Epidemiol. Infect. (2004), **132**, 467–484. © 2004 Cambridge University Press
DOI: 10.1017/S0950268804001979 Printed in the United Kingdom

Disease burden in The Netherlands due to infections with Shiga toxin-producing *Escherichia coli* O157

A. H. HAVELAAR[1]*, Y. T. H. P. VAN DUYNHOVEN[1], M. J. NAUTA[1], M. BOUWKNEGT[1], A. E. HEUVELINK[2], G. A. DE WIT[1], M. G. M. NIEUWENHUIZEN[3] AND N. C. A. J. VAN DE KAR[4]

[1] *National Institute for Public Health and the Environment (RIVM), PO Box 1, 3720 BA Bilthoven, The Netherlands*
[2] *Inspectorate for Health Protection and Veterinary Public Health, PO Box 202, 7200 AE Zutphen, The Netherlands*
[3] *Renine, Renal Replacement Registry of The Netherlands, PO Box 4097, 3006 AB Rotterdam, The Netherlands*
[4] *Department of Paediatric Nephrology, University Hospital Nijmegen, PO Box 9101, 6500 HB Nijmegen, The Netherlands*

(*Accepted 22 December 2003*)

SUMMARY

Surveys carried out between 1990 and 2000 indicated that the incidence of STEC O157-associated gastroenteritis in The Netherlands was 1250 cases/year (median), of which 180 visited a general practitioner, 40 are reported and 0·6 are fatal, mainly in the elderly. There are approximately 20 cases of STEC O157-associated haemolytic–uraemic syndrome (HUS) per year, mainly in children. There are 2·5 HUS patients per year who develop end-stage renal disease (ESRD). There are an estimated 2 HUS-related and 0·5 ESRD-related fatalities per year. The mean disease burden associated with STEC O157 in the Dutch population is 116 (90% confidence interval 85–160) Disability Adjusted Life Years (DALYs) per year. Mortality due to HUS (58 DALYs), and ESRD (21 DALYs) and dialysis due to ESRD (21 DALYs) constitute the main determinants of disease burden. Sensitivity analysis indicates that uncertainty associated with model assumptions did not have a major effect on these estimates.

INTRODUCTION

To quantify the public health burden of foodborne disease, a wide spectrum of illnesses must be accounted for. Most frequently, mild and self-limiting gastroenteritis (GE) occurs but more severe manifestations of acute disease or complications may also occur [1]. Case-fatality ratios of complications may be as high as 10%, but GE can also be life threatening, particularly in the elderly [2]. To account for the widely different clinical manifestations of foodborne illness, a uniform health measure is needed. In a previous study [3], the public health indicator 'Disability Adjusted Life Years' (DALYs)[1] was used to integrate the disease burden of different clinical syndromes associated with thermophilic *Campylobacter* species. In this study, we present epidemiological data on the incidence of illness associated with Shiga toxin (Stx)-producing *E. coli* serogroup O157 (STEC O157) in The Netherlands, and use the DALY concept to estimate the disease burden in the Dutch population.

* Author for correspondence.

[1] See Appendix A for definition of symbols, subscripts and probability distribution functions used in the text.

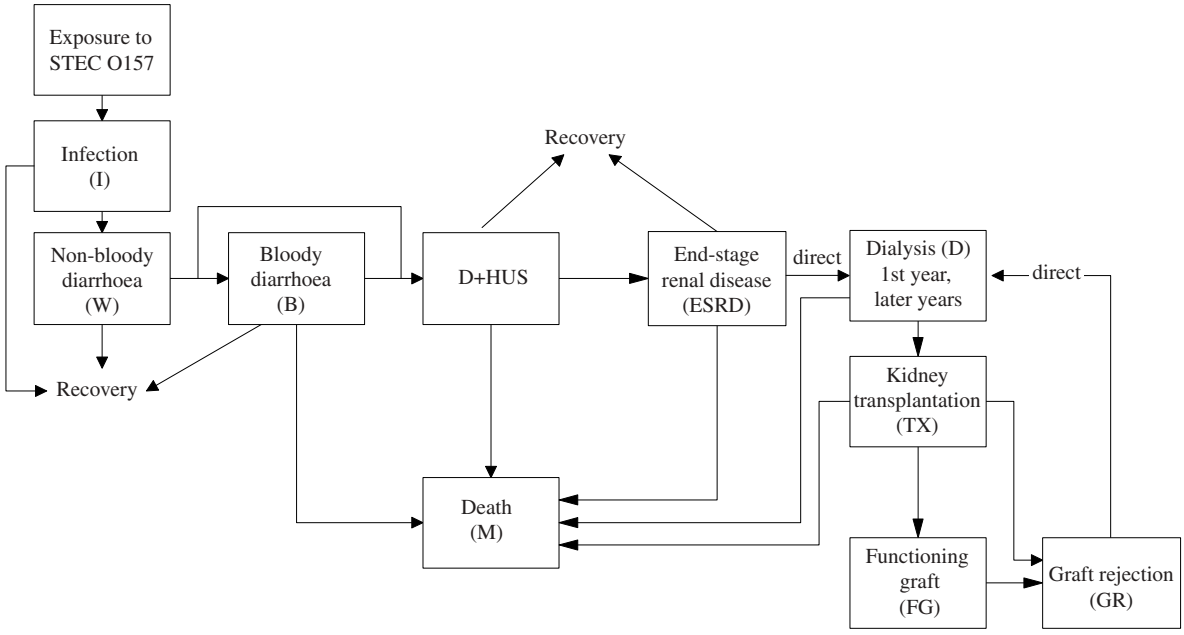

Fig. 1. Disease model for infection with Shiga toxin-producing *Escherichia coli* O157.

In The Netherlands, only two small outbreaks have been recognized [4, 5] and the number of endemic cases is low [6]. However, because the bacteria are present in the food chain, the possibility of outbreaks, with severe disease, cannot be ruled out and preventive measures are indicated.

Disease model

Infections with STEC O157 may be asymptomatic, or may lead to diarrhoeal illness (see Fig. 1) including haemorrhagic colitis. In 1985, Karmali and colleagues [7] reported an association between STEC and post-diarrhoeal haemolytic–uraemic syndrome (D+ HUS). This illness occurs mainly in young children and may lead to death during the acute phase, to end-stage renal disease (ESRD) or other sequelae [8]. Numerous studies have demonstrated that STEC, and particularly serogroup O157, is the major aetiological agent of D+ HUS. Acute renal failure is the most prominent feature, leading to oliguria or anuria in the majority of patients [8]. Patients are routinely treated with peritoneal dialysis or haemodialysis. Some patients develop ESRD directly, but renal damage may also become manifest after a period of apparently normal kidney function (late ESRD). Patients with ESRD are initially treated with one of the different forms of dialysis (D) and may later be eligible for kidney transplantation (TX). Siegler [9] reviewed the extrarenal involvement associated with D+ HUS such as diabetes mellitus. Given the limited duration and/or severity of the non-renal sequelae, their contribution to the overall disease burden is small and will not be evaluated further.

DALYs

The different outcomes of (infectious) disease can be combined in one single measure, the Disability Adjusted Life Year (DALY), following the methodology proposed by Murray and Acharya [10]:

$$\mathrm{DALY} = \mathrm{YLL} + \mathrm{YLD},$$

YLL is the number of years of life lost due to mortality and YLD is the number of years lived with a disability, weighted with a factor that reflects the severity of the disease on a scale from 0 (perfect health) to 1 (death). YLL is calculated by the accumulation over all fatal cases and all diseases of the individual lifespan (e) had illness and death not occurred. Thus:

$$\mathrm{YLL} = \sum_{\text{all diseases}} \sum_{\text{all fatal cases}} (e).$$

We derived the expected lifespan of fatal cases from the standard life-table as reported by Statistics Netherlands, including an additional module to simulate the variability in individual lifespans (see [11]).

https://doi.org/10.1017/S0950268804001979 Published online by Cambridge University Press

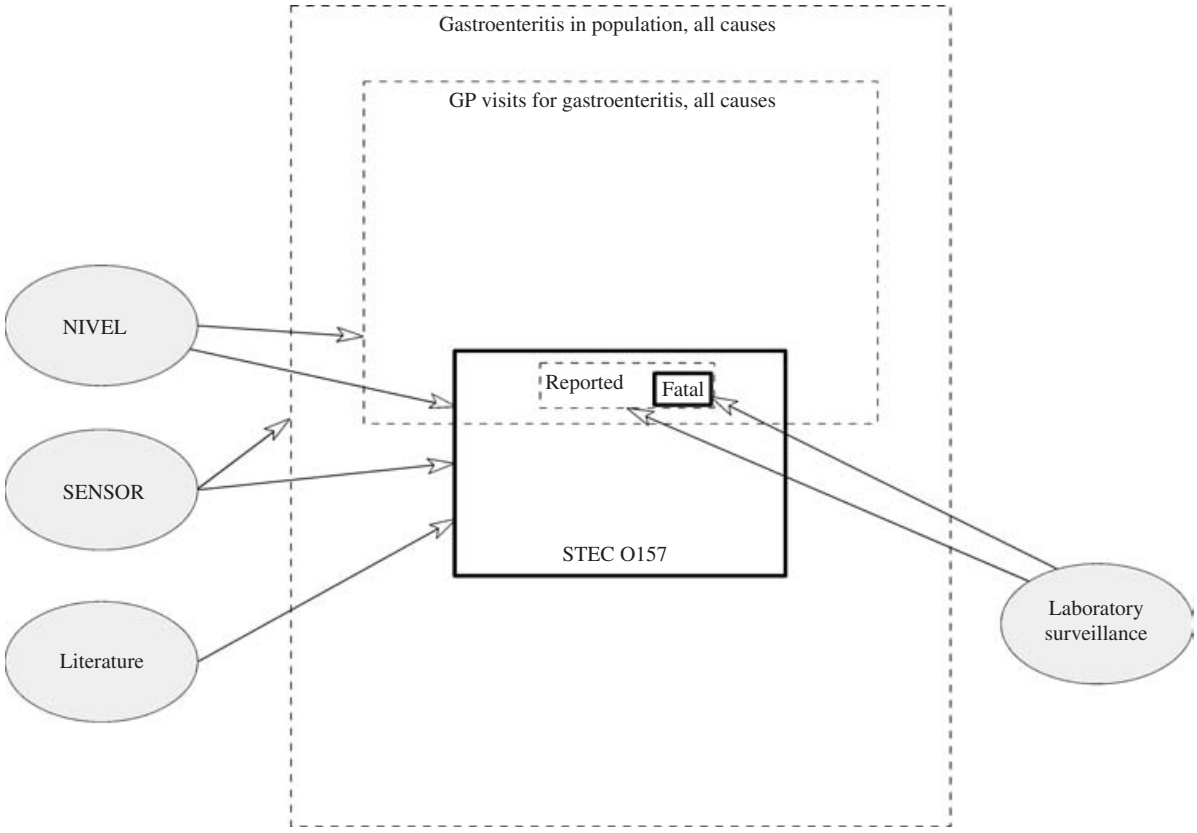

Fig. 2. Data sources for estimating the incidence of gastroenteritis (GE) and related mortality (not to scale). Variables of interest were the total number of STEC O157-associated GE and the number of fatalities (bold boxes). The total number of cases was estimated indirectly from the number of GP visits for GE and the fraction attributable to STEC O157 (NIVEL study). The multiplier was based on data on consultation behaviour in the SENSOR study and literature data. The number of fatal cases was based on reports from intensified laboratory surveillance, a subset of GP visits.

