie H Havelaar[4,5], Russell John Brooke[2], Piotr Kramarz[6], Karen L Peterson[1], Anke L Stuurman[4,7], Alessandro Cassini[6], Eric M Fèvre[8,9], Mirjam EE Kretzschmar[2,4], On behalf of the Burden of Communicable diseases in Europe (BCoDE) consortium

Abstract

Background: Efficient and reliable surveillance and notification systems are vital for monitoring public health and disease outbreaks. However, most surveillance and notification systems are affected by a degree of underestimation (UE) and therefore uncertainty surrounds the 'true' incidence of disease affecting morbidity and mortality rates. Surveillance systems fail to capture cases at two distinct levels of the surveillance pyramid: from the community since not all cases seek healthcare (under-ascertainment), and at the healthcare-level, representing a failure to adequately report symptomatic cases that *have* sought medical advice (underreporting). There are several methods to estimate the extent of under-ascertainment and underreporting.

Methods: Within the context of the ECDC-funded Burden of Communicable Diseases in Europe (BCoDE)-project, an extensive literature review was conducted to identify studies that estimate ascertainment or reporting rates for salmonellosis and campylobacteriosis in European Union Member States (MS) plus European Free Trade Area (EFTA) countries Iceland, Norway and Switzerland and four other OECD countries (USA, Canada, Australia and Japan). Multiplication factors (MFs), a measure of the magnitude of underestimation, were taken directly from the literature or derived (where the proportion of underestimated, under-ascertained, or underreported cases was known) and compared for the two pathogens.

Results: MFs varied between and within diseases and countries, representing a need to carefully select the most appropriate MFs and methods for calculating them. The most appropriate MFs are often disease-, country-, age-, and sex-specific.

Conclusions: When routine data are used to make decisions on resource allocation or to estimate epidemiological parameters in populations, it becomes important to understand when, where and to what extent these data represent the true picture of disease, and in some instances (such as priority setting) it is necessary to adjust for underestimation. MFs can be used to adjust notification and surveillance data to provide more realistic estimates of incidence.

Keywords: Underestimation, Underreporting, Under-ascertainment, Surveillance, Infectious diseases

* Correspondence: C.L.Gibbons@sms.ed.ac.uk
[1]Centre for Immunity, Infection and Evolution, Ashworth Laboratories, Kings Buildings, University of Edinburgh, Edinburgh, UK
Full list of author information is available at the end of the article

Background

Efficient and reliable surveillance and notification systems are vital for monitoring public health trends and disease outbreaks. They also often form the backbone of evidence-based decision-making processes, as well as infectious disease (ID) public health policies that deal with prioritisation, and the planning of intervention measures and healthcare services [1]. However, there are limitations associated with the use of data from surveillance and notification systems since most systems are affected by a degree of underestimation and therefore uncertainty surrounds the 'true' incidence of disease [2]. IDs are considered particularly prone to underestimation due to their specific characteristics (e.g. asymptomatic or self-limiting disease courses) and are therefore represented inadequately by raw surveillance data. Thus, when routine data are used to inform decisions relating to resource allocation or to estimate epidemiological parameters in a population, it becomes important to understand when, where and to what extent these data do or do not comprehensively represent the true picture of disease. Furthermore, in certain circumstances, such as priority setting, it is appropriate to adjust infectious disease datasets in order to account for the portion not captured by the surveillance system. There are several metrics that can be employed in priority setting with Disability-Adjusted Life Years (DALYs) being just one composite health metric that combines and measures adverse health effects and premature mortality in a single unit. DALYs were chosen by the European Centre for Disease Prevention and Control (ECDC) and used within the Burden of Communicable Diseases in Europe (BCoDE)-project to generate evidence-based and comparable burden of disease (BoD) estimates for 32 IDs across European Member States (MS) [3-6]. A major prerequisite of DALY calculations is 'true' incidence data but since data are often obtained from (inter)national-level routine surveillance datasets that are frequently incomplete, data must be adjusted before serving as input for computing disease burden.

Here we present an overview of why, where and in what form underestimation occurs within the morbidity surveillance pyramid (Figure 1A) and we give several disease-specific examples from the literature of the methods that can be used to estimate the extent of underestimation. Furthermore, we compare the extent of underestimation and multiplication factors to adjust for it using key examples from the literature for two diseases, salmonellosis and campylobacteriosis. This body of work was a core aspect of the BCoDE-project.

Definitions

Underestimation (UE), as defined here, can be understood as the many ways in which surveillance systems fail or are unable to reflect all infections in a given population. Mathematically, UE is the number of infections estimated to have occurred in a population that have not been captured by the surveillance system for every reported case over a given time period. UE can be split into two distinct levels as represented by the surveillance pyramid for IDs (Figure 1A); under-ascertainment (UA) of infections occurring at the community-level and underreporting (UR) of infections occurring at the healthcare-level. Under-ascertained infections occur in individuals that do not seek healthcare and hence cannot be captured by surveillance systems which are typically designed to capture cases that do seek healthcare. UA can be estimated as the number of infections occurring in individuals that do not attend healthcare services for every case that attends. There is a symptomatic fraction of all under-ascertained cases that do not attend healthcare due to mild symptoms and/or the knowledge that the illness is self-limiting or for some other reason, and an asymptomatic fraction that do not seek healthcare as they are not aware of their infection status due to lack of symptoms [10]. Underreported infections are infections in individuals that do seek healthcare, but whose health event is not captured by the surveillance system and not notified through the notification system [7,8,11,12]. UR can be estimated as the number of infected individuals attending healthcare services whose health event is not reported to the appropriate public health body for every attending case whose health event is reported. UR can be due to under-diagnosis which accounts for the cases attending healthcare but whose infection or pathogen is not diagnosed or misdiagnosed [7,8], and under-notification which accounts for the failure to report (using correct International Classification of Diseases (ICD) codes [13,14]) all positive diagnoses through the notification system [15,16]. Reporting completeness refers to the proportion of cases attending healthcare whose health event was correctly diagnosed and appropriately reported [17]. These technical terms are used frequently in the literature, however often with varying definitions. The definitions of UA, UR and UE as stated here were developed during the BCoDE-project [3,18] and will be used as such for the remainder of the article.

Factors influencing UA in morbidity datasets

Not all people who are infected with a pathogen seek healthcare [19]. One important reason for this is that when symptoms are absent, mild or self-limiting, there is a lack of urgency to seek healthcare [8]. Therefore, surveillance systems can only capture cases with symptoms that are severe enough to motivate infected individuals to attend healthcare services. Health literacy also influences the decision to attend healthcare services or not. If a community has a poor understanding of when to seek healthcare and lacks knowledge of the severity or duration of an illness; then the uptake of healthcare services and levels of case ascertainment could be lower than expected. Such
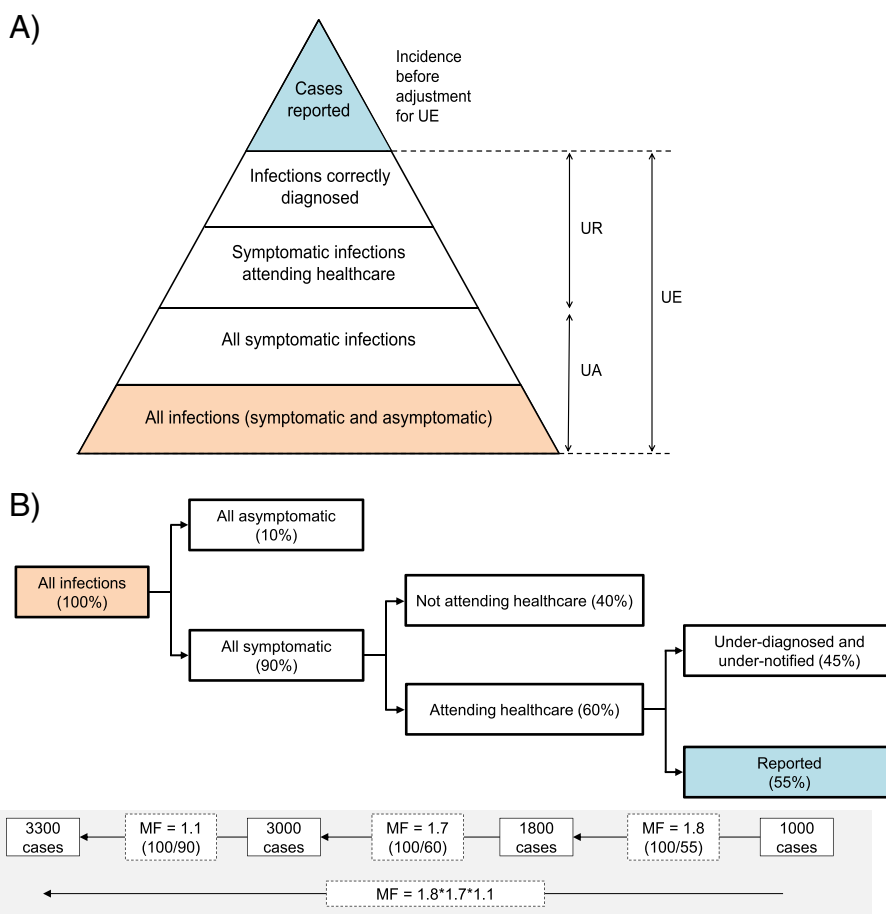
Figure 1 Deriving multiplication factors from the morbidity surveillance pyramid. A: The morbidity surveillance pyramid is often used to illustrate the availability of morbidity data at each surveillance level. With each ascending level (from the community, to healthcare institutions (GPs, hospital, laboratory), to regional and national public health agencies); data availability shrinks and only a fraction of cases from the level below is captured [7-9]. In contrast to the narrow tip of the pyramid which represents data held by national public health agencies, the base is wide as it holds all infections in the community. The difference between the number at the tip and base can be considered cases lost to 'underestimation' (UE). **B**: The proportions of infections that are symptomatic, that attend healthcare, and that are reported are represented in this decision tree model. Here, only 55% of all infected individuals attending healthcare are reported through the notification system. If 1000 cases were reported then a MF of 1.8 (=100/55) could be derived and would correct for those underreported cases. The true number attending healthcare would be 1800 cases. Likewise, if only 60% of symptomatic cases attended healthcare, then a MF of 1.7 (=100/60) would correct for under-ascertainment of symptomatic cases. The true number of cases attending healthcare would be 3000 symptomatic cases (=1.7*1800). Finally, since 90% of infections were symptomatic, a MF of 1.1 (=100/90) would correct for under-ascertainment of asymptomatic cases. The true number of infections would be 3300 (=1.1*3000). A MF to correct for total underestimation of symptomatic cases in one step would be 3.06 (=1.8*1.7) and for all infections 3.4 (=1.8*1.7*1.1). 'All infections' shaded in orange in Figure 1A represents the same population as the orange box in Figure 1B. 'Cases reported' in blue in Figure 1A represents the same population as the blue box in Figure 1B.

awareness and recognition of disease as well as the urgency or perceived need to seek healthcare can vary through time and space, particularly during outbreak years and especially if there is enhanced surveillance or widespread campaigns and intensive media coverage [20]. Compared with non-outbreak years, the proportion of cases in a population that is ascertained is expected to be greater. Health literacy and perceived need for healthcare may also explain often observed differences between age- and sex-specific ascertainment rates with, for example, children aged less than 15 years being statistically more likely to seek healthcare for gastroenteritis compared to adults (30–64 years) [21].

