Frequency of five *Escherichia Coli* pathotypes in Iranian adults and children with acute diarrhea

Running title

Frequency of *E. coli* pathotypes in Iran

Authors

Sana Eybpoosh¹ ¶, Saeid Mostaan² ¶, Mohammad Mehdi Gouya³, Hossein Masoumi-Asl³, Parviz Owlia⁴, Babak Eshrati⁵, Mohammad Reza Montazer Razavi Khorasan³, Saeid Bouzari^{2*}

1 Department of Epidemiology and Biostatistics, Research Centre for Emerging and Reemerging infectious diseases, Pasteur Institute of Iran, Tehran, Iran

2 Department of Molecular Biology, Pasteur Institute of Iran, Tehran, Iran

3 Centre for Communicable Diseases Control, Ministry of Health and Medical Education, Tehran, Iran.

4 Molecular Microbiology Research Center, Shahed University, Tehran, Iran

5 Center For Preventive Medicine, Department of Social Medicine, Iran University of Medical Sciences, Tehran, Iran

*** Corresponding author**

E-mail address: bouzari@pasteur.ac.ir (SB)

¶ These authors contributed equally to this work.

Abstract

Background: knowledge about the distribution of *Escherichia Coli* (*E. coli*) pathotypes in Iran is limited to studies with small scale and limited scope. This nation-wide survey aims to provide a more generalizable estimate of pathogenic *E. coli* distribution in Iran.

Methods: During January 2013 and January 2014, stool samples were collected from 1306 acute diarrhea cases of 15 provinces. Culture-positive *E. coli* samples were tested with PCR for detection of five *E. coli* pathotypes (STEC, ETEC, EPEC, EAEC, and EIEC). Frequency of these pathotypes was estimated for different provinces, age groups, and months/seasons.

Results: Of 1305 diarrheal samples, 979 were *E. coli*-positive (prevalence: 75.0%; 95% CI: 72.6, 77.3%). Pathogenic *E. coli* was detected in 659 diarrheal samples (prevalence: 50.5%; 95% CI: 47.8, 53.2%). STEC and EIEC was the most and the least frequent pathotypes (35.4% and 0.3%, respectively). ETEC (14.0%) and EPEC (13.1%) were the second and the third frequent pathotypes, respectively. EAEC was not highly prevalent (4.3%). Fars and Razavi Khorasan provinces had the highest and lowest frequencies (88.7% and 34.8%, respectively). *E. coli* pathotypes were more frequent in warmer (i.e., spring and summer) than cooler (i.e., fall and winter) seasons. The highest frequency of pathogenic *E. coli* was observed in infants and children under 5 years (73% each). There was no association between sex and pathogenic *E. coli* infection.

Conclusions: Diarrheagenic *E. coli* may be an important cause of acute diarrhea in adults and children in Iran. STEC and ETEC seem to be widespread and show a peak in warmer seasons. This finding could impact the recommended use of STEC and ETEC vaccines during warmer

seasons, especially for infants, young children and elderlies. Monitoring the rate of diarrheagenic *E. coli* infection, *E. coli* serotypes, and their antibiotic resistance is recommended for evaluations of time-trends and effectiveness of interventions.

Keywords: *Escherichia coli*, Bacterial Typing, Diarrhea, Epidemiology, Iran

Author summary

Escherichia coli, also known as *E. coli* is a bacterium of the genus *Escherichia* that is normally found in the lower intestine of human. Most *E. coli* strains are harmless, but some can cause infection in the gastrointestinal tract, causing diarrhea. These pathogenic *E. coli* strains are classified based on their mechanism of pathogenesis. In this regard, five important *E. coli* strains include Shiga toxin-producing *E. coli* (STEC), enterotoxigenic *E. coli* (ETEC), enteropathogenic *E. coli* (EPEC), enteroaggregative *E. coli* (EAEC), and enteroinvasive *E. coli* (EIEC). In a national survey conducted in Jan 2013 till Jan 2014, we collected 1305 diarrheal samples from 15 (out of 31) provinces of Iran. Of these, 979 samples (75%) were *E. coli*-positive in the culture test. Molecular tests showed that 659 samples were pathogenic *E. coli*, suggesting that 50.5% of the diarrhea cases were induced due to pathogenic *E. coli* infection. The most prevalent pathogenic *E. coli* strains in Iran were STEC (35.4%) and ETEC (0.3%), and were more commonly detected in warmer seasons, infants, and children less than five years. So, the use of vaccines, especially for STEC and ETEC, during warmer seasons and for infants, young children and elderlies are recommended.

Introduction

Diarrheal diseases are among major cause of illness and death, especially among children (1). Deaths are mainly caused by loss of water and electrolytes resulting from intestinal malabsorption or increased secretion (2). Although the condition is considered as a major health concern worldwide, its prevalence and adverse effects are more prominent in lower income and less developed countries. Projections have estimated that by 2030, diarrheal diseases remain as one the top 10 causes of death in low income countries (1).

Escherichia coli (*E. coli*) is a gram-negative bacterium found naturally in the intestinal tract of human. A subset of *E. coli* is capable of causing diarrheal disease (known as diarrheagenic *E. Coli* [DEC]), with some of its strains leading to severe diarrhea and even death. Different DEC strains (known as pathotypes) have been determined relatively recently, including Shiga toxin-producing *E. coli* (STEC), enteropathogenic *E. coli* (EPEC), enterohemorrhagic *E. coli* (EHEC), enterotoxigenic *E. coli* (ETEC), enteroaggregative *E. coli* (EAEC), enteroinvasive *E. coli* (EIEC) and diffusely adherent *E. coli* (DAEC) (3).

