IDSA — Infectious Diseases Society of America   hivma — hiv medicine association   OXFORD

Estimating the Burden of Medically Attended Norovirus Gastroenteritis: Modeling Linked Primary Care and Hospitalization Datasets

Thomas Verstraeten,[1] Tom Cattaert,[1] John Harris,[2] Ben Lopman,[4] Clarence C. Tam,[3,5] and Germano Ferreira[1]

[1]P95 Pharmacovigilance and Epidemiology, Heverlee, Belgium; [2]National Institute for Health Research Health Protection Research Unit in Gastrointestinal Infections, University of Liverpool, and [3]Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine, United Kingdom; [4]Rollins School of Public Health, Emory University, Atlanta, Georgia; [5]Saw Swee Hock School of Public Health, National University of Singapore

(See the editorial commentary by Riddle, on pages 929–31.)

Background. Norovirus is the leading cause of community-acquired and nosocomial acute gastroenteritis. Routine testing for norovirus is seldom undertaken, and diagnosis is mainly based on presenting symptoms. This makes understanding the burden of medically attended norovirus-attributable gastroenteritis (MA-NGE) and targeting care and prevention strategies challenging.

Methods. We used linked population-based healthcare datasets (Clinical Practice Research Datalink General Practice OnLine Database linked with Hospital Episode Statistics Admitted Patient Care) to model the incidence of MA-NGE associated with primary care consultations or hospitalizations according to age groups in England in the period July 2007–June 2013.

Results. Mean annual incidence rates of MA-NGE were 4.9/1000 person-years and 0.7/1000 person-years for episodes involving primary care or hospitalizations, respectively. Incidence rates were highest in children aged <5 years: 34.0 consultations/1000 person-years and 3.3 hospitalizations/1000 person-years. Medically attended norovirus-attributable gastroenteritis hospitalization rates were second highest in adults aged >65 years (1.7/1000 person-years).

Conclusions. In this particular study, the burden of MA-NGE estimated from healthcare datasets was higher than previously estimated in small cohort studies in England. Routinely collected primary care and hospitalization datasets are useful resources to estimate and monitor the burden of MA-NGE in a population over time.

Keywords. norovirus; gastroenteritis; electronic health records; Clinical Practice Research Datalink (CPRD); Hospital Episode Statistics (HES).

Norovirus is the leading cause of acute gastroenteritis worldwide, causing both community-acquired disease and healthcare-associated outbreaks. Incidence rates are higher in young children, and hospitalization is common in both the very young and old [1]. Globally, norovirus-attributable gastroenteritis is estimated to cost society $60.3 billion each year, of which 62% is borne by high-income countries and 66% is related to illness in children aged <5 years [2]. In the United States, medical costs due to norovirus infection in children aged <5 years alone have been estimated to approach $300 million per year [3]. In a population-based longitudinal study in the United Kingdom, norovirus was found to be the most common cause of community-acquired infectious intestinal disease (IID), resulting in nearly 3 million annual cases [4] and ₤81 million annually in costs to patients and the health system, exceeding costs of *Campylobacter* and rotavirus combined (estimated costs of ₤50 and ₤25 million, respectively) [5].

Paradoxically, unless there is an institutional outbreak of gastroenteritis, testing for norovirus is rarely performed. Knowledge of the burden of medically attended norovirus-attributable gastroenteritis (MA-NGE) is generated mainly through infrequently conducted prospective cohort studies. These studies are labor-intensive and costly and lack precision for individual pathogens due to the limitation on the size of the study population. Alternative investigative methods that are less expensive, can be conducted regularly, and cover longer study periods, thereby allowing researchers to account for the 2–4 yearly pandemic cycles seen for norovirus [6], should therefore be explored and considered to create more in-depth and temporal knowledge on the burden of MA-NGE.

The aim of this study was to use linked population-based healthcare datasets to assess the incidence of medically attended gastroenteritis (MA-AGE) and model the incidence of MA-NGE based on seasonal patterns of MA-AGE for unspecified causes [7, 8]. Using linked primary care and hospitalization datasets for England for the

period July 2007–June 2013, we estimated the burden of MA-AGE and MA-NGE associated with primary care or hospitalization for different age groups, according to calendar year and month.

METHODS

Data Sources

Data were extracted from the Clinical Practice Research Datalink General Practice Online Database (CRPD-GOLD), linked with the Hospital Episode Statistics Admitted Patient Care (HES APC) data. the Clinical Practice Research Datalink General Practice Online Database is a primary care dataset of computerized general practice medical records in the United Kingdom, managed by the Department of Health [9]. Data collection into CPRD GOLD began in 1987, and it currently holds data on 8% of the UK population. Hospital Episode Statistics Admitted Patient Care data contain details of all admissions to or attendances at English National Health Service (NHS) healthcare providers, including private patients treated in NHS hospitals and NHS hospital patients treated elsewhere (Northern Ireland, Scotland, and Wales). Information has been collected since 1989 and includes >125 million inpatient, outpatient, and accident and emergency records each year [10]. Approximately 75% of English practices participate in the CPRD linkage scheme.

For this study, we used data collected during the 6-year period between July 2007 and June 2013, starting at a time since sensitive norovirus diagnostics were introduced in the United Kingdom until the introduction of the rotavirus vaccine in the national childhood immunization program in July 2013.

Ethical Considerations

Specific patient informed consent was not required for this study. The CPRD has been granted Multiple Research Ethics Committee approval (05/MRE04/87) for purely observational studies, with external data linkages, including HES APC data. The work of CPRD is also covered by National Information Governance Board-Ethics and Confidentiality Committee (ECC) approval ECC 5-05 (a) 2012.

Data linkage was carried out by NHS Digital, and anonymized data were provided to the study team following approval of study protocol (Independent Scientific Advisory Committee protocol 16_063R) by the Medicines and Healthcare Products Regulatory Agency Independent Scientific Advisory Committee.