In addition, cultural and religious factors could prevent individuals from seeking healthcare if, for example, there is stigma or negative beliefs associated with healthcare services, illness and treatment [22,23]. There may also be legal, administrative and financial barriers to attending healthcare if individuals are not registered or are unable to register and if healthcare is dependent on the legal status of an individual or ability to pay. Migrants or marginalised groups may be particularly affected by this [24] and in addition to

not having their disease episode captured by the surveillance system, they may not be enumerated at all and hence do not contribute to the country population count or denominator. Individuals from remote communities and their illness may also be uncounted due to healthcare services being physically unreachable. Overall, ascertainment rates are thought to vary significantly within and between disease groups, population groups and countries.

Factors influencing UR in morbidity datasets

Not all cases that attend healthcare will have their health status correctly diagnosed and reported to the appropriate health authorities. This break in the surveillance chain can occur within clinics, hospitals or laboratories due to healthcare workers lacking in ability, capacity or knowledge of how and when to act. Under-diagnosis may arise when biological samples are not requested from or provided by patients, where there are budget restrictions forcing healthcare professionals to limit their requests for testing samples, lack of knowledge of which tests to perform, inadequate diagnostic tools, or due to restrictions of laboratory testing regimes (regulations on which tests to apply routinely, and lack of availability of more specialised tests). Under-notification may result from an inadequate reporting system or lack of knowledge of when, for which diseases and how to report correctly including knowledge of ICD codes [25-29]. The proportion of cases reported is often higher where there is a legal requirement to report (some diseases or pathogens have mandatory reporting statuses) [30] or where there are incentives for healthcare workers to request or test biological samples from patients or to report results [26]. Furthermore, where there is a perceived urgency to request biological samples or report cases, for example during outbreak years where there is higher awareness and chance of recognition and motivation for testing [31,32], reporting rates are likely to increase. In contrast, UR may be greater for rarer diseases, those with only occasional outbreaks or those without mandatory reporting statuses. This perceived urgency or necessity to test or report tends to increase for more serious conditions (severity and duration of illness [8,33]) and can also be age- or sex-dependent [34]. Incomplete reporting of additional information, such as age and sex of the patient, concurrent infections or sequelae following an initial infection [35] becomes particularly important when UR for a particular disease is age- or sex-specific.

Identifying areas and extent of UE

Various study designs can be used to determine the extent of UR and UA in the surveillance system.

Community-based studies

Community-based studies (CBS) aim to generate new estimates of pathogen carriage or infection in a (representative) sample of the population. This alone is useful and interesting but in addition, if this new estimate is considered the 'true' incidence in the community, it can be compared to notification data and the magnitude of UE deduced. This order of magnitude can then be used as a multiplication factor (MF) (Figure 1B) to adjust disease datasets assuming that the base value to which the MF is being applied was created using the same data type (i.e. MF calculated by comparing CBS data with notification data is used to adjust other notification datasets, and not laboratory data). These observational studies can also be used to produce incidence rates of symptomatic and asymptomatic cases as well as estimate the specific proportion of symptomatic cases presenting to healthcare facilities (by asking about attendance in the questionnaire). Furthermore, the proportion of cases underreported can be estimated by a variation of CBS in which healthcare professionals are surveyed to gain information on propensity to request biological samples and reporting habits.

CBS can take many forms but generally involve active searching within the community for disease episodes, pathogen carriage or infection, with questionnaire-based data acquisition often accompanied by biological sampling. Active searching can be conducted face-to-face, by telephone, internet or post, with several possible study designs e.g. based on probability samples, prospective or retrospective cohorts, population cross-sections, involving representative samples of the whole population or certain interest or high-risk groups only. CBS are especially useful for diseases commonly under-ascertained (i.e. those with many mild and asymptomatic cases of mostly self-limiting illness) and where an unknown burden exists within the community, e.g. sexually transmitted infections with *Chlamydia trachomatis* [36-40] or *Neisseria gonorrhoea* [36-39,41-44]; influenza and influenza-like illnesses [45,46]; and food and water-borne diseases [2,47-54]). The first and second Infectious Intestinal Disease studies (IID1 and IID2) were prospective community cohort studies that estimated overall incidence of infectious intestinal disease (IID) in the UK community, the proportion seeking healthcare (ascertained), and the proportion reported to the national public health agency [2,11,47,55]. Weekly surveys (by email or prepaid postcard) recorded if participants had experienced diarrhoea and/or vomiting and if they had, they were asked additional questions and requested to provide a stool sample. Similarly, several studies have employed statistical and mathematical methods to estimate incidence and UE using data collected during a Dutch prospective community-based cohort study of gastroenteritis (SENSOR) and a Dutch GP-based cohort of gastroenteritis (NIVEL) (e.g. [56-63]).

CBS are not without limitations as bias can arise at numerous points. Sampling bias, owing to non-random

sampling of a population, can result in a study that is not representative of the entire population with certain groups (such as ethnic, migrant, age or occupational) inadvertently excluded from the study because they are unregistered, not easily locatable, do not have access to a telephone (in the case of telephone surveys), have language barriers or are marginalised for other reasons. Responder bias can also lead to unrepresentative samples since only certain groups of people will agree to participate, and measurement bias can result from case definitions that are undefined, too general, too strict or simply not used consistently. Additionally, interviewers may ask questions or interpret responses in a leading manner or respondents may induce bias during disease occurrence recall. To minimise the effects of bias, Wilking *et al.* [54] took several steps in a recent population-based telephone survey of acute gastrointestinal illness in Germany including; contacting listed and unlisted telephone numbers and using the 'last-birthday method' to reduce responder bias, using the computer-assisted telephone interview (CATI) method to minimise interviewer bias and applying study weights to improve representation of the target population. Telephone surveys, however, have a further limitation associated with them since disease occurrence tends to be based on self-reported symptoms and so without clinically-determined or laboratory-confirmed diagnoses, there is uncertainty surrounding the causative agent and incidence of infection. To address this, Kubota *et al.* [50] used data from population-based telephone surveys to adjust the number of cases for each pathogen from active laboratory-based surveillance (whilst modelling for uncertainty). Other telephone surveys focus on general conditions, such as gastroenteritis, rather than specific pathogens and these have been successfully applied in many countries (e.g. [54,64-66]) with the results often the basis for pyramid reconstruction activities (see below).

Serological surveys
Serological surveys are a specific type of CBS that measure sero-incidence (the rate of new infections) or sero-prevalence (the total number of infections in the community or cohort) as quantified by antigen or antibody positivity. This CBS can capture asymptomatic and symptomatic, historical and acute infections but if biological sampling is combined with a questionnaire asking about disease episodes, or if the antibody or antigen threshold at which symptoms manifest is known; then the symptomatic fraction can be obtained (e.g. hepatitis B [67], hepatitis B and C [68], pertussis [69], measles [70], HIV [71]). This is crucial for BoD studies since, for the majority of diseases, it is only clinical manifestations and not asymptomatic infections that contribute to burden in terms of DALYs. The exceptions to this are the few infectious diseases with a possible asymptomatic

acute stage that can result in sequelae (and death) at a later time and hence contribute to disease burden (including hepatitis B, hepatitis C, chlamydia, Invasive Meningococcal Disease, and Q-fever). When calculating DALYs for these diseases, symptomatic cases serve as input to estimate the numbers of asymptomatic infections and thus the calculated burden attributable to asymptomatic infections is included in the final burden estimate [4].

It is often difficult to differentiate between historic and current infections and therefore it is important to test for recognised serological markers of recent infection and to have full knowledge of antibody decay rates in different populations [72,73]. Furthermore, antibodies resulting from natural exposure versus vaccination cannot be distinguished and thus serological surveys of diseases with universal vaccine coverage, including tuberculosis (BCG vaccine), measles, rubella, and other childhood vaccine-preventable diseases, may have limited use.

Returning traveller studies (RTS)
Returning traveller studies (RTS) are further examples of CBS where individuals returning from abroad represent sentinel populations for the reported national incidence of infection in a traveller's destination of travel [74]. In RTS, the risk of infection for travellers from country A visiting country B is calculated by taking the number of infected travellers returning home from country B from surveillance records as a numerator, and the total number of travellers from country A visiting country B from travel pattern databases as the denominator. This measure of risk can then be used to generate a new estimate of incidence in country B (risk multiplied by the population size), which when compared to the national notification records of country B, can generate a MF of underestimation. Using this method, the incidence and proportion of underestimated cases of salmonellosis in several European countries was calculated by de Jong and Ekdahl [75] by comparing the incidence of infection in Swedish returning travellers to the national incidence in the countries the travellers had returned from, using Norway as a reference country. This not only produced national-level comparable estimates of incidence, but also a MF of UE (or in this case 'under-detection index'). More recently, Havelaar *et al.* [76] calculated incidence rates of salmonellosis and campylobacteriosis and MFs based on disease risks of returning Swedish travellers, anchored to data from SENSOR, the Dutch population-based study on gastroenteritis (see Tables 1 and 2).

