Community incidence of pathogen-specific gastroenteritis: reconstructing the surveillance pyramid for seven pathogens in seven European Union member states

J. A. HAAGSMA[1,2]*, P. L. GEENEN[1], S. ETHELBERG[3], A. FETSCH[4],
F. HANSDOTTER[5], A. JANSEN[6,7], H. KORSGAARD[8], S. J. O'BRIEN[9],
G. SCAVIA[10], H. SPITZNAGEL[4,6], P. STEFANOFF[11], C. C. TAM[12],
A. H. HAVELAAR[1,13], on behalf of a Med-Vet-Net Working Group†

[1] *Centrum voor Infectieziektebestrijding, Rijksinstituut voor Volksgezondheid en Milieu, Bilthoven,
The Netherlands;* [2] *Erasmus Medisch Centrum, Erasmus Universiteit, Rotterdam, The Netherlands;*
[3] *Statens Serum Institut, Copenhagen, Denmark;* [4] *Bundesinstitut für Risikobewertung, Berlin, Germany;*
[5] *Smittskyddsinstitutet, Solna, Sweden;* [6] *Robert Koch Institut, Berlin, Germany;* [7] *European Centre for Disease
Prevention and Control, Stockholm, Sweden;* [8] *National Food Institute, Technical University of Denmark,
Søborg, Denmark;* [9] *University of Liverpool, South Wirral, UK;* [10] *Instituto Superiore di Sanità, Rome, Italy;*
[11] *Narodowy Instytut Zdrowia Publicznego – Panstwowy Zaklad, Warsaw, Poland;* [12] *London School of Hygiene
& Tropical Medicine, London, UK;* [13] *Institute for Risk Assessment Sciences, Utrecht Universiteit, Utrecht,
The Netherlands*

SUMMARY

By building reconstruction models for a case of gastroenteritis in the general population moving through different steps of the surveillance pyramid we estimated that millions of illnesses occur annually in the European population, leading to thousands of hospitalizations. We used data on the healthcare system in seven European Union member states in relation to pathogen characteristics that influence healthcare seeking. Data on healthcare usage were obtained by harmonized cross-sectional surveys. The degree of under-diagnosis and underreporting varied by pathogen and country. Overall, underreporting and under-diagnosis were estimated to be lowest for Germany and Sweden, followed by Denmark, The Netherlands, UK, Italy and Poland. Across all countries, the incidence rate was highest for *Campylobacter* spp. and *Salmonella* spp. Incidence estimates resulting from the pyramid reconstruction approach are adjusted for biases due to different surveillance systems and are therefore a better basis for international comparisons than reported data.

Key words: *Campylobacter*, gastrointestinal infections, incidence, public health, *Salmonella*.

INTRODUCTION

There is an increasing interest in assessing the burden of foodborne disease [1–3]. Valid and representative data on the incidence of specific pathogens causing gastroenteritis are key components of these assessments. However, reported data, which are largely based on passive surveillance, underestimate the true incidence. Underreporting and under-diagnosis contribute to this problem. Under-diagnosis is the failure of the healthcare system to capture cases in the community that do not seek medical advice (i.e. the number of cases seeking medical attention divided by the total number of cases in the population) [1].

* Author for correspondence: J. A. Haagsma, Erasmus MC, PO Box 2040, 3000 CA Rotterdam, The Netherlands.
(Email: j.haagsma@erasmusmc.nl)
† Additional members of the Working Group are given in the Appendix.

Health outcomes caused by infectious gastroenteritis vary from mild to very severe and recorded diseases often represent only the tip of the iceberg (i.e. surveillance pyramid) of all disease in a particular region. Underreporting refers to cases that have sought medical advice but are not correctly diagnosed, classified, notified, or disseminated to surveillance authorities (i.e. the number of cases correctly diagnosed, classified, notified, or disseminated to surveillance authorities divided by the total number of cases seeking medical attention) [1]). The degree of underreporting and under-diagnosis varies by pathogen and by country, due to differences in routinely investigated pathogens, differences in healthcare use and differences in laboratory practice and surveillance.

A few cohort studies provide insight into the degree of underreporting and under-diagnosis (the 'multiplier') of disease caused by gastrointestinal pathogens. In England and Wales, the IID1 study estimated that in 1993–1996 there were 3·2 cases of salmonellosis in the population for every case reported to national surveillance [4]. The recently published IID2 study indicated that in 2008–2009, this multiplier had increased to 4·7. Possible reasons for this include a general decrease in use of General Practitioner (GP) services by gastroenteritis patients in particular because of increased self-management and possibly a decrease in symptom severity [5]. In The Netherlands, based on the Sensor study in 1999, the multiplier for salmonellosis was estimated as 13·4 [6, 7]. A recent study, based on disease risks in Swedish travellers estimated that for the European Union (EU) as a whole, the multiplier for salmonellosis was 58, but ranged between 0·4 and 2000 in different countries [8]. Even though such cohort studies provide valuable data for the calibration of surveillance data, they are highly demanding on resources and this prevents them from being performed in more countries.

As an alternative, several research groups have conducted studies that explicitly reconstruct the surveillance pyramid to estimate the degree of underreporting and under-diagnosis by estimating the proportion of patients who would visit their GP, and would submit a faecal specimen [9, 10]. The reconstruction models also include estimates of laboratory practice (i.e. if particular pathogens are analysed, and if positive samples are reported to national surveillance authorities). These studies have typically addressed a specific pathogen, i.e. Shiga toxin-producing *E. coli* O157 (STEC) in Canada [11]. Some studies have taken multiple pathogens into account in a region or country, i.e. Ontario, Canada [12], Australia [13] or the USA [1]. Due to differences in details of the reconstruction approach, the results of these studies cannot directly be compared. The scope of this study is to develop a transparent model to reconstruct the surveillance pyramid for seven pathogens that cause gastroenteritis in seven EU member states, which can serve as a basis for prioritization of foodborne and zoonotic hazards in the EU.

METHODS

Countries

Germany (DE), Denmark (DK), Italy (IT), The Netherlands (NL), Poland (PL), Sweden (SE) and the United Kingdom (UK) participated in the study, as these countries had previously conducted telephone surveys allowing for estimation of key parameters that are necessary for reconstruction of the surveillance pyramid.

Pathogens

Based on availability of routine surveillance data in the participating countries, *Campylobacter* spp. (Camp), *Salmonella* spp. (Salm), *Yersinia enterocolitica* (Yers), *Shigella* spp. (Shig), Shiga-toxin producing *Escherichia coli* O157 (STEC), enteropathogenic *Escherichia coli* (EPEC), and *Cryptosporidium* spp. (Cryp) were selected for analysis. Not all countries had data for all pathogens.