Case Definition

Using International Classification of Diseases (ICD)-10 codes and Read codes, MA-AGE events were extracted from the linked dataset (see Supplementary Data). Read codes are the standard clinical terminology system used in general practices in the United Kingdom since 1985. Events were included when any ICD-10 or Read code associated with a diagnosis relating to gastroenteritis, gastrointestinal infection, or gastrointestinal pathogen was recorded, regardless of the diagnostic position of the indication. Medically attended gastroenteritis–related events in

the HES APC data were derived from patients' discharge records (the HES APC dataset does not distinguish between ICD-10 positions as primary cause of admission or secondary diagnosis), including primary and nonprimary reasons for admission and hospital-acquired infections. An MA-AGE episode was defined as at least 1 MA-AGE event with the lag time between a successive event not exceeding 14 days. Hence, a recurrence with a lag time of >14 days was considered a new MA-AGE episode. Any episode was only accounted for once and classified as primary care only or hospitalization based on the level of care required over the course of the episode: if an episode contained at least 1 MA-AGE event from HES APC Inpatient, it was classified as associated with hospitalization (and not counted under primary care even if the episode involved a primary care event).

Episodes were grouped into MA-AGE episodes unspecified cause (not otherwise specified [NOS]) or cause-specific episodes, with the latter being further categorized into norovirus, rotavirus, other viral cause, *Clostridium difficile*, other bacterial cause, or parasitic. When a specific virus or bacterium was not specified, the episode was classified as 'NOS'.

Statistical Methods

Observed MA-AGE incidence rates per 1000 person-years and exact Poisson 95% confidence intervals (Cis) were calculated for the total population, as well as separately by sex, age group, and time (year, month).

Because only a small proportion of MA-AGE episodes are cause-specified, the proportion of norovirus-attributable MA-AGE was estimated using an indirect modeling method developed and published for US datasets, with minor adaptations [7, 8]. Briefly, the model assumes that among the episodes with unspecified cause, the number due to a specified pathogen or pathogen group in a given month is proportional to the number of episodes actually coded for that same pathogen in the month. By using regression models assuming a negative binomial distribution to account for overdispersion, the monthly number of unspecified episodes was modeled as a function of cause-specified episodes. Separate models were fitted for each age group x and health care setting y, and an identity link function was used, which allows for an additive interpretation. Instead of assuming a constant population denominator as applied in the US studies, the denominator was adapted according to changes in population sizes over time.

Hence, the expected number of episodes "not otherwise specified" is modeled as

$$E(N\_NOS_{x,y,t}) = (\alpha_{x,y} \times pt_{x,t}) + (\beta1_{x,y} \times N\_CDIFF_{x,y,t}) + (\beta2_{x,y} \times N\_OTHER\ BACTERIAL_{x,y,t}) + (\beta3_{x,y} \times (N\_ROTA_{<5,y,t} / pt_{<5,t}) \times pt_{x,t}) + (\gamma_{x,y} \times t \times pt_{x,t}),$$

with N representing "counts" of AGE episodes (E) due to causes "not otherwise specified" (NOS), *C. difficile* (CDIFF), other bacterial infections (OTHER BACTERIAL) or rotavirus (ROTA), for age group x, health care setting y, and month t. α represents the background episode rate not explained by infections due to one of the specified pathogen categories; β

represents specific causes; and γ represents a secular linear time trend adding to the background rate. Person-time (in days) for age group x during month t is expressed as $pt_{x,t}$.

The interpretation of the background allows for other non-identified causes but also captures partly the nonseasonal component of the causes captured.

For rotavirus infections, information from children aged <5 year was used for all age groups due to lack of data in other age groups, multiplying the rate in the group of <5 years with the denominator in the respective age groups. The classifications "parasitic infections" and "other viral infections" were not included in the models due to insufficient data. Backward step-wise regression was used; variables that were not significant at the 5% level or those with negative coefficients were removed (as these are implausible) for each age group x and healthcare setting y, separately.

Because norovirus disease is rarely laboratory confirmed outside of institutional outbreaks, norovirus-confirmed episodes were not included in the equation. Instead, norovirus-attributable MA-AGE episodes were estimated for each month by calculating the difference between observed cause-unspecified and model-predicted counts (ie, the model residuals $r_{x,y,t} = N\_NOS_{x,y,t} - E(N\_NOS_{x,y,t})$, assuming that any remaining seasonality not otherwise captured in the model is due to norovirus). From these residuals the (negative) quantity $pt_{x,t} \times \min(r_{x,y,t}/pt_{x,t})$ was subtracted (while adding it to the background), assuming that there is a month in which there are no MA-NGE

episodes. The minimum represents the seasonal (July to June) minimum norovirus rate based on the raw residuals (ie, without subtracting). Estimated MA-NGE episodes modeled from non-specific episodes were then added to the few coded MA-NGE episodes to get the total number of MA-NGE episodes.

Confidence intervals were obtained by bootstrapping based on the Poisson distribution of means from the original data using 10 000 replicates. Because the residuals procedure led to a systematic overestimation of norovirus rates in the bootstrap replicates, the method was adapted by centering the distribution of the bootstrap estimates on the original estimate: 95% confidence intervals of the total rates were obtained from the 2.5 and 97.5 percentiles of the bootstrapped results minus the difference between the 50 percentiles of the bootstrapped results and the original results.