While RTS generate comparable estimates of incidence and multipliers across different countries, there are several assumptions and limitations. Even if the travel patterns database is representative of the whole population and that the origin of infection, as reported in surveillance data, is correct [74]; in calculating risk of infection, the numerator

Table 1 A comparison of multiplication factors (MFs) for salmonellosis in several countries

Country	Underestimation MF	Primary study type	Study	Under-ascertainment MF	Primary study type	Study	Underreporting MF	Primary study type	Study
Austria	3	RTS	[75]						
Austria	11	RTS	[76]						
Belgium	1.9	RTS	[75]						
Belgium	3.5	RTS	[76]						
Bulgaria	271	RTS	[75]						
Bulgaria	718.5	RTS	[76]						
Croatia	30.6	RTS	[75]						
Cyprus	71.2	RTS	[75]						
Cyprus	173.2	RTS	[76]						
Czech Republic	3	RTS	[75]						
Czech Republic	28.9	RTS	[76]						
Denmark	1.8	RTS	[75]						
Denmark	4.4	RTS	[76]						
Denmark	17	PRC	[9]						
Denmark	NT UEinf: 325 (5-95% quartiles: 190–505)	Sero/MOD	[77]						
Estonia	1.3	RTS	[75]						
Estonia	16.9	RTS	[76]						
Finland	0.6	RTS	[75]						
Finland	0.4	RTS	[76]						
France	8.3	RTS	[75]				*S.entd.* $: Data1: 7.1 (95% CI: 6.7-7.7) Data2: 12.5 (95% CI: 11.1-14.3) Data3: 2 (95% CI: 1.9-2.2) *Spp* other than *S.entd.& S.typh.* Data1: 12.5 (95% CI: 7.1-16.7) Data2: 16.6 (95% CI: 10–25) Data3: 2.0 (95% CI: 1.1-2.8)	CRS	[78]
France	26.9	RTS	[76]						
Germany	1.8	RTS	[75]						
Germany	9.8	RTS	[76]						
Germany	6.7	PRC	[9]						
Greece	97.7	RTS	[75]				1.75	MOD/RC	[79]
Greece	1228.5	RTS	[76]						
Greece	51.45 (PERT: 3.2; 99.7)	BoD/CRS	[80]						
Hungary	5.5	RTS	[75]						
Hungary	66.8	RTS	[76]						
Ireland	4.3	RTS	[75]						
Ireland	5.4	RTS	[76]						
Italy	13.1	RTS	[75]						

Table 1 A comparison of multiplication factors (MFs) for salmonellosis in several countries (Continued)

Italy	71.7	RTS	[76]						
Italy	17	PRC	[9]						
Latvia	11.7	RTS	[75]						
Latvia	44.3	RTS	[76]						
Lithuania	10	RTS	[75]						
Lithuania	59.1	RTS	[76]						
Luxembourg	4.5	RTS	[76]						
Malta	92.6	RTS	[75]						
Malta	222.7	RTS	[76]						
Poland	16.2	RTS	[75]						
Poland	114.1	RTS	[76]						
Poland	18	PRC	[9]						
Portugal	378	RTS	[75]						
Portugal	2082.9	RTS	[76]						
Romania	332	RTS	[75]						
Romania	349.9	RTS	[76]						
Slovakia	3.5	RTS	[75]						
Slovakia	53.2	RTS	[76]						
Slovenia	10	RTS	[75]						
Slovenia	40.3	RTS	[76]						
Spain	103	RTS	[75]						
Spain	214.2	RTS	[76]				NT: Data 1 = 2.0 (95% CI: 2.0 - 2.1)	CRS	[81]
							Data 2 = 1.5 (95% CI: 1.4 - 1.5)		
Sweden	0.5	RTS	[76]						
Sweden	10	PRC	[9]				Data 1 =1.05, Data 2 =1.02	CRS	[82]
The Netherlands	7.7	RTS	[75]						
The Netherlands	26.3 (ref)	RTS	[76]						
The Netherlands	24.7 (5-95% quartiles: 5.2 - 64.7)	BRI/BoD	[57]	5.8 (5-95% quartiles: 0.8 - 25.6)	BRI	[57]	4.3 (5-95% quartiles: 2.5 - 6.5)	BRI	[57]
The Netherlands	14 (5-95% quartiles: 3.6 − 56)	CBS/BoD	[56]	6.5 (5-95% quartiles: 0.0 - 20)	CBS/BoD	[56]			
The Netherlands	14.3	LAB	[62]						
The Netherlands	20	PRC	[9]						
United Kingdom	4.3	RTS	[75]						
United Kingdom	7.3	RTS	[76]						
United Kingdom	4.7 (95% CI : 1.2 - 18.2)	CBS	[47]	3.4 (95% CI: 0.4 - 32.2)	CBS	[47]	1.4 (95% CI: 0.6 - 3.3)	CBS	[47]
United Kingdom	3.2 (95% CI : 1.4 - 12.0)	CBS	[2]	GP only, 1.4 (95% CI: 0.7 - 2.8)	CBS	[2]			
United Kingdom	40	PRC	[9]						
United Kingdom	NT, UElab: 3.9	CBS/BoD	[55]						
EU-27 (excl.Croatia)	57.5 (11–140)	RTS	[76]						
Iceland	0.4	RTS	[75]						
Norway	1.0 (ref)	RTS	[75]						

Table 1 A comparison of multiplication factors (MFs) for salmonellosis in several countries (Continued)

Norway	1.2	RTS	[76]						
Switzerland	7.1	RTS	[76]						
USA	NT, UElab: BD 9.8, NBD 67.7, total 38.6	CBS/BoD	[49]	NT, BD 6.8, NBD 8.6	[49]				
USA	NT, 38 (taken from [49])	BoD	[83]	NT, BD - 2.86 (PERT 1.96 – 5.26)	BoD	[84]	NT UN, 1		[84]
				NBD – 5.56 (PERT 5–6.67)					
Canada	13 - 37	PM	[85]						
Australia	‡BD:1-2d: 11.39 (95% CrI: 8.49–16.36)	PM	[33]	‡BD: 1-2d: 10 (95% CrI: 7.1-14.3)	[33]				
	3-4d: 2.82 (95% CrI: 2.17–3.98)			3-4d: 2.3 (95% CrI: 1.9-3.2)					
	≥5d: 1.81 (95% CrI: 1.33–2.72)			≥5d: 1.5 (95% CrI: 1.1-2.2)					
	NBD: 1-2d 143.29 (95% CrI: 83.3–371)			NBD; 1-2d: 10 (95% CrI: 7.1-14.3)					
	3-4d 13.06 (95% CrI: 6.37–67.83)			3-4d: 2.3 (95% CrI: 1.9-3.2)					
	≥5d 3.93 (95% CrI: 2.10–11.92)			≥5d: 1.5 (95% CrI: 1.1-2.2)					
	Overall: 7 (95% CrI: 4–16)								
Japan	74.0 (5-95% quartiles: 35.8, 140.7)	CBS	[50]	*S.brae.* Age < 10 years: 1.2,	CBS	[48]			
				> = 10 years: 1.7, Overall:1.6,					

1. This table lists all extracted or derived MFs (with variance shown as 95% CI, 95% CrI, PERT distribution (max, min, mode), or 5-95% quartiles where available) from relevant studies found during the extensive literature review. MFs give an estimation of the extent of UE (combined UA and UR), UA and UR for salmonellosis in a particular country; the higher the MF, the higher the proportion of cases not captured by the surveillance system. These MFs could be applied to official figures as reported by public health agencies to adjust for UE and give a new estimate of total symptomatic infections occurring in a population at a given time. Exceptions include; "UEinf" where the MF can adjust official figures from public health agencies and give a new estimate of total infections (both symptomatic and asymptomatic) occurring in a population at a given time, and "UElab" where the MF can adjust official laboratory figures of laboratory confirmed infections and give a new estimate of total infections (both symptomatic and asymptomatic) occurring in a population at a given time. MFs of UA and UR can be multiplied together to make one MF of UE.
2. Study types abbreviations: CBS: Community-based study, RTS: Returning traveller study, CRS: Capture-recapture study, PRC: Pyramid reconstruction model, BRI: Bayesian risk of infection model, BoD: Burden of disease calculation, Sero: Analysis of serology data, LAB: Analysis of laboratory surveillance, RC: Analysis of reporting completeness, PM: Probability model, OUT: Outbreak analysis, MOD: Modelling other. Symptoms abbreviations: NT: Non-typhoidal salmonellosis; NBD: Non-bloody diarrhoea; BD: Bloody diarrhoea; severity of diarrhoea (d = days). *Salmonella* species abbreviations: *S.entd.* : *S. enteritidis*; *S.typh.* : *S.typhimurium*; *S.brae.* : *S.braenderup*. Other abbreviations: UN: Under-notification of laboratory confirmed infection; GP only: cases attending GP surgeries (not hospitals) only; $ Estimates corrected by the positive predictive value of one data source where (unlike the other two sources) notifications are not validated by a systematic procedure; ‡ No. cases in the community for every 100 reported.
3. For CRS, a MF is given to correct for UR for each data source (i.e. MF for 'Data 1' will estimate the underreporting in data source 1). For MFs estimated in the same RTS, one country will be listed as the reference country (i.e. 'ref') and all other countries compared to this.

(number of infections caught abroad) will still be affected by UR and UA. This in part is due to general reasons affecting UE, but in addition travellers have different health-seeking behaviour than that of the general population and there may be bias in requesting samples and reporting by health professionals following travel to certain destinations [74]. Furthermore, in the case of Swedish travellers the duration and reason for travel differs by country, with many short business trips to neighbouring Scandinavian countries but longer holidays to Mediterranean countries [76] and therefore there is a bias in the risk estimates towards countries with the most tourists [75]. The risk of infection differs given the destination of travel but in addition the probability of the infection being ascertained also differs since on a longer trip, the case may have recovered before returning home and attending healthcare [76]. Lastly, a traveller's risk is not the same as the native population's risk due to differences in behaviours and activities, as well as immunity to local pathogen populations [76].

Capture-Recapture Studies (CRS)
Capture-recapture studies (CRS) utilise the ecological principle for studying populations of wildlife by marking subjects on initial release or first capture and recovering information from them on subsequent captures [87-89]. In terms of human disease surveillance, a personal identifier

Table 2 A comparison of multiplication factors (MFs) for campylobacteriosis in several countries

Country	Underestimation MF	Primary study type	Study	Under-ascertainment MF	Primary study type	Study	Underreporting MF	Primary study type	Study
Austria	15	RTS	[74]						
Austria	29	RTS	[76]						
Belgium	25	RTS	[74]						
Belgium	11	RTS	[76]						
Bulgaria	39,000	RTS	[76]						
Cyprus	310	RTS	[76]						
Czech Republic	11	RTS	[76]						
Denmark	4	RTS	[74]						
Denmark	4.1	RTS	[76]						
Denmark	29	PRC	[9]						
Estonia	13	RTS	[76]						
Finland	1.0 (ref)	RTS	[74]	GP *C.jejuni*, 5.2	CBS/OUT	[86]			
Finland	0.4	RTS	[76]						
France	3,958	RTS	[74]						
France	280	RTS	[76]						
Germany	6	RTS	[74]						
Germany	4.4	RTS	[76]						
Germany	9.3	PRC	[9]						
Greece	47,191	RTS	[74]						
Hungary	52	RTS	[76]						
Ireland	46	RTS	[74]						
Ireland	29	RTS	[76]						
Italy	660	RTS	[76]						
Italy	100	PRC	[9]						
Lithuania	40	RTS	[76]						
Luxembourg	19	RTS	[74]						
Luxembourg	3.9	RTS	[76]						
Malta	90	RTS	[76]						
Poland	4,100	RTS	[76]						
Poland	72	PRC	[9]						
Romania	6,900	RTS	[76]						
Slovakia	35	RTS	[76]						
Slovenia	14	RTS	[76]						
Spain	270	RTS	[76]						
Sweden	0.4	RTS	[76]						
Sweden	17	PRC	[9]						
The Netherlands	31	RTS	[74]						
The Netherlands	22 (ref)	RTS	[76]						
The Netherlands	22.9 (5-95% quartiles: 8.2 - 50)	BRI/BoD	[57]	4.1 (5-95% quartiles: 9.3 – 56.7)	BRI/BoD	[57]	5.6	BRI/BoD	[57]
The Netherlands	10.9 - 21.4	Sero	[58]	5.0 - 5.4	Sero	[58]	2.0 - 4.3	Sero	[58]