Reconstruction model parameters

The model consists of sets of country-specific and pathogen-specific parameters. All but one of these parameters are represented by beta distributions and were based on observed data in the different countries. Table 1 provides an overview of the sets of country-specific and pathogen-specific parameters. Country-specific parameters in Table 1 will result in different multipliers. It is assumed that the probability of visiting a GP, and of submitting a sample for laboratory analysis differs for patients with bloody and non-bloody diarrhoea. Other potential determinants of GP visits or sample submission, such as duration of illness and age of the patient, were not taken into account, because no data were consistently available. Pathogen-specific parameters were assumed not to vary between countries, but result in different

Table 1. *Parameters used in the pyramid reconstruction model*

Symbol	Description	Distribution
Country-specific parameters		
	Probability of visiting a GP with:	
a	• Bloody diarrhoea	Beta($a1$;$a2$)
b	• Non-bloody diarrhoea	Beta($b1$;$b2$)
	Probability of submitting a stool sample for a consulting patient with:	
c	• Bloody diarrhoea	Beta($c1$;$c2$)
d	• Non-bloody diarrhoea	Beta($d1$;$d2$)
e	Probability of submitting a stool sample for a hospitalized patient	Beta($e1$;$e2$)
	Probability of analysing a pathogen* in samples for:	
f	• Patients visiting a GP	Beta($f1$;$f2$)
g	• Hospitalized patients	Beta($g1$;$g2$)
	Probability of reporting a positive laboratory result for:	
h	• Patients visiting a GP	Beta($h1$;$h2$)
i	• Hospitalized patients	Beta($i1$;$i2$)
Pathogen-specific parameters		
j	Sensitivity of laboratory analysis	
	• *Campylobacter*, *Salmonella*, *Yersinia*, *Shigella*	Triang($j3$;$j4$;$j5$)
	• STEC, EPEC, *Cryptosporidium*	Beta($j1$;$j2$)
k	Proportion of bloody diarrhoea in population cases	Beta($k1$;$k2$)

GP, General Practitioner; STEC, Shiga toxin-producing *E. coli* O157; EPEC enteropathogenic *Escherichia coli*.
* Different within a country for each pathogen.

multipliers for pathogens within one country. The probability of submitting a stool sample for a hospitalized patient was modelled separately, and was not assumed to depend on the nature of the diarrhoeal illness. The probability of analysing a submitted stool sample for a specific pathogen was assumed to vary between countries as well as between pathogens, and also to depend on the origin of the sample. Reporting positive laboratory results was assumed to vary depending on the origin of the sample. The sensitivity of laboratory methods and the proportion of patients with bloody diarrhoea were assumed to vary between pathogens, but not by country.

Model equations

The reconstruction model estimates the probability for a case of gastroenteritis by a specific pathogen in the general population to move through different steps of the surveillance pyramid (i.e. the probability of visiting a GP, the probability of submitting a stool sample for a consulting patient/hospitalized patient, the probability of analysing a pathogen in samples of a consulting patient/hospitalized patient, and the probability of reporting a positive laboratory result for a consulting patient/hospitalized patient). Ultimately, the reconstruction model estimates the multiplier for each reported case to the number of

symptomatic cases in the population. The model equations are presented in Table 2. The source of data for the model are laboratory-diagnosed cases reported to national surveillance and the number of these who are hospitalized. The reconstruction model estimates the probability of a case at different layers of the surveillance pyramid. Uncertainty in the results of the reconstruction model was explored by Monte Carlo simulations and the reconstruction model was validated by comparison to the results of previously performed independent studies that used other study designs (population-based surveys).

Data

The annual number of reported cases per pathogen and per country was obtained from national surveillance systems. To moderate the effect of annual variation in incidence, the average number of reported cases in the years 2001–2005 was used. Data were corrected for the coverage of the surveillance network if the surveillance network did not cover the entire population of a country. In Italy, the surveillance network covered three provinces. Hence, it was decided to limit the study to these three provinces. In The Netherlands, Camp and Salm data were extrapolated from the whole country by applying scaling factors which corrected for surveillance

Table 2. *Model for reconstructing the surveillance pyramid for a specific pathogen in a specific country**

Symbol	Description	Formula
n_R	Number of reported cases per year†	Data
n_H	Number of hospitalized cases per year	Data
n_{GP}	Number of cases who are not hospitalized, but visit a GP	$n_R - n_H$
p	Probability of visiting a GP with gastroenteritis	$k \cdot a + (1-k)b$
m	Probability of submitting a stool sample when visiting a GP	$k \cdot c + (1-k)d$
	Probability of reporting a case for	
n	• Patients visiting a GP	$m \cdot f \cdot j \cdot h$
o	• Hospitalized patients	$e \cdot g \cdot j \cdot i$
N_{GP}	Total number of cases visiting a GP	n_{GP}/n
N_H	Total number of hospitalized cases	n_H/o
N_P	Total cases in the population	$(N_{GP} + N_H)/p$
N_{GP-}	Cases in the population who do not visit a GP	$N_P - (N_{GP} + N_H)$
M	Multiplier	N_P/n_R

GP, General Practitioner.

* Model parameters are further described in Table 1.

† Average for 2001–2005 when available.

network coverage. The data were obtained from the following sources:

- *DE*: Surveillance data were obtained from the national level database SurvNet (accessed 23 April 2008), hosted at the federal national health authority, the Robert Koch Institute (RKI) in Berlin. With implementation of the Protection against Infection Act (Infektionsschutzgesetz, IfSG; http://bundesrecht.juris.de/ifsg/index.html) in 2001, the seven pathogens belong to a list of 47 pathogens for which a laboratory-confirmed acute infection is notifiable to the local health department (LHD) by the investigating laboratory. After verification of a case according to national case definitions the LHD forwards the case electronically via the state health department to the RKI [14].
- *DK*: Case numbers were obtained from national surveillance from the Register of Enteric Pathogens [15]. The number of hospitalized cases was based on Helms *et al.* [16].
- *IT*: Case numbers for Salm were obtained from the national official surveillance. Data for Camp, available only for limited areas of the country, were obtained from the laboratory surveillance network for enteric pathogens (Enter-Net Italia; http://www.salute.gov.it/malattieInfettive/pagina InternaMenuMalattieInfettive. jsp?id = 812&menu = strumentiservizi.). The number of hospitalized cases were obtained from the hospital discharge database.
- *NL*: Case numbers and number of hospitalized cases per pathogen were obtained from laboratory surveillance [17].

- *PL*: Case numbers and number of hospitalized cases per pathogen were obtained from the national comprehensive surveillance system based on mandatory notifications of physicians of all suspected and diagnosed cases; the reported five pathogens belonged to 79 diseases and syndromes under surveillance during 2002–2008.
- *SE*: Case numbers and number of hospitalized cases per pathogen were obtained from surveillance data [18] and database of closed medical care records.
- *UK*: Case numbers and number of hospitalized cases per pathogen were obtained from surveillance data [4].

The country-specific parameters on GP visits and sample submission were derived from harmonized cross-sectional surveys of acute gastroenteritis incidence in the community. All studies used a similar design and the same case definition. A case of gastroenteritis was defined as a person with at least three loose stools, or any vomiting, in 24 h, in the 4 weeks prior to completion of the questionnaire, but excluding those (*a*) with cancer of the bowel, irritable bowel syndrome, Crohn's disease, ulcerative colitis, celiac disease, or another chronic illness with symptoms of diarrhoea or vomiting, or (*b*) who report their symptoms were due to drugs, alcohol, or pregnancy. In each of these surveys, randomly selected residents of private households were contacted either by telephone (DE, DK, IT, PL, UK) or by mail (NL, SE). The studies were carried out in 2008 and 2009 and were reported in detail elsewhere ([5, 19–22]; unpublished

Table 3. *Reported cases per year by pathogen and by country* (*average 2001–2005*)