YLD is calculated by the accumulation over all cases and all diseases of the product of the duration of the illness (t) and the severity weight (w):

$$\mathrm{YLD} = \sum_{\text{all diseases}} \sum_{\text{all cases}} (t \times w).$$

To estimate the disease burden of STEC O157 in the Dutch population, we needed to translate the available data on disease incidence, symptoms, duration and mortality into estimates of e, t and w. There is uncertainty (lack of knowledge about a system) and variability (inherent randomness of a system) in the model used to calculate the disease burden. Both uncertainty and variability can be expressed in a statistical distribution function, but require a different strategy to account for in the analysis. In this study, severity and duration of disease were considered to be variable parameters, whereas all other parameters (related to the incidence of illness and death) were considered to be uncertain. Second-order uncertainty (i.e. uncertainty in the parameter estimates of the distribution functions of the variable quantities) was also considered.

Incidence of STEC O157 related GE in The Netherlands

We based our estimates on several recently completed studies on the incidence and aetiology of GE in The Netherlands (SENSOR [12], a population-based study; NIVEL [13], a general practice-based study) and on intensified laboratory surveillance [6] (see Fig. 2). In the general-practice-based NIVEL study the incidence of consultations of a general practitioner (GP) for GE[2] by all causes was estimated as 8/1000 person years (pyr) (age standardized and

[2] Case definition: three or more loose stools in 24 h; or diarrhoea with two additional GE symptoms (vomiting, nausea, fever, abdominal cramps, abdominal pain, blood in stool, mucus in stool) or vomiting with two additional GE symptoms (diarrhoea, nausea, fever, abdominal cramps, abdominal pain, blood in stool, mucus in stool) preceded by a symptom-free period of 2 weeks.

https://doi.org/10.1017/S0950268804001979 Published online by Cambridge University Press

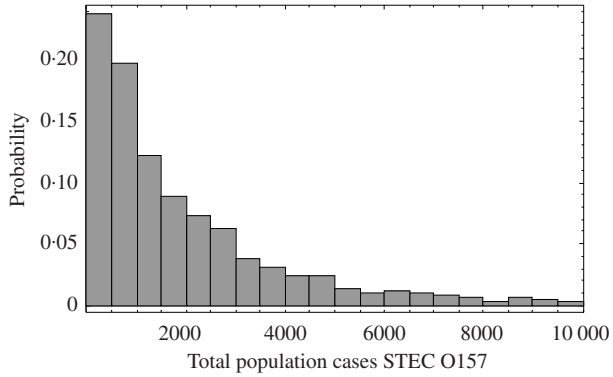

Fig. 3. Uncertainty of the incidence of total population cases of STEC O157-associated gastroenteritis (left histogram, right cumulative frequency distribution). Cases/year: 5-percentile, 83; median, 1251; mean, 2111; 95-percentile, 7157; standard deviation, 2620.

partially adjusted for incomplete participation of cases and GPs, and for list inflation) [13]. In a nested case-control study, 4 out of 798 faecal samples tested positive for STEC of which one was *E. coli* O157 K^- H^-. Direct enumeration of GP visits, reported by patients in the population-based SENSOR study resulted in a standardized incidence of 13·8/1000 pyr [14]. This estimate resulted in a mean incidence of 274 GP consultations for STEC O157-associated GE per year ($13{\cdot}8/1000\ \text{pyr} \times 15{\cdot}8 \times 10^6$ persons $\times$ 1/798 STEC O157). The uncertainty of this estimate (quantified as described in Appendix B) was substantial and followed a highly skewed distribution. The median incidence was 182 [90% confidence interval (CI) 12–859] GP consultations per year.

Based on the approach of Michel et al. [15] and Dutch data on GP consultation for bloody and non-bloody diarrhoea [13], we estimated that 14% (median, 90% CI 7–27) of all cases with STEC O157-associated GE reported to their GP (see Appendix B). In a GP practice, only a very small proportion (median 0·09%, 90% CI 0·006–0·37) of all GE cases was due to STEC O157.

Combining this information, we estimated the median incidence of STEC O157-associated GE in the general population as 1251 (mean 2114, 90% CI 83–2620) cases per year. The uncertainty in this estimate is shown in Figure 3. To estimate the proportion of bloody and non-bloody cases in the population, outbreak data were the more pertinent source of information as laboratory-confirmed cases may be positively selected for bloody stools. Michel et al. [15] summarized data from 10 outbreaks, leading to an overall estimate of 48·3% of symptomatic patients who developed bloody diarrhoea.

Age distribution

Since 1999 all medical microbiological laboratories report all positive findings of STEC O157 to the Regional Health Services, and submit an isolate for typing to the National Institute for Public Health and the Environment (RIVM) [6]. In 1999, 36 patients with O157 STEC infection were identified, or 0·2 cases/100 000 pyr. The number of cases in laboratory surveillance was sufficient to obtain information about age distribution. We used this information to infer the age distribution of cases in the general population, assuming that there was no age-related selection bias in submission of stool specimens to clinical laboratories. In general, this assumption is not valid for gastrointestinal illness because there is an increased likelihood of consulting a GP and submitting a faecal specimen if the illness occurs in children [14]. However, STEC O157 frequently results in bloody diarrhoea, which is another determinant of consultation and faecal examination, reducing the age-related selection bias. Combining the above data for bloody/non-bloody diarrhoea, age distribution and incidence of diarrhoeal cases leads to the summary presented in Table 1.

Duration of GE

Several outbreak studies have demonstrated that the duration of bloody diarrhoea is longer than that of non-bloody diarrhoea (see [11] for a complete list of references). We based our estimates on the report by Belongia et al. [16], who investigated a hamburger-associated outbreak in the United States, affecting students at a junior high school (age 9–15 years). These authors reported a median duration of 5 days (range 2–12 days) for bloody diarrhoea and 3 days (range 1–7 days) for non-bloody diarrhoea. We used Gamma distributions to describe the variability in the duration (see Table 3).

Mortality due to STEC O157-related GE

STEC O157-associated mortality by causes other than HUS is rare and only very large outbreaks provide information. In the Walkerton outbreak [17] involving both STEC O157 and *Campylobacter* spp., 1 fatal case was recorded among 2321 patients.

https://doi.org/10.1017/S0950268804001979 Published online by Cambridge University Press

Table 1. *Incidence estimates* (*median values*) *for STEC O157-associated gastroenteritis* (*GE*) *and haemolytic–uraemic syndrome* (*HUS*) *in The Netherlands*

	Age group (years)			
	0–4	5–14	15+	Total*
GE				
All cases	300	188	720	1251
Bloody diarrhoea	140	90	338	590
Non-bloody diarrhoea	160	98	382	661
Mortality	n.a.§	n.a.	n.a.	0·56
HUS	12	3	6	21
$\pi_{H\|G,S}$†	0·04	0·016	0·008	0·017
$\pi_{H\|B,S}$‡	0·086	0·033	0·018	0·036

* Note that due to the stochastic nature of the model, not all summations necessarily tally.
† Probability of developing HUS as a consequence of STEC O157-associated GE (all cases).
‡ Probability of developing HUS as a consequence of STEC O157-associated bloody diarrhoea.
§ n.a., Not available.

Applying this estimate to the above-mentioned number of cases in The Netherlands resulted in a predicted mean of 0·66 deaths per year (2 per 3 years). Laboratory surveillance was another source of information. Among 120 cases recorded in The Netherlands during 1999–2001, one death in a woman aged 85 years was attributed to the consequences of diarrhoeal illness. We based our estimate of the case-fatality ratio of diarrhoeal illness on this single case, and included the uncertainty as described in Appendix B. The low number of recognized fatal cases implied that there is no accurate information on age distribution. We assumed that the age distribution was similar as for other cases of GE and used data from Statistics Netherlands, as previously reported [3].

Incidence of STEC O157-related HUS

In The Netherlands, the incidence of HUS is 2·0/100 000 children <5 years [18], i.e. 20 cases/year.[3] HUS does occur in children $\geqslant 5$ years, but less frequently. Data from the University Hospital Nijmegen [19] indicated that for each case of HUS in the 5–14 years age group 8·2 cases are expected in the 0–4 years age group. Hence, the incidence rate of HUS in children <15 years was estimated at $20+20/8{\cdot}2 \approx 22$ cases/year. Van de Kar et al. [20] found an attributable proportion of 77% for STEC infection. In some cases with evidence for STEC infection, another serovar other than O157 may have been involved. However, the majority (81%) of cases were associated with serovar O157. Thus, the minimum attributable proportion for STEC O157 was 62%. This leads to an estimated incidence of 14–17 paediatric cases/year. The incidence at ages >15 years was estimated indirectly from laboratory surveillance data. Among 120 cases during 1999–2001, there were 18 HUS patients. Of these, 13 (72%) were <15 years, and 5 were older. Combining these data resulted in an incidence of 19–24 cases of STEC O157-associated HUS per year.