Table 2 A comparison of multiplication factors (MFs) for campylobacteriosis in several countries (Continued)

The Netherlands	9.7 (5-95% quartiles: 4.1 - 23.0)	CBS / BoD	[56]	4.2 (5-95% quartiles: 0.0 - 7.4)	CBS/BoD	[56]				
The Netherlands	18.9	LAB	[62]							
The Netherlands	49	PRC	[9]							
The Netherlands	UElab, 13.94 (PERT 4.96 – 28.67)	MOD	[61]							
United Kingdom	11	RTS	[74]							
United Kingdom	4.4	RTS	[76]							
United Kingdom	9.3 (95% CI: 6–14.4)	CBS	[47]	7.2 (95% CI: 3.3 - 15.9)	CBS	[47]	1.3 (95% CI: 0.9 - 1.8)	CBS	[47]
United Kingdom	52	PRC	[9]							
United Kingdom	UElab 10.3	CBS /BoD	[55]							
United Kingdom	7.6 (95% CI: 3.6 - 17.4)	CBS	[2]	2.1 (95% CI: 5–3.0)	CBS	[2]				
EU-27	47 (range 0.4 - 39,000)	RTS	[76]							
Norway	4	RTS	[74]							
Norway	2.4	RTS	[76]							
Switzerland	3.3	RTS	[76]							
USA	38 (taken from [49])	BoD	[83]	NT, BD: 2.9 (PERT 2.0-5.3) NBD: 5.6 (PERT 5–6.7)	BoD	[84]	NT UN, 1	BoD	[84]
USA										
Canada	23 - 49	PM	[85]							
Australia	‡BD: 1-2d 12.40 (95% CrI: 9.16– 17.82) 3-4d 3.06 (95% CrI: 2.32–4.33) ≥5d 1.97 (95% CrI: 1.42–2.95) NBD: 1-2d 154.2 (95% CrI: 89.3–397.6) 3-4d 14.15 (95% CrI: 6.80–73.32) ≥5d 4.25 (95% CrI: 2.25–13.36) Overall: 10 (95% CrI: 6.6–22)	PM	[33]	‡BD: 1–2 days: 10 (95% CrI: 7.1-14.3) 3-4 days: 2.3 (95% CrI: 1.9-3.2) ≥5 days: 1.5 (95% CrI: 1.1-2.2) NBD: 1–2 days: 10 (95% CrI: 7.1-14.3) 3–4 days: 2.3 (95% CrI: 1.9-3.2) ≥5 days: 1.5 (95% CrI: 1.1-2.2)	PM	[33]				

1. This table lists all extracted or derived MFs (with variance shown as 95% CI, 95% CrI, PERT distribution (max, min, mode), or 5-95% quartiles where available) from relevant studies found during the extensive literature review. MFs give an estimation of the extent of UE (combined UA and UR), UA and UR for campylobacteriosis in a particular country; the higher the MF, the higher the proportion of cases not captured by the surveillance system. These MFs could be applied to official figures as reported by public health agencies to adjust for UE and give a new estimate of total symptomatic infections occurring in a population at a given time. Exception is; "UElab" where the MF can adjust official laboratory figures of laboratory confirmed infections and give a new estimate of total infections (both symptomatic and asymptomatic) occurring in a population at a given time. MFs of UA and UR can be multiplied together to make one MF of UE.
2. For abbreviations and symbols, see footnote 2 for Table 1.
3. For MFs estimated in the same RTS, one country will be listed as the reference country (i.e. 'ref') and all other countries compared to this.
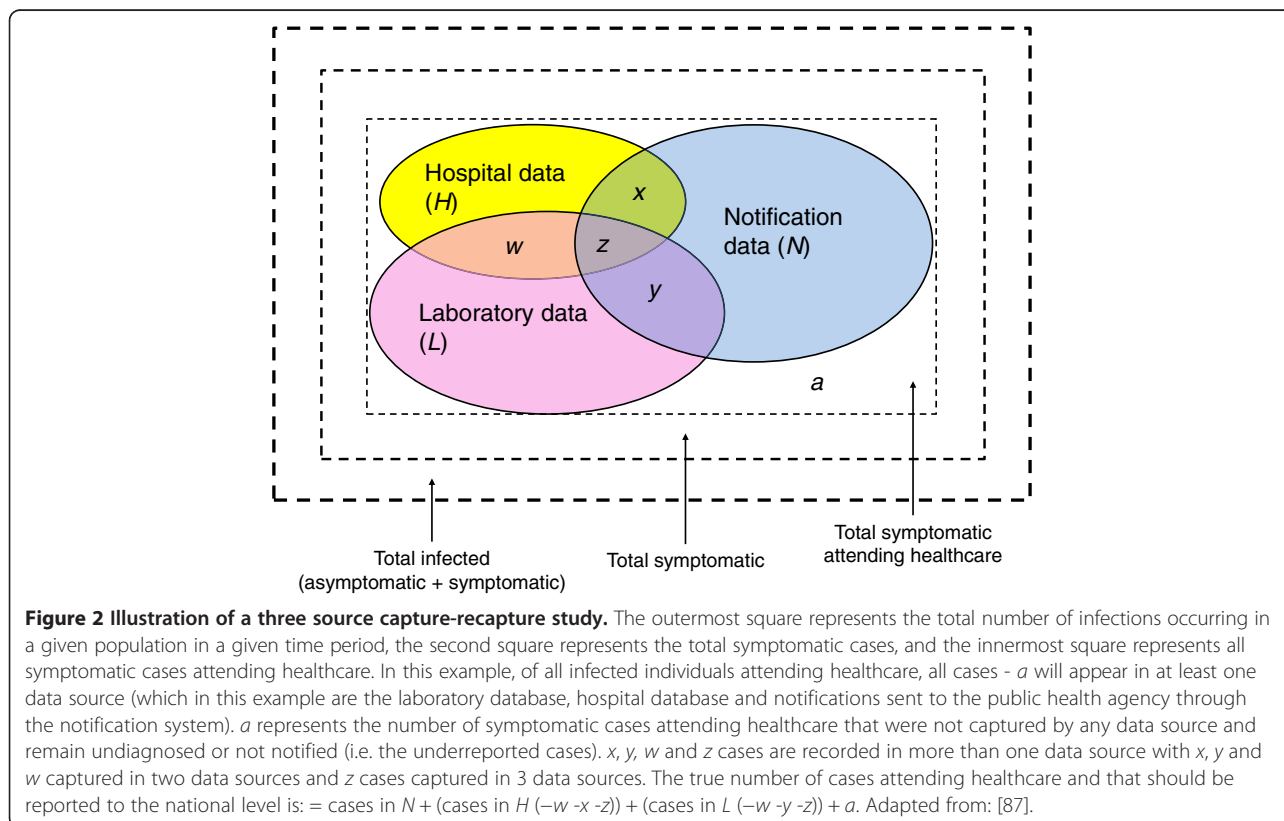
number or code usually represents the 'marker' and the 'captures' are records of disease episodes, colonisations or infections found in data sources including national notifications of morbidity and death, hospital and GP records, laboratory reports, as well as other public health registries [87,90]. Two or more data sources are compared (Hest *et al.* stated that at least three are preferred to avoid correlations [90]) or cross-linked (through personal identifiers) and duplicates removed to approximate reporting completeness of each data source, identify the cases that would have been missed if using only a single data source and calculate a new estimate of incidence (Figure 2). There are several examples of CRS studies in the literature spanning a wide range of both communicable and non-communicable diseases (e.g. CRS of tuberculosis in several countries; Greece [79,91], Italy [92,93], the Netherlands [94,95], Romania [96], UK [97-99] and USA [17,100,101]). Despite the usefulness of this method, some cases attending healthcare are not captured in any single data source or are not correctly recorded and hence the new incidence of cases will still be affected by UR [89]. In addition, CRS only correct for UR of infections and do not account for UA occurring in the community.

Modelling

There are numerous mathematical and statistical methods that can generate new estimates of incidence in the population as well as calculating the predicted proportion of UE occurring at several steps of the reporting chain, which can then be used to generate country- and disease-specific MFs. As well as using simulated data, these methods often utilise data from national surveillance records, CBS, CRS and several other study designs and therefore it is difficult to consider modelling as an independent method for identifying UE (many studies use a combination of methods, e.g. statistical modelling is often used to analyse results of CBS).

Attack rates, that estimate the proportion infected (from which cumulative incidences can be estimated and MFs generated), are calculated using data from CBS or national surveillance data (although these are still subject to UE), (e.g. influenza [102,103]). Vaccine coverage, when below a certain threshold and when the basic reproduction number is known for a pathogen, can be used to estimate the number of susceptible individuals in the community and therefore expected incidence (and hence MFs), (e.g. measles [104]). Scallan *et al.* [105] and Thomas *et al.* [106] used statistical methods to 'scale-up' counts of laboratory-confirmed cases to an estimated number of illnesses in the United States and Canada respectively, therefore adjusting for UE. Serological data can be modelled statistically to estimate incidence (e.g. HIV [107]) and past incidence of infection calculated from the current prevalence of antibody in a population using a catalytic model (e.g. hepatitis A

Figure 2 Illustration of a three source capture-recapture study. The outermost square represents the total number of infections occurring in a given population in a given time period, the second square represents the total symptomatic cases, and the innermost square represents all symptomatic cases attending healthcare. In this example, of all infected individuals attending healthcare, all cases - *a* will appear in at least one data source (which in this example are the laboratory database, hospital database and notifications sent to the public health agency through the notification system). *a* represents the number of symptomatic cases attending healthcare that were not captured by any data source and remain undiagnosed or not notified (i.e. the underreported cases). *x, y, w* and *z* cases are recorded in more than one data source with *x, y* and *w* captured in two data sources and *z* cases captured in 3 data sources. The true number of cases attending healthcare and that should be reported to the national level is: = cases in *N* + (cases in *H* (−*w* -*x* -*z*)) + (cases in *L* (−*w* -*y* -*z*)) + *a*. Adapted from: [87].