The main virulence factor of STEC is a phage-encoded potent cytotoxin. The cell toxicity effect of this pathotype was also demonstrated on Vero cells, resulting in a parallel nomenclature of Shiga/Vero toxin-producing *E.coli* (STEC) and VTEC, respectively (4). ETEC is characterized by its ability to elaborate the heat-stable and/or heat-labile enterotoxins, and is noted to be the major cause of travelers' diarrhea globally (5). Like many strains of STEC, EPEC harbors the *Eae* gene, encoding the outer membrane protein "intimin". This enables both pathotypes to attach epithelial cells of the intestinal tract (6). STEC are distinguished from EPEC strains by harboring

the Shiga toxin–encoding gene. While EPEC strains lack chromosomal Shiga toxin-encoding genes, they harbor the EPEC adherence factor (EAF) virulence plasmid, which encodes the *Bfp* fimbriae. EAEC is recently defined as a DEC, and is increasingly reported in lower income countries (7). Recent identification of virulence factors under the control of the AggR regulator have categorized EAEC among diarrheal pathogens (8). EIEC are closely related to *Shigella* species, in terms of their biochemical and physiological properties and their genetic structure. Both *Shigella* and EIEC encode common putative virulence genes and secret toxins that allow these pathogens to invade intestinal tract mucosa, move within epithelial cells, and penetrate neighboring cells (9, 10).

Existing evidence suggest a location-specific distribution of *E. coli* pathotypes. For example, some of the serotypes of STEC, including O157:H7, are well recognized in the United States and Canada as the major cause of hemolytic uremic syndrome. The non-O157:H7, however, is prevalent in Latin America, Australia, and Europe. ETEC is characterized as a major cause of travelers' diarrhea, in almost all parts of the world (5). In Iran, very limited data is available about the existence and distribution of DEC in human. This may be the result of the paucity of large-scale and well-designed systematic epidemiological studies, and the absence of a surveillance system (11). For example, many of the available studies have focused on infants and children under 5 years of age (12-17). The very few studies conducted on adolescents and adults are limited to small samples, few geographical areas, and limited *E. Coli* pathotypes (18-21). The sampling procedure in all studies available from Iran is non-probabilistic. The sampling location of most available studies is also limited to one city. These factors limit the generalizability of the findings of existing studies.

Given current knowledge gap, we performed this nationally-representative survey to elucidate the role of five *E. Coli* pathotypes in the epidemiology of diarrheal diseases, and to provide more recent estimates of the prevalence of DEC pathotypes among adults and children of Iran. It is hoped that the results of this survey be used for policy making for diarrhea management in Iranian population and provide a basic insight for vaccine design.

Material and Method

Ethics statement

This study was approved by the ethics committee in Pasteur Institute of Iran. All participants were informed about study objectives, assured about the confidentiality of their information, and gave their written informed consent to be included. For collecting sample from children and infants with diarrhea, a parent or guardian of any child participant provided informed consent on their behalf.

Study design

A cross-sectional survey was conducted from January1, 2013 to January 30, 2014 to determine the prevalence and populations of *E. coli* in cases with acute diarrhea.

Study population

The study population consisted of all Iranian nationals who were residents of target provinces of Iran and referred to the primary health centers with a chief complain of acute diarrhea (with or without bleeding). Acute diarrhea was defined as the acute onset of liquid and watery bowel movements more than 2-3 times a day. In order to identify the real distribution of *E. coli* in

selected provinces, we excluded patients who reported to travel out of the province in the past two weeks. We also excluded diarrhea cases that were under chemotherapy, antibiotic therapy, or corticosteroid therapy as well as all immunologically compromised patients. Patients with chronic diarrhea were not included either.

Sampling procedure

Of 31 provinces of Iran, 15 provinces were selected (sampling fraction= 50%). Provinces were sampled in a way to provide an even distribution of all geographic areas across the country. In each of the provinces three cities were selected. This resulted in the selection of 45 cities across the country. In each city, stool sample from diarrhea cases were collected at the second half of each month, from which a random sample were cultured for *E. coli* detection. *E. coli*-positive samples identified in culture method were then sent to the National *E. Coli* Reference Laboratory (NECRL) in Pasteur Institute of Iran (PII) for molecular analyses. Collecting samples from diarrhea cases that were occurred at the second half of each month was to homogenize the sampling time in all cities. Finally, a total of 1305 diarrhea samples were collected over a period of one year.

To ensure data quality and standardize data collection and classification in all field centers, we held a series of workshops in which participating staffs were thoroughly trained about the procedures under their responsibility. Trainings involved how to collect stool samples, do the cultures, interpret and record culture results, and transfer the samples to NECRL.

Isolation and Identification of *E. coli*

Culture

Stool samples were cultured in the laboratories of the field centers. Samples that were confirmed as *E. coli*-positive, were transferred to NECRL for molecular analysis. For the culture test, stool samples were inoculated on MacConkey agar (Merck, Catalog No. 105465) at 37 °C for 24 hours. After O/N incubation (incubator: Vision Korea), results were checked, and five typical colonies on MacConkey agar (pink color) were selected and transferred to triple sugar iron agar media (Merck, Catalog No. 104728). Strains with *E. coli* characteristics were selected and transferred to SIM medium (Merck, Catalog No. 105470), Simmons citrate (Merck, Catalog No.102501) and MRVP broth (Merck, Catalog No. 105712) to check for the presence of *E. coli* characteristics. Indol-positive (Kovacs, Merck, Catalog No. 109293), MR-positive (Methyl red, Merck, Catalog No. 106076), VP-negative (KOH, SIGMA USA, Catalog No. P5958 and alpha naphthol-1 Catalog No. 70480) and citrate-negative samples were identified as *E. coli*. Identified *E. coli*-positive samples were incubated on LB (Merck, Catalog No. 110285). After O/N incubation, samples were centrifuged (Eppendorf Germany) and the final product was stored at –70 °C until used (22).

PCR

Eight *E. coli* virulence genes indicative of five *E. coli* pathotypes were selected and checked for using PCR test (Table 1). These included genes coding toxins (Shiga toxins [*Stxs*] *Stx1* and *Stx2*, heat-labile toxin [*LT*], and human heat-stable toxin [*ST*]), adhesion factors (*Eae* and *Bfp*), and invasion genes (*invE*). The primers used for probe amplification were either chosen from

existing literature on virulence gene detection or designed from available gene sequences. Target genes in this study, were used to identify five *E. coli* pathotypes, including: (i) STEC, identified by *Eae*, as well as *Stx1* or *Stx2* or both; (ii) ETEC, identified by *LT* and *ST* genes; (iii) EPEC, identified by *Eae* and *Bfp* genes; (iv) EAEC, identified by *AA* gene; and (v) EIEC, identified by *invE* gene.