RESULTS

Demographics

The linked dataset contained records of 4 487 649 individuals, with 15.5% having a follow-up time of <1 year; 49.1% having a follow-up time of >1 but <6 years, and 35.4% having a complete follow-up time of 6 years. At the time of inclusion, 10.7% of the population was aged 0–4 years, 13.3% was aged 5–17 years, 62.3% was aged 18–64 years, and 13.8% was aged ≥65 years. There were similar proportions of males (49.3%) and females (50.7%) in the dataset.
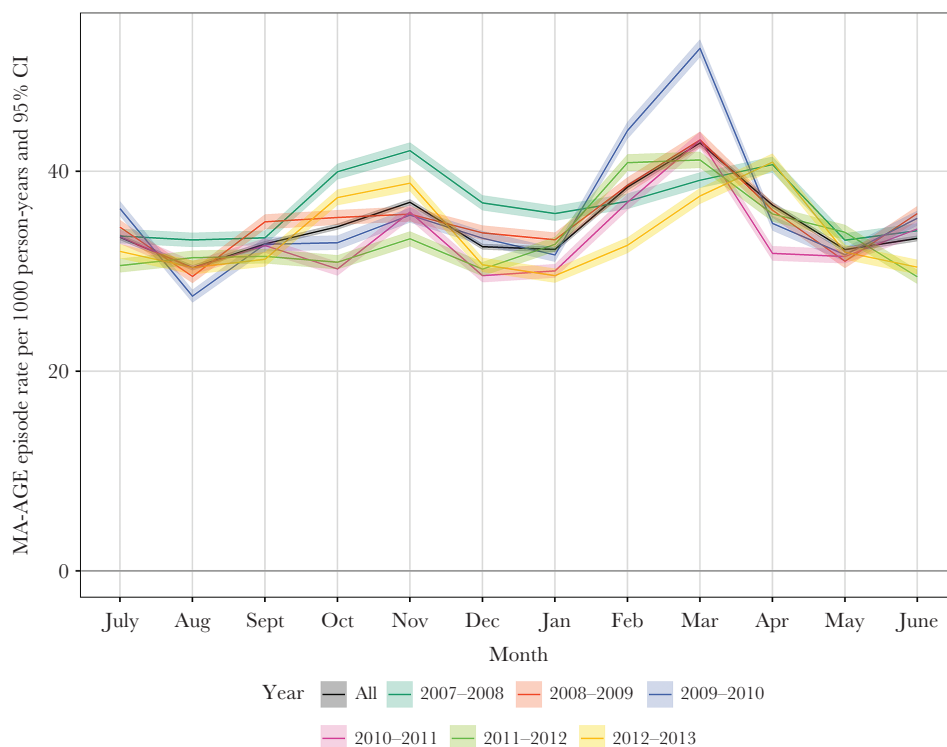
Figure 1. Observed incidence of medically attended gastroenteritis per calendar month in the period July 2007–June 2013 in England. Abbreviations: CI, confidence interval; MA-AGE, medically attended gastroenteritis.

Incidence of Medically Attended Gastroenteritis According to Health Care, Season, Age, Sex, and Specified Cause

A total of 592 696 MA-AGE episodes were recorded in the period July 2007–June 2013, corresponding to a mean annual rate of 34.62 episodes per 1000 person years (95% CI, 34.54–34.71). Although incidence rates varied year by year (Figure 1), a pattern of 2 seasonal peaks was observed: a smaller one in October–November and a larger one in January–April.

Most episodes (84%; n = 497 089 episodes) were associated with primary care consultations only. The mean annual incidence rate of MA-AGE associated with primary care was 29.04/1000 person-years (95% CI, 28.96–29.12), and the mean annual incidence rate of MA-AGE associated with hospitalization was 5.59/1000 person-years (95% CI, 5.55–5.62).

Children aged <5 years and adults aged ≥65 years had the highest rates of MA-AGE associated with primary care (age 0–4 years: 120.53/1000 person-years, 95% CI, 119.84–121.22; age ≥65: 37.31/1000 person-years, 95% CI, 37.09–37.53) or hospitalization (age 0–4 year: 13.40/1000 person-years, 95% CI, 13.17–13.63; age ≥65: 13.95/1000 person-years, 95% CI, 13.81–14.08) compared with age groups 5–17 years or 18–64 years (Figure 2). Overall the incidence of MA-AGE was higher for

females (38.01 /1000 person years; 95% CI, 37.88–38.15) than for males (31.19/1000 person years; 95% CI, 31.08–31.31), except in boys aged <18 years for whom the incidence of both primary care and hospitalization associated MA-AGE was higher than in girls.

A pathogen was specified for 17 825 (3%) of the MA-AGE episodes. The proportion of episodes with unknown infectious cause was higher for those associated with primary care (98.3%) than hospitalization (90.0%) (Table 1). *Clostridium difficile* was the most frequently identified pathogen in hospitalization-associated episodes, and bacterial pathogens other than *C. difficile* were the most frequently identified specified cause in primary care–associated episodes. Norovirus and rotavirus were more often identified in MA-AGE episodes associated with hospitalization than with primary care.