[108]). Decision tree models, (e.g. for Sexually Transmitted Infections (STIs) [109], hepatitis A [110]), and similarly probability models and pyramid reconstruction models (e.g. food and water borne disease [9,33,58,85,111,112], influenza [19]) estimate the country- and pathogen-specific probabilities of action at each incremental stage of the surveillance pyramid (e.g. attending healthcare versus not, submitting a sample versus not, reporting versus non-reporting). Further modelling techniques include Bayesian synthesis of multiple evidence sources that estimate the 'true' incidence of an infection at several steps of the surveillance pyramid, as well as changes in contact patterns and health-seeking behaviour (e.g. H1N1 influenza pandemic [113,114]). Simulation models, based on outcome trees of disease progression are also tools that can estimate expected incidence, (e.g. hepatitis B [115]).

Methods

An extensive literature review was conducted to identify studies (of designs presented above) that estimate ascertainment or reporting rates for salmonellosis and campylobacteriosis in European Union Member States (MS), plus European Free Trade Area (EFTA) countries Iceland, Norway and Switzerland and all other OECD countries. Articles were considered relevant if they: measured the sensitivity of reporting or reported the rate of UA, UR or UE; reported MFs, measured a new incidence or prevalence of infection from which a MF could be derived; or used any alternative methodology to correct surveillance or notification data. To identify appropriate studies, a literature review for each disease (salmonellosis and campylobacteriosis) and each pathogen (*Salmonella* spp. and *Campylobacter* spp.) was conducted in PubMed using the search terms: burden, cost-of-illness, cost of disease, cost-effectiv*, cost-analys*, cost-benefit, cost-utility, disability-adjusted, mathematical model*, multiplication factor*, multiplier*, outbreak*, prospective stud*, quality of life, quality-adjusted, serological stud*, serological survey*, serosurveillance, sero-surveillance, seroprevalence, statistical model*, telephone (*denotes any ending to the search term); linked by 'OR'. The search was restricted to articles written in English and to the years 1990–2011 since surveillance systems, reporting protocols and epidemiological patterns may have been different in the years preceding 1990, hence MFs would be less appropriate for adjusting current surveillance and notification data.

Following identification of these studies, MFs were either taken directly from the literature or derived where the proportion of underestimated, under-ascertained, or underreported cases was known (MF = 100/(percentage reported or ascertained or estimated), Figure 1B). MFs for salmonellosis and campylobacteriosis were compared to gain an understanding of variation between and within countries when using different methods to estimate UR and UA.

Results

MFs were found or derived for all European Union Member States (MS) plus European Free Trade Area (EFTA) countries Iceland, Norway and Switzerland and four other OECD countries (USA, Canada, Australia and Japan) for salmonellosis from 22 references, and similarly for campylobacteriosis (excluding Croatia, Greece, Iceland, Latvia and Portugal) from 18 references. Table 1 (salmonellosis) and Table 2 (campylobacteriosis) present MFs for adjusting surveillance data for UE in one step, and MFs for adjusting for UA and UR in these countries. By multiplying together one MF of UR and one MF of UA, this results in one single MF of UE. MFs were found to vary widely between study types, countries and diseases with MFs for UE ranging from 0.4 (suggesting over-reporting) [75] in Iceland to 2082.9 [76] in Portugal for salmonellosis and from 0.4 [76] in Sweden and Finland to 39,000 [76] in Bulgaria for campylobacteriosis. In countries with mandatory notification of infection (for salmonellosis this includes all EU countries other than Belgium, France, Luxembourg and Spain which are voluntary, and the UK which requires reporting of the pathogen rather than disease), reporting rates were expected to be higher (and hence MFs lower). Unfortunately, there are too few MFs of UR to verify this from these studies. The most common study type for generating MFs of UE for both diseases identified by the literature review was RTS which can provide comparable multipliers for several countries. Furthermore, most studies provide a single MF to adjust for UE in one step, suggesting this is the most straight-forward approach. Few studies (12 for salmonellosis and 10 for campylobacteriosis) provide any measure of uncertainty surrounding the MF. While age-stratified MFs are preferred since ascertainment and reporting are affected by age; stratification into age bands was only found for UA of salmonellosis (*S.braenderup*) in Japan in one study [48]. For individuals less than 10 years of age, the MF was lower than for individuals over 10 years and hence the older age group was ascertained less often. No study described sex-specific stratification of MFs for either disease. Two studies stratified MFs by severity of clinical symptoms based on duration of symptoms [33] and if there was blood in stool samples [33,49]. Where duration of illness was short, cases were less likely to seek healthcare leading to higher UA and higher MFs. With bloody diarrhoea due to *Salmonella* in USA, Voetsch *et al.* [49] estimated higher ascertainment (lower MF) and lower overall underestimation compared to non-bloody diarrhoea. Few studies make a distinction between strain types. Where MFs of UR were estimated from CRS (two studies of salmonellosis), a MF for each data source is listed showing the degree of UR per data source which is consistent for both studies. The United Kingdom and the Netherlands had the most number of MFs of UE for both salmonellosis and

campylobacteriosis. For each country, all studies were reasonably consistent in terms of order of magnitude. This may be due to the overlap of data sources used in the studies. However, the highest MF estimates for both countries and each disease were estimated by the pyramid reconstruction model. These higher estimates of MFs may represent the thorough nature of accounting for each incremental step of the surveillance pyramid which could lead to double-counting of cases.

Discussion

Here, we discuss the advantages and disadvantages of different methods for identifying UE in the surveillance pyramid and compare MFs resulting from those methods. MFs show considerable between-country and -disease variation which may reflect true differences in reporting and ascertainment rates. However, study design (which often depends on the disease, data type, data quality and availability of resources) and hence the method used to estimate UE likely accounts for the presented within-country variability of MFs. It remains difficult to select the most appropriate study (and corresponding MFs) as there are limitations associated with each. CBS that are representative of the whole population are often favoured for approximating UE, UA, and UR but the quality of MFs derived depends highly on the study design; CRS are very good at estimating UR in the surveillance pyramid but some reported cases may remain undetected and UA is not considered; and RTS and serological surveys also estimate UE effectively but the many associated limitations must be realised. In addition, an important limitation is the uncertainty that surrounds estimates which is relevant to all study types. Many of the studies from the literature review do not report uncertainty for MFs which gives a false impression that we are sure that these point estimates are correct. In fact, a great level of uncertainty is expected [33]. Since the factors contributing to UE are multiplicative (Figure 1B), they explode rapidly and the resulting MF can be very large (this is clearly observed in pyramid reconstruction studies which aim to account and correct for UE at each incremental step of the surveillance chain). Therefore, even small degrees of uncertainty in measuring individual components of UE can lead to wide ranges in incidence estimates and MFs. To estimate predictive intervals, uncertainty can be modelled by incorporating, for example, either uniform or pert probability distributions (rather than fixed point estimates), and using techniques such as Monte Carlo simulations [4,50,80,84].

One systematic way to decide on the best method for estimating UE or choosing MFs is to use the Delphi method or expert consensus. In the next phase of the BCoDE-project, internal ECDC experts with final input from the consortium and other external experts will create lists of the most appropriate country- and disease-specific MFs for 32 IDs. In general we assume that the most appropriate

MFs should be disease-, country-, age-, and sex-specific because underestimation rates are disproportionately distributed between diseases, countries with differing surveillance systems and reporting procedures, and between demographic groups. While our literature review returned only one result with age- or sex-stratification of MFs, there are other studies that provide age- (at least for given age bands) [12,21,34,116-119] and sex-specific [117,118] MFs for general gastroenteritis and diarrhoea (i.e. unspecified pathogen). However, there remains a paucity of age- and sex-specific MFs in the literature.

Where no MF exists, it is (under certain conditions) possible to 'borrow' or extrapolate from a disease of similar epidemiology or from the same disease in a country with a similar surveillance system, or likewise apply the same MF to a group of diseases (as demonstrated by Mead *et al.* [83]). However, it must be acknowledged that the base value of "cases reported" (Figure 1A and B) that we seek to adjust for UE by applying an appropriate MF, may not always capture the same proportion of infections that have occurred or provide comparable information of disease incidence estimates for different diseases or countries. Therefore, borrowing MFs (particularly from different countries and especially from different disease groups (e.g. taking MFs for STIs and applying to gastroenteric disease data)) is not a favoured method owing to the inherent heterogeneity of national surveillance systems in terms of population covered, test sensitivity and specificity, the source of data (physician, laboratory, hospital or other) and surveillance type (whether compulsory versus voluntary reporting of positive results or cases, comprehensive versus sentinel, active versus passive surveillance, case-based versus aggregated reporting [120]).

Here we did not address UE in the mortality reporting chain. Similar to the surveillance pyramid for morbidity data, the tip of mortality pyramid represents the cases correctly reported. The wide base of the pyramid contains data of all deaths including those that are ascertained and those that are not. Whilst it is expected that in a European setting under-ascertainment of deaths is rare (if not irrelevant), underreporting or over-reporting of mortality events due to certain diseases or conditions is not. The number of deaths may be well reported, but there is considerable misclassification of the cause of death. This misclassification may be deliberate in countries without nationalised healthcare such as the United States, where reimbursement by private insurance may be related to the ICD code used for the primary cause of death. Elsewhere, there may be other reasons to misclassify the cause of death, such as government targets and pressure to reduce the number of deaths due to a certain cause. In addition, often lacking are additional details relating to underlying conditions and sequelae that an individual died *with* (e.g. secondary and tertiary causes) but not necessarily *of* (i.e. the primary cause of death)

[13,121,122]. For example, chronic conditions with infectious causes (e.g. liver cirrhosis) are often not counted as sequelae deaths and therefore the surveillance system may underestimate the long-term burden due to the infection that led to the sequelae.

Conclusion

UE masks the true magnitude of disease incidence and reduces the efficiency of the notification system and surveillance potential [123]. In some instances, such as BoD estimates for the BCoDE-study and for comparing the impact of diseases between countries, it is necessary to quantify and adjust for UE. After correction for UE, preferably by age and sex, surveillance and notification data become a better estimate for evidence-based and comparable disease burden estimations. However, since adjusting for UE results in higher disease burden estimates and can result in diseases with differing ranks of public health importance compared with unadjusted surveillance data; care should be taken to clearly communicate both the need for such adjustment and the methodologies applied to adjust the raw data. The results presented here confirm that UR and UA have a significant impact resulting in UE of surveillance and notification data in our examples for salmonellosis and campylobacteriosis. To a varying extent, this is also true for all other pathogens in the BCoDE-study. The BCoDE-project is currently compiling and verifying estimates of UE and MFs derived from extensive literature reviews for 32 IDs. Here, we have presented several viable approaches for estimating UE and MFs of salmonellosis and campylobacteriosis although the best option will undoubtedly vary between countries.