Table 1. Target genes and their characteristics for isolation of different pathotypes

Pathotype	Target Gene(s)	Gene Size (kb)	Primers	
EPEC	<i>Eae</i>	544	Primer (Forward)	CTGAACGGCGATTACGCGAA
			Primer (Reverse)	CGAGACGATACGATCCAG
	<i>Bfp</i>	910	Primer (Forward)	GACACCTCATTGCTGAAGTCG
			Primer (Reverse)	CCAGAACACCTCCGTTATGC
ETEC	<i>Lt</i>	655	Primer (Forward)	GAACAGGAGGTTTCTGCGTTAGGTG
			Primer (Reverse)	CTTCAATGGCTTTTTTTTGGGAGTC
	<i>St</i>	157	Primer (Forward)	CCTCTTTTAGCCAGACAGCTGAATCACTTG
			Primer (Reverse)	CAGGCAGGATTACAACAAAGTTCACAG
STEC	<i>Eae</i>	544	Primer (Forward)	CAGGCAGGATTACAACAAAGTTCACAG
			Primer (Reverse)	CAGGCAGGATTACAACAAAGTTCACAG
	<i>Stx1</i>	244	Primer (Forward)	CGATGTTACGTTTGTACTGTGACAGC
			Primer (Reverse)	AATGCCACGCTTCCCAGAATTG
	<i>Stx2</i>	324	Primer (Forward)	GTTTTGACCATCTTCGTCTGATTATTGAG
			Primer (Reverse)	AGCGTAAGGCTTCTGCTGTGAC
EAEC	<i>AA</i>	629	Primer (Forward)	CTGGCGAAAGACTGTATCAT
			Primer (Reverse)	CAATGTATAGAAATCCGCTGTT
EIEC	<i>invE</i>	766	Primer (Forward)	CGATAGATGGCGAGAAATTATATCCCG
			Primer (Reverse)	CGATCAAGAATCCCTAACAGAAGAATCAC

For the preparation of samples, 10 μ l Master Mix 2X (Fermentas, Catalog No. K0171) plus 7 μ l of DDW and forward and reverse primers (1 μ l of each) was added to 1 μ l of sample. For positive and negative controls, 1 and kb DNA ladder and ladder mix were used. The processes of denaturation, annealing and extension were performed in the Eppendorf thermo cycler (Germany).

Statistical analysis

Data were collected, revised, and analyzed by Stata software (version 14). Frequency of *E. coli* pathotypes was estimated in the overall population and in subgroups of age, location (province/city), and time (season/month). Frequency of *E. coli* pathogenic strains and target *E. coli* genes were estimated as the number of pathogenic *E. coli* samples identified by PCR tests to the total number of diarrhea samples with positive culture results for *E. coli* (presented as number and percent). Ninety-five percent confidence intervals for frequency measures were calculated with Wilson Score interval method. We grouped seasons into warmer (spring and summer) and cooler (fall and winter) seasons, between which the frequency of each *E. coli* pathotype was compared using the Chi Square test. Association of each *E. coli* pathotype with the type of diarrhea (with/without bleeding) was also assessed using the Chi Square test. Statistical tests were considered as significant at 0.05 levels.

Results

In this study 1305 diarrheal samples were collected from 15 out of 31 provinces of Iran between January 2013 and January 2014. Of these, a total of 979 samples were *E. coli*-positive (75.0%; 95% CI: 72.6, 77.3%), and were subjected to molecular assays. Pathogenic *E. coli* strains were detected in 659 out of 1305 diarrheal samples, resulting in a prevalence of 50.5% (95% CI: 47.8, 53.2%).

STEC was the most and EIEC was the least frequent pathotypes among *E. coli*-positive diarrhea samples (frequency: 35.4% and 0.3%, respectively). The most frequent virulence genes identified included *stx1* and *Eae* with a frequency of 26.1% and 25.9%, respectively. The lowest

197 frequent virulence genes included *bfp* and *invE*, each having a frequency of less than 1% (Table
198 2).

199

Table 2. prevalence of five *E. coli* pathotypes and their virulence genes in Iran

Pathotype	n	Prevalence in diarrheal samples*		Frequency in <i>E. coli</i> -positive samples**	
		% (95% CI)		% (95% CI)	
STEC	347	26.6	(24.2, 29.1)	35.4	(32.4, 38.5)
ETEC	137	10.5	(8.9, 12.3)	14.0	(11.9, 16.3)
EPEC	129	9.9	(8.3, 11.6)	13.2	(11.1, 15.5)
EAEC	43	3.3	(2.4, 4.4)	4.4	(3.2, 5.9)
EIEC	3	0.2	(0.1, 0.7)	0.3	(0.001, 0.9)
Total	659	50.5	(47.7, 53.2)	67.3	(64.3, 70.3)
Virulence genes					
Stx1	255	19.5	(17.4, 21.8)	26.1	(23.3, 28.9)
Eae	254	19.5	(17.3, 21.7)	25.9	(23.2, 28.8)
Stx2	201	15.4	(13.5, 17.5)	20.5	(18.0, 23.2)
LT	80	6.1	(4.9, 07.6)	8.2	(6.5, 10.1)
ST	78	6.0	(4.8, 7.4)	8.0	(6.5, 9.8)
AA	43	3.3	(2.4, 4.4)	4.4	(3.2, 5.9)
Bfp	5	0.4	(0.1, 0.9)	0.5	(0.2, 1.2)
invE	4	0.3	(0.1, 0.8)	0.4	(0.1, 1.0)

* Prevalence is calculated by dividing the numbers to the total number of diarrheal cases sampled (n= 1305). **

Frequency is calculated by dividing the numbers to the total number of *E. coli*-positive samples identified in culture (n= 979)