Seasonal peaks of pathogen-specific MA-AGE episodes were observed in all age groups, except for *C. difficile*–related episodes, which were identified continuously throughout the year and predominantly in older, hospitalized patients (Figure 3). Seasonal peaks of rotavirus-associated MA-AGE were observed in children aged <5 years, the only age group with a recommendation to routinely test for rotavirus [11]. Seasonal peaks of MA-NGE were observed in the group aged ≥65 years.
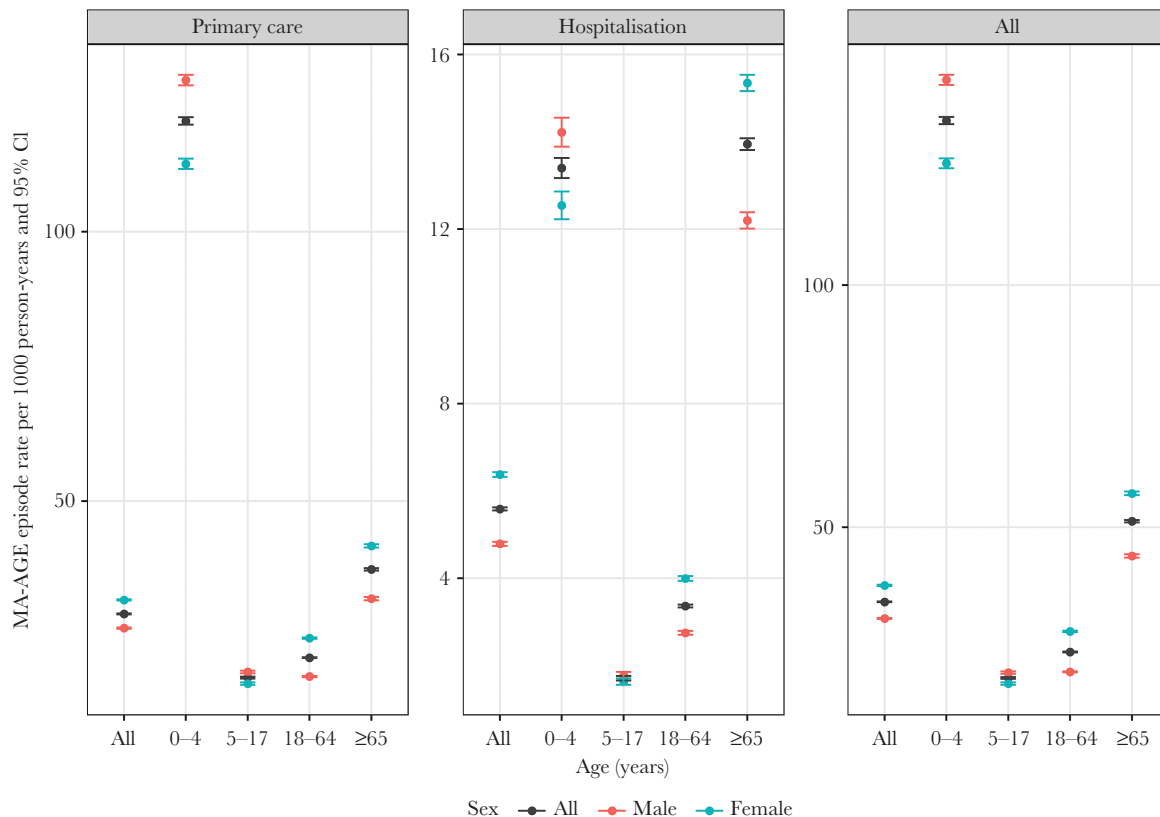
Figure 2. Observed incidence of medically attended gastroenteritis, overall or associated with primary health care or hospitalization, according to age groups and sex. Abbreviations: CI, confidence interval; MA-AGE, medically attended gastroenteritis.

Table 1. Pathogen-Specified and Unspecified Acute Gastroenteritis Episodes Associated With Primary Care or Hospitalization
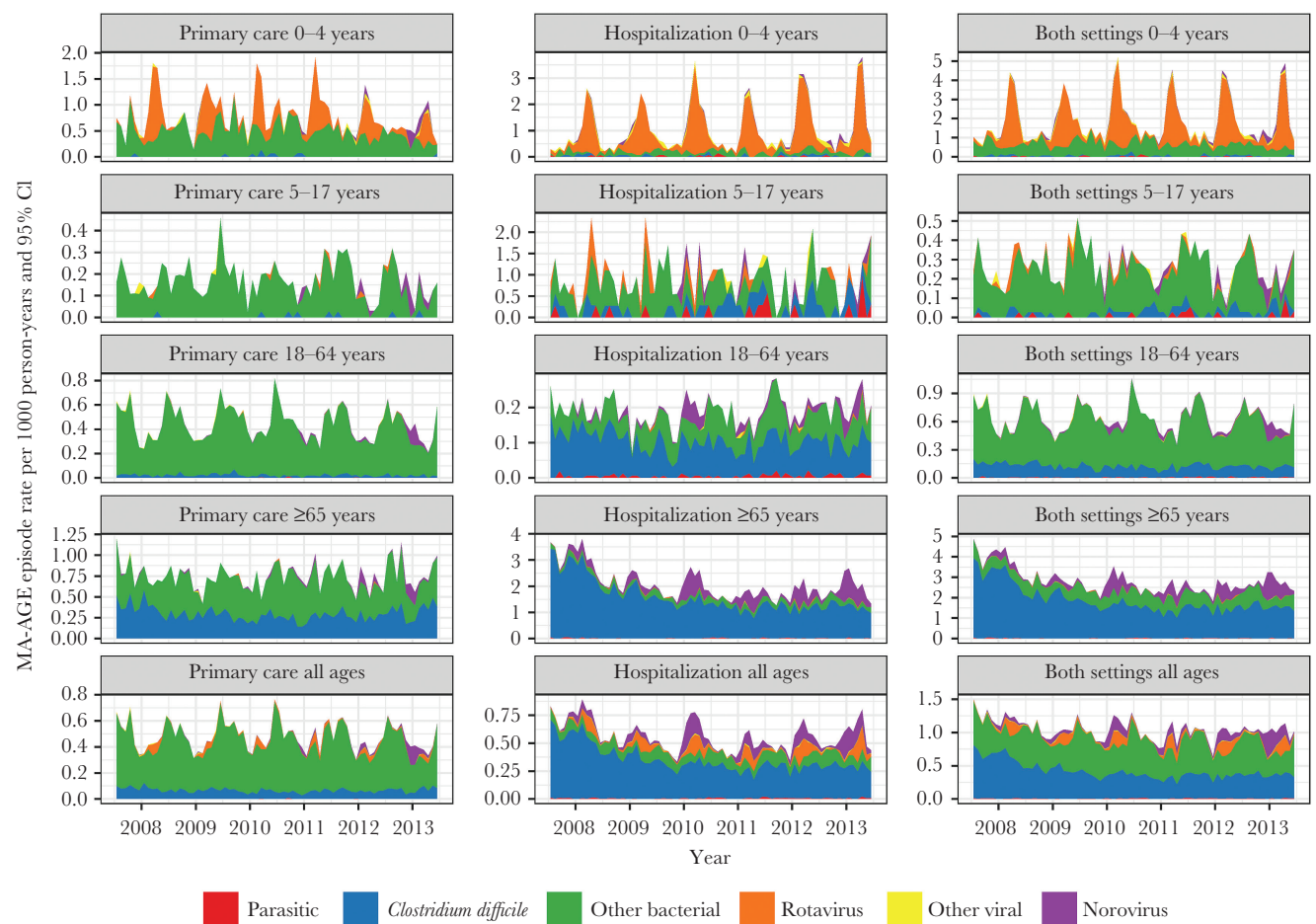
	Primary care			Hospitalization		
	No.	All episodes, %	Pathogen- specified episodes, %	No.	All episodes, %	Pathogen- specified episodes, %
Total episodes	497 089			95 607		
Cause unspecified	488 825	98.3 %		86 046	90.0%	
Norovirus	181	<0.1%	2.2%	1088	1.1%	11.3%
Rotavirus	279	<0.1%	3.4%	834	0.9%	8.6%
Other viral	38	<0.1%	0.5%	68	<0.1%	0.7%
Clostridium difficile	1177	0.2%	14.2%	6004	6.3%	62.1%
Other bacterial	6598	1.3%	79.7%	1546	1.6%	16.0%
Parasitic	4	<0.1%	<0.1%	121	0.1%	1.3%