Author details
[1]Centre for Immunity, Infection and Evolution, Ashworth Laboratories, Kings Buildings, University of Edinburgh, Edinburgh, UK. [2]Julius Centre for Health Sciences and Primary Care, University Medical Centre Utrecht, Utrecht, the Netherlands. [3]Department of Public Health Medicine, School of Public Health, University of Bielefeld, Bielefeld, Germany. [4]Centre for Infectious Disease Control, National Institute for Public Health and the Environment, Bilthoven, the Netherlands. [5]Institute for Risk Assessment Sciences, Utrecht University, Utrecht, the Netherlands. [6]European Centre for Disease Prevention and Control, Stockholm, Sweden. [7]Pallas, Health Research and Consultancy BV, Rotterdam, the Netherlands. [8]International Livestock Research Institute, Nairobi, Kenya. [9]Institute of Infection and Global Health, University of Liverpool, Liverpool, UK.

References
1. Keramarou M, Evans MR: **Completeness of infectious disease notification in the United Kingdom: a systematic review.** *J Infect* 2012, **64**(6):555–564.
2. Wheeler JG, Sethi D, Cowden JM, Wall PG, Rodrigues LC, Tompkins DS, Hudson MJ, Roderick PJ: **Study of infectious intestinal disease in England: rates in the community, presenting to general practice, and reported to national surveillance.** *BMJ* 1999, **318**(7190):1046–1050.
3. Kretzschmar M, Mangen M-JJ, Pinheiro P, Jahn B, Fèvre EM, Longhi S, Lai T, Havelaar AH, Stein C, Cassini A, et al: **New methodology for estimating the burden of infectious diseases in Europe.** *PLoS Med* 2012, **9**(4):e1001205.
4. Mangen M-JJ, Plass D, Havelaar AH, Gibbons CL, Cassini A, Mühlberger N, van Lier A, Haagsma JA, Brooke RJ, Lai T, et al: **The Pathogen- and Incidence-Based DALY Approach: An Appropriated Methodology for Estimating the Burden of Infectious Diseases.** *PLoS ONE* 2013, **8**(11):e79740.
5. Plass D, Mangen M-JJ, Havelaar AH, Gibbons C, Haagsma J, Jahn B, Lai T, van Lier A, Longhi S, McDonald SA, et al: **The incidence-based and pathogen-based disability-adjusted life-years approach for measuring infectious disease burden in Europe: the burden of communicable diseases in Europe (BCoDE) project.** *Lancet* 2013, **381**:S114.
6. McDonald SA, van Lier A, Plass D, Kretzschmar ME: **The impact of demographic change on the estimated future burden of infectious diseases: examples from hepatitis B and seasonal influenza in the Netherlands.** *BMC Public Health* 2012, **12**(1046):1471–2458.
7. Hardnett FP, Hoekstra RM, Kennedy M, Charles L, Angulo FJ, for the Emerging Infections Program FoodNet Working Group: **Epidemiologic issues in study design and data analysis related to FoodNet activities.** *Clin Infect Dis* 2004, **38**(Supplement 3):S121–S126.
8. MacDougall L, Majowicz S, Dore K, Flint J, Thomas K, Kovacs S, Sockett P: **Under-reporting of infectious gastrointestinal illness in British Columbia, Canada: who is counted in provincial communicable disease statistics?** *Epidemiol Infect* 2008, **136**(02):248–256.
9. Haagsma JA, Geenen PL, Ethelberg S, Fetsch A, Hansdotter F, Jansen A, Korsgaard H, O'Brien SJ, Scavia G, Spitznagel H, et al: **Community incidence of pathogen-specific gastroenteritis: reconstructing the surveillance pyramid for seven pathogens in seven European Union member states.** *Epidemiol Infect* 2012, **27**:1–15.
10. European Centre for Disease Prevention and Control (ECDC): *Report: Surveillance and Prevention of Hepatitis B and C in Europe.* Stockholm, Sweden: ECDC; 2010.
11. O'Brien S, Rait G, Hunter P, Gray J, Bolton F, Tompkins D, McLauchlin J, Letley L, Adak G, Cowden J, et al: **Methods for determining disease burden and calibrating national surveillance data in the United Kingdom: the second study of infectious intestinal disease in the community (IID2 study).** *BMC Med Res Methodol* 2010, **10**(1):39.
12. Sethi D, Wheeler J, Rodrigues LC, Fox S, Roderick P: **Investigation of under-ascertainment in epidemiological studies based in general practice.** *Int J Epidemiol* 1999, **28**(1):106–112.
13. Khosravi A, Rao C, Naghavi M, Taylor R, Jafari N, Lopez AD: **Impact of misclassification on measures of cardiovascular disease mortality in the Islamic Republic of Iran: a cross-sectional study.** *Bull World Health Organ* 2008, **86**(9):688–696.
14. Crowcroft NS, Andrews N, Rooney C, Brisson M, Miller E: **Deaths from pertussis are underestimated in England.** *Arch Dis Child* 2002, **86**(5):336–338.
15. Martin-Ampudia M, Mariscal A, Lopez-Gigosos RM, Mora L, Fernandez-Crehuet J: **Under-notification of cryptosporidiosis by routine clinical and laboratory practices among non-hospitalised children with acute diarrhoea in Southern Spain.** *Infection* 2012, **40**(2):113–119.
16. Yuguero O, Serna MC, Real J, Galvan L, Riu P, Godoy P: **[Using treatment compliance to determine the under-notification of tuberculosis in a health region for the years 2007–2009].** *Aten Primaria* 2012, **44**(12):703–708.
17. Doyle TJ, Glynn MK, Groseclose SL: **Completeness of notifiable infectious disease reporting in the united states: an analytical literature review.** *Am J Epidemiol* 2002, **155**(9):866–874.
18. European Centre for Disease Prevention and Control (ECDC): *Report: Methodology Protocol for Estimating Burden of Communicable Diseases.* Stockholm, Sweden: ECDC; 2010:20–25.

19. Reed C, Angulo FJ, Swerdlow DL, Lipsitch M, Meltzer MI, Jernigan D, Finelli L: **Estimates of the prevalence of pandemic (H1N1) 2009, United States, April-July 2009.** *Emerging infectious diseases* 2009, **15**(12):2004–2007.

20. del Beccaro MA, Brownstein DR, Cummings P, Goldoft MJ, Quan L: **Outbreak of Escherichia coli O157:H7 hemorrhagic colitis and hemolytic uremic syndrome: effect on use of a pediatric emergency department.** *Ann Emerg Med* 1995, **26**(5):598–603.

21. van Cauteren D, de Valk H, Vaux S, le Strat Y, Vaillant V: **Burden of acute gastroenteritis and healthcare-seeking behaviour in France: a population-based study.** *Epidemiol Infect* 2012, **140**(4):697–705.

22. Szczepura A: **Access to health care for ethnic minority populations.** *Postgrad Med J* 2005, **81**(953):141–147.

23. Thomas VN, Saleem T, Abraham R: **Barriers to effective uptake of cancer screening among Black and minority ethnic groups.** *Int J Palliat Nurs* 2005, **11**(11):564–571.

24. European Centre for Disease Prevention and Control: *Technical Report - Migrant Health: Background Note to the 'ECDC Report on Migration and Infectious Diseases in the EU'.* Stockholm, Sweden: ECDC; 2009.

25. Abdool Karim SS, Dilraj A: **Reasons for under-reporting of notifiable conditions.** *S Afr Med J* 1996, **86**(7):834–836.

26. Durrheim DN, Thomas J: **General practice awareness of notifiable infectious diseases.** *Public Health* 1994, **108**(4):273–278.

27. Spedding RL, Jenkins MG, O'Reilly SA: **Notification of infectious diseases by junior doctors in accident and emergency departments.** *J Accid Emerg Med* 1998, **15**(2):102–104.

28. Voss S: **How much do doctors know about the notification of infectious diseases?** *BMJ* 1992, **304**(6829):755–755.

29. Hsieh Y-H, Kuo M-J, Hsieh T-C, Lee H-C: **Underreporting and underestimation of gonorrhea cases in the Taiwan National Gonorrhea Notifiable Disease System in the Tainan region: evaluation by a pilot physician-based sentinel surveillance on Neisseria gonorrhoeae infection.** *Int J Infect Dis* 2009, **13**(6):e413–e419.

30. Riley LW, Finch MJ: **Results of the first year of national surveillance of campylobacter infections in the United States.** *J Infect Dis* 1985, **151**(5):956–959.

31. Yoder JS, Blackburn BG, Craun GC, Hill V, Levy DA, Chen N, Lee SH, Calderon RL, Beach MJ: **Surveillance for waterborne-disease outbreaks associated with recreational water –– United States, 2001–2002.** *MMWR Surveill Summ* 2004, **53**:1–22.

32. Hoxie NJ, Davis JP, Vergeront JM, Nashold RD, Blair KA: **Cryptosporidiosis-associated mortality following a massive waterborne outbreak in Milwaukee, Wisconsin.** *Am J Public Health* 1997, **87**(12):2032–2035.

33. Hall G, Yohannes K, Raupach J, Becker N, Kirk M: **Estimating community incidence of Salmonella, Campylobacter, and Shiga toxin-producing Escherichia coli infections, Australia.** *Emerg Infect Dis* 2008, **14**(10):1601–1609.

34. van den Brandhof WE, Bartelds AI, Koopmans MP, van Duynhoven YT: **General practitioner practices in requesting laboratory tests for patients with gastroenteritis in the Netherlands, 2001–2002.** *BMC Fam Pract* 2006, **7**:56.

35. Scholten JN, de Vlas SJ, Zaleskis R: **Under-reporting of HIV infection among cohorts of TB patients in the WHO European Region, 2003–2004.** *Int J Tuberc Lung Dis* 2008, **12**:S85–S91.

36. Borges-Costa J, Matos C, Pereira F: **Sexually transmitted infections in pregnant adolescents: prevalence and association with maternal and foetal morbidity.** *J Eur Acad Dermatol Venereol* 2012, **26**(8):972–975.

37. Adams OP, Carter AO, Prussia P, McIntyre G, Branch SL: **Risk behaviour, healthcare access and prevalence of infection with Chlamydia trachomatis and Neisseria gonorrhoeae in a population-based sample of adults in Barbados.** *Sex Transm Infect* 2008, **84**(3):192–194.

38. Lewis DA, Pillay C, Mohlamonyane O, Vezi A, Mbabela S, Mzaidume Y, Radebe F: **The burden of asymptomatic sexually transmitted infections among men in Carletonville, South Africa: implications for syndromic management.** *Sex Transm Infect* 2008, **84**(5):371–376.

39. Farley TA, Cohen DA, Elkins W: **Asymptomatic sexually transmitted diseases: the case for screening.** *Prev Med* 2003, **36**(4):502–509.