Pathogen	DE	DK	IT	NL	PL	SE	UK*
Inhabitants ($\times 10^6$)	83·5	5·4	57·7†	16·2	38·2	9·0	50·0
Total cases							
Campylobacter	55 355	3987	475†	6541	10 600	7176	47 968
Salmonella	64 380	2,007	9888	2747	13 531	3916	14 025
Yersinia	6624	248	—	430	135	692	25
Shigella	1181	146	—	329	60	468	1190
STEC	104	138	—	140	—	102	738
EPEC	3497	457	—	—	1129	—	—
Cryptosporidium	1084	—	—	314	—	—	4128
Hospitalized							
Campylobacter	6459	429	69†	1177	6400	619	2399
Salmonella	13 162	356	4253	694	9548	393	420
Yersinia	1093	23	—	5	91	65	0·1
Shigella	156	20	—	24	38	36	29
STEC	26	11	—	21	2	44	280
EPEC	566	46	—	—	866	—	—
Cryptosporidium	83	—	—	10	—	—	10

DE, Germany; DK, Denmark; IT, Italy; NL, The Netherlands; PL, Poland; SE, Sweden; UK, United Kingdom.
STEC, Shiga toxin-producing *E. coli* O157; EPEC enteropathogenic *Escherichia coli*.
—, No data.
* Data for England only.
† For *Campylobacter*, data from three provinces only (2 million inhabitants) were available.

data, Germany). The probability of submitting a sample for hospitalized patients was based on the cross-sectional survey of gastroenteritis in DE and IT, expert opinion in DK, SE, and IT [23], the GEOPS study in NL [24, 25], a health utilization survey in PL [26], a cross-sectional survey in SE (F. Hansdotter, personal communication) and the IID studies in UK. As the study period for this paper (2001–2005) was in between the time of execution of the IID1 [27] and IID2 [28] studies, the mean value of the results from the two studies was used. The probability of analysing a sample for a specific pathogen was based on expert opinion, except for NL and IT where survey data were available [29]. Data on the probability of reporting a pathogen was based on expert opinion for all countries. When using expert opinion, subjective beta distributions were defined with parameters chosen to reflect the degree of confidence in the estimates.

To assess the proportion of bloody diarrhoea per pathogen a literature review was performed. Details are provided in Annex B (available online). The sensitivity of laboratory tests was based on Ethelberg *et al.* [30] for Camp, Salm, Yers and Shig; expert opinion for STEC and EPEC and Weber *et al.* [31] for Cryp. Uncertainty was expressed by beta or triangular distributions.

Uncertainty analysis

Stochastic models were built using @RISK 5.7 (Palisade Corporation, USA), a Monte Carlo simulation add-in to Microsoft Excel. The model was run for 10 000 iterations to stabilize the output distributions. Median values from the output distributions and 95% credible intervals were reported.

RESULTS

Reported cases

The average annual number of reported cases per pathogen and country is shown in Table 3. For IT, data on Camp were only available for three provinces, and the results are also reported for these provinces (with 2 million inhabitants). For the UK, data were used for England only.

Country-specific parameters

The median values of the country-specific parameters are shown in Table 4*a*. Details of distribution functions are shown in Annex A (available online). The probability of visiting a GP and submitting a stool sample was greater for cases with bloody diarrhoea

Table 4*a*. *Median values of country-specific parameters*

Parameter	Short description	DE	DK	IT	NL	PL	SE	UK
a	Visit GP, bloody	0·51	0·61	0·50	0·31	0·63	0·61	0·22
b	Visit GP, non-bloody	0·34	0·13	0·36	0·08	0·28	0·08	0·16
c	Submit sample, bloody	0·50	0·29	0·50	0·82	0·23	0·39	0·30
d	Submit sample, non-bloody	0·34	0·18	0·01	0·10	0·02	0·48	0·19
e	Submit sample, hospital	0·75	0·93	0·82	0·89	0·70	0·84	0·29
f	Analyse sample, GP							
	Campylobacter	0·99	0·99	0·52	0·87	0·61	0·99	1·00
	Salmonella	0·99	0·99	0·91	0·84	0·99	0·99	1·00
	Yersinia	0·61	0·99	—	0·56	0·29	0·87	0·02
	Shigella	0·99	0·99	—	0·78	0·61	0·99	1·00
	STEC	0·29	0·39	—	0·09	0·50	0·50	1·00
	EPEC	0·29	0·29	—	—	—	—	—
	Cryptosporidium	0·18	—	—	0·11	—	—	1·00
g	Analyse sample, hospital							
	Campylobacter	0·99	0·99	0·52	0·99	0·61	0·99	0·99
	Salmonella	0·99	0·99	0·91	0·99	0·99	0·99	0·99
	Yersinia	0·61	0·99	—	0·20	0·29	0·87	0·02
	Shigella	0·99	0·99	—	0·99	0·61	0·99	0·99
	STEC	0·29	0·39	—	0·82	0·50	0·60	0·99
	EPEC	0·29	0·29	—	—	—	—	—
	Cryptosporidium	0·18	—	—	0·18	—	—	0·99
h	Report result, GP	0·99	0·99	0·71	1·00	0·71	0·99	0·69
i	Report result, hospital	0·99	0·99	0·82	1·00	0·93	0·99	0·98

DE, Germany; DK, Denmark; IT, Italy; NL, The Netherlands; PL, Poland; SE, Sweden; UK, United Kingdom.
GP, General Practitioner; STEC, Shiga toxin-producing *E. coli* O157; EPEC enteropathogenic *Escherichia coli*.
—, No data.

Table 4*b*. *Median values of pathogen-specific parameters*

| Pathogen	Sensitivity of diagnostic method	Proportion of patients with bloody diarrhoea
Campylobacter	76%	17%
Salmonella	88%	37%
Yersinia	78%	33%
Shigella	63%	25%
STEC	71%	85%
EPEC	71%	25%
Cryptosporidium	50%	1%

STEC, Shiga toxin-producing *E. coli* O157; EPEC enteropathogenic *Escherichia coli*.

than for cases with non-bloody diarrhoea (except submission of stool samples in SE; there was considerable uncertainty in this estimate due to a small sample size). The probability of visiting a GP was smaller in NL than in other countries, but the probability of submitting a sample was notably higher in that country for cases with bloody diarrhoea. The probability of submitting a stool sample in PL was considerably lower than in other countries. Salm and Shig were commonly tested for in samples from the GP (with lower estimates for NL), whereas testing for Camp was common in some (DE, DK, SE) but less so in other (IT, NL, PL) countries. Other pathogens were typically tested for in less than 50% of samples, except testing for Yers in SE. In samples from hospitals, pathogen testing was as frequent as or more frequent than in samples from GPs. Reporting of positive results was assumed to take place with high probability, except for IT.