A pathogen was specified for 8264 primary care–associated acute gastroenteritis (AGE) episodes, including 13 episodes for which 2 pathogens were identified, bringing the total number of identified pathogens to 8277. For hospitalization-associated AGE, a pathogen was specified for 9561 episodes, including 100 episodes with 2 identified pathogens, bringing the total number of identified pathogens to 9661.

Modelled Incidence Rates of Medically Attended Norovirus- and Rotavirus-Attributable Gastroenteritis

Norovirus and rotavirus were identified as the main causes of MA-AGE in all age groups, in addition to "bacterial pathogens other than *C. difficile*" in the group aged 17–64 years (Figure 4).

Medically attended rotavirus-attributable gastroenteritis (MA-RGE) occurred predominantly between January and May, and this was true for all age groups (Figure 5). Medically attended norovirus-attributable gastroenteritis occurred more continuously throughout the year, with indications of 2 peaks

Figure 3. Observed temporal incidence of medically attended gastroenteritis, overall or associated with primary health care or hospitalization, according to age groups. Abbreviations: CI, confidence interval; MA-AGE, medically attended gastroenteritis.
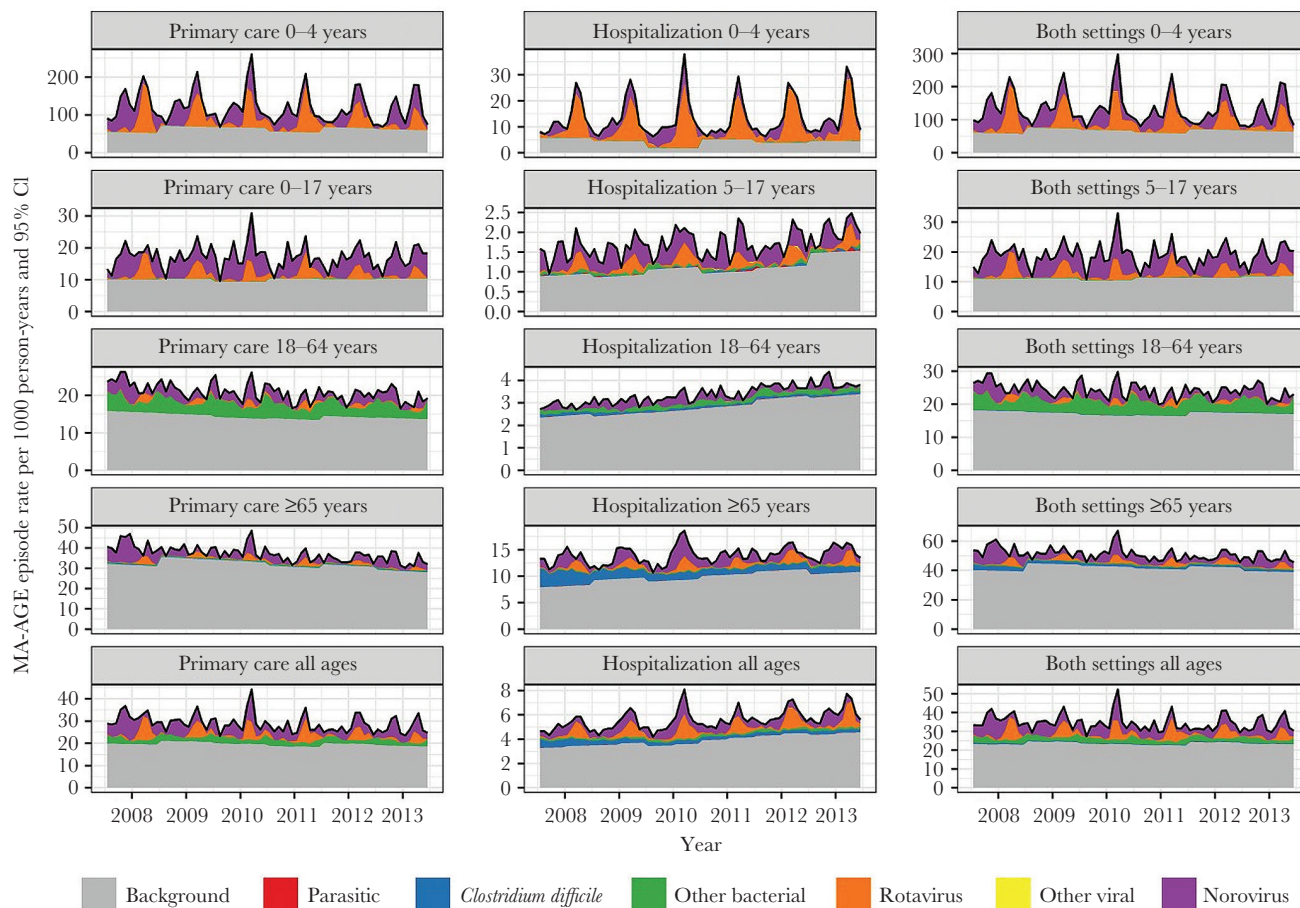
Figure 4. Estimated temporal incidence of medically attended gastroenteritis, overall or associated with primary health care or hospitalization, according to age groups. Abbreviations: CI, confidence interval; MA-AGE, medically attended gastroenteritis.

occurring in the younger age groups (0–4 y and 5–17 y) during October–December and February–May.