40. Schachter J, Stoner E, Moncada J: **Screening for chlamydial infections in women attending family planning clinics.** *West J Med* 1983, **138**(3):375–379.

41. Klouman E, Masenga EJ, Sam NE, Klepp KI: **Asymptomatic gonorrhoea and chlamydial infection in a population-based and work-site based sample of men in Kilimanjaro, Tanzania.** *Int J STD AIDS* 2000, **11**(10):666–674.

42. Benzaken AS, Galban EG, Antunes W, Dutra JC, Peeling RW, Mabey D, Salama A: **Diagnosis of gonococcal infection in high risk women using a rapid test.** *Sex Transm Infect* 2006, **82**:V26–V28.

43. Forni J, Miles K, Hamill M: **Microscopy detection of rectal gonorrhoea in asymptomatic men.** *Int J STD AIDS* 2009, **20**(11):797–798.

44. Hong Y, Fang X, Zhou Y, Zhao R, Li X: **Factors associated with sexually transmitted infection underreporting among female sex workers in China.** *J Womens Health (Larchmt)* 2011, **20**(1):129–136.

45. Sypsa V, Bonovas S, Tsiodras S, Baka A, Efstathiou P, Malliori M, Panagiotopoulos T, Nikolakopoulos I, Hatzakis A: **Estimating the disease burden of 2009 pandemic influenza A(H1N1) from surveillance and household surveys in Greece.** *PLoS One* 2011, **6**(6):e20593.

46. Centers for Disease Control and Prevention (CDC): **Self-reported influenza-like illness during the 2009 H1N1 influenza pandemic–United States, September 2009 - March 2010.** *MMWR Morb Mortal Wkly Rep* 2011, **60**(2):37–41.

47. Tam CC, Rodrigues LC, Viviani L, Dodds JP, Evans MR, Hunter PR, Gray JJ, Letley LH, Rait G, Tompkins DS, *et al*: **Longitudinal study of infectious intestinal disease in the UK (IID2 study): incidence in the community and presenting to general practice.** *Gut* 2012, **61**(1):69–77.

48. Mizoguchi Y, Suzuki E, Tsuchida H, Tsuda T, Yamamoto E, Nakase K, Doi H: **Outbreak of Salmonella Braenderup infection originating in boxed lunches in Japan in 2008.** *Acta Med Okayama* 2011, **65**(2):63–69.

49. Voetsch AC, van Gilder TJ, Angulo FJ, Farley MM, Shallow S, Marcus R, Cieslak PR, Deneen VC, Tauxe RV: **FoodNet estimate of the burden of illness caused by nontyphoidal Salmonella infections in the United States.** *Clin Infect Dis* 2004, **38**(Suppl 3):S127–134.

50. Kubota K, Kasuga F, Iwasaki E, Inagaki S, Sakurai Y, Komatsu M, Toyofuku H, Angulo FJ, Scallan E, Morikawa K: **Estimating the burden of acute gastroenteritis and foodborne illness caused by campylobacter, salmonella, and vibrio parahaemolyticus by using population-based telephone survey data, Miyagi prefecture, Japan, 2005 to 2006.** *J Food Prot* 2011, **74**(10):1592–1598.

51. Kuusi M, Aavitsland P, Gondrosen B, Kapperud G: **Incidence of gastroenteritis in Norway–a population-based survey.** *Epidemiol Infect* 2003, **131**(1):591–597.

52. Arias C, Sala MR, Dominguez A, Bartolome R, Benavente A, Veciana P, Pedrol A, Hoyo G: **Waterborne epidemic outbreak of Shigella sonnei gastroenteritis in Santa Maria de Palautordera, Catalonia, Spain.** *Epidemiol Infect* 2006, **134**(3):598–604.

53. Leder K, Sinclair M, Forbes A, Wain D: **Household clustering of gastroenteritis.** *Epidemiol Infect* 2009, **137**(12):1705–1712.

54. Wilking H, Spitznagel H, Werber D, Lange C, Jansen A, Stark K: **Acute gastrointestinal illness in adults in Germany: a population-based telephone survey.** *Epidemiol Infect* 2013, **1**:1–11.

55. Adak GK, Long SM, O'Brien SJ: **Trends in indigenous foodborne disease and deaths, England and Wales: 1992 to 2000.** *Gut* 2002, **51**(6):832–841.

56. Kemmeren JM, Mangen M-JJ, van Duynhoven YTHP, Havelaar AH: *Report: Priorization of foodborne pathogens: disease burden and costs of selected enteric pathogens.* Bilthoven: National Institute for Public Health and the Environment; 2006.

57. Haagsma JA, Siersema PD, de Wit NJ, Havelaar AH: **Disease burden of post-infectious irritable bowel syndrome in The Netherlands.** *Epidemiol Infect* 2010, **138**(11):1650–1656.

58. Havelaar AH, van Pelt W, Ang CW, Wagenaar JA, van Putten JP, Gross U, Newell DG: **Immunity to Campylobacter: its role in risk assessment and epidemiology.** *Crit Rev Microbiol* 2009, **35**(1):1–22.

59. Vijgen SMC, Mangen M-JJ, Koortbeek LM, van Duynhoven YTHP, Havelaar AH: *Disease Burden and Related Costs of two Protozoan Pathogens.* Bilthoven: National Institute for Public Health and the Environment; 2007.

60. Havelaar AH, van Duynhoven YT, Nauta MJ, Bouwknegt M, Heuvelink AE, de Wit GA, Nieuwenhuizen MG, van de Kar NC: **Disease burden in The Netherlands due to infections with Shiga toxin-producing Escherichia coli O157.** *Epidemiol Infect* 2004, **132**(3):467–484.

61. Mangen M-JJ, Havelaar AH, A.J.A.M. Bernsen R, Van Koningsveld R, De Wit GA: **The costs of human Campylobacter infections and sequelae in the Netherlands: A DALY and cost-of-illness approach.** *Food Economics - Acta Agriculturae Scandinavica, Section C* 2005, **2**(1):35–51.

62. van Pelt W, de Wit MAS, Wannet WJB, Ligtvoet EJJ, Widdowson MA, van Duynhoven Y: **Laboratory surveillance of bacterial gastroenteric pathogens in The Netherlands, 1991–2001.** *Epidemiol Infect* 2003, **130**(3):431–441.

63. Havelaar AH, Haagsma JA, Mangen MJ, Kemmeren JM, Verhoef LP, Vijgen SM, Wilson M, Friesema IH, Kortbeek LM, van Duynhoven YT, *et al*: **Disease**

burden of foodborne pathogens in the Netherlands, 2009. *Int J Food Microbiol* 2012, **156**(3):231–238.

64. Muller L, Korsgaard H, Ethelberg S: Burden of acute gastrointestinal illness in Denmark 2009: a population-based telephone survey. *Epidemiol Infect* 2012, **140**(2):290–298.

65. Baumann-Popczyk A, Sadkowska-Todys M, Rogalska J, Stefanoff P: Incidence of self-reported acute gastrointestinal infections in the community in Poland: a population-based study. *Epidemiol Infect* 2012, **140**(7):1173–1184.

66. Fitzgerald M, Scallan E, Collins C, Crowley D, Daly L, Devine M, Igoe D, Quigley T, Smyth B: Results of the first population based telephone survey of acute gastroenteritis in Northern Ireland and the Republic of Ireland. *Euro Surveill* 2004, **8**(18):2456.

67. McMahon BJ, Holck P, Bulkow L, Snowball M: Serologic and clinical outcomes of 1536 Alaska Natives chronically infected with hepatitis B virus. *Ann Intern Med* 2001, **135**(9):759–768.

68. Hagan H, Snyder N, Hough E, Yu T, McKeirnan S, Boase J, Duchin J: Case-reporting of acute hepatitis B and C among injection drug users. *J Urban Health* 2002, **79**(4):579–585.

69. de Melker HE, Versteegh FGA, Schellekens JFP, Teunis PFM, Kretzschmar M: The incidence of Bordetella pertussis infections estimated in the population from a combination of serological surveys. *J Infect* 2006, **53**(2):106–113.

70. de Melker H, Pebody RG, Edmunds WJ, Lévy-Bruhl D, Valle M, Rota MC, Salmaso S, van den Hof S, Berbers G, Saliou P, *et al*: The seroepidemiology of measles in Western Europe. *Epidemiol Infect* 2001, **126**(02):249–259.

71. Simms I, Rogers P, Catchpole M, McGarrigle CA, Nicoll A: Trends in undiagnosed HIV-1 infection among attenders at genitourinary medicine clinics, England, Wales, and Northern Ireland: 1990–6. *Sex Transm Infect* 1999, **75**(5):332–336.

72. Teunis PF, van Eijkeren JC, Ang CW, van Duynhoven YT, Simonsen JB, Strid MA, van Pelt W: Biomarker dynamics: estimating infection rates from serological data. *Stat Med* 2012, **31**(20):2240–2248.

73. Mangen M-JJ, Batz MB, Kasbohrer A, Hald T, Morris JG Jr, Taylor M, Havelaar AH: Integrated approaches for the public health prioritization of foodborne and zoonotic pathogens. *Risk Anal* 2010, **30**(5):782–797.

74. Ekdahl K, Giesecke J: Travellers returning to Sweden as sentinels for comparative disease incidence in other European countries, campylobacter and giardia infection as examples. *Euro Surveill* 2004, **9**(9):6–9.

75. de Jong B, Ekdahl K: The comparative burden of salmonellosis in the European Union member states, associated and candidate countries. *BMC Public Health* 2006, **6**:4.

76. Havelaar AH, Ivarsson S, Lofdahl M, Nauta MJ: Estimating the true incidence of campylobacteriosis and salmonellosis in the European Union, 2009. *Epidemiol Infect* 2013, **13**:1–10.

77. Simonsen J, Molbak K, Falkenhorst G, Krogfelt KA, Linneberg A, Teunis PF: Estimation of incidences of infectious diseases based on antibody measurements. *Stat Med* 2009, **28**(14):1882–1895.

78. Gallay A, Vaillant V, Bouvet P, Grimont P, Desenclos JC: How many foodborne outbreaks of Salmonella infection occurred in France in 1995? Application of the capture-recapture method to three surveillance systems. *Am J Epidemiol* 2000, **152**(3):171–177.

79. Jelastopulu E, Merekoulias G, Alexopoulos EC: Underreporting of communicable diseases in the prefecture of Achaia, western Greece, 1999–2004 - missed opportunities for early intervention. *Euro Surveill* 2010, **15**(21):19579.

80. Gkogka E, Reij MW, Havelaar AH, Zwietering MH, Gorris LGM: Risk-based estimate of effect of foodborne diseases on public health. *Greece Emerg Infect Dis* 2011, **17**(9):1581–1590.