Pathogen-specific parameters

The median values of the pathogen-specific parameters are shown in Table 4*b*. There is very little information on the performance characteristics of routine methods in medical microbiological laboratories. The sensitivity of isolating Camp, Salm, Yers and Shig was between 63% and 88%, based on data from DK. Such data were not available for STEC and EPEC and was assumed to be in the same range as the bacteria, i.e. median value of 71% (see Annex A, Table A2). To stress the uncertainty in this estimate, a beta distribution with a wide confidence interval was
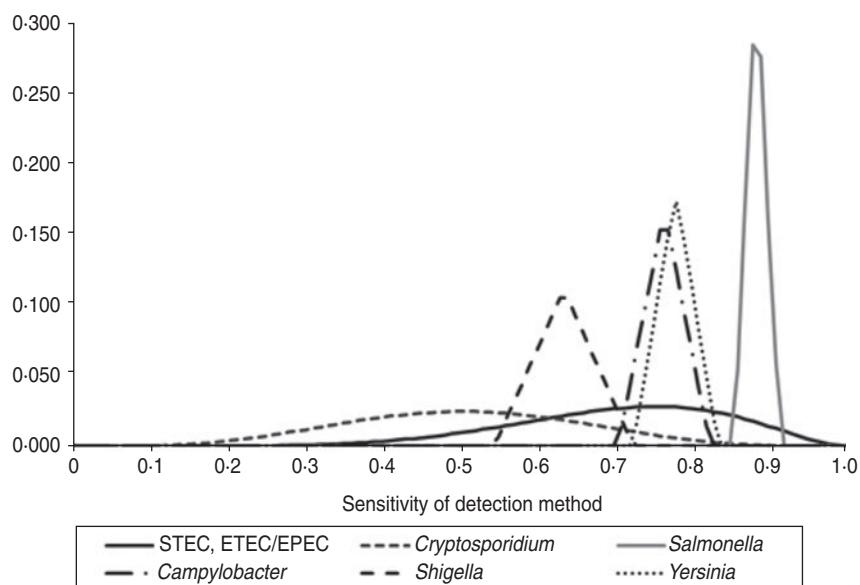
Fig. 1. Uncertainty of the sensitivity of isolating gastrointestinal pathogens. STEC, Shiga toxin-producing *E. coli* O157; ETEC/EPEC, enterotoxigenic *Escherichia coli*/enteropathogenic *Escherichia coli*.
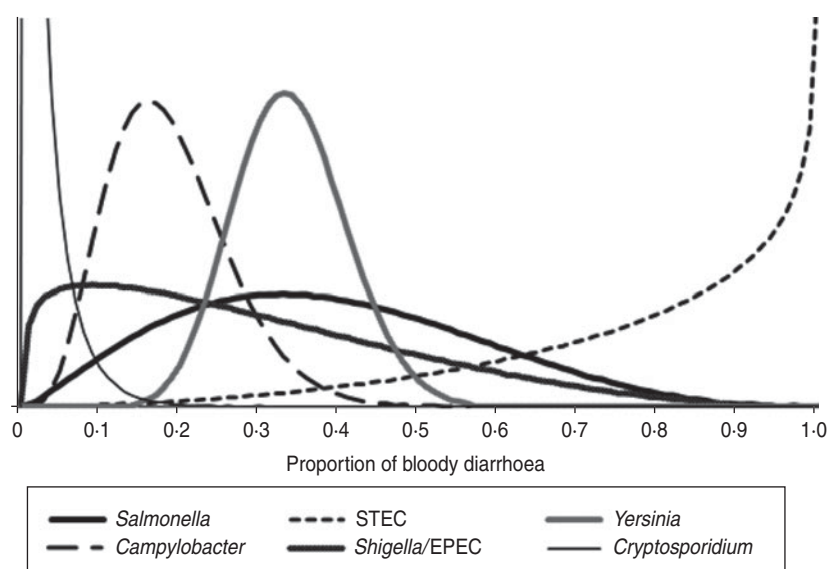
Fig. 2. Uncertainty distribution of the proportion of bloody diarrhoea in patients with gastroenteritis, per pathogen. STEC, Shiga toxin-producing *E. coli* O157; EPEC enteropathogenic *Escherichia coli*.

generated. The microscopy methods used for Cryp are generally less sensitive and an average of 50% was used. Uncertainty distributions are presented in Figure 1 (details can be found in Annex A).

The proportion of bloody diarrhoea in population cases, based on the literature survey is presented in Figure 2. As no data were available for EPEC, data for Shig were used as a proxy. With a median of 85% of cases with bloody diarrhoea, the highest proportion of bloody diarrhoea was caused by STEC.

The median proportions of bloody diarrhoea due to Salm, Yers, Shig/EPEC, and Camp were 37%, 33%, 25% and 17%, respectively. The lowest proportion of bloody diarrhoea was associated with Cryp (1%). Further details of the review are given in Annex B.

Pyramid reconstruction

In Figure 3*a*, the stepwise reconstruction of the pyramid for Salm is presented. Only the reporting
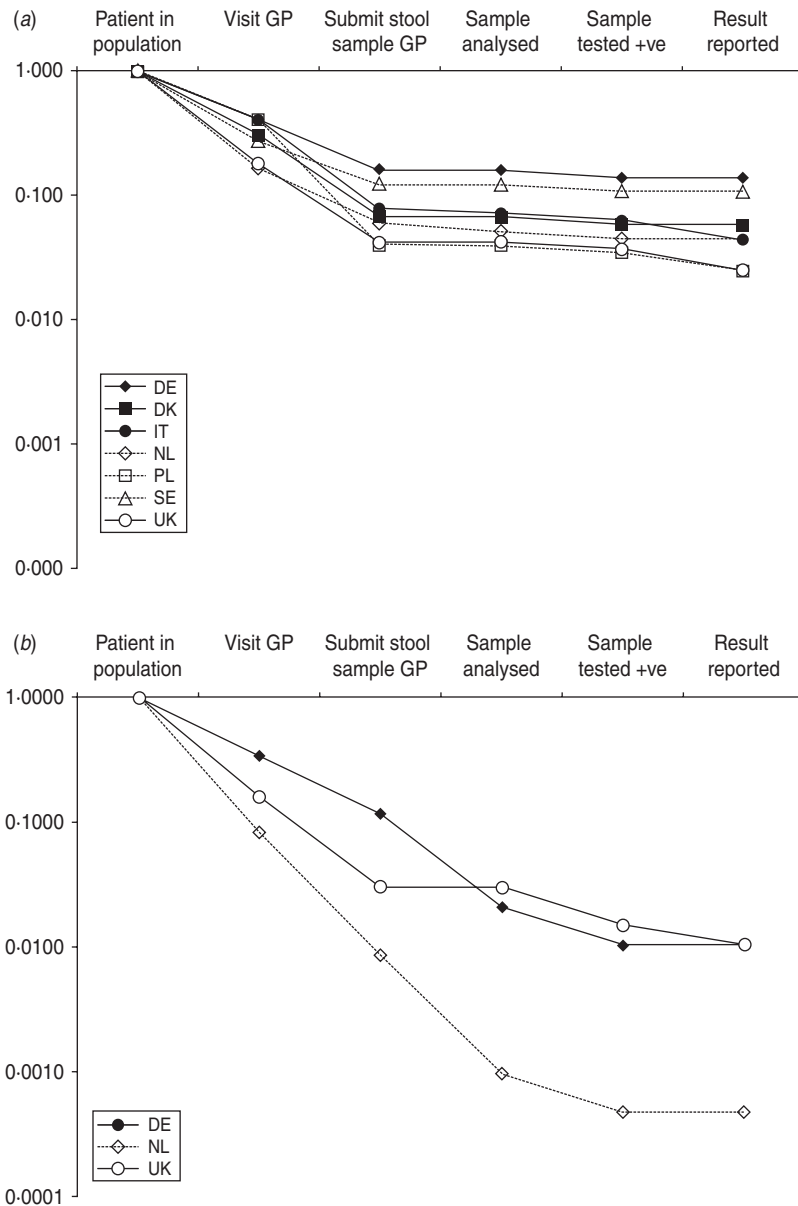
Fig. 3. (*a*) Stepwise reconstruction of the surveillance pyramid for *Salmonella* spp. (*b*) Stepwise reconstruction of the surveillance pyramid for *Cryptosporidium* spp. GP, General Practitioner; DE, Germany; DK, Denmark; IT, Italy; NL, The Netherlands; PL, Poland; SE, Sweden; UK, United Kingdom.

process for patients in the GP stream was reconstructed. This results in multipliers that are slightly different from those reported elsewhere, as the hospital stream is not taken into account. In the figures, the number of patients in different layers of the pyramid is presented as a fraction of the total number of cases in the general population, which is set to 1. It can be seen that there are substantial differences in the fraction of patients who visit their GP. The fraction is lowest in NL and UK (17% and 18%, respectively) and highest in DE, IT and PL and (40%, 41% and 41%, respectively). This difference is propagated in the next steps of the pyramid. To take into account the difference between countries in further steps, the slope of the connecting lines should be evaluated. Hence, the largest impact of the proportion of GP patients who submit a stool sample is observed in PL and UK (4%).

