Sub-clinical enterovirus infections in Norwegian infants: A prospective cohort study on viral circulation and predictors of infection

Doctoral thesis by

Elisabet Witsø

Department of Genes and Environment
Division of Epidemiology
&
Department of Virology
Division of Infectious Disease Control
Norwegian Institute of Public Health
Oslo 2008
TABLE OF CONTENTS

1 Preface ............................................................................................................................. 4
1.1 Acknowledgement ......................................................................................................... 4
1.2 Summary ...................................................................................................................... 6
1.3 Selected abbreviations ................................................................................................. 8
1.4 List of papers .............................................................................................................. 9

2 Introduction to the enteroviruses .............................................................................. 10
2.1 Background ................................................................................................................ 10
2.2 Viral taxonomy .......................................................................................................... 11
2.3 Pathogenesis and immunity ..................................................................................... 11
2.4 Clinical disease ....................................................................................................... 13
2.5 Methods for detection of enteroviruses ................................................................ 14
2.6 Characteristics and genome organization ............................................................. 15
2.7 Entervirus 71 ........................................................................................................... 17
       Clinical disease and molecular epidemiology .................................................... 17
2.8 Viral structural elements associated with neurovirulence .................................. 18
2.9 Epidemiology .......................................................................................................... 19
       Transmission routes ............................................................................................ 19
       Descriptive epidemiology .................................................................................. 19
2.10 Epidemiologic surveillance data .......................................................................... 20
2.11 Enteroviruses as causative agents of type 1 diabetes .......................................... 21

3 Subjects and methods ............................................................................................. 24
3.1 Study design ............................................................................................................ 24
3.2 Subjects ................................................................................................................... 24
3.3 Methods ................................................................................................................... 25
       PCR protocols for detection and ‘molecular serotyping’ of enteroviruses .... 25
       Analysis of determinants of neurovirulence in enterovirus 71 ..................... 25
       Repeated observations within a child (statistical methods) ....................... 26

4 Aims of the study ....................................................................................................... 27

5 Brief summary of papers ......................................................................................... 28

6 Discussion of main results ..................................................................................... 31
6.1 Methodological considerations ............................................................................ 31
       Study design ....................................................................................................... 31
       Measurement error ............................................................................................ 31
       Systematic error (bias) ...................................................................................... 32
       Can we generalize our results to the general population? ......................... 32
       Confounding ....................................................................................................... 33
       Molecular typing of enteroviruses ................................................................ 33
       Breastfeeding and protection against enterovirus infections .................... 34
       Reliable diagnosis of enterovirus infections in type 1 diabetes ................ 34
6.2 Theoretical considerations ..................................................................................... 35
       Natural circulation of enteroviruses ............................................................... 35
       Sub-clinical circulation of enterovirus 71 in Norway and determinants of neurovirulence ..... 37
       Predictors of sub-clinical enterovirus infections ............................................. 38

7 Conclusions and future research .......................................................................... 40

8 Reference list ............................................................................................................ 41

9 Papers I-IV ................................................................................................................ 54
1 Preface

1.1 Acknowledgement

This work was performed at the Department of Genes and Environment (EPAM), Division of Epidemiology, Norwegian Institute of Public Health (NIPH), Oslo in the period 2004-2008. The work was also performed as collaboration with the Department of Virology (SMVI), Division of Infectious Disease Control, NIPH, and Jerome L. and Dawn Greene Infectious Disease Laboratory, Mailman School of Public Health, Columbia University, New York. I want to thank each of these institutions for excellent working conditions, especially Greene lab for the laboratory analysis part, and Division of Epidemiology for skills in advanced statistics. This study was supported by a Ph.D. grant received from The Research Council of Norway. Parts of the study was further supported by other grants given to the MIDIA project, the most important being from The Research Council of Norway, and from grants offered by the Ministry of Education of the Czech Republic, The National Institutes of Health, United States, and New Generesis. I have also received funding from The Norwegian Diabetes Association, The Norwegian Biochemical Society, and Sigurd K. Thoresen foundation.

Many people have been indispensable throughout the process of writing this thesis. I would like to express my deepest gratitude to all those who have contributed me in completing this work; My main supervisor, Kjersti Skjold Rønningen and co-supervisor Bjørn Grinde, NIPH, have both given me valuable input during this study, and generously shared their knowledge and expertise. I am grateful to Kjersti for initiating this project, for giving me the opportunity, believing in me, motivating me and trusting me with responsibilities, and for taking responsibility when most needed. Bjørn, for his support, for having clear and realistic perspectives when others did not, for his constructive inputs and fruitful discussions.