Geographical distribution of pathogenic *E. coli* in Iran

Pathogenic *E. coli* were detected in all investigated provinces. In this regard, frequency of *E. coli* pathotypes was highest in Fars province (S), where 88.7% of received *E. coli* samples were pathogenic (95% CI: 78.5, 94.4%). On the other hand, Razavi Khorasan Province (NE) had the lowest frequency among investigated provinces, although the prevalence in this province was still considerably high (34.8%; 95% CI: 35.5, 71.2%; Fig 1, Panel A, and S1 File, Table S1). At the city level, 100% of the received samples from Gorgan (NE; with 3 received sample), Tarem (WN; with 3 received sample), and Abadeh (S; with 9 received sample) were pathogenic *E. coli*. Pathogenic *E. coli* were not detected in three out of forty-three cities. These cities included Saravan (SE; with 2 received samples), Damghan (CN; with 1 received sample), and Najaf Abad

(C; with 5 samples; S2 File). Geographic distribution of five *E. coli* pathotypes is described below.

Fig 1. Frequency of *E. coli* pathotypes in 979 *E. coli*-positive culture samples collected in 15 provinces of Iran

STEC: was detected in all investigated provinces (Fig 1, Panel B). With an estimated frequency of 35.4% (95% CI: 32.4, 38.5%), STEC was identified as the most frequent pathotype. STEC constituted 52.7% of all identified pathogenic *E. coli* (S1 File, Table S2). The highest frequency of STEC was seen in Fars province (S; 72.5%). On the other hand, Esfahan had the lowest frequency of STEC (C; 5.5%; Fig 1, Panel B and S1 File, Table S2). At the city-level, STEC was detected in 33 out of 43 sampled cities (76.7%). In the majority of these cities (n= 21; 63.6%), STEC was also the predominant pathotype (Fig 2 panel A, and S2 File).

Fig 2. Distribution of five pathogenic *E. coli* in 15 selected provinces of Iran. The maps are originally developed by the authors of this manuscript.

ETEC: was also detected in all investigated provinces, although the number of identified ETEC isolates was quite lower than STEC (Fig 1, Panel C). With an estimated frequency of 14.0% (95% CI: 11.9, 16.3%), ETEC was identified as the second most frequent pathotype. ETEC constituted 20.8% of all identified pathogenic *E. coli* (S1 File, Table S3). At the provincial level, Esfahan

province (C) had the highest ETEC frequency (31.4%), while Fars province (S) had the lowest frequency (1.6%) for this pathotype (S1 File, Table S3 and Fig 1, Panel C). At the city-level, ETEC was detected in 29 out of 43 cities (67.4%), and was the predominant pathotype in 9 of them (Fig 2 Panel B, and S2 File).

EPEC: was the third most common *E. coli* pathotype among all identified *E. coli* species in Iran, with a frequency of 13.1% (95% CI: 11.1, 15.4%). EPEC strains also constituted 19.6% of all identified 'pathogenic *E. coli*' (S1 File, Table S4). The highest frequency of EPEC (31.0%) was observed in Khuzestan province (SW). This pathotype was not observed in Esfahan province (C). Its frequency among *E. coli*-positive samples was also less than 5% in three provinces, including Razavi Khorasan (NE; 4.3%), Gilan (N; 4.0%), and Semnan (CN; 1.2%; S1 File, Table S4 and Fig 1, Panel D). At the city-level, EPEC was detected in 28 out of 43 cities (65.1%), and was the predominant pathotype in 4 of them (17.4%; Fig 2, Panel C, and S2 File).

EAEC: was not highly prevalent in Iran (4.3%; 95% CI: 3.1, 5.8%). This pathotype also constituted a small proportion of all identified pathogenic *E. coli* (6.5%; S1 File, Table S5). Except Kurdistan province (W) which had moderate number of positive EAEC pathotypes (n= 16), all other provinces identified few EAEC isolates. In fact, nine out of fifteen provinces identified ≤ 5 EAEC isolates, and five provinces did not identify this pathotype at all (S1 File, Table S5 and Fig 1, Panel E). At the city level, EAEC was detected in 19 out of 43 cities (44.2%). However, in most of these cities (n= 13) only one EAEC isolate was detected (Fig 2, Panel D, and S2 File).

EIEC: was a rare pathotype in Iran (Fig 1, Panel F). Only three EIEC strains were identified in our sample, which yielded an overall frequency of 0.3% (95% CI: 0.1, 0.9%; S1 File, Table S6). The

three positive strains were detected in Sirjan (SE; 0.9%), Andimeshk (SW; 0.6%), and Semnan (CN; 0.2%) cities (Fig 2, Panel E, and S2 File).

Temporal variation in the prevalence of *E. coli* pathotypes

STEC was more prevalent during summer and fall, and ETEC usually showed a pick during spring and summer. Notifiable seasonal trends could not be observed for other pathotypes, probably due to very few samples (Fig 3, Panel A). Analysis of monthly trends showed a pick for STEC frequency in March and a pick for ETEC frequency in June. Notifiable monthly trends could not be identified for other pathotypes, probably due to very few samples (Fig 3, Panel B). Our results also showed that *E. coli* pathotypes were generally more frequent in warmer (*i.e.*, spring and summer) than cooler seasons (*i.e.*, fall and winter). However, the difference was not statistically significant (69.1 vs. 65.9%, P value= 0.321). This seasonal pattern was observed for the ETEC pathotype and the seasonal trend was also statistically significant (17.6 vs. 11.3%, P value= 0.047). The same pattern was observed for EAEC and EIEC pathotypes, as well. However, the observed seasonal difference was negligible in size and was not statistically significant for these pathotypes. A reverse seasonal pattern was observed for STEC and EPEC, in a way that they were more frequently detected in cooler seasons than warmer seasons. The difference, however, was small in size and was not statistically significant (Fig 3, Panel C). The S3 File provides further details on the seasonal trend of five pathotypes in different provinces investigated in this study.