The mean annual incidence of MA-NGE associated with primary care was estimated at 4.93/1000 person-years (Table 2). The highest rate of MA-NGE was in children aged <5 years, with mean annual incidence rates of 34.0/1000 person-years for episodes associated with primary care and 3.3/1000 person-years for episodes associated with hospitalization. The incidence rate of MA-NGE associated with hospitalization was second highest in adults aged ≥65 years. Apart from episodes associated with hospitalization in children aged 0–4 years, MA-NGE incidence rates associated with primary care or hospitalization were higher than for MA-RGE.

DISCUSSION

Using linked data from the CPRD-GOLD and HES APC datasets, we were able to estimate age group–specific annual incidence rates of MA-NGE in England in the period July 2007–June 2013, finding rates as high as 34.0 consultations/1000 person-years and 3.3 hospitalizations/1000 person-years for norovirus-attributable gastroenteritis in children aged <5 years.

Previously, 2 prospective cohort studies have been conducted in the United Kingdom estimating the burden of IID in the community and associated general practice care, covering the periods 1993–1996 (IID1) [12] and 2008–2009 (IID2) [4]. These IID cohort data enabled us to compare the results from our study with a more conventional approach. We found a mean annual incidence rate of primary care–associated MA-AGE of 28.8/1000 person-years in England between July 2007 and June 2013. This is comparable to the 25.3/1000 person-years estimated from the IID2 population cohort for the period April 2008–August 2009 [4]. In both the IID2 and our study, incidence rates were the highest for children aged <5 years, corresponding to a mean rate of 120/1000 person-years in our study and 133/1000 person-years in the IID2 cohort study.

In our analysis of hospitalization-associated gastroenteritis, all discharge records that included a gastroenteritis indication were taken into account, irrespective of whether this was a primary or nonprimary reason for admission or a hospital-acquired infection. A similar approach was taken in an earlier study in the United States [7]. Compared with an incidence rate of 5.5/1000 person-years found in our study, the US study
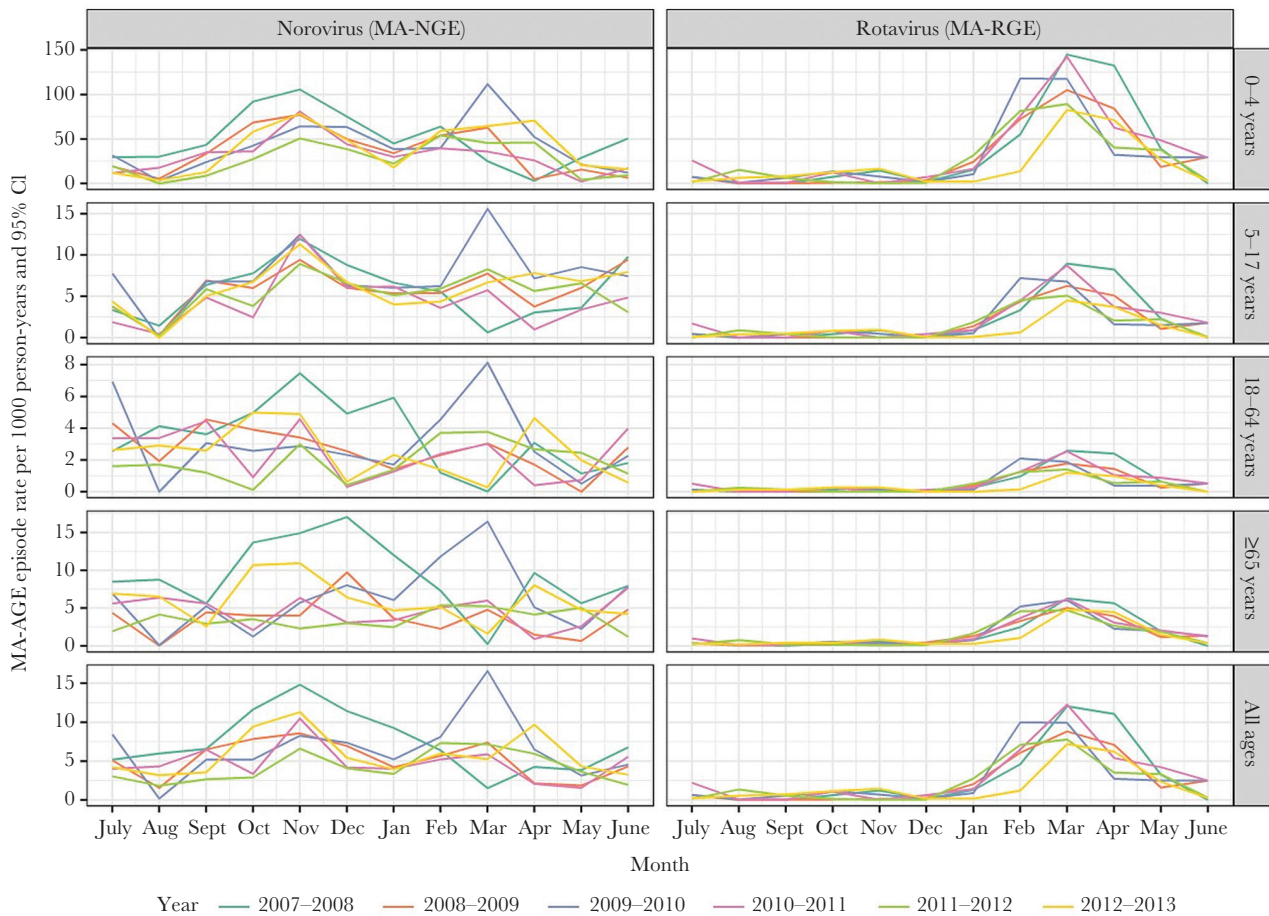
Figure 5. Estimated incidence of medically attended gastroenteritis attributable to norovirus or rotavirus per calendar month for each year in the period July 2007–June 2013 in England. Abbreviations: CI, confidence interval; MA-AGE, medically attended gastroenteritis; MA-NGE, medically attended norovirus-attributed gastroenteritis; MA-RGE, medically attended rotavirus-attributed gastroenteritis.

reported a rate of 2.5/1000 person-years for the period 1996–2007 [7]. A rise in incidence rates was observed in the US study during the study period; it is therefore possible that after 2007 rates may have been higher and approaching the rate we found in England during 2007–2013. Together these studies confirm that routinely collected hospital discharge records are useful

resources to assess the incidence of hospitalization-associated gastroenteritis.