The contributions from all my co-authors are greatly acknowledged. I want to express particular gratitude to Ondrej Cinek, Charles University, Prague, for taking responsibility, for his unreservedly willingness to help, and for his excellent and precise work. Also thanks to technical assistants at his lab. I am particular grateful to Gustavo Palacios, Columbia University, who has been an invaluable supervisor in molecular virology and bio-informatics. The active role he played for achieving these goals, and his interest on enteroviruses, has been of vital importance for my work. For his friendship, and for letting me get to know his nice family, I am grateful. A special thanks to Ian W. Lipkin, Columbia University, New York,
who let me stay at his laboratory, and who together with his staff willingly, shared their expertise with me. Also a special thanks to Diana Janowitz who helped me in the lab. I am also most grateful to Lars Christian Stene, NIPH, who introduced me to the field of epidemiology and statistical methods, for always having valuable comments to my work, and for his support and friendship. I have also benefited from collaboration with Magne Aldrin, Norwegian Computing Center. I want to thank him for his collaboration, for his teaching and patient guidance. I would like to thank Trond Rasmussen, NIPH, for excellent data management and IT solutions, and for always being helpful and efficient in the delivery of data, and also thanks to Stig Jeansson, Ullevål University Hospital, for his collaboration and contribution. Turid Wetlesen, NIPH, and the other public health care nurses for their continuous effort in recruitment to MIDIA and follow-up of high-risk children. The staff at the Biobank, NIPH, for DNA extraction and genotyping. Håkon Gjessing and Anders Skrondal, NIPH, for help with advanced statistical methods. I would also thank the secretarial staff at EPAM, in particular Liv Stene-Larsen, and colleagues at EPAM who have given their sympathy throughout the difficult period MIDIA went through. I would also like to thank good colleagues at SMVI. Bjørg Guri Gutigard, Dolores Labay, and Synnøve Hestad Myhre for technical assistance, and Gabriel Ånestad, Tom Øystein Jonassen, Kathrine Stene-Johansen, Olav Hungnes, and Einar Sverre Berg, for their collaboration and constructive input.

Most of all, I am deeply grateful to the parents in the MIDIA study, for their great contribution and belief in the study. This thesis belongs to them.

Last but not least, to my family, in particular my beloved husband Kjetil, for his never-ending patience and support, to our children Oskar and Marte for showing me what is really important in life. To my parents, Oddbjørg and Tore for always being there for me, and my sisters and friends for their support.

Oslo, July 2008
1.2 Summary

*Background:* Enteroviruses are common in infancy, but usually sub-clinical and self-limiting. Most previous data on enterovirus circulation derive from analyses of specimens from individuals with disease. Studies of enterovirus circulation in healthy populations antedate the advent of molecular technologies. Population-based studies that use molecular approaches for diagnosing are needed to obtain unbiased estimates of enterovirus circulation.

*Objectives:* The objectives of the present studies were to estimate the prevalence of enterovirus infections in Norwegian children, and to gain insight into the molecular epidemiology of natural circulating enteroviruses. Other aims included studying associations of infection with disease, and investigating possible risk factors of infection.

*Subjects and methods:* Newborns were recruited as part of a prospective cohort study in Norway aimed at identification of environmental risk factors for type 1 diabetes (T1D). The infant’s parents submitted monthly stool samples (collected from their children) for viral analysis from 3 months of age. Data on symptoms of disease, as well as putative predictors of infection, were based on parental reports at 3-months intervals. Three papers (paper I-III) were based on data from 113 children with HLA high genetic risk for T1D and enrolled from 2001 to 2003. The majority of the 1255 stool samples were collected when the children were aged 3-15 months. The last paper (paper IV) is based upon data from 639 children. 4279 samples were collected when the children were age 3-12 months from 2001 to 2006. Among these children, 394 did not have increased genetic risk for T1D.

Enteroviruses were detected and quantified using a one-step real-time reverse transcription (RT)-PCR targeting the 5’ untranslated region. Enterovirus positive samples were typed directly by nucleotide sequencing of the VP1 region.

*Main results:* Enterovirus infections were common in infancy. Strains of species human enterovirus A (HEV-A) and HEV-B were approximately equally represented among the children, while strains of HEV-C were rare. Poliovirus and HEV-D were not detected. Widespread circulation of a single strain of serotype enterovirus 71 (EV71) belonging to genotype C1 was discovered during a restricted period of time, but was not associated with central nervous system disease (CNS) among Norwegian infants. Complete genome
sequencing of this strain revealed differences in the 5’ nontranslated region and RNA dependent polymerase with respect to more pathogenic viral genotypes that may explain its reduced virulence. Logistic regression indicated a linear decrease in risk of enterovirus infection with more intensive breastfeeding. The most pronounced effect was at 3 months of age. The protective effect gradually faded until the age of about 11 months. Interestingly, carrying the HLA high risk genotype for T1D