Fig 3. Seasonal trend of *E. coli* pathotypes in Iran. (a) Frequency of five *E. coli* pathotypes in four seasons. STEC was more prevalent during summer and fall, and ETEC usually showed a pick during spring

and summer. Notifiable seasonal trends could not be observed for other pathotypes, probably due to very few samples. **(B) Frequency of five *E. coli* pathotypes in different months of a year.** STEC showed a pick in March and ETEC showed a pick in June. Notifiable monthly trends could not be observed for other pathotypes, probably due to very few samples. **(C) Difference in the frequency of five *E. coli* pathotypes in warmer (spring and summer) versus cooler seasons (fall and winter).** In overall, *E. coli* pathotypes were more frequently detected in warmer than cooler seasons but the difference was not statistically significant (P value= 0.321). This pattern was observed for the ETEC and EAEC pathotypes and was statistically significant for ETEC (P value= 0.04). STEC and EPEC were more frequently detected in cooler than warmer seasons but the difference, was negligible and not statistically significant.

Age-specific prevalence of *E. coli* pathotypes

The highest frequency of pathogenic *E. coli* infections was observed in infants and children of 1-5 years of age (73% each). Most of the five *E. coli* pathotypes also had a high frequency among children. In this regard, STEC (40.5% and 41.1%, respectively) and EIEC (2.7% and 0.4%, respectively) pathotypes had the highest frequency among infants and children of 1-5 years. EPEC was highly frequent in children of 1-5 years, as well as in adults of 30-40 years (16.7% each). ETEC had the highest frequency in children of 5-10 years (21.4%). Although the frequency of pathogenic *E. coli* was higher in children, these pathotypes was also common in diarrhea samples received from adults and geriatrics (Table 3).

Table 3. Age-specific frequency of *E. coli* pathotypes in 979 *E. coli*-positive culture samples received from 15 provinces of Iran

Age group	All pathotypes	STEC	EPEC	ETEC	EIEC	EAEC	Received samples
Year	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	N
0-1	27 (73.0)	15 (40.5)	4 (10.8)	4 (10.8)	1 (2.7)	0 (0.0)	40
1-5	197 (73.0)	111 (41.1)	45 (16.7)	28 (10.4)	1 (0.4)	12 (4.4)	273
5-10	45 (64.3)	19 (27.1)	8 (11.4)	15 (21.4)	0 (0.0)	3 (4.3)	70
10-20	86 (71.1)	39 (32.2)	16 (13.2)	23 (19.0)	0 (0.0)	8 (6.6)	121
20-30	88 (67.2)	47 (35.9)	18 (13.7)	18 (13.7)	0 (0.0)	5 (3.8)	131
30-40	65 (67.7)	31 (32.3)	16 (16.7)	13 (13.5)	0 (0.0)	5 (5.2)	96
40-50	46 (54.1)	27 (31.8)	7 (8.2)	9 (10.6)	1 (1.2)	2 (2.4)	85
50-60	33 (64.7)	19 (37.3)	3 (5.9)	10 (19.6)	0 (0.0)	1 (2.0)	51

60-70	32 (66.7)	15 (31.3)	4 (8.3)	9 (18.8)	0 (0.0)	4 (8.3)	48
> 70	40 (57.1)	21 (30.0)	8 (11.4)	8 (11.4)	0 (0.0)	3 (4.3)	70
Total	659 (67.3)	347 (35.1)	129 (13.2)	137 (14.0)	3 (0.3)	43 (4.4)	979

Association with Sex

Our results showed no association between sex and infection with pathogenic *E. coli* pathotypes (data not shown).

Discussion

In this large, nationally-representative survey in Iran, it was estimated that 31.6% of acute diarrhea cases were culture-positive for *E. coli*, among which 67.3% were infected with an *E. coli* pathotype. The study has yielded several findings about the frequency and spatial and temporal variation of *E. coli* pathotypes among cases with acute diarrhea in Iran.

STEC was the predominant pathotype in our sample, both at the national and provincial levels. STEC was also the most prevalent pathotype in infants and children under five years of age. Historically, STEC is among one of the first *E. coli* pathotypes isolated in Iran. The first report of this pathotype in Iran was published in 1998, where 2008 fecal samples were randomly selected from the general population of Ilam province (western Iran), and resulted in a prevalence of 4.9% for this pathotype (23). In a similar study conducted a few years later in the general population of two Northern provinces of Iran (Golestan and Mazandaran), the prevalence was estimated at 0.7%, which was quite lower than that reported from Ilam province (24). Both studies used cytotoxic methods for identification of STEC pathotypes in the Vero cells. The heterogeneity observed in the prevalence of STEC in these studies is consistent with our results, as we observed a considerable amount of heterogeneity in DEC frequency between

investigated provinces (see Fig 1 and 2). Regional heterogeneity in the distribution of VTEC have been reported in other countries as well (25, 26). It has been well acknowledged that STEC pathotype has a strong negative association with age, with children under the five years of age being more susceptible to the infection. Therefore, the heterogeneity of STEC prevalence observed between 15 provinces investigated in this study could be attributed to heterogeneous age distribution in samples collected from these provinces.

Frequency of STEC pathotype in diarrheal children under five years of age was estimated at 7.8%, which is quite lower than similar estimates available from Iran (27-31). Molecular tests were used as the diagnostic method in all these studies. These results suggest that factors other than age also predict the extent of *E. coli* pathotype infection in the communities. In a study in Brazil that hierarchically analyzed data from 3725 children less than five years of age, Vasconcelos, et al. reported that environmental factors associated with the geographical area of residence, number of people per room, maternal age, and the age of the child are the factors that independently associate with diarrhea in the children (32). A similar study conducted later in Kenya (2013), reported that hand hygiene of the care-givers and the child, drinking untreated water from the river and lack of exclusive breastfeeding were significant predictors of child diarrhea (33). These factors would not be homogenous across provinces of Iran, and may justify the observed heterogeneity between different provinces and difference studies in the country.