Mortality associated with D+ HUS

Due to earlier recognition and improved management in the acute phase of the disease, the mortality associated with paediatric cases of D+ HUS has decreased to <5%. A summary of the literature (see [11] for complete list of references) resulted in a pooled estimate of the case-fatality ratio of 32/867 (3·7%). On the basis of limited data, we assumed that

[3] During 1999 and 2000, telephone enquiries to academic paediatric nephrology units in The Netherlands by Professor L. A. Monnens (Department of Paediatrics, University Hospital Nijmegen) resulted in a total of 19 cases/year. Hence, the incidence appeared to be relatively constant throughout the years.

https://doi.org/10.1017/S0950268804001979 Published online by Cambridge University Press

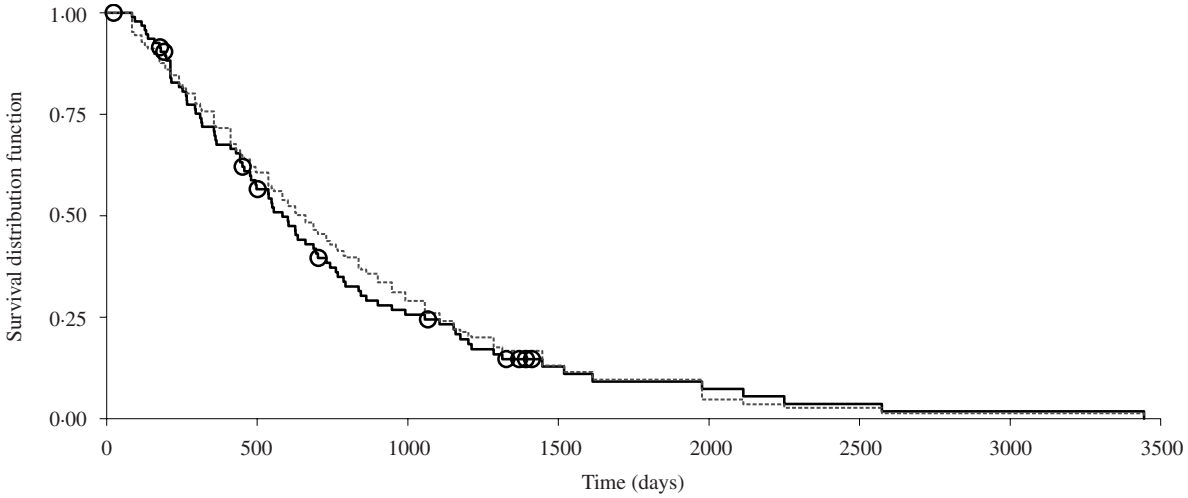

Fig. 4. Kaplan–Meier curves for time to transplantation, age group 0–15 years, diagnosis HUS and PYN. ○, Censored data; —, observed; - - -, fitted.

this low case-fatality ratio was valid for cases up to 65 years of age. At higher ages, we estimated a considerably higher case-fatality ratio of 56%, based on Scottish outbreak data [21]. The late fatalities that were reported in some papers were accounted for in this study as mortality associated with ESRD.

ESRD

A summary of the literature (see [11] for list of references) resulted in a pooled estimate of the probability of developing direct ESRD of 23/734 (2·9%). Here, direct ESRD was defined as a direct outcome of HUS, without recovery of the renal function. The probability of late ESRD (with temporary recovery of renal function) was estimated at 2–3% in the first 10 years. However, from long-term follow-up studies performed in Paris, France, it appeared that the hazard of developing ESRD after initial recovery of renal function might well extend over a period of more than 20 years [22]. Based on this study, we estimated the probability of late ESRD as 8/76 (10·5%), with onset uniformly distributed between 0 and 40 years after D+ HUS. Most information in the literature is on paediatric cases, no long-term follow-up for adult patients has been reported. Conlon et al. [23] reported that of their 51 surviving patients, 2 (3·9%) required long-term dialysis. We assumed that the probability to develop ESRD in adults was similar to that in children.

The clinical history of ESRD patients was derived from the Renine (Renal Replacement Registry of The Netherlands, Rotterdam, The Netherlands) database. This database recorded data from all Dutch dialysis and/or transplantation centres on patients who were treated for chronic, terminal renal failure. From the database, information was extracted on 31 new patients, starting with dialysis in the period 1980–2000 inclusive with primary diagnosis HUS, who were <16 years of age at the start of treatment. This relatively small number of patients would lead to large uncertainties in parameter estimates. Therefore, data on 75 patients with primary diagnosis pyelonephritis (PYN) were included in the analysis, because they were expected to have similar prognoses (L. A. H. Monnens, personal communication). Patients were grouped in five age categories, 0–15, 16–44, 45–64, 65–74 and 75+ years. The time to transplantation was first analysed by Kaplan–Meier analysis (SAS version 8.2, SAS Institute, Cary, NC, USA: Proc Lifetest). Differences between age or diagnostic groups were evaluated by the log-rank test with a critical P value of 0·05. For both HUS and PYN, there were significant differences between age categories. There were no significant differences between HUS and PYN for any of the age categories. For the 75+ years age group, only censored data were available. As transplantations in this age group were very rare, the data were combined with the 65–74 years age group to create a 65+ years group. A parametric analysis of the data was carried out by fitting a Weibull model (SAS Institute: Proc Lifereg). Goodness of fit was tested by comparing the predicted survival curve with the curve from the non-parametric Kaplan–Meier analysis. Figure 4 shows an example of the observed data and the

https://doi.org/10.1017/S0950268804001979 Published online by Cambridge University Press

fitted model. The waiting time to transplantation increased considerably with age (Table 5). The median waiting times computed from the Weibull model corresponded well with the medians from the non-parametric Kaplan–Meier method (data not shown). The high median value in the 65+ years age group suggested that it is uncommon for a patient in this age group to be eligible for transplantation, which is in accordance with the finding in the 75+ years age group that only censored cases were observed. Inspection of the cumulative density functions (not shown), and comparison of standard deviations between and within simulations, suggested that uncertainty was of minor importance for ages up to 64 years, but was of major importance for the 65+ years age group.

The time to graft failure was analysed according to the same strategy as time to transplantation. The Kaplan–Meier method demonstrated that there were significant differences between HUS and PYN patients in all age categories, hence pooling was not allowed. Within the HUS category, there were no significant differences between age groups. Therefore, all data for HUS patients were pooled and a Weibull model was fitted. The median duration of life with a functioning graft was 4 years, but this was highly variable between patients. The scale parameter was significantly smaller than 1 (see Table 5), indicating that the probability of graft rejection was greatest directly after the implantation. This resulted in the 90% CI for the time to graft failure to range from 0·016 years (6 days) to 96 years (effectively lifelong). Median waiting times of the Weibull model and the non-parametric Kaplan–Meier method were similar (data not shown).

The Renine database also allowed for analysis of case-fatality ratios in dialysis patients and after renal transplantation. In the first year after starting dialysis, mortality ratios were relatively high and different between age groups, ranging from 5% in the 0–15 years age group to 79% in the 65+ years age group (see Table 5). These estimates were based on pooled data for HUS and PYN patients. After the first year, there was no significant excess mortality in dialysis patients.

There were only few fatalities after renal transplantation, and no information on age-specific risks could be obtained. In the first year after transplantation, 3 out of 39 HUS patients died. These fatalities were considered to be directly related to the transplantation, resulting in a mean case-fatality ratio of 7·7%.

Summary of incidence data

From Table 1, it is possible to infer that the probability of developing HUS after STEC O157-associated GE was 4% in the 0–4 years age group, 1·6% in the 5–14 years age group and 0·8% for ages above 15 years. These probabilities were considerably smaller than those reported in outbreak investigations. This may be related to incomplete case assurance of non-bloody diarrhoea in outbreak studies. However, the probability of HUS given bloody diarrhoea in our study (see Table 1) was also at the low end of estimates from outbreak studies. Other sources of under-ascertainment may (partly) explain this difference. It is also possible that outbreak strains were more virulent than strains involved in endemic cases, or that our incidence estimate for STEC O157-associated GE was too high.

Simulation model for disease burden

We used a second-order stochastic simulation model to quantify the disease burden of STEC O157-associated illness. One iteration of the model represented 1 year, in which a particular number of cases of GE and HUS occurred. For each patient, the severity and duration of the illness was simulated, as well as possible complications (mortality, ESRD). These outcomes were variable between cases. A total of 1500 iterations of the model constituted one simulation that represented the variability between different years. For each simulation, a random sample was obtained from all distributions of uncertain parameters. Thus, running the model for 250 simulations represented the effects of parameter uncertainty on the results. Analysis of other sources of uncertainty was done by sensitivity analysis.

The GE model

A stepwise summary of the GE model is given in Table 2, parameter values are shown in Table 3. In each iteration, a Poisson distribution represented the variability in the annual number of cases of GE (bloody and non-bloody diarrhoea), with the incidence rate based on estimates shown in Table 1. The disease burden for all cases of GE was then simulated from a distribution representing the total population.

https://doi.org/10.1017/S0950268804001979 Published online by Cambridge University Press

Table 2. *The gastroenteritis model*

(i)	Simulate the incidence of bloody diarrhoea	$N_B \sim \text{Poisson}(\nu_B)$
(ii)	Simulate the incidence of non-bloody diarrhoea	$N_W \sim \text{Poisson}(\nu_W)$
(iii)	Simulate disease burden of bloody diarrhoea	$YLD_B = \text{Gamma}(N_B \times 1{\cdot}232, 1{\cdot}792)/365$*
(iv)	Simulate disease burden of non-bloody diarrhoea	$YLD_W = \text{Gamma}(N_W \times 1{\cdot}065, 0{\cdot}211)/365$
(v)	Simulate incidence of laboratory-confirmed cases	$N_{S,LS} \sim \text{Poisson}(\nu_{S,LS})$
(vi)	Simulate incidence of mortality	$N_{M,S} = \text{Binomial}(N_{S,LS}, \pi_{M,S,LS})$
(vii)	Simulate expected lifespan of each fatal case	$yll_G \sim e_{M,G}$
(viii)	Calculate disease burden for one iteration	$DALY = YLD_B + YLD_W + \Sigma yll_G$
(ix)	Repeat 1500 iterations = 1 simulation	
(x)	Repeat simulation 250 times with random samples from uncertain parameters	

* For computational purposes, the disease burden per case was not simulated individually but for all cases together. First, for one episode of diarrhoea, the disease burden was simulated as the product of samples from the variability distributions for severity and duration:

$yld_{ui} = t_{ui} \times w_{ui}$,

with u = iteration counter and i = B (bloody diarrhoea) or W (non-bloody diarrhoea).
Then, 10 000 samples of yld_{ui} were obtained and fitted to new Gamma distributions yielding:

$yld_B = \text{Gamma}(1{\cdot}232, 1{\cdot}792)$ and $yld_W = \text{Gamma}(1{\cdot}065, 0{\cdot}211)$.