For comparison, Figure 3*b* presents the pyramid for Cryp. Only three countries have reconstructed the surveillance pyramid for this pathogen. Multipliers are relatively high because Cryp rarely causes bloody diarrhoea. Moreover, limited testing and lower test sensitivity have an impact on the multipliers.

Table 5a. *Median values of multipliers and incidence rates per country and per pathogen*

Country …		DE	DK	IT*	NL	PL	SE	UK	Median
Campylobacter	Multiplier	9·3[a]	29[b]	100[b]	49[b]	72[c]	17[b]	52[a]	—
	Population	620[a]	2200[b]	2300[b]	2000[b]	2000[c]	1400[b]	5000[a]	2200
	GP	220[a]	430[b]	870[b]	220[a]	617[c]	220[a]	830[a]	430
	Hospital	14[a]	12[a]	14[b]	11[a]	50[c]	11[b]	24[a]	14
Salmonella	Multiplier†	6·7[a]	17[b]	17[b]	20[b]	18[c]	10[b]	40[a]	—
	Population‡	520[a]	620[b]	290[b]	340[b]	633[c]	440[b]	1100[a]	520
	GP‡	180[a]	160[b]	100[b]	47[b]	189[c]	100[b]	200[a]	160
	Hospital‡	25[a]	8·4[a]	16[b]	5·5[a]	58[c]	6·2[b]	5·2[a]	8·4
Yersinia	Multiplier	13[b]	20[b]	—	46[b]	71[c]	13·6[b]	2200[b]	—
	Population	100[b]	93[b]	—	120[b]	406[c]	110[b]	110[b]	110
	GP	37[b]	25[b]	—	18[a]	154[c]	24[b]	20[a]	24
	Hospital	3·9[b]	0·5[a]	—	0·2[c]	2·4[b]	1·4[b]	0·1[d]	1·0
Shigella	Multiplier	11[a]	30[b]	—	53[b]	65[c]	18[b]	61[a]	—
	Population	15[a]	82[b]	—	110[b]	10[c]	94[b]	150[a]	88
	GP	5·3[a]	18[b]	—	14[b]	3[c]	17[b]	25[a]	17
	Hospital	0·4[a]	0·7[a]	—	0·3[a]	0·4[b]	0·8[b]	0·4[a]	0·45
STEC	Multiplier	23[b]	33[d]	—	87[b]	—	13[b]	34[b]	—
	Population	2·2[b]	85[d]	—	76[b]	—	15[b]	50[b]	50
	GP	1·1[b]	37[d]	—	19[b]	—	5·1[c]	6·8[b]	6·8
	Hospital	0·2[b]	0·9[b]	—	0·3[b]	—	1·6[b]	3·1[b]	0·9
EPEC	Multiplier	34[b]	100[c]	—	—	—	—	—	—
	Population	140[b]	870[c]	—	—	—	—	—	505
	GP	48[b]	200[c]	—	—	—	—	—	124
	Hospital	4·8[b]	5·0[b]	—	—	—	—	—	4·9
Cryptosporidium	Multiplier	100[c]	—	—	2100[b]	—	—	93[b]	—
	Population	130[c]	—	—	4000[b]	—	—	770[b]	770
	GP	43[c]	—	—	330[b]	—	—	130[b]	130
	Hospital	1·6[c]	—	—	0·8[c]	—	—	0·2[b]	0·8

DE, Germany; DK, Denmark; IT, Italy; NL, The Netherlands; PL, Poland; SE, Sweden; UK, United Kingdom.
GP, General Practitioner; STEC, Shiga toxin-producing *E. coli* O157; EPEC enteropathogenic *Escherichia coli*.
* Data from three provinces in Italy.
† CV (coefficient of variation = standard deviation/mean): [a] $CV < 0.25$, [b] $0.25 \leqslant CV < 1.0$, [c] $1.0 \leqslant CV < 4.0$, [d] $CV \geqslant 4.0$.
‡ Incidence rate/100 000 person-years.

Population incidence and multipliers

Table 5 shows the median number of reported cases and multipliers that result from the reconstruction of the surveillance pyramid per country and pathogen. The data are shown graphically in Figure 4 (see graphical reconstruction of the surveillance pyramid for all pathogens in online Annex C). Multipliers differ widely. For example, in DE in the period 2001–2005 an average 64 380 Salm cases were reported annually. The multiplier for Salm in DE was 6·7. This means that an estimated 430 000 cases of salmonellosis occurred in the general population (520 cases/100 000 inhabitants). In PL, only 13 531 cases of Salm were reported, but the multiplier for Salm was approximately ten times higher than in DE. Hence, 832 000 cases of salmonellosis were estimated to occur in the general population (2200 cases/100 000

inhabitants). Overall, multipliers were lowest for DE and SE, followed by DK, NL, IT, UK and PL.

Based on medians of the (median) results for all countries, the incidence rate was highest for Camp, followed by Salm while incidence rates of Yers, Shig and STEC were lower and of the same magnitude. Between 16% and 32% of all cases in the population consulted their GP, and 1–3% were hospitalized. These proportions were highest for Salm and lowest for Camp and Shig. In individual countries, results usually differed by no more than a factor of 2 from the median across countries, with some notable exceptions. The incidence rate of Salm in PL was clearly larger than the median at all levels of the pyramid due to high multipliers. The incidence rate of Camp in DE, and of Shig in DE and PL at the population and GP level, but not at hospital level, was clearly lower than the median. The incidence rates of STEC

Table 5*b*. *Median annual number of infections per country and per pathogen*

Country ...		DE	DK	IT	NL	PL	SE	UK
Inhabitants ($\times 10^6$)		83·5	5·4	57·7*	16·2	38·2	9	50
Campylobacter	Population	515 000	116 000	156 000	322 000	765 000	122 000	2 500 000
	GP	178 000	23 000	54 000	36 000	236 000	19 000	416 000
	Hospital	12 000	640	8700	1700	19 000	1000	12 000
Salmonella	Population	430 000	33 000	34 000	55 000	242 000	40 000	563 000
	GP	151 000	8600	8900	7700	72 000	9300	98 000
	Hospital	20 000	606	250	890	22 000	560	2600
Yersinia	Population	86 000	5000	—	20 000	9600	9500	55 000
	GP	30 000	1300	—	2900	3400	2200	9800
	Hospital	3200	30	—	40	310	120	40
Shigella	Population	12 000	4400	—	17 000	3900	8400	72 000
	GP	4400	970	—	2300	1300	1600	12 000
	Hospital	440	40	—	40	140	70	180
STEC	Population	2400	4600	—	12 000	—	1360	25 000
	GP	890	2000	—	3000	—	450	3400
	Hospital	180	50	—	40	—	140	1500
EPEC	Population	117 000	47 000	—	—	—	—	—
	GP	40 000	11 000	—	—	—	—	—
	Hospital	4000	270	—	—	—	—	—
Cryptosporidium	Population	109 000	—	—	651 201	—	—	385 961
	GP	35 000	—	—	54 178	—	—	62 698
	Hospital	1400	—	—	134	—	—	77