ETEC is more prevalent in developing countries and is a major pathogenic strain in travelers' diarrhea. This is consistent to our results, because ETEC was the second most prevalent pathotype in our study. This pathotype, however, had a moderate prevalence among children

less than five years of age who participated in our study. Its prevalence was highest among adolescents and adults. Isolation and detection of ETEC is more difficult than other DEC strains, which makes it more difficult to isolate and detect it in the fecal samples of children and infants with diarrhea (16).

EPEC was the third most frequent pathotype in our study. It was also highly frequent in children under five years of age. EPEC is known to be among major causes of diarrhea in children, especially those under two years of age, with a prevalence of about 5-10% (34). Prevalence of EPEC in Iranian children has been reported by few studies. In two studies conducted in 2001 and 2006 on diarrheal children who were less than five years of age and referred to the community health centers of Eslam-Shahr (a suburban area of Tehran), EPEC prevalence was reported around 7% (35, 36). The same estimate was also reported by Salmanzadeh, et al (2005) in diarrhea children of Tehran (37). Two other studies conducted on diarrheal children of the same age who referred to four hospitals of Iran, reported an EPEC prevalence of 12.6 and 23% (28, 29). Higher prevalence of EPEC observed in the two latter studies, which selected hospitalized diarrheal children, might be indicative of the role of EPEC pathotype in severe diarrhea diseases. Prevalence of EPEC among Iranian adults has been less regarded in the literature. The very few available studies have focused on serological detection of EPEC. For example, Alikhani, et al. (2006 and 2007), reported an isolation rate of 44.9% and 36.4% for EPEC among diarrheal adults living in Ilam, Mazandaran, and Tehran provinces (38, 39). Since these estimates are based on serological tests, it might be impractical to compare them against our results which are based on PCR methods.

In our sample, EAEC was not frequently observed in adults and children. It constituted a small proportion of DEC pathotypes (about 6%). This pathotype has been recently acknowledged as a DEC, and is mainly regarded because of its role in travelers' diarrhea. As we did not include travelers in our study, the lower frequency of EAEC in our sample can be justified. Very few studies have investigated the prevalence of this pathotype in Iran. The study conducted by Miri, et al. (2017) in Northern provinces of Iran, reported zero prevalence of EAEC in diarrheal patients of all ages who referred to community outpatients (40). These results are in line with our findings. Another study which was conducted by Jafari, et al. (2009), reported an EAEC prevalence of 16.3% in acute diarrhea children who were less than five years of age and hospitalized in children's hospitals of Tehran (29). The rate is considerably higher than our estimates, and suggests that the prevalence of EAEC is higher in severe cases of diarrhea that lead to hospitalization. The role of EPEC and EAEC in severe diarrhea, especially among young children, would be an area that warrants further investigation.

There is no report on the prevalence and distribution of EIEC in Iran. As the first study in Iran, we found a frequency of 0.3% for this pathotype. Two out of three identified EIEC isolates identified in our study were detected in children less than five years of age. Human is the sole reservoir for the EIEC pathotype. Its transmission also requires high load of the pathogen, which decreases the chance of human-to-human transmission. Given these features of the pathotype and the very low prevalence of this pathotype in our study (which possesses a representative sample from Iran); it seems that infection with EIEC is a less important issue in Iran.

Our results showed that the overall frequency of pathogenic *E. coli* was higher in warmer seasons than cooler seasons. This was also the case for the ETEC and EAEC pathotypes, but a reverse pattern was observed for STEC and EPEC, in a way that their prevalence was slightly lower during warmer seasons than cooler seasons. These results are in line with previous studies in Mexico (41) and Kenya (42). These epidemiological findings could impact the recommended use of *E. coli* vaccines during warmer and cooler months. However, additional studies using pathotype-specific vaccines would be needed to further illuminate the possible benefits during lower acquisition rate seasons. The difference between ETEC-EAEC and STEC-EPEC rates in terms of seasonality suggests that the two pathotype groups have different pathways of transmission and reservoirs in Iran.

This study has notable strengths. Most studies on this topic in Iran are limited to one or two DEC pathotypes, in a specific age group and particular geographical areas. To the best of our knowledge, this is the first study that provides a nationally representative sample of diarrhea cases, which also includes all age groups and five major DEC pathotypes including STEC, ETEC, EPEC, EAEC, and EIEC. Half of the provinces in Iran were included in this study leading to a sampling proportion of 1:2 for provinces. We also sampled provinces in a way that an even distribution of sampled provinces is reached throughout the country. Our sampling also covered a time interval of 12 months, providing the opportunity to assess the rate of DEC pathotypes over a year. We hope that these attempts make the results more representative of the distribution and frequency of DEC pathotypes in the country.

This study has also a number of limitations that need to be considered. Due to financial constraints and large-scale nature of the study, we did not identify the serotype of isolated *E. coli*. Hence, the results represent the frequency estimates that are based on molecular detection of *E. coli* pathotypes. Further studies on the O, K, and H antigens of isolated pathotypes would provide a clearer picture of the distribution of major *E. coli* serotypes in the country. Antibiotic resistance profile of isolated pathogenic *E. coli* was not also investigated in this study. In our future work, we plan to determine the antibiotic resistance profile of *E. coli* pathotypes isolated in the current study. April coincides with the new year in Iran when most Iranians travel to other cities/countries. Therefore, diarrhea cases who refer to health centers during April are usually a mix of travelers and the local people. Given that we did not aim to study travelers' diarrhea in this study, we did not include diarrheal cases referred to health centers during April. In order to provide a comprehension about DEC prevalence during April, we intrapolated the trend in this period. We also did not include hospitalized diarrhea cases as well as cases with chronic diarrhea. So, the distribution of pathogenic agents responsible for these types of diarrhea in Iran remains unknown and warrants future research. Finally, existing prevalence studies in Iran are limited to few DEC strains, specific age groups, and limited geographical areas. The sampling method as well as pathogen isolation techniques are also heterogenous among these studies. These issues limited our ability to compare our results to those of previous studies. So, we could not derive conclusions about trends in the prevalence of DEC pathotypes. This issue highlights the need for development of similar national-level evaluations in the future. The results of such evaluations can then be used for the evaluation of time-trends and effectiveness of interventions and policies.