The product of N samples from a Gamma(a, b) distribution followed a Gamma(Na, b) distribution. Therefore, the variability in the disease burden at population level can be modelled as indicated [(iii), (iv)].

Table 3. *Parameters of the gastroenteritis (GE) model*

	Description	Unit	Type*	Distribution	Mean	5-perc.	Median	95-perc.
ν_B	Incidence rate bloody diarrhoea	$year^{-1}$	U	Simulation results (see Appendix B)	992	39	590	3335
ν_W	Incidence rate non-bloody diarrhoea	$year^{-1}$	U	Simulation results (see Appendix B)	1120	43	667	3755
t_B	Duration of bloody diarrhoea	day	V	Gamma(3·2, 1·75)	5·6	1·6	5·0	11·5
t_W	Duration of non-bloody diarrhoea	day	V	Gamma(2·8, 1·2)	3·4	0·9	3·0	7·2
w_B	Severity of bloody diarrhoea	—	V	Beta(1·23, 1·9)†	0·39	0·05	0·37	0·82
w_W	Severity of non-bloody diarrhoea	—	V	Beta(1·5, 21)†	0·067	0·008	0·054	0·168
$\nu_{S,LS}$	Incidence rate lab-confirmed STEC O157 GE	$year^{-1}$	U	Gamma(120, 1/3)	40·0	34·2	39·9	46·2
$\pi_{M,S,LS}$	Case-fatality ratio lab-confirmed STEC O157 GE	—	U	Beta(2, 118)	0·0067	0·003	0·0014	0·039
$e_{M,G}$	Expected lifespan of a fatal case of STEC O157 GE	year	U + V	Discrete†‡	10·4	0·5	4·4	47·2

* U, uncertain; V, variable.
† See reference [3].
‡ Characteristic values of the variability distribution were presented with uncertain parameters at their expected value.

Mortality estimates were based on cases identified in laboratory surveillance. A Poisson distribution represented the variability in reported cases, and the number of fatal cases was sampled from a binomial process. For each individual fatal case, the expected lifespan at the time of death was sampled. Finally, the disease burden related to GE was calculated by accumulation of yld and yll over all cases and all diseases in a year.

The HUS model

A stepwise summary of the HUS model is given in Table 4, the parameter values are described in

https://doi.org/10.1017/S0950268804001979 Published online by Cambridge University Press

Table 4. *The haemolytic–uraemic syndrome (HUS) model*

(*a*) Incidence of HUS in The Netherlands per year (one iteration)		
(i)	Simulate the incidence of HUS <15 years	$N_{H<15} \sim \text{Poisson}(\nu_{H<15})$
(ii)	Simulate attributable proportion STEC O157	$\psi_{S\|H<15}$
(iii)	Calculate incidence STEC O157-associated HUS <15 years	$\psi_{H<15,S} \sim \text{Binomial}(\psi_{H<15}, \psi_{S\|H<15})$
(iv)	Simulate proportion of cases <15 years	$\pi_{H<15}$
(v)	Calculate incidence STEC O157-associated HUS all ages	$N_{H,S} \sim \text{NegBin}(\psi_{H<15,S}, \pi_{H<15})$
(*b*) For each individual STEC O157-associated HUS patient in one iteration		
(vi)	Simulate age at onset of HUS	a_H
(vii)	Simulate individual expected lifespan at HUS	e_H
(viii)	Calculate expected age at death *Note*: For any individual, the simulation is terminated if $a_i \geqslant a_M$	$a_M = a_H + e_H$
(ix)	Survive HUS?	$\text{Binomial}(1, \pi_{M\|H})$
(x)	No	$yll_H = e_H$; stop
(xi)	Yes: Simulate duration of clinical HUS	t_H
(xii)	Simulate severity of clinical HUS	w_H
(xiii)		$yld_H = t_H \times w_H$
(xiv)	Develop direct ESRD?	$\text{Binomial}(1, \pi_{E\|H})$
(xv)	Yes: go to step (xxi)	
(xvi)	Develop late ESRD?	$\text{Binomial}(1, \pi_{lE\|H})$
(xvii)	No: stop	
(xviii)	Yes: Simulate time to late ESRD	t_{lE}
(xix)	Calculate age at onset of (late) ESRD	$a_{lE} = a_H + t_{lE}$
(xx)		$yld_E = t_{lE} \times w_D/2$
(*c*) For each individual ESRD patient in one iteration		
(xxi)	Simulate time to first transplantation	$t_{TX1} \sim t_{TX}$
(xxii)	Calculate case-fatality ratio dialysis	if $t_{TX1} < 1$, $\pi_{M\|D1} = \pi_{M\|D*}$, or, $\pi_{M\|D1} = \pi_{M\|D*}/t_{TX1}$
(xxiii)	Survive to first transplantation?	$\text{Binomial}(1, \pi_{M\|D})$
(xxiv)	No	$yll_E = a_M - e_{ESRD\frac{1}{2}}$; stop
(xxv)	Yes: Simulate severity of dialysis	$w_{D1} \sim w_D$
(xxvi)		$yld_{D1} = t_{TX1} \times w_{D1}$
(xxvii)	Calculate age at first transplantation	$a_{TX1} = a_{(l)E} + t_{TX1}$
(xxviii)	Survive first transplantation?	$\text{Binomial}(1, \pi_{M\|TX})$
(xxix)	No	$yll_{TX1} = a_M - a_{TX1}$; stop
(xxx)	Yes: Simulate severity	$yld_{TX1} \sim w_{TX}$
(xxxi)	Simulate graft survival	$t_{FG1} \sim t_{FG}$
(xxxii)	Simulate severity functioning graft	$w_{FG1} \sim w_{FG}$
(xxxiii)		$yld_{FG1} = t_{FG1} \times w_{FG1}$
(xxxiv)	Calculate age at graft rejection	$a_{GR1} = a_{TX1} + t_{FG1}$
(xxxv)	Repeat steps (xxi)–(xxiii) for second graft	
(xxxvi)	Calculate age at second graft rejection	$a_{GR2} = a_{TX2} + t_{FG2}$
(xxxvii)	Calculate remaining lifespan	$t_{D3} = a_M - a_{GR2}$
(xxxviii)	Calculate case-fatality ratio dialysis	if $t_{D3} < 1$, $\pi_{M\|D3} = \pi_{M\|D*}$, or, $\pi_{M\|D3} = \pi_{M\|D*}/t_{D3}$
(xxxix)	Premature death by dialysis?	$\text{Binomial}(1, \pi_{M\|D3})$
(xl)	Yes: Simulate age at death	$a_{M\|D3} = a_{GR2} + \text{Uniform}(0, t_{D3})$
(xli)		$yll = a_M - a_{M\|D3}$
(xlii)	Simulate severity dialysis	$w_{D3} \sim w_D$
(xliii)		$yld_{D3} = (a_{M\|D3} - a_{GR2}) \times w_{D3}$; stop
(xliv)	No	$yld_{D3} = t_{D3} \times w_{D3}$
(*d*) For all HUS patients in one iteration		
(xlv)	Calculate disease burden for one iteration:	$DALY = \Sigma_i\, yld_i + \Sigma_i\, yll_i$
(xlvi)	Repeat (*a*)–(*d*) for 1500 iterations = 1 simulation	
(xlvii)	Repeat simulation 250 times with random samples from uncertain parameters	

https://doi.org/10.1017/S0950268804001979 Published online by Cambridge University Press

Table 5. *Parameters of the haemolytic–uraemic syndrome (HUS) model*

Param.	Description	Unit	Type*	Distribution	Mean	5-perc.	Median	95-perc.
$\nu_{H<15}$	Incidence rate HUS <15 yr	year^{-1}	C	Constant	22	—	—	—
$\psi_{S\|H}$	Attr. prop. STEC O157 <15 yr	—	U	Uniform($\psi_{S\|H1}$, $\psi_{S\|H2}$)	0·69	0·62	0·69	0·76
$\psi_{S\|H1}$	Only conclusive evidence	—	U	Bootstrapping	0·62	0·53	0·62	0·69
$\psi_{S\|H2}$	All VTEC is O157	—	U	Bootstrapping	0·77	0·70	0·77	0·84
$\pi_{H<15}$	Proportion of HUS <15 yr	—	U	Beta(13, 5)	0·72	0·54	0·73	0·88
a_H	Age at onset HUS	—						
	<15 yr		V+U	Discrete†‡	3·4	0·5	2·5	9·5
	>15 yr		V	Uniform(15, 100)	57·5	19·3	57·5	97·8
e_M	Expected lifespan at HUS	—	V	Discrete	Dependent on age at onset HUS			
$\pi_{M\|H}$	CFR of HUS	—	U					
	<65 yr			Beta(32, 835)	0·037	0·027	0·037	0·048
	>65 yr			Beta(10, 8)	0·56	0·36	0·56	0·74
t_H	Duration of clinical HUS	days	V	Uniform(14, 28)	21	15	21	27
w_H	Severity of clinical HUS	—	V	Discrete	0·93	0·73	1·00	1·00
$\pi_{E\|H}$	Prob. direct ESRD	—	U	Beta(23, 711)	0·031	0·022	0·031	0·043
$\pi_{lE\|H}$	Prob. late ESRD	—	U	Beta(8, 68)	0·11	0·05	0·10	0·17
t_{lE}	Time to late ESRD	year	V	Uniform(0, 40)	20	2	20	38
w_D	Severity weight dialysis	—	V	Discrete	0·18	0·00	0·13	0·57
t_{TX}	Time to transplantation	year	V+U					
	0–15 yr			Weibull(1·31, 2·31)†	2·1	0·2	1·7	5·3
	16–44 yr			Weibull(1·07, 3·51)†	3·4	0·2	2·5	9·6
	45–64 yr			Weibull(0·98, 9·79)†	9·9	0·5	6·7	30
	65+ yr			Weibull(1·04, 86·8)†	85	5	61	250
$\pi_{M\|D*}$	CFR 1st year dialysis	—	U					
	0–15 yr			Beta(2, 40)	0·048	0·009	0·041	0·111
	16–44 yr			Beta(18, 185)	0·089	0·058	0·087	0·124
	45–64 yr			Beta(61, 102)	0·37	0·31	0·37	0·44
	65–74 yr			Beta(81, 44)	0·65	0·58	0·65	0·72
	75+ yr			Beta(48, 13)	0·79	0·70	0·79	0·87
$\pi_{M\|TX}$	CFR transplantation	—	U	Beta(3, 36)	0·077	0·022	0·070	0·16
w_{TX}	Severity weight transplantation	—	V	Beta(18, 82)	0·18	0·12	0·18	0·25
w_{FG}	Severity weight functioning graft	—	V	Beta(12, 88)	0·12	0·07	0·12	0·18
t_{FG}	Graft survival	year	V+U	Weibull(0·47, 8·75)†	19·8	0·02	4·0	90·3

yr, Years.