DE, Germany; DK, Denmark; IT, Italy; NL, The Netherlands; PL, Poland; SE, Sweden; UK, United Kingdom.
GP, General Practitioner; STEC, Shiga toxin-producing *E. coli* O157; EPEC enteropathogenic *Escherichia coli*.
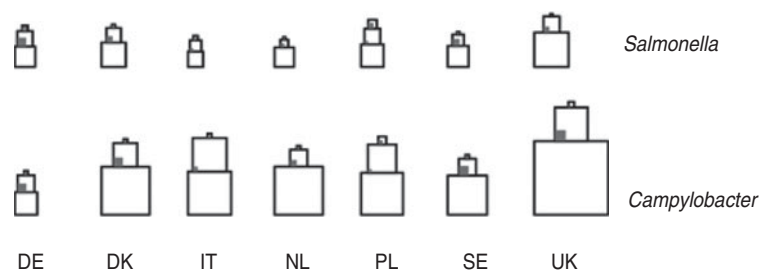* Data from three provinces in Italy.

Fig. 4. Incidence of pathogen-specific gastroenteritis per 100 000 person-years, per layer of the surveillance pyramid. Upper block: hospitalized cases; middle block: cases visiting a General Practitioner; lower block: non-consulting cases. Shaded areas indicate reported cases. Figures are drawn to scale of the incidence rate in the general population. DE, Germany; DK, Denmark; IT, Italy; NL, The Netherlands; PL, Poland; SE, Sweden; UK, United Kingdom.

differed strongly between countries, with low values in DE, intermediate values in SE and high values in DK and NL. The number of reported cases in DE (0·12/100 000) was clearly lower than for NL (0·86/100 000) and in particular DK (26/100 000). Furthermore, the multiplier in DE (30) was also lower than in DK (74) and NL (120). In NL, only 9% of all STEC patients submitted a faecal sample for analysis (see Table 4*a*). Results for EPEC were only available for two countries.

Uncertainty analysis

The uncertainty in the model outputs is indicated in Table 5 by subscripts that indicate the magnitude of the coefficient of variation (CV, details are provided in Annex C). For many results, the CV was between typically between 0·25 and 1·0 for incidence rates at population and GP level, as well as for the multiplier. The CV of the incidence rate in hospitals was typically <0·25, because in this estimate less uncertain

parameters are included. The CV in results for DE was lower than for other countries, because the uncertainty in the country-specific parameters was smaller due to a larger sample size in the telephone survey (see Annex A). The uncertainty in estimates for STEC in DK was relatively large. This can be related to relatively high uncertainty in the probability of patients with bloody diarrhoea visiting a GP or submitting a sample (see detailed results in Annex A).

The output distributions typically have a right-hand tail, as the upper band is larger than the lower band. Generally, means of the output distributions were higher than medians, and even the median of the output distributions was higher than a simple point estimate of a deterministic calculation using the mean values of all parameter uncertainty distributions. Hence, including parameter uncertainty in the model tends to influence the estimates for incidence rates to higher values than deterministic estimates.

DISCUSSION

Overview of results

Annually, EFSA and ECDC report on routine surveillance data from different (European) countries [32]. Such data are extracted from 'Eavesdrop on' the healthcare system which functions differently in each country and is not designed to provide harmonized information on disease incidence. It is widely acknowledged that comparing such routinely collected data needs to be done with utmost care. Reconstructing the surveillance pyramid provides important information for interpretation of reported data, and provision of incidence estimates that can be meaningfully compared. Such studies are attractive as they are relatively easy to perform in comparison to population-based surveys such as the IID studies in the UK [4, 5] and the Sensor study in The Netherlands [6]. Our results confirm that the degree of under-reporting varies widely by pathogen and by country. The main factors that contribute to these varying degrees of underreporting and under-diagnosis are differences in healthcare usage and differences in laboratory practice. Multipliers varied from 6·7 (Salm in DE) to 2300 (Yers in UK). Overall, multipliers were lowest for DE and SE and highest for PL.

Comparison with other reconstruction studies

Other authors have previously reported similar studies to reconstruct the surveillance pyramid for specific pathogens or for gastrointestinal illness in general [1, 9, 11–13, 33]. The current study is more comprehensive than previous studies, considering seven enteric pathogens in seven countries. There are, however, differences in the level of detail in the actual calculations. As in the current study, three published studies [9, 13, 33] applied differential rates for patients with bloody and non-bloody diarrhoea. One study also accounted for the duration of illness as a determinant of seeking medical care and of stool examination. None of the previously published studies considered differences in reporting rates for hospitalized and non-hospitalized patients, and only one study attempted full uncertainty analysis [13].

Limitations of this study

This study is the first to attempt a multi-country reconstruction of the surveillance pyramid. The multipliers vary widely by country and by pathogen. Further, the incidence rates vary widely between countries, and there is considerable uncertainty. This may at least, in part, be due to the uncertainties in the parameters used to reconstruct the surveillance pyramid. Nevertheless, the transparency of the model allowed consideration of the different steps in the pyramid in more detail, in order to understand the observed differences in multipliers. This demonstrated that there are substantial differences in the proportion of patients who visit their GP and that this factor alone explains to a large extent differences in the multiplier. This difference is propagated in the next step of the pyramid, namely the proportion of GP patients who submit a stool sample. These two steps in the pyramid clearly have the largest impact on the final results. The impact of sample analysis and test sensitivity is not large (and is more similar between countries).

Our model is based on the key assumption that bloody diarrhoea is an important determinant of healthcare-seeking behaviour, and of seeking laboratory diagnosis. The differences between parameter estimates as based on cross-sectional surveys in participating countries and literature data [5, 19–22] confirm this hypothesis. However, case numbers with bloody diarrhoea in the telephone surveys were quite small, leading to considerable uncertainty in our results. Furthermore, other factors including presence of pain, fever, duration of symptoms, education and recent history of foreign travel, are reported to be at

least as important and are not included in our study [4, 34, 35]. The telephone surveys do not distinguish between viral and bacterial gastroenteritis. Viral gastroenteritis typically has a milder and shorter course than bacterial gastroenteritis, hence the frequency of GP consultation and stool submission may be lower than for bacterial (i.e. more severe) gastroenteritis. This may lead to a general inflation of all the multipliers and case estimates for all countries.

Even though our parameter estimates for healthcare use and sample submission were based on a series of coordinated cross-sectional surveys, other factors in the reconstruction model were less well known and were based on expert opinion. This resulted in considerable uncertainty and potential biases in the incidence estimates, that should be reduced by further surveys, e.g. on laboratory practice. We also noted that there is very little information on the sensitivity of test methods applied in routine diagnostic laboratories. Diagnostic algorithms may vary between countries. Especially for STEC O157 differences in laboratory practice and methods (culture and serotyping vs. Shiga toxin detection only) may have a greater impact on correctly diagnosed and reported cases compared to other pathogens. This may in part explain differences in reported cases and, consequently, differences in the incidence estimates between countries.