Conclusion

The current study suggests that diarrheagenic *E. coli* may be an important cause of acute diarrhea both in adults and children in Iran, perhaps accounting for 7.2% and 10% of all acute diarrhea cases. Of five investigated pathotypes, STEC and ETEC seem to be widespread in Iran, which show a peak in warmer seasons (*i.e.*, spring and summer, respectively). This epidemiological finding could impact the recommended use of STEC and ETEC vaccines during warmer seasons, especially for infants, young children and elderlies. EPEC and EAEC seem to be less prevalent, and EIEC seems to be a rare pathotype in outpatients with acute diarrhea. The low EPEC and EAEC rates observed in our study and the high rate of these pathotypes reported among hospitalized children, suggests that the two important causes of diarrhea may be positively associated with severe diarrhea in children. Further investigations are warranted in this area. Finally, continued national evaluations of the prevalence of DEC pathotypes, as well as circulating strains and their antibiotic resistance profile, is recommended for evaluations of time-trends and effectiveness of interventions.

Acknowledgements

The authors would like to thank the staff of health centers in selected provinces for their collaboration with this study. We also would like to thank laboratory staff at Molecular Biology Department in Pasteur Institute of Iran, for their important technical assistance.

References

1. Bryce J, Boschi-Pinto C, Shibuya K, Black RE, Group WCHER. WHO estimates of the causes of death in children. *The Lancet*. 2005;365(9465):1147-52.
2. Salazar-Lindo E, Santisteban-Ponce J, Chea-Woo E, Gutierrez M. Racecadotril in the treatment of acute watery diarrhea in children. *New England Journal of Medicine*. 2000;343(7):463-7.
3. Kaper J, Nataro J, Mobley H. Pathogenic *Escherichia coli*. *Nature reviews Microbiology*. 2004;2(2):123.
4. Gyles C. Shiga toxin-producing *Escherichia coli*: an overview. *Journal of animal science*. 2007;85(suppl_13):E45-E62.
5. Keskimäki M, Mattila L, Peltola H, Siitonen A. Prevalence of diarrheagenic *Escherichia coli* in Finns with or without diarrhea during a round-the-world trip. *Journal of clinical microbiology*. 2000;38(12):4425-9.
6. Donnenberg MS. Enteropathogenic *Escherichia coli*. In: , eds. *Infections of the gastrointestinal tract*. Blaser MJ, Smith PD, Ravdin JI, Greenberg HB, Guerrant RL, editors. New York: Raven Press; 1995.
7. Huang DB, Okhuysen PC, Jiang Z-D, DuPont HL. Enteropathogenic *Escherichia coli*: an emerging enteric pathogen. *The American journal of gastroenterology*. 2004;99(2):383.
8. Nataro JP. Enteropathogenic *Escherichia coli* pathogenesis. *Current opinion in gastroenterology*. 2005;21(1):4-8.
9. Marteyn B, Gazi A, Sansonetti P. Shigella: a model of virulence regulation in vivo. *Gut microbes*. 2012;3(2):104-20.
10. Croxen MA, Law RJ, Scholz R, Keeney KM, Wlodarska M, Finlay BB. Recent advances in understanding enteric pathogenic *Escherichia coli*. *Clinical microbiology reviews*. 2013;26(4):822-80.
11. Jafari A, Aslani M, Bouzari S. *Escherichia coli*: a brief review of diarrheagenic pathotypes and their role in diarrheal diseases in Iran. *Iranian journal of microbiology*. 2012;4(3):102.
12. Kalantar E, SOHEYLI F, Salimi H, SOLTAN DMM. Frequency, antimicrobial susceptibility and plasmid profiles of *Escherichia coli* pathotypes obtained from children with acute diarrhea. 2011.
13. Pourakbari B, Heydari H, Mahmoudi S, Sabouni F, Teymuri M, Ferdosian F, et al. Diarrhoeagenic *E. coli* pathotypes in children with and without diarrhoea in an Iranian referral paediatrics centre. 2013.
14. Khoshvaght H, Haghi F, Zeighami H. Extended spectrum betalactamase producing Enteropathogenic *Escherichia coli* from young children in Iran. *Gastroenterology and Hepatology from bed to bench*. 2014;7(2):131.
15. Mitra M, Mehdi R, Hosein A, Ahmad K. Multiple drug resistance of enteropathogenic *Escherichia coli* isolated from children with diarrhea in Kashan, Iran. *African Journal of Microbiology Research*. 2011;5(20):3305-9.
16. Heidary M, Momtaz H, Madani M. Characterization of diarrheagenic antimicrobial resistant *Escherichia coli* isolated from pediatric patients in Tehran, Iran. *Iranian Red Crescent Medical Journal*. 2014;16(4).
17. Alizade H, Ghanbarpour R, Aflatoonian MR. Molecular study on diarrheagenic *Escherichia coli* pathotypes isolated from under 5 years old children in southeast of Iran. *Asian Pacific Journal of Tropical Disease*. 2014;4:S813-S7.
18. Alikhani MY, Hashemi SH, Aslani MM, Farajnia S. Prevalence and antibiotic resistance patterns of diarrheagenic *Escherichia coli* isolated from adolescents and adults in Hamedan, Western Iran. *Iranian journal of microbiology*. 2013;5(1):42.
19. Bafandeh S, Haghi F, Zeighami H. Prevalence and virulence characteristics of enteropathogenic *Escherichia coli* in a case-control study among patients from Iran. *Journal of medical microbiology*. 2015;64(5):519-24.