* U, uncertain; V, variable; C, constant.

† For all factors that are both variable and uncertain, characteristic values of the variability distribution were presented with uncertain parameters at their expected value.

‡ Data from reference [20].

Table 5. In each iteration, a Poisson distribution represented the variability in the annual incidence of HUS for patients <15 years (all causes). The uncertainty in the attributable proportion of STEC O157 was introduced at two levels. First, the uncertainty related to the actual proportion of all STEC-positive results that are due to serogroup O157 was represented by a uniform distribution between the two extreme positions: only conclusive evidence for O157 was considered ($\psi_{S|H1}=0{\cdot}62$) or all STEC are O157 ($\psi_{S|H2}=0{\cdot}77$). Secondly, the uncertainty in the extreme values of $\psi_{S|H}$ was obtained by bootstrapping from the original case-control data. The model then simulated the total number of cases (all ages) by sampling from a negative binomial distribution. Because of the low number of cases, the disease burden model was set up as a micro-simulation, in which the life history was simulated for each individual patient.

https://doi.org/10.1017/S0950268804001979 Published online by Cambridge University Press

First, the model simulated an individual expected lifespan at the age of developing HUS. At each step of the model, the age of the patient was compared with this lifespan, and the simulation was halted when the predicted age at death by other causes was reached. The age distribution of HUS patients was sampled from a discrete distribution, based on observed data. The model then simulated if the patient survives the clinical phase of HUS by a Bernoulli trial. All transition probabilities (e.g. mortality ratios) were represented by Beta distributions, obtained by multiplying a Beta(0, 0) prior with the likelihood function of the observed data. If the patient died from HUS, the number of life years lost was calculated by subtracting the current age from the expected age at death. If the patient survived, the morbidity burden was sampled from a distribution of severity weights for 1 year with an episode of clinical HUS.

If the patient survived HUS, the model simulated whether the patient developed direct ESRD. If this was not the case, the model simulated if the patient developed late ESRD and if so, at what age. Patients who developed late ESRD experienced a gradual decline of their renal function that negatively affected their quality of life before the diagnosis was made. The model approximated the disease burden by assuming that the severity weight increased linearly from 0 to an individually simulated severity weight for dialysis. Then, the morbidity burden was calculated as 50% of the product of the time to late ESRD and the severity weight for dialysis. For patients who developed (late or direct) ESRD, the time to transplantation was then simulated. While on dialysis, there was an increased risk of death in the first year, which was simulated by a Bernoulli trial. If the patient died, the mortality burden was calculated by subtracting the age halfway during the time to transplantation from the lifespan. If the patient survived, the morbidity burden was calculated as the product of the time to transplantation and the severity weight. Then, a Bernoulli trial was performed to simulate whether the patient survived the transplantation. If not, the mortality burden was calculated as the difference between the age at transplantation and the expected age at death. If the patient survived, the morbidity burden was sampled from the distribution of severity weights for a year including transplantation. The survival time of the functioning graft was then simulated, the age at graft rejection and the morbidity burden were calculated. If graft rejection occurred before the simulated individual lifespan, the patient entered another cycle of dialysis–transplantation–graft rejection. In the absence of data, parameter values for this cycle were the same as for the first cycle (but all simulations were based on new samples from these distributions). If the age at second graft rejection was less than the expected age at death, patients were assumed to be dependent on dialysis for the rest of their life.

Severity weights

A preliminary analysis of the data indicated that for STEC O157, the disease burden was mainly determined by life years lost due to HUS-related mortality. Therefore, the sensitivity of the outcome for severity weights was limited, and no attempts were made to produce specific estimates for this study. We chose the most appropriate weights from published work.

GE

In a previous study [3], two severity weights were used for GE. The weight for watery diarrhoea (five episodes per day without major pain or cramps) was based on the Global Burden of Disease study [10]. We used this weight (median 0·054, mean 0·067) for non-bloody diarrhoea by STEC O157 to describe the variability of severity per case. A more severe case definition was developed for bacterial GE. We used this weight (median 0·37, mean 0·39) for bloody diarrhoea. As in the previous study, these weights were applied to the period of acute disease.

HUS and renal replacement therapy

There were no published severity weights for HUS. We therefore based our estimates on expert opinion, i.e. the experts used the EuroQol-5D instrument [24] to describe the health status of HUS patients and translated these descriptive health states into a single indicator for quality of life[4] by a published a regression model [25]. The mean severity weight for the clinical phase of HUS (duration 2–4 weeks) was 0·90 (range 0·73–1·00). After discharge from hospital, patients may take from weeks to months to fully recover function. In the absence of data on this

[4] The model of Dolan produces a quality of life estimate w' which ranges from 1·000 for a EuroQol score of 11 111 to $-0{\cdot}594$ for a EuroQol score of 33 333. To convert this weight to the 0–1 scale used for DALYs, we applied the formula $w=(1-w')/1{\cdot}594$.

https://doi.org/10.1017/S0950268804001979 Published online by Cambridge University Press

Table 6. *Burden of STEC O157-associated illness in The Netherlands* (*mean values per year*)

	Cases	YLD	Fatal cases	YLL	DALY
Total		***29·4***		***86·6***	***116·0***
Gastroenteritis	**2114**	**6·7**	**0·56**	**7·4**	**14·1**
Non-bloody	1118	0·7			0·7
Bloody	996	6·0	0·56	7·4	13·4
HUS	**21·7**	**1·0**	**2·2**	**57·8**	**58·8**
ESRD	**2·6**	**21·7**	**0·6**	**21·4**	**43·1**
Dialysis		16·1	0·3	7·1	23·2
Transplantation		0·6	0·3	14·3	14·9
Functioning graft		5·0	—		5·0

Note that due to the stochastic nature of the model, not all summations necessarily tally.
Bold, total per disease; bold italic, grand total.

phase, we did not account for reduced quality of life, which led to a slight underestimation of the disease burden.

For ESRD patients, we used data from the study of De Wit et al. [26]. These authors evaluated the health status of 165 dialysis (D) patients in The Netherlands. The quality of life of these patients was evaluated by different means, including the EuroQol-5D instrument. The EuroQol-5D scores were converted into severity weights as described above, resulting in a mean weight of 0·17 (range 0·00–0·68).

Patients on dialysis were eligible for renal transplantation. The severity weight for the transplantation stage also included the first year after surgery. If the graft was not rejected within that year it was assumed to have a normal function, albeit with a finite probability of rejection in following years. The average severity weight for renal transplantation (first year), based on data from a German study [27], was 0·18. The same study provided information on the average severity weight for life with a functioning graft as 0·12. Variability in the severity weights was characterized by either a discrete distribution (dialysis) or a continuous Beta distribution (transplantation, functioning graft).

Disease burden

Results of the baseline model

The mean incidence of HUS (all ages) due to infection with STEC O157 was 22 cases/year (Table 6), of which 15 were in children <15 years. Of these, 2·2 HUS cases had a fatal outcome, of which 1·5 were $\geqslant 65$ years. The renal damage due to HUS resulted in an annual mean of 2·6 cases of ESRD, of which 0·6 occurred immediately after the clinical HUS episode, and 2·0 at a later age. Of these 2·6 ESRD patients, 0·6 eventually died as a result of the effects of either dialysis or kidney transplantation.

The mean disease burden of STEC O157-associated illness was 116 DALYs/year. Death in the clinical phase of HUS was the largest single contributory factor, with 59 DALYs (50%), followed by death due to ESRD (21 DALYs = 18%). Overall, fatal outcomes accounted for 87 DALYs or 75% of the total disease burden. Morbidity accounted for 25% of the total disease burden, with dialysis due to ESRD and bloody diarrhoea as the most important factors.

Uncertainty and variability

The second-order Monte Carlo model used to calculate the disease burden allowed separate evaluation of uncertainty and variability in the average disease burden estimates. The model was set up as a series of 250 simulations, with 1500 iterations each. Each simulation calculated the variability of the disease burden, for a particular set of uncertain input parameters. Hence, the difference between results of different simulations quantified the uncertainty in the model results. Represented graphically (Fig. 5), the results of 250 simulations were sorted according to the mean result for total DALYs, that ranged between 68 and 223. The 90% CI for the mean burden was 84–159 (see Table 7). The figure also shows the median of each simulation, which was very similar to the mean, indicating that the output distributions were relatively symmetrical. The 5- and 95-percentiles of each simulation indicated the variability of the disease burden. It can be seen that the variability increased slightly with increasing mean and that variability was larger than uncertainty. The relative importance of uncertainty and variability was quantified by the variance ratio:

$$F = \sigma^2_{\text{var}} / \sigma^2_{\text{unc}} = (76{\cdot}6/23{\cdot}3)^2 = 10{\cdot}8.$$

In other words, the effect of variability was approximately 11 times bigger than the effect of uncertainty.