The overall and age-specific rates of primary care-associated MA-NGE found in our models were comparable with rates reported in the US between 2001 and 2009 [8] but were nearly double of what was reported for the IID2 study [13]. However, the

Table 2. Modeled Mean Annual Incidence Rate of Medically Attended Gastroenteritis, Norovirus-Attributed Gastroenteritis, and Rotavirus-Attributable Gastroenteritis According to Medical Care Category and Age Groups

	MA-AGE		MA-NGE		MA-RGE	
	Primary care	Hospitalization	Primary care	Hospitalization	Primary care	Hospitalization
	Incidence per 1000 person-years (95% CI)		Incidence per 1000 person-years (95% CI)		Incidence per 1000 person-years (95% CI)	
Age group						
0–4 y	120.53 (119.85–121.21)	13.41 (13.18–13.64)	34.00 (27.46–42.60)	3.30 (2.41–4.44)	23.79 (21.56–25.85)	5.60 (5.29–5.92)
5–17 y	17.22 (17.06–17.39)	1.71 (1.66–1.76)	5.43 (4.90–6.06)	0.40 (0.30–0.50)	1.55 (1.31–1.78)	0.14 (.10–.19)
18–64 y	20.93 (20.84–21.01)	3.36 (3.33–3.40)	2.42 (2.10–2.76)	0.26 (0.20–0.33)	0.48 (.10–0.64)	0.0008 (.0003–.0013)
≥65 y	37.31 (37.09–37.52)	13.97 (13.84–14.10)	3.87 (3.26–4.49)	1.69 (1.41–2.00)	0.80 (.09–1.08)	0.70 (.57–.83)
All	29.04 (28.96–29.12)	5.59 (5.56–5.63)	4.93 (4.45–5.49)	0.71 (.62–0.80)	2.03 (1.63–2.27)	0.47 (.44–.50)

Abbreviations: MA-AGE, medically attended gastroenteritis; MA-NGE, medically attended norovirus-attributed gastroenteritis; MA-RGE, medically attended rotavirus-attributed gastroenteritis.

population burden of norovirus infection in the IID2 study was recently reported to be around 26% higher than reported originally when using a less stringent cutoff for the norovirus diagnostic polymerase chain reaction that included asymptomatic carriage and subclinical disease [14]. It is not clear whether the same difference could apply to the estimation of primary-care associated MA-NGE in the IID2 study, and this might more closely resemble the findings of the modeling approach. Prospective cohort studies are associated with a number of difficulties that could result in under-reporting, as well as relatively small sample sizes [13]. Alternatively, our model may have overestimated MA-NGE rates by falsely attributing cases to norovirus. Because we did not have access to full medical charts, we could not validate the specificity of the codes. However, we expect that the modeling approach filters out most noninfectious cases of gastroenteritis.

In our model, children aged <5 years had the highest estimated rates for primary care–associated and hospitalization-associated MA-NGE. The same was found for primary care–associated MA-NGE in the IID2 study [13]. Adults aged ≥65 years had the second highest estimated rate of hospitalization-associated MA-NGE in our model. To our knowledge, only 1 other study has reported on hospitalization-associated MA-NGE in older adults in the United Kingdom; however, reported rates were considerably lower (10–43/100 000 person-years in the period 2000–2006 [15]) than in our study (169/100 000 person-years), most likely because the earlier study only took into account emergency admissions with gastroenteritis as a primary diagnosis.

When stratifying for sex, the incidence of MA-AGE was found to be 30%–40% higher in adult females than males, but 10%–15% higher in boys than girls aged <18 years. The same was true for MA-RGE, but not for MA-NGE, for which rates were higher in females across all age groups (data not shown). Thus the observed sex effect could be partly explained by a generally higher healthcare usage of adult women than men [16], as well as a sex difference in susceptibility to specific pathogens [17].

Our models showed that, although MA-RGE predominantly occurred in children aged <5 years, MA-NGE was found across all age groups. Childhood rotavirus vaccination was introduced in the United Kingdom in July 2013, with early reports finding a rapid decline in rotavirus infections since then [18–20]. It can therefore be expected that after July 2013 relatively more hospitalization-associated gastroenteritis in the group aged <5 years will be attributable to norovirus than rotavirus. Indeed, a recent study reported that in the period 2014–2015, 44% of patients in a UK pediatric hospital positive for a gastrointestinal virus were positive for norovirus [21]. Our modeling approach, which used linked primary care and hospitalization datasets to estimate temporal age group–specific MA-NGE and MA-RGE incidence rates, would be a very suitable tool to assess the impact of rotavirus vaccine implementation and reassess the incidence of MA-NGE in future studies.