81. Perez-Ciordia I, Ferrero M, Sanchez E, Abadias M, Martinez-Navarro F, Herrera D: Salmonella enteritis in Huesca. 1996–1999. *Enferm Infecc Microbiol Clin* 2002, **20**(1):16–21.

82. Jansson A, Arneborn M, Ekdahl K: Sensitivity of the Swedish statutory surveillance system for communicable diseases 1998. *Epidemiol Infect* 2005, **133**(03):401–407.

83. Mead PS, Slutsker L, Dietz V, McCaig LF, Bresee JS, Shapiro C, Griffin PM, Tauxe RV: Food-related illness and death in the United States. *Emerg Infect Dis* 1999, **5**(5):607.

84. Scallan E, Hoekstra RM, Angulo FJ, Tauxe RV, Widdowson MA, Roy SL, Jones JL, Griffin PM: Foodborne illness acquired in the United States–major pathogens. *Emerg Infect Dis* 2011, **17**(1):7–15.

85. Thomas MK, Majowicz SE, Sockett PN, Aamir F, Pollari F, Doré K, Flint JA, Edge VL: Estimated numbers of community cases of illness Due to salmonella, campylobacter and verotoxigenic escherichia coli: pathogen-specific community rates. *Can J Infect Dis Med Microbiol* 2006, **17**(4):229–234.

86. Kuusi M, Nuorti JP, Hanninen ML, Koskela M, Jussila V, Kela E, Miettinen I, Ruutu P: A large outbreak of campylobacteriosis associated with a municipal water supply in Finland. *Epidemiol Infect* 2005, **133**(4):593–601.

87. McCarty DJ, Tull ES, Moy CS, Kwoh CK, LaPorte RE: Ascertainment corrected rates: applications of capture-recapture methods. *Int J Epidemiol* 1993, **22**(3):559–565.

88. Trotter C, Samuelsson S, Perrocheau A, de Greeff S, de Melker H, Heuberger S, Ramsay M: Ascertainment of meningococcal disease in Europe. *Euro Surveill* 2005, **10**(12):247–250.

89. Hook EB, Regal RR: Capture-recapture methods in epidemiology: methods and limitations. *Epidemiol Rev* 1995, **17**(2):243–264.

90. van Hest NA, Grant AD, Smit F, Story A, Richardus JH: Estimating infectious diseases incidence: validity of capture-recapture analysis and truncated models for incomplete count data. *Epidemiol Infect* 2008, **136**(1):14–22.

91. Jelastopulu E, Alexopoulos EC, Venieri D, Tsiros G, Komninou G, Constantinidis TC, Chrysanthopoulos K: Substantial underreporting of tuberculosis in West Greece: implications for local and national surveillance. *Euro Surveill* 2009, **14**(11):19152.

92. Farchi S, Mantovani J, Borgia P, Giorgi Rossi P: Tuberculosis incidence, hospitalisation prevalence and mortality in Lazio, Italy, 1997–2003. *Int J Tuberc Lung Dis* 2008, **12**(2):193–198.

93. Baussano I, Bugiani M, Gregori D, van Hest R, Borraccino A, Raso R, Merletti F: Undetected burden of tuberculosis in a low-prevalence area. *Int J Tuberc Lung Dis* 2006, **10**(4):415–421.

94. van Hest NA, Smit F, Baars HW, de Vries G, de Haas PE, Westenend PJ, Nagelkerke NJ, Richardus JH: Completeness of notification of tuberculosis in The Netherlands: how reliable is record-linkage and capture-recapture analysis? *Epidemiol Infect* 2007, **135**(6):1021–1029.

95. van Loenhout-Rooyackers JH, Leufkens HGM, Hekster YA, Kalisvaart NA: Pyrazinamide use as a method of estimating under-reporting of tuberculosis. *Int J Tuberc Lung Dis* 2001, **5**(12):1156–1160.

96. Cojocaru C, van Hest NA, Mihaescu T, Davies PD: Completeness of notification of adult tuberculosis in Iasi County, Romania: a capture-recapture analysis. *Int J Tuberc Lung Dis* 2009, **13**(9):1094–1099.

97. Teo SSS, Alfaham M, Evans MR, Watson JM, Riordan A, Sonnenberg P, Clark J, Hayward A, Sharland M, Moore-Gillon J, *et al*: An evaluation of the completeness of reporting of childhood tuberculosis. *Eur Respir J* 2009, **34**(1):176–179.

98. van Hest NA, Story A, Grant AD, Antoine D, Crofts JP, Watson JM: Record-linkage and capture-recapture analysis to estimate the incidence and completeness of reporting of tuberculosis in England 1999–2002. *Epidemiol Infect* 2008, **136**(12):1606–1616.

99. Pillaye J, Clarke A: An evaluation of completeness of tuberculosis notification in the United Kingdom. *BMC Public Health* 2003, **3**(1):31.

100. Mancuso JD, Tobler SK, Eick AA, Olsen CH: An evaluation of the completeness and accuracy of active tuberculosis reporting in the United States military. *Int J Tuberc Lung Dis* 2010, **14**(10):1310–1315.

101. Curtis AB, McCray E, McKenna M, Onorato IM: Completeness and timeliness of tuberculosis case reporting: a multistate study. *Am J Prev Med* 2001, **20**(2):108–112.

102. Dijkstra F, Donker GA, Wilbrink B, van Gageldonk-Lafeber AB, van der Sande MA: Long time trends in influenza-like illness and associated determinants in The Netherlands. *Epidemiol Infect* 2009, **137**(4):473–479.

103. McDonald SA, Presanis AM, de Angelis D, van der Hoek W, Hooiveld M, Donker G, Kretzschmar ME: An evidence synthesis approach to estimating the incidence of seasonal influenza in the Netherlands. *Influenza Other Respir Viruses* 2014, **8**(1):33–41.

104. Stein CE, Birmingham M, Kurian M, Duclos P, Strebel P: The global burden of measles in the year 2000 - a model that uses country-specific indicators. *J Infect Dis* 2003, **187**(SUPPL. 1):S8–S14.

105. Scallan E, Mahon BE, Hoekstra RM, Griffin PM: Estimates of illnesses, hospitalizations and deaths caused by major bacterial enteric pathogens in young children in the United States. *Pediatr Infect Dis J* 2013, **32**(3):217–221.

106. Thomas MK, Murray R, Flockhart L, Pintar K, Pollari F, Fazil A, Nesbitt A, Marshall B: Estimates of the burden of foodborne illness in Canada for 30 specified pathogens and unspecified agents, circa 2006. *Foodborne Pathog Dis* 2013, **10**(7):639–648.

107. le Vu S, le Strat Y, Barin F, Pillonel J, Cazein F, Bousquet V, Brunet S, Thierry D, Semaille C, Meyer L, et al: Population-based HIV-1 incidence in France, 2003–08: a modelling analysis. Lancet Infect Dis 2010, 10(10):682–687.

108. Armstrong GL, Bell BP: Hepatitis A virus infections in the United States: model-based estimates and implications for childhood immunization. Pediatrics 2002, 109(5):839–845.

109. Aledort JE, Ronald A, Rafael ME, Girosi F, Vickerman P, le Blancq SM, Landay A, Holmes K, Ridzon R, Hellmann N, et al: Reducing the burden of sexually transmitted infections in resource-limited settings: the role of improved diagnostics. Nature 2006, 1:59–72.

110. O'Connor JB, Imperiale TF, Singer ME: Cost-effectiveness analysis of hepatitis A vaccination strategies for adults. Hepatology 1999, 30(4):1077–1081.

111. Albert I, Espié E, de Valk H, Denis J-B: A Bayesian evidence synthesis for estimating campylobacteriosis prevalence. Risk Anal 2011, 31(7):1141–1155.

112. Bender JB, Smith KE, McNees AA, Rabatsky-Ehr TR, Segler SD, Hawkins MA, Spina NL, Keene WE, Kennedy MH, van Gilder TJ, et al: Factors affecting surveillance data on Escherichia coli O157 infections collected from FoodNet sites, 1996–1999. Clin Infect Dis 2004, 15(38):S157–164.

113. Birrell PJ, Ketsetzis G, Gay NJ, Cooper BS, Presanis AM, Harris RJ, Charlett A, Zhang X-S, White PJ, Pebody RG, et al: Bayesian modeling to unmask and predict influenza A/H1N1pdm dynamics in London. Proc Natl Acad Sci 2011, 108(45):18238–18243.

114. Presanis AM, Pebody RG, Paterson BJ, Tom BD, Birrell PJ, Charlett A, Lipsitch M, de Angelis D: Changes in severity of 2009 pandemic A/H1N1 influenza in England: a Bayesian evidence synthesis. BMJ 2011, 8(343):d5408.

115. Beutels P, Musabaev EI, van Damme P, Yasin T: The disease burden of hepatitis B in Uzbekistan. J Infect 2000, 40(3):234–241.

116. Ziv T, Heymann AD, Azuri J, Leshno M, Cohen D: Assessment of the underestimation of childhood diarrhoeal disease burden in Israel. Epidemiol Infect 2011, 139(9):1379–1387.

117. Imhoff B, Morse D, Shiferaw B, Hawkins M, Vugia D, Lance-Parker S, Hadler J, Medus C, Kennedy M, Moore MR, et al: Burden of self-reported acute diarrheal illness in FoodNet surveillance areas, 1998–1999. Clin Infect Dis 2004, 15(38):S219–226.

118. Herikstad H, Yang S, van Gilder TJ, Vugia D, Hadler J, Blake P, Deneen V, Shiferaw B, Angulo FJ: A population-based estimate of the burden of diarrhoeal illness in the United States: FoodNet, 1996–7. Epidemiol Infect 2002, 129(1):9–17.

119. de Wit MA, Kortbeek LM, Koopmans MP, de Jager CJ, Wannet WJ, Bartelds AI, van Duynhoven YT: A comparison of gastroenteritis in a general practice-based study and a community-based study. Epidemiol Infect 2001, 127(3):389–397.

120. Pebody RG, Hellenbrand W, D'Ancona F, Ruutu P: Pneumococcal disease surveillance in Europe. Euro Surveill 2006, 11(9):171–178.

121. Lilienfeld DE, Stolley PD: Foundations of Epidemiology. 3rd edition. New York: Oxford University Press; 1994.

122. Lozano R, Murray CJL, Lopez AD, Satoh T: Miscoding and misclassification of ischaemic heart disease mortality. In Global Programme on Evidence for Health Policy Working Paper No 12. Geneva: World Health Organization; 2001.

123. Brabazon ED, O'Farrell A, Murray CA, Carton MW, Finnegan P: Under-reporting of notifiable infectious disease hospitalizations in a health board region in Ireland: room for improvement? Epidemiol Infect 2008, 136(02):241–247.