The relative contribution of variability was largest in YLL (Table 7) which was related to the low numbers of fatal cases and the wide age range for these cases. For example, for direct mortality due to HUS,

https://doi.org/10.1017/S0950268804001979 Published online by Cambridge University Press

Table 7. *Summary statistics of output distributions*

	Mean of means	5-percentile of means	Median of means	95-percentile of means	s.d. of means (unc.)	Mean of s.d.s (var.)	*F**	Median of medians
DALY	116·0	84·9	116·0	158·5	23·3	76·6	10·8	101·8
YLD	29·4	16·7	29·4	47·9	10·3	21·8	4·5	22·7
YLL	86·6	62·5	86·6	122·6	19·0	72·1	14·4	75·2

* Variance ratio (variability/uncertainty).

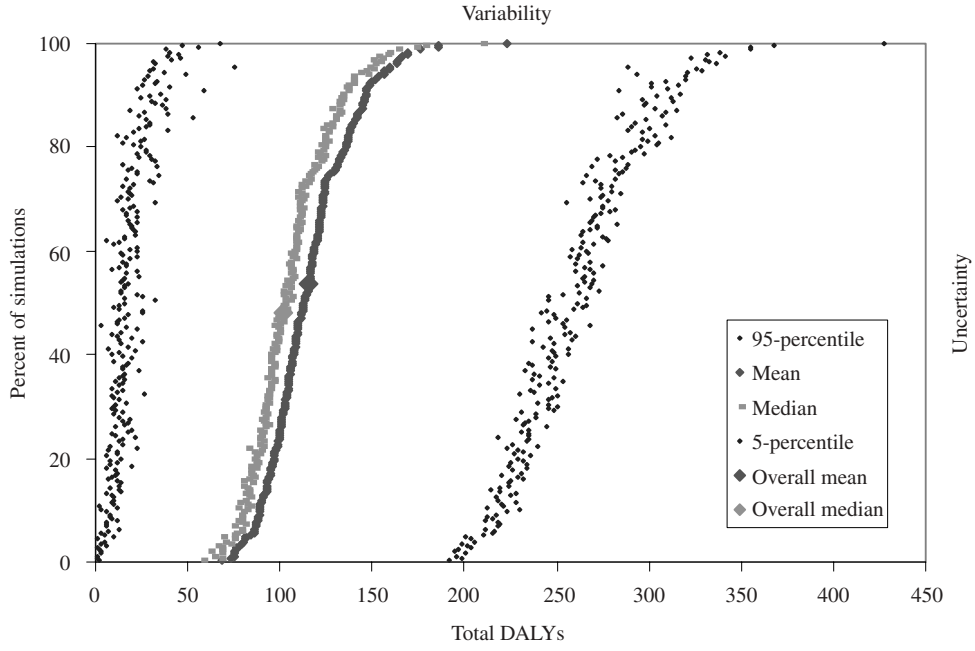

Fig. 5. Uncertainty and variability in total DALYs estimate. The graph shows characteristic values of 250 simulations, consisting of 1500 iterations each and ordered by the simulation mean.

$F = 23{\cdot}1$ and for all mortality related to HUS and ESRD, $F = 15{\cdot}2$. In comparison, the F ratio for YLD related to GE was 0·00033, confirming that the incidence estimates for GE are highly uncertain. However, as the contribution of YLD-GE to the overall disease burden was small, this uncertainty was not reflected in the overall estimate.

Sensitivity analysis

Here we examine the effect of several alternative choices of parameter values on the outcomes of the model. Univariate scenario analyses were performed by setting all uncertain parameters in the model at their median value, and replacing the alternative values as described below. The model was then run with 1500 iterations to simulate the disease burden. The results (Fig. 6) show that the mean in the baseline scenario (108 DALYs/year) was lower than that shown in Table 6. This is because these simulations did not include possible high values for uncertain parameters.

Scenario 1. The baseline scenario was based on an incidence of 13·8 consultations/1000 pyr. This estimate was corrected for under-reporting on the basis of reported GP visits in the SENSOR study. The uncorrected figure was 8·0 consultations/1000 pyr, whereas a maximum estimate was 35 consultations/1000 pyr. In the baseline scenario, $2{\cdot}18 \times 10^5$ consultations/year were expected in The Netherlands, in the alternative scenarios these figures were 1·26 and $5{\cdot}53 \times 10^5$/year respectively. Alternative estimates for the incidence of STEC O157-associated bloody/non-bloody diarrhoea were 360/410 (scenario 1a) and 1570/1760 (scenario 1b) cases per year.

https://doi.org/10.1017/S0950268804001979 Published online by Cambridge University Press

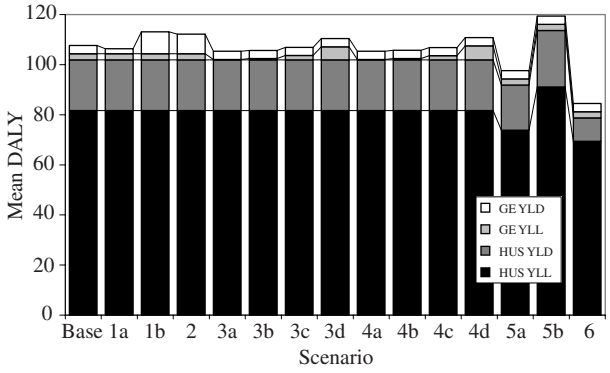

Fig. 6. Results of scenario analysis. The graph shows the mean disease burden and different building blocks for different scenarios, representing the effect of uncertain model assumptions (for details see text).

Scenario 2. Duration of diarrhoea was based on relatively few data and the duration in children (the age group with the highest incidence) could be higher. In the alternative scenario, we used data as observed in an outbreak at a day-care centre: 12 days for bloody and 7 days for non-bloody diarrhoea [28].

Scenario 3. The case-fatality ratio for GE was highly uncertain. In this scenario, we used four different values [0·1 (3a), 0·3 (3b), 1 (3c) and 3 (3d) %] around the baseline estimate of 1·4%.

Scenario 4. In the baseline model it was assumed that the life expectancy of fatal cases of GE was the same as for the general population. It is possible that those who died from GE had underlying diseases and consequently a lower life expectancy. The life expectancy of fatal cases was varied in this scenario between 0·3 (4a), 1 (4b), 3 (4c) and 10 (4d) years. Alternatively, the outcomes of this scenario could be considered to account for a lower quality of life of fatal cases, had they not died (i.e. the severity weight for death was <1).

Scenario 5. In this scenario, we evaluated the effect of uncertainty in the proportion of HUS cases that was attributable to infection with STEC O157, by assuming that only culture-positive evidence was a valid indication of an aetiological role of STEC O157 ($\psi_{S|H}=0{\cdot}62$) (5a) or that all cases with Stx in the faeces were attributable to STEC O157 ($\psi_{S|H}=0{\cdot}77$) (5b).

Scenario 6. In this scenario, we did not use the long-term follow-up data on the probability of developing late ESRD, as reported in France, but only used follow-up data, as reported in The Netherlands and Belgium [29]. These data represented a shorter period of approximately 10 years and a considerably lower probability of developing ESRD (2·7 *vs.* 10%).

Most scenarios related to parameters that influenced YLD or YLL by GE, and these factors had a relatively small effect on the overall disease burden. The figure shows that GE YLD was relatively high in Scenarios 1b and 2, but the total disease burden only increased by 5 DALYs/year.

The most important effects on the model output were observed in Scenarios 5 and 6. Increasing the fraction of HUS cases that was attributable to STEC O157 infection to 0·77 increased YLL HUS from 82 to 91 and YLD HUS from 20 to 23, resulting in an increase of the total disease burden from 108 to 120 DALYs. It must be noted that it is well established in the literature that other STEC serogroups are also able to induce HUS, so this output may overestimate the disease burden attributable to STEC O157. On the other hand, the method to detect Stx and/or Stx-encoding genes is limited in sensitivity and accounting for this factor may increase the estimate of the true attributable fraction. In Scenario 6, YLL due to HUS (including ESRD) decreased to 69 and YLD decreased to 9. This reduced the total disease burden to 85 DALYs (a decrease of 21%). Comparing these results with the 5-percentile value of the uncertainty of mean burden in the complete model (84·9–158·5, Table 7) indicated that the probability to develop ESRD was one of the most important sources of uncertainty in the model.

The most important contribution to the total disease burden was made by HUS-related mortality. We therefore separately evaluated the effect of uncertainty in the case-fatality ratio of HUS on the model results. The parameter $\pi_{M|H}$ for age groups under and over 65 years was set at the 1-, 5-, 50-, 95- and 99-percentile values of the distributions defined in Table 5 (ranging between 0·024–0·053 for <65 years and 0·291–0·800 for >65 years). This range of parameter values reflected the larger uncertainty in the case-fatality ratio for the 65+ years age group, due to a lower number of observations. However, the results (data not shown) demonstrated that the total disease burden was most sensitive to the (relatively small) uncertainty in the case-fatality ratio for <65 years. This can be explained by the high number of life years lost by one fatal case among children. In these scenarios, the total disease

burden ranged between 90 and 125 DALYs, less than the range of uncertainty in the complete baseline model.

DISCUSSION

Diarrhoeal illness associated with infection by STEC O157 is relatively severe, and a high proportion of infected children suffer from HUS. To quantify the burden of disease by STEC O157 in The Netherlands, and to compare this burden with that caused by other illnesses including by other enteropathogens, we used the concept of DALYs. The incidence of STEC O157-related disease in The Netherlands was low but highly uncertain with a median of 1250 and a mean of 2100 cases of GE and 22 cases of HUS per year. Nevertheless, we estimated that approximately 120 DALYs/year are lost due to these illnesses. A previous study [3] estimated the disease burden of campylobacteriosis in The Netherlands in the early 1990s as 1400 DALYs/year. On an absolute basis, campylobacteriosis is more significant from a public health perspective. It was also estimated that the mean incidence of *Campylobacter*-associated GE was 318 000 cases/year. Thus, on a case per case basis, STEC O157 has a >12-fold higher health impact (55 DALYs/1000 cases) than *Campylobacter* (4·6 DALYs/1000 cases). Any outbreak or increase in incidence of STEC O157 will be significant for public health. A further indication of the serious nature of STEC O157-associated illness can be derived from comparison with preliminary disease burden estimates made for several key waterborne pathogens: *Cryptosporidium parvum* 1·47, rotavirus in high-income regions 14, and hepatitis A virus in developed countries 191 DALYs/1000 cases [30]. Compare this with a preliminary estimate of 482 DALYs/1000 cases of rotavirus-associated illness in low-income regions, which reflects the significance of child mortality.