An important finding from our study is that linked routinely collected primary care and hospitalization datasets can be used to study temporal and subgroup-specific incidence rates of MA-AGE and, more important, can be used to estimate the burden attributable to main causative pathogens not routinely tested for, such as norovirus. This approach has a number of advantages above the more conventional prospective cohort studies. First, the analysis of population-based healthcare datasets can be repeated frequently for relatively low costs and provide updated "real-time" and temporal data. This can be particularly relevant to measure or predict the impact of prevention and intervention programs. By linking primary care data with hospitalization data, we were able to stratify by the type of healthcare provided, which generates information relevant when studying the economic burden of disease and health interventions. Furthermore, due to the large size of the population included in these datasets, incidence rates can be assessed for subgroups of specific interest—for example, age groups as done here, but ultimately also for risk groups with specific underlying conditions. The use of routinely collected electronic health data for infectious disease burden assessment will be limited by the testing performed as routine. We overcame this limitation by modeling the seasonality of the pathogen of interest. To apply this modeling approach, seasonality is needed combined with a minimal level of testing for the pathogen of interest and for the other pathogens that are also seasonal. If testing information is lacking for only 1 pathogen, as was the case in our dataset, the residuals approach can be used.

In summary, this is the first study using the UK CPRD-GOLD primary care research dataset linked with England's Hospital Episode Statistics inpatient dataset to estimate the incidence rates of MA-AGE and MA-NGE in England, thereby differentiating between primary care and hospitalizations. Although the infectious cause of MA-AGE is rarely specified in clinical records, we demonstrate here that incidence rates of MA-NGE can be estimated using routinely collected primary care and hospitalization datasets. Datasets such as the CRPD can thus provide a useful resource to estimate and monitor the burden of MA-NGE in a population and provide baseline data for considering the potential health and economic benefit of new interventions such as vaccines and potentially to measure the impact of their introduction.

Supplementary Data

Supplementary materials are available at *The Journal of Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

Notes

References

1. Lopman BA, Steele D, Kirkwood CD, Parashar UD. The vast and varied global burden of norovirus: prospects for prevention and control. PLoS Med **2016**; 13:e1001999.

2. Bartsch SM, Lopman BA, Ozawa S, Hall AJ, Lee BY. Global economic burden of norovirus gastroenteritis. PLoS One **2016**; 11:e0151219.

3. Payne DC, Vinjé J, Szilagyi PG, et al. Norovirus and medically attended gastroenteritis in U.S. children. N Engl J Med **2013**; 368:1121–30.

4. Tam CC, Rodrigues LC, Viviani L, et al; IID2 Study Executive Committee. Longitudinal study of infectious intestinal disease in the UK (IID2 study): incidence in the community and presenting to general practice. Gut **2012**; 61:69–77.

5. Tam CC, O'Brien SJ. Economic cost of campylobacter, norovirus and rotavirus disease in the United Kingdom. PLoS One **2016**; 11:e0138526.

6. Karst SM, Baric RS. What is the reservoir of emergent human norovirus strains? J Virol **2015**; 89:5756–9.

7. Lopman BA, Hall AJ, Curns AT, Parashar UD. Increasing rates of gastroenteritis hospital discharges in US adults and the contribution of norovirus, 1996–2007. Clin Infect Dis **2011**; 52:466–74.

8. Gastanaduy PA, Hall AJ, Curns AT, Parashar UD, Lopman BA. Burden of norovirus gastroenteritis in the ambulatory setting—United States, 2001–2009. J Infect Dis **2013**; 207:1058–65.

9. Herrett E, Gallagher AM, Bhaskaran K, et al. Data resource profile: Clinical Practice Research Datalink (CPRD). Int J Epidemiol **2015**; 44:827–36.

10. Sinha S, Peach G, Poloniecki JD, Thompson MM, Holt PJ. Studies using English administrative data (Hospital Episode Statistics) to assess health-care outcomes–systematic review and recommendations for reporting. Eur J Public Health **2013**; 23:86–92.

11. Atchison CJ, Lopman BA, Harris CJ, Tam CC, Iturriza Gomara M, Gray JJ. Clinical laboratory practices for the detection of rotavirus in England and Wales: can surveillance based on routine laboratory testing data be used to evaluate the impact of vaccination? Euro Surveill **2009**; 14(20):pii=19217.

12. Wheeler JG, Sethi D, Cowden JM, 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. BMJ **1999**; 318:1046–50.

13. O'Brien SJ, Donaldson AL, Iturriza-Gomara M, Tam CC. Age-specific incidence rates for norovirus in the community and presenting to primary healthcare facilities in the United Kingdom. J Infect Dis **2016**; 213(suppl 1):S15–8.

14. Harris JP, Iturriza-Gomara M, O'Brien SJ. Re-assessing the total burden of norovirus circulating in the United Kingdom population. Vaccine **2017**; 35:853–5.

15. Haustein T, Harris JP, Pebody R, Lopman BA. Hospital admissions due to norovirus in adult and elderly patients in England. Clin Infect Dis **2009**; 49:1890–2.

16. Hippisley-Cox J, Vinogradova Y, The NHS Information Centre. Trends in consultation rates in general practice 1995 to 2008: analysis of the QResearch database. http://content.digital.nhs.uk/catalogue/PUB01077/tren-cons-rate-gene-prac-95-09-95-08-rep.pdf. Accessed 20 June 2017.

17. Ruggieri A, Anticoli S, D'Ambrosio A, Giordani L, Viora M. The influence of sex and gender on immunity, infection and vaccination. Ann Ist Super Sanita **2016**; 52:198–204.

18. Atchison CJ, Stowe J, Andrews N, et al. Rapid declines in age group-specific rotavirus infection and acute gastroenteritis among vaccinated and unvaccinated individuals within 1 year of rotavirus vaccine introduction in England and Wales. J Infect Dis **2016**; 213:243–9.

19. Inns T, Trindall A, Dunling-Hall S, Shankar AG. Introduction of a new rotavirus vaccine: initial results of uptake and impact on laboratory confirmed cases in Anglia and Essex, United Kingdom, July 2015. Hum Vaccin Immunother **2016**; 12:1040–4.

20. Hungerford D, Read JM, Cooke RP, et al. Early impact of rotavirus vaccination in a large pediatric hospital in the UK. J Hosp Infect **2016**; 93:117–20.

21. Brown JR, Shah D, Breuer J. Viral gastrointestinal infections and norovirus genotypes in a paediatric UK hospital, 2014–2015. J Clin Virol **2016**; 84:1–6.