In this study we combined surveillance data from 2001–2005 with survey data from 2008–2009. This discrepancy may have affected the incidence estimates resulting from the pyramid reconstruction. However, surveillance data have not changed much over the time period between 2001–2009 [32]. Hence, we expect that the results would be similar using more recent surveillance data.

Furthermore, in this study a proxy data source for all cases in the population outbreak data was used to estimate the proportion of cases with bloody diarrhoea. However, outbreak cases are only a small fraction of all cases and they could easily overrepresent persons with more severe illness who visited a GP.
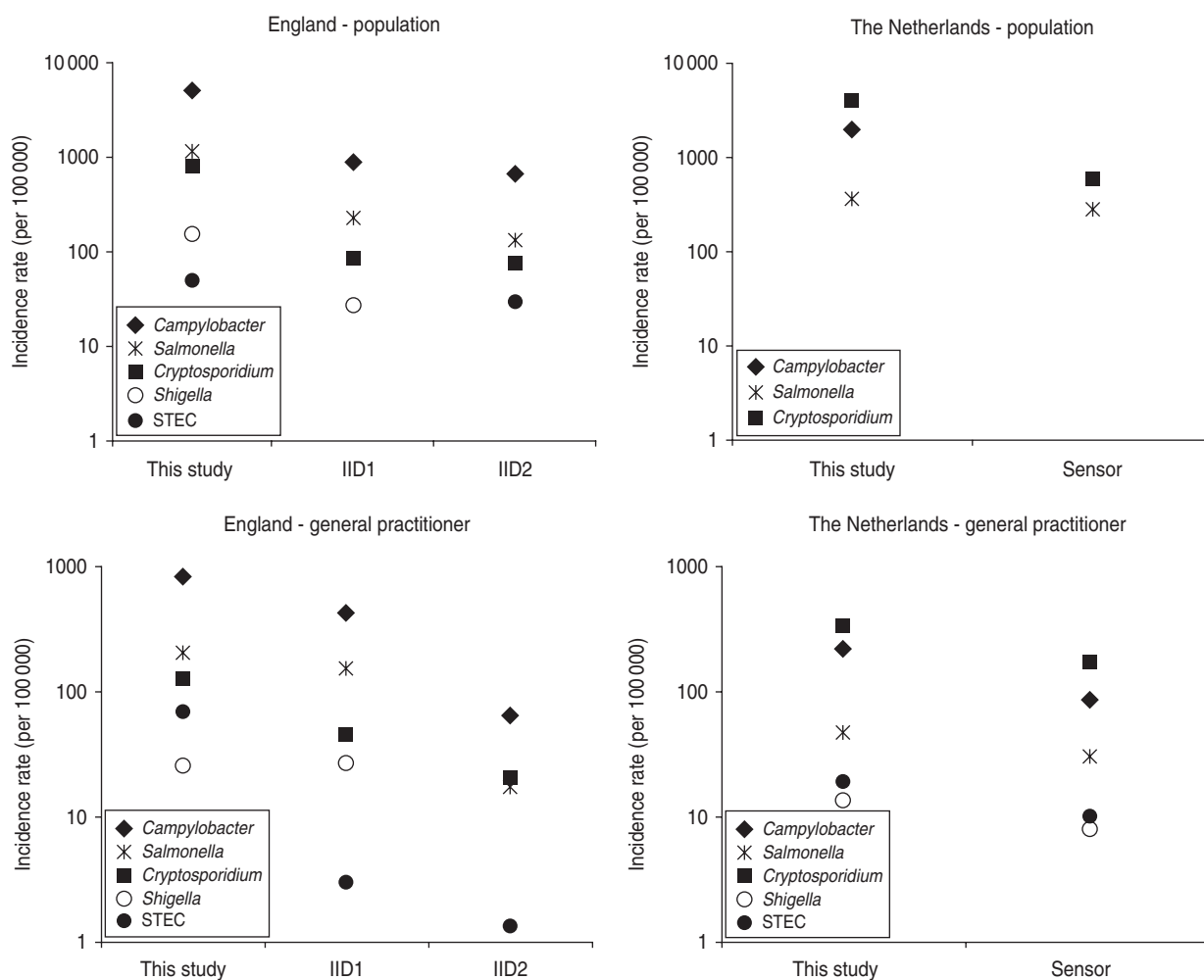
Validity of the pyramid reconstruction

The validity of the pyramid reconstruction approach can be explored by comparison with datasets from independent studies: the IID studies in England (IID1, 1993–1996 [27]; IID2, 2008 [5] and the Sensor/NIVEL studies in The Netherlands (1996–1999) [6, 36]). Both countries organized large-scale population-based cohort studies to observe the incidence of infectious gastroenteritis and to analyse its aetiology. Studies were performed in the general population and in patients with gastroenteritis who consulted their GP. Figure 5 gives a summary of incidence data from these studies, compared to the incidence estimates from the current study. Overall, the incidence rates estimated by the pyramid reconstruction approach are higher than those estimated by prospective cohort studies. The correction for sensitivity of laboratory analysis, which is included in the pyramid reconstruction model explains a 1·1–2 times higher incidence rate compared to the cohort studies, where this correction is not made. Nevertheless, estimates at population level for Camp, Salm, Shig, and Cryp in the UK and Camp and Cryp in NL are considerably higher than the estimates for cohort studies. Possibly, the probability of visiting a GP is underestimated for patients with bacterial gastroenteritis if only bloody diarrhoea is used as an explanatory factor. Other factors, such as duration of illness, severity of symptoms and a high proportion of young children among symptomatic cases, may affect consultation behaviour and should be accounted for in future reconstruction studies. The appropriateness of the model for the other countries awaits further verification by similar population-based cohort studies or alternative study designs.

Implications of this study for public health

Despite the remaining uncertainties, the pyramid reconstruction approach provides more realistic information on the community incidence of gastroenteritis due to specific pathogens in EU member states. This study shows that each year millions of illnesses occur in the European population, which lead to thousands of hospitalizations. These estimates are a better basis for comparing disease incidence in different countries than reported data, because the incidence estimates obtained with the surveillance pyramid reconstruction approach adjust for biases due to different surveillance systems.

The burden of illness is shown to be particularly high for pathogens Camp and Salm. To a large extent, these two pathogens are transmitted through food. Hence, these estimates can also be used to assess food safety management, which requires a consistent,

Fig. 5. Comparison of pyramid reconstruction and population-based studies (incidence rates per 100 000) in England and The Netherlands. STEC, Shiga toxin-producing *E. coli* O157.

quantitative assessment of the relative public health impact of foodborne disease.

Furthermore, the incidence estimates obtained with the pyramid reconstruction approach provide a better basis for burden-of-disease estimates, quantified in terms of disability-adjusted life years [7, 37], and for estimating the public health benefits and cost–benefit ratios of interventions supported by active surveillance, aimed at reducing disease incidence.

APPENDIX. Additional members of the Working Group

G. Adak (HPA), Y. Andersson (SMI), S. Ivarsson (SMI), A. Käsbohrer (BfR), A. Maçiag (PZH), K. Mølbak (SSI), K. Stark (RKI), K. Sundström (SLI), A. Thébault (ANSES), H. Wahlström (SVA), B. van der Zanden (RIVM), A. Zielinski (PZH).

SUPPLEMENTARY MATERIAL

For supplementary material accompanying this paper visit http://dx.doi.org/10.1017/S0950268812002166.