As in the previous study, the DALY concept proved to be a flexible and robust tool to estimate the impact of infectious intestinal illness on public health. The robustness of the final estimates was based in part on the aggregate nature of the estimates. Different disease end-points contributed to the overall disease burden, and it was not likely that all estimates were simultaneously biased in the same direction. This does not put aside the fact that there was considerable uncertainty in the underlying estimates and hence in the estimated disease burden. The major sources of uncertainty were the (lack of) epidemiological data, and in particular the incidence of GE due to STEC O157, the fraction of HUS that was attributable to STEC O157 and the probability of developing ESRD.

The second-order Monte Carlo approach used in this report allowed separate evaluation of the effects of variability and uncertainty in the model parameters. Overall, uncertainty was only found to be of major importance for the incidence estimate of STEC O157-associated GE, and consequently also YLD due to GE. For all other factors in the model, and for the aggregated estimates of YLD, YLL and DALY, variability was much greater than uncertainty, even though several parameter estimates were highly uncertain. This is mainly related to the fact that there were only a low number of cases of HUS and related death or ESRD per year. Hence, even if incidence rates are constant over time (as was assumed in this model), the actual number of cases per year will show important variation. Sensitivity analysis indicated that alternative choices of, for example a particular study for parameter estimation did not have a major impact on the estimate of the total disease burden.

The uncertainty in the incidence estimate for STEC O157-associated GE was mainly caused by the low number of positive stools in the SENSOR and NIVEL studies (0 and 1 respectively) and the absence of direct information on the GP consultation pattern of STEC O157 cases with GE. As a consequence, the uncertainty distribution of the incidence was extremely skewed with a mode of 0, a median of 1200 and a mean of 2000 cases/year. The probability that the mean underestimated the true incidence was only 33% and it is therefore more likely that the true incidence is lower than the simulation mean.

ACKNOWLEDGEMENTS

The authors thank the following individuals for their advice and providing data: L. A. H. Monnens (Nijmegen), clinical aspects of haemolytic–uraemic syndrome; M. A. S. de Wit, W. van Pelt (RIVM), epidemiology of gastroenteritis; N. Nagelkerke, R. Hoogenveen (RIVM), D. de Wet (Johannesburg, South Africa), statistical analysis; F. de Charro (Rotterdam), renal replacement registry; M. A. Gagnadoux (Hôpital Necker-Enfants-Malades, Paris, France), long-term follow-up of HUS patients.

APPENDIX A: Symbols, subscripts and probability density functions

Symbol list

Parameters

e	individual lifespan	year
a	age	year
ν	incidence rate	year^{-1}
N	incidence	year^{-1}
M	population size	
t	observation or person time	year
w	severity weight	
π	binomial probability	
ψ	attributable proportion	

Level of observation

No subscript = population; subscript GP = general practitioner; subscript LS = laboratory surveillance.

Other subscripts

- i index for age class
- j index for type of illness. No index = total Dutch population 1999 (CBS); G = gastroenteritis; S = STEC O157-associated illness; B = bloody diarrhoea; W = non-bloody diarrhoea; H = haemolytic–uraemic syndrome; E = end-stage renal disease; M = mortality; TX = kidney transplantation; FG = functioning graft; D = dialysis; R = recovered renal function; * = participants in case-control study, ′ = raw (non-standardized) data
- u index for iteration number
- s index for simulation number
- z dummy variable

Abbreviations

DALYs	Disability Adjusted Life Years*
ESRD	end-stage renal disease
GE	gastroenteritis
(D+) HUS	(diarrhoea-associated) haemolytic–uraemic syndrome
NIVEL	Netherlands Institute of Primary Health Care
PYN	pyelonephritis
Renine	Renal Replacement Registry of The Netherlands
SENSOR	population based cohort study on gastroenteritis
Stx	Shiga toxin
STEC O157	Shiga toxin-producing *Escherichia coli* serogroup O157
TX	renal transplantation
YLD	years lived with a disability*
YLL	years of life lost*

* Upper-case letters indicate estimates at population level; lower-case letters indicate estimates at individual level.

Probability density functions

$$\text{Beta}(\alpha,\beta)=f(x)=\frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)},$$

$$\text{where } B(\alpha,\beta)=\int_0^1 t^{\alpha-1}(1-t)^{\beta-1}\,\mathrm{d}t,$$

$$\text{Binomial}(n,\pi): f(x)=\binom{n}{x}\pi^x(1-\pi)^{n-1},$$

$$\text{Gamma}(\alpha,\beta): f(x)=\frac{\beta^{-\alpha}x^{\alpha-1}\exp\left(-\frac{x}{\beta}\right)}{\Gamma(\alpha)},$$

$$\text{where } \Gamma(\alpha)=\text{ Euler's Gamma function},$$

$$\text{NegBin}(s,\pi): f(x)=\binom{s+x-1}{x}\pi^s(1-\pi)^x,$$

$$\text{Poisson}(\lambda): f(x)=\frac{\mathrm{e}^{-\lambda}\lambda^x}{x!},$$

$$\text{Uniform}(\min,\max): f(x)=\frac{1}{\max-\min},$$

$$\text{Weibull}(\alpha,\beta): f(x)=\alpha\beta^{-\alpha}x^{\alpha-1}\exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right),$$

APPENDIX B: Statistical analysis of data on gastroenteritis (GE)

Uncertainty in the incidence of GP consultations, based on the SENSOR study and the fraction attributable to STEC O157

The Appendix B Table shows the standardization of incidence rates for all cases of GE and GP consultations, based on raw data from the SENSOR study. In contrast to De Wit et al. [12] we only standardized for age and not for cohort and sex. This is mainly because otherwise, there would be subgroups with no GP consultations in the middle-age classes. Comparison of the standardized incidence rate in the SENSOR study (0·283/pyr) with the standardized incidence rate in this study (0·295/pyr) shows that the difference was small. The uncertainty in the standardized incidence rates was obtained by Bayesian methods [31]. The observed number of $N'_{G,i}$ cases in each age group was assumed to arise from a Poisson process with underlying incidence rate $\nu_{G,i}$ and observation time t_i for each age group i. Then, $\nu_{G,i}\sim$ Gamma($N'_{G,i}$, $1/t_i$) and $N_{G,i}\sim$Poisson($\nu_{G,i}$. M_i). The observed number of $N'_{GP|G^*,i}$ consultations among $N'_{G^*,i}$ respondents to questionnaires was assumed to result from a binomial process with the underlying probability of consulting a GP being $\pi_{GP|G^*,i}$. Then, $\pi_{GP|G^*,i}\sim$Beta($N'_{GP|G^*,i}$, $N'_{G^*,i}-N'_{GP|G^*,i}$). The incidence of GP consultations for GE was estimated using the normal approximation of the

https://doi.org/10.1017/S0950268804001979 Published online by Cambridge University Press

Appendix B Table. *Simulation model of the uncertainty in the incidence of gastroenteritis (GE) (means)*

Age group (years)	M_i	t_i	$N'_{G,i}$	$\nu_{G,i}$	$N_{G,i}$	$N'_{G^*,i}$	$N_{GP\|G^*,i}$	$\pi_{GP\|G^*,i}$	$N_{G,GP,i}$	$\nu_{G,GP}$
0	$2{\cdot}00\times10^5$	311	237	0·762	$1{\cdot}52\times10^5$	182	29	0·159	$2{\cdot}43\times10^4$	
1–4	$7{\cdot}76\times10^5$	419	372	0·888	$6{\cdot}89\times10^5$	253	21	0·083	$5{\cdot}72\times10^4$	
5–11	$1{\cdot}38\times10^6$	444	224	0·505	$6{\cdot}96\times10^5$	119	7	0·059	$4{\cdot}09\times10^4$	
12–17	$1{\cdot}11\times10^6$	307	49	0·160	$1{\cdot}77\times10^5$	15	1	0·067	$1{\cdot}18\times10^4$	
18–64	$1{\cdot}02\times10^7$	420	104	0·248	$2{\cdot}52\times10^6$	51	1	0·020	$4{\cdot}94\times10^4$	
65+	$2{\cdot}13\times10^6$	328	64	0·195	$4{\cdot}16\times10^5$	24	2	0·083	$3{\cdot}46\times10^4$	
Crude	$1{\cdot}58\times10^7$	2229	1050	0·471	$7{\cdot}42\times10^6$	644	61	0·095	$7{\cdot}03\times10^5$	
Stand.				0·295	$4{\cdot}65\times10^6$				$2{\cdot}18\times10^5$	0·014

Subscript i refers to different age groups. M_i, number of persons in the Dutch population, 1999 (CBS); t_i, observation time (pyr); $N'_{G,i}$, observed cases of GE; $\upsilon_{G,i}$, inferred incidence rate of GE in the general population (pyr^{-1}); $N_{G,i}$, inferred incidence of GE (yr^{-1}); $N_{G^*,i}$, number of cases with GE who participated in the case-control study; $N_{GP|G^*,i}$, number of cases with GE who participated in the case-control study and consulted a GP; $\pi_{GP|G^*,i}$, probability of consulting a GP for a case with GE; $N_{G,GP,i}$, inferred incidence of GP consultations for GE (yr^{-1}); $\nu_{G,GP,i}$, inferred incidence rate of GP consultations for GE (yr^{-1}).

binomial distribution: $N_{G,GP,i}\sim$ normal($N'_{G,i}\,.\,\pi_{GP|G^*,i}$, $\sqrt{\{N'_{G,i}\,.\,\pi_{GP|G^*i}(1-\pi_{GP|G^*,i})\}}$) and the standardized incidence rate of GP consultations was estimated as $\nu_{G,GP}=N_{G,GP}/M$. The uncertainty in the standardized estimates was estimated by Latin hypercube sampling from these distributions.

The uncertainty in the fraction of all GP consultations, attributable to STEC O157 was simulated as $\pi_{GP,S|G}\sim$ Beta(1, 797), i.e. a prior Beta(0, 0) distribution[5] multiplied by the binomial likelihood of observing 1/798 cases. Finally, the incidence of GP consultations due to GE by STEC O157 was estimated as $\nu_{S,GP,i}\sim$ Poisson($\nu_{G,GP,i}\,.\,\nu_{GP,S|G}$).