20. Alizade H, Sharifi H, Naderi Z, Ghanbarpour R, Bamorovat M, Aflatoonian MR. High frequency of diarrheagenic *Escherichia coli* in HIV-infected patients and patients with thalassemia in Kerman, Iran. *Journal of the International Association of Providers of AIDS Care (JIAPAC)*. 2017;16(4):353-8.
21. Alizadeh A, Behrouz N, Salmanzadeh S, Ranjbar M, Azimian M, Habibi E, et al. *Escherichia coli*, *Shigella* and *Salmonella* species in acute diarrhoea in Hamedan, Islamic Republic of Iran. 2007.
22. Quinn PJ, Carter ME, Markey B, Carter GR *Clinical Veterinary Microbiology*. London: Mosby; 2002.
23. Aslani MM, Badami N, Mahmoodi M, Bouzari S. Verotoxin-producing *Escherichia coli* (VTEC) infection in randomly selected population of Ilam Province (Iran). *Scandinavian journal of infectious diseases*. 1998;30(5):473-6.
24. Aslani MM, Bouzari S. An epidemiological study on Verotoxin-producing *Escherichia coli* (VTEC) infection among population of northern region of Iran (Mazandaran and Golestan provinces). *European journal of epidemiology*. 2003;18(4):345-9.
25. Willshaw GA, Cheasty T, Smith HR, O'Brien SJ, Adak GK. Verotoxin-producing *Escherichia coli* (VTEC) O157 and other VTEC from human infection in England and Wales: 1995-1998. *J Med Microbiol*. 2001;50(2):135-42.
26. Payne SM, Finkelstein RA. Detection and differentiation of iron-responsive avirulent mutants on Congo red agar. *Infection and immunity*. 1977;18(1):94-8.
27. Jafari F, Shokrzadeh L, Hamidian M, Salmanzadeh-Ahrabi S, Zali MR. Acute diarrhea due to enteropathogenic bacteria in patients at hospitals in Tehran. *Jpn J Infect Dis*. 2008;61(4):269-73.
28. Jafari F, Garcia-Gil L, Salmanzadeh-Ahrabi S, Shokrzadeh L, Aslani M, Pourhoseingholi M, et al. Diagnosis and prevalence of enteropathogenic bacteria in children less than 5 years of age with acute diarrhea in Tehran children's hospitals. *Journal of infection*. 2009;58(1):21-7.
29. Jafari F, Hamidian M, Rezadehbashi M, Doyle M, Salmanzadeh-ahrabi S, Derakhshan F, et al. Prevalence and antimicrobial resistance of diarrheagenic *Escherichia coli* and *Shigella* species associated with acute diarrhea in Tehran, Iran. *Canadian journal of infectious diseases and medical microbiology*. 2009;20(3):e56-e62.
30. Gupta S, Keck J, Ram P, Crump J, Miller M, Mintz E. Part III. Analysis of data gaps pertaining to enterotoxigenic *Escherichia coli* infections in low and medium human development index countries, 1984–2005. *Epidemiology & Infection*. 2008;136(6):721-38.
31. Qadri F, Svennerholm A-M, Faruque A, Sack RB. Enterotoxigenic *Escherichia coli* in developing countries: epidemiology, microbiology, clinical features, treatment, and prevention. *Clinical microbiology reviews*. 2005;18(3):465-83.
32. Vasconcelos MJdOB, Rissin A, Figueiroa JN, Lira PICd, Batista Filho M. Factors associated with diarrhea in children under five years old in the state of Pernambuco, according to surveys conducted in 1997 and 2006. *Revista de saude publica*. 2018;52:48.
33. Karambu S, Matiru V, Kiptoo M, Oundo J. Characterization and factors associated with diarrhoeal diseases caused by enteric bacterial pathogens among children aged five years and below attending Igembe District Hospital, Kenya. *Pan African Medical Journal*. 2014;16(1).
34. Ochoa TJ, Barletta F, Contreras C, Mercado E. New insights into the epidemiology of enteropathogenic *Escherichia coli* infection. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. 2008;102(9):852-6.
35. SOLTAN DMM. Diarrhea caused by enteropathogenic bacteria in children. 2001.
36. Dallal MS, Khorramizadeh M, MoezArdalan K. Occurrence of enteropathogenic bacteria in children under 5 years with diarrhoea in south Tehran. 2006.
37. Salmanzadeh-Ahrabi S, Habibi E, Jaafari F, Zali M-R. Molecular epidemiology of *Escherichia coli* diarrhoea in children in Tehran. *Annals of tropical paediatrics*. 2005;25(1):35-9.

38. Alikhani MY, Mirsalehian A, Aslani MM. Detection of typical and atypical enteropathogenic *Escherichia coli* (EPEC) in Iranian children with and without diarrhoea. *Journal of medical microbiology*. 2006;55(9):1159-63.
39. Alikhani MY, Mirsalehian A, Fatollahzadeh B, Pourshafie MR, Aslani MM. Prevalence of enteropathogenic and Shiga toxin-producing *Escherichia coli* among children with and without diarrhoea in Iran. *Journal of health, population, and nutrition*. 2007;25(1):88.
40. Miri ST, Dashti A, Mostaan S, Kazemi F, Bouzari S. Identification of different *Escherichia coli* pathotypes in north and north-west provinces of Iran. *Iranian journal of microbiology*. 2017;9(1):33.
41. Cohen MB, Nataro JP, Bernstein DI, Hawkins J, Roberts N, Staat MA. Prevalence of diarrheagenic *Escherichia coli* in acute childhood enteritis: a prospective controlled study. *The Journal of pediatrics*. 2005;146(1):54-61.
42. Shah M, Kathiiko C, Wada A, Odoyo E, Bundi M, Miringu G, et al. Prevalence, seasonal variation, and antibiotic resistance pattern of enteric bacterial pathogens among hospitalized diarrheic children in suburban regions of central Kenya. *Tropical medicine and health*. 2016;44(1):39.

Supporting information

S1 File. Frequency of *E. coli* pathotypes in 15 selected provinces of Iran

S2 File. Frequency of *E. coli* pathotypes in 43 selected cities of Iran

S3 File. Seasonal trend of *E. coli* pathotypes in 15 selected provinces of Iran

Fig 1.tif

A

All identified pathogenic *E. coli*

bioRxiv preprint doi: <https://doi.org/10.1101/2019.08.05.273952>; this version posted August 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

B

STEC

C

ETEC

A

EPEC

B

EAEC

C

EIEC

Fig 3.tiff