DECLARATION OF INTEREST

None.

REFERENCES

1. **Scallan E, et al.** Foodborne illness acquired in the United States – major pathogens. *Emerging Infectious Diseases* 2011; **17**: 7–15.
2. **WHO.** WHO Initiative to Estimate the Global Burden of Foodborne Diseases. Second formal meeting of the Foodborne Disease Burden Epidemiology Reference Group (FERG). Geneva: World Health Organization; 2008.
3. **Lake RJ, et al.** Risk ranking for foodborne microbial hazards in New Zealand: burden of disease estimates. *Risk Analysis* 2010; **30**: 743–752.
4. **Adak GK, Long SM, O'Brien SJ.** Trends in indigenous foodborne disease and deaths, England and Wales: 1992 to 2000. *Gut* 2002; **51**: 832–841.
5. **Tam CC, et al.** Longitudinal study of infectious intestinal disease in the UK (IID2 study): incidence in the community and presenting to general practice. *Gut* 2011; **61**: 69–77.
6. **de Wit MA, et al.** Sensor, a population-based cohort study on gastroenteritis in the Netherlands: incidence and etiology. *American Journal of Epidemiology* 2001; **154**: 666–674.
7. **Havelaar AH, et al.** Disease burden of food-borne pathogens in the Netherlands, 2009. *International Journal of Food Microbiology* 2012; **156**: 231–8.
8. **Havelaar AH, et al.** Estimating the true incidence of campylobacteriosis and salmonellosis in the EU, 2005–2009. *Epidemiology and Infection* (in press).
9. **Thomas MK, et al.** Estimated numbers of community cases of illness due to *Salmonella. Campylobacter* and verotoxigenic *Escherichia coli*: pathogen-specific community rates. *Canadian Journal of Infectious Diseases & Medical Microbiology* 2006; **17**: 229–234.
10. **Vally H, et al.** Higher rate of culture-confirmed *Campylobacter* infections in Australia than in the USA: is this due to differences in healthcare-seeking behaviour or stool culture frequency? *Epidemiology and Infection* 2009; **137**: 1751–1758.
11. **Michel P, et al.** Estimation of the under-reporting rate for the surveillance of *Escherichia coli* O157:H7 cases in Ontario, Canada. *Epidemiology and Infection* 2000; **125**: 35–45.
12. **Majowicz SE, et al.** Estimating the under-reporting rat for infectious gastrointestinal illness in Ontario. *Canadian Journal of Public Health* 2005; **96**: 178–181.
13. **Hall G, et al.** Estimating community incidence of *Salmonella. Campylobacter*, and Shiga toxin-producing *Escherichia coli* infections, Australia. *Emerging Infectious Diseases* 2008; **14**: 1601–1609.
14. **Faensen D, et al.** SurvNet@RKI – a multistate electronic reporting system for communicable diseases. *Eurosurveillance* 2006; **11**: 100–103.
15. **Ethelberg S, et al.** Spatial distribution and registry-based case-control analysis of *Campylobacter* infections in Denmark, 1991–2001. *American Journal of Epidemiology* 2005; **162**: 1008–1015.
16. **Helms M, Simonsen J, Molbak K.** Foodborne bacterial infection and hospitalization: a registry-based study. *Clinical Infectious Diseases* 2006; **42**: 498–506.
17. **van Pelt W, et al.** Laboratory surveillance of bacterial gastroenteric pathogens in The Netherlands, 1991–2001. *Epidemiology and Infection* 2003; **130**: 431–441.
18. **Rolfhamre P, et al.** SmiNet-2: Description of an internet-based surveillance system for communicable diseases in Sweden. *Eurosurveillance* 2006; **11**: 103–107.
19. **Muller L, Korsgaard H, Ethelberg S.** Burden of acute gastrointestinal illness in Denmark 2009: a population-based telephone survey. *Epidemiology and Infection* 2012; **140**: 290–298.
20. **Doorduyn Y, van Pelt W, Havelaar AH.** The burden of infectious intestinal disease (IID) in the community: a survey of self-reported IID in The Netherlands. *Epidemiology and Infection* 2012; **140**: 1185–1192.
21. **Baumann-Popczyk A, et al.** Incidence of self-reported acute gastrointestinal infections in the community in Poland: a population-based study. *Epidemiology and Infection* 2012; **140**: 1173–1184.
22. **Scavia G, et al.** The burden of self-reported acute gastrointestinal illness in Italy: a retrospective survey, 2008–2009. *Epidemiology and Infection* 2012; **140**: 1193–1206.
23. **Del Manso M, et al.** Food-borne illnesses: a survey of regional reference laboratories in Italy [in Italian]. *Igiene e Sanita Pubblica* 2009; **65**: 573–582.
24. **Friesema IH, et al.** Etiology of acute gastroenteritis in children requiring hospitalization in the Netherlands. *European Journal of Clinical Microbiology & Infectious Disease* 2012; **31**: 405–415.
25. **Friesema IH, et al.** Geographical association between livestock density and human Shiga toxin-producing *Escherichia coli* O157 infections. *Epidemiology and Infection* 2011; **139**: 1081–1089.
26. **Stefanoff P, et al.** Antibacterial prescriptions for acute gastrointestinal infections: uncovering the iceberg. *Epidemiology and Infection*. Published online: 15 June 2012. doi:10.1017/S0950268812001173.
27. **Wheeler JG, et al.** Study of infectious intestinal disease in England: rates in the community, presenting to general practice, and reported to national surveillance. The Infectious Intestinal Disease Study Executive. *British Medical Journal* 1999; **318**: 1046–1050.
28. **Tam CC, et al.** Longitudinal study of infectious intestinal disease in the UK (IID2 study): incidence in the community and presenting to general practice. *Gut* 2011; **61**: 69–77.
29. **van den Brandhof WE, et al.** General practitioner practices in requesting laboratory tests for patients with gastroenteritis in the Netherlands, 2001–2002. *BMC Family Practice* 2006; **7**: 56.
30. **Ethelberg S, et al.** The significance of the number of submitted samples and patient-related factors for faecal bacterial diagnostics. *Clinical Microbiology and Infection* 2007; **13**: 1095–1099.
31. **Weber R, et al.** Threshold of detection of *Cryptosporidium* oocysts in human stool specimens: evidence for

low sensitivity of current diagnostic methods. *Journal of Clinical Microbiology* 1991; **29**: 1323–1327.

32. **Anon.** The Community Summary Report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in the European Union in 2008. Stockholm, Parma: European Centre for Disease Prevention and Control (ECDC) and the European Food Safety Authority (EFSA); 2010.

33. **Voetsch AC, *et al.*** FoodNet estimate of the burden of illness caused by nontyphoidal Salmonella infections in the United States. *Clinical Infectious Diseases* 2004; **38**: S127–134.

34. **Food Standards Agency.** A report of the study of infectious intestinal disease in England. London: The Stationery Office, 2000.

35. **Scallan E, *et al.*** Acute gastroenteritis in northern Ireland and the Republic of Ireland: a telephone survey. *Communicable Disease and Public Health* 2004; **7**: 61–67.

36. **de Wit MA, *et al.*** Gastroenteritis in sentinel general practices, The Netherlands. *Emerging Infectious Diseases* 2001; **7**: 82–91.

37. **Kretzschmar M, *et al.*** New methodology for estimating the burden of infectious diseases in Europe. *PLoS Medicine* 2012; **9**: e1001205.