Estimation of cases in the total population from GP consultations

Let N_S be the annual incidence of symptomatic cases (bloody and non-bloody diarrhoea) in the population, with $\pi_{B|S}$ the proportion of bloody diarrhoea and $\pi_{W|S}=(1-\pi_{B|S})$ the proportion of non-bloody diarrhoea. Michel et al. [15, table 1] show 253/538 unselected cases of STEC O157-associated diarrhoea are bloody. Then $\pi_{B|S}$ followed a Beta(254, 286) distribution [using a uniform Beta(1, 1) prior].

Let $\pi_{GP|B}$ be the proportion of patients with bloody diarrhoea who seek medical attention and $\pi_{GP|W}$ the proportion of patients with watery diarrhoea who do so. Estimates come from De Wit et al. [14] who reported that in the SENSOR study, 2 out of 9 patients with bloody diarrhoea and 59 out of 635 patients with watery diarrhoea consulted their GP. Using uniform priors, $\pi_{GP|B}\sim$ Beta(3, 8) and $\pi_{GP|W}\sim$ Beta(60, 577). The overall proportion reporting to a GP is $\pi_{GP|S}=\pi_{B|S}\,.\,\pi_{GP|B}+(1-\pi_{B|S})\,.\,\pi_{GP|W}$. Let $\pi_{S,GP}$ be the inferred incidence of GP consultations due to GE by STEC O157. This can be considered the number of 'successes' in a binomial process with a total of N_S cases and a probability of success $\pi_{GP|S}$. The number of cases who did not consult a GP followed a negative binomial distribution and the incidence of all STEC O157 cases in the population can be estimated as $N_S\sim$ NegBin($N_{S,GP}+1$, $\pi_{GP|S}$) $+N_{S,GP}$ [31].

Uncertainty in the incidence of laboratory-confirmed STEC O157-associated GE and the case-fatality ratio

Let $N_{S,LS}$ be the number of cases of STEC O157-associated GE in laboratory surveillance, and t_{LS} be the observation time. The uncertainty in the incidence rate can be quantified as $\nu_{S,LS}\sim$ Gamma($N_{S,LS}$, $1/t_{LS}$) [31]. For 1999–2001; $N_{S,LS}=120$ and $t_{LS}=3$ years, hence $\pi_{S,LS}\sim$ Gamma(120, 1/3). Let $N_{M,S,LS}$ be the number of fatal cases observed in laboratory surveillance. Then, using a Beta(0, 0) prior, $\pi_{M,S,LS}\sim$ Beta($N_{M,S,LS}$, $N_{S,LS}-N_{M,S,LS}$). For 1999–2001, $N_{M,S,LS}=2$, hence $\pi_{M,S,LS}\sim$ Beta(2, 118).

[5] Let π be the parameter of a binomial distribution. x positives are observed among a total of n tests. Using a Beta(a, b) prior, the posterior distribution for πi followed a Beta($x+a$, $n-x+b$) distribution. A uniform Beta(1, 1) prior is recommended [31]. In the case of small numbers of (positive) observations, the choice of a and b significantly influences the posterior distribution and a uniform prior may shift the expected value away from the maximum likelihood estimate from $\pi(=x/n)$ to $(x+a)/(n+b)$. The Beta(0, 0) distribution, while mathematically undefined, is an acceptable alternative prior distribution that does not affect the expected value of the posterior distribution [32]. Note that the 90% credible interval of this posterior distribution is smaller than the MLE-based interval.

https://doi.org/10.1017/S0950268804001979 Published online by Cambridge University Press

REFERENCES

1. Lindsay JA. Chronic sequelae of foodborne disease. Emerg Infect Dis 1997; **3**: 443–452.
2. Smith JL. Foodborne illness in the elderly. J Food Prot 1998; **61**: 1229–1239.
3. Havelaar AH, De Wit MAS, Van Koningsveld R, Van Kempen E. Health burden in the Netherlands due to infection with thermophilic *Campylobacter* spp. Epidemiol Infect 2000; **125**: 505–522.
4. Cransberg K, Van den Kerkhof JHCT, Bänffer JRJ, et al. Four cases of hemolytic uremic syndrome – source contaminated swimming water? Clin Nephrol 1996; **46**: 45–49.
5. Heuvelink AE, Tilburg JJHC, Herbes RG, et al. Een explosie van *E. coli* O157-infectie binnen een gezin [An outbreak of *E. coli* O157 infection in a family]. Infectieziekten Bulletin 1998; **9**: 174–175.
6. Van Duynhoven YTHP, De Jager CM, Heuvelink AE, et al. Enhanced laboratory-based surveillance of Shiga-toxin producing *Escherichia coli* O157 in The Netherlands. Eur J Clin Microbiol Infect Dis 2002; **21**: 513–522.
7. Karmali MA, Petric M, Lim C, Fleming PC, Arbus GS, Lior H. The association between idiopathic hemolytic uremic syndrome and infection by verotoxin-producing *Escherichia coli*. J Infect Dis 1985; **151**: 775–782.
8. Siegler RL. The hemolytic uremic syndrome. Pediatr Clin North Am 1995; **42**: 1505.
9. Siegler RL. Spectrum of extrarenal involvement in postdiarrheal hemolytic-uremic syndrome. J Pediatr 1994; **125**: 511–518.
10. Murray CJ, Acharya AK. Understanding DALYs (disability-adjusted life years). J Health Econ 1997; **16**: 703–730.
11. Havelaar AH, Van Duynhoven YTHP, Nauta MJ, et al. Disease burden in the Netherlands due to infections with Shiga-toxin producing *Escherichia coli* O157. Bilthoven: National Institute for Public Health and the Environment, 2003. Report no. 284550008. (http://www.rivm.nl/bibliotheek/rapporten/284550008.html). Accessed April 2003.
12. De Wit MAS, Koopmans MPG, Kortbeek LM, et al. Sensor, a population-based cohort study on gastroenteritis in the Netherlands: incidence and etiology. Am J Epidemiol 2001; **154**: 666–674.
13. De Wit MAS, Koopmans MPG, Kortbeek LM, Van Leeuwen NJ, Vinje J, Van Duynhoven YTHP. Etiology of gastroenteritis in sentinel general practices in the Netherlands. Clin Infect Dis 2001; **33**: 280–288.
14. De Wit MAS, Kortbeek LM, Koopmans MP, et al. A comparison of gastroenteritis cases in a general practice based study and a community based study. Epidemiol Infect 2001; **127**: 389–397.
15. Michel P, Wilson JB, Martin SW, Clarke RC, McEwen SA, Gyles CL. Estimation of the under-reporting rate for the surveillance of *Escherichia coli* O157:H7 cases in Ontario, Canada. Epidemiol Infect 2000; **125**: 35–45.
16. Belongia EA, MacDonald KL, Parham GL, et al. An outbreak of *Escherichia coli* O157:H7 colitis associated with consumption of precooked meat patties. J Infect Dis 1991; **164**: 338–343.
17. Anonymous. The investigative report of the Walkerton outbreak of waterborne gastro-enteritis. Owen Sound, Ont. Canada: Bruce-Grey-Owen Sound Health Unit, 2000.
18. Van de Kar NCAJ, Heuvelink AE, de Boer E, Monnens LAH. Infecties met verocytotoxine-producerende *Escherichia coli* en het hemolytisch-uremisch syndroom [Infections with verocytotoxin-producing *Escherichia coli* and the haemolytic–uraemic syndrome]. Ned Tijdschr Geneeskd 1996; **140**: 134–137.
19. Chart H, Van der Kar NCA, Tolboom JJM, Monnens LMH, Rowe B. Serological detection of verocytotoxin-producing *Escherichia coli* in patients with haemolytic uraemic syndrome in western Europe. Eur J Clin Microbiol Infect Dis 1993; **12**: 707–709.
20. Van de Kar NCAJ, Roelofs HGR, Muytjens HL, et al. Verocytotoxin-producing *Escherichia coli* infection in hemolytic uremic syndrome in part of western Europe. Eur J Pediatr 1996; **155**: 592–525.
21. Dundas S, Murphy J, Soutar RL, Jones GA, Hutchinson SJ, Todd WTA. Effectiveness of therapeutic plasma exchange in the 1996 Lanarkshire *Escherichia coli* O157:H7 outbreak. Lancet 1999; **354**: 1327–1330.
22. Gagnadoux MF, Habib R, Gubler MC, Bacri JL, Broyer M. Long-term (15–25 years) outcome of childhood hemolytic-uremic syndrome. Clin Nephrol 1996; **46**: 39–41.
23. Conlon PJ, Howell DN, Macik G, Kovalik EC, Smith SR. The renal manifestations and outcome of thrombotic thrombocytopenic purpura/hemolytic uremic syndrome in adults. Nephrol Dial Transplant 1995; **10**: 1189–1193.
24. Brooks R with the EuroQol Group. EuroQol: the current state of play. Health Policy 1996; **37**: 53–72.
25. Dolan P. Modeling valuations for EuroQol health states. Med Care 1997; **35**: 1095–1108.
26. De Wit GA, Ramsteijn PG, de Charro FTh. Economic evaluation of end stage renal disease treatment. Health Policy 1998; **44**: 215–232.
27. Greiner W. Ökonomische Evaluationen von Gesundheitsleistungen [Economic evaluations of health care provisions]. Baded-Baden: Nomos Verlagsgesellschaft, 1999.
28. Spika JS, Parsons JE, Nordenberg D, Wells JG, Gunn RA, Blake PA. Hemolytic uremic syndrome and diarrhea associated with *Escherichia coli* O157:H7 in a day care center. J Pediatr 1986; **109**: 287–291.
29. De Jong M, Monnens L. Haemolytic–uraemic syndrome: a 10-year follow-up study of 73 patients. Nephrol Dial Transplant 1988; **3**: 379–382.
30. Havelaar AH, Melse JM. Quantifying public health risks in the WHO Guidelines for Drinking-Water Quality. Bilthoven: National Institute for Public Health and the Environment, 2003. Report no. 734301022.
31. Vose D. Risk analysis. A quantitative guide. 2nd edn. Chichester: John Wiley & Sons, 2000.
32. Gelman A, Carlin JB, Stern HS, Rubin DB. Bayesian data analysis. London: Chapman and Hall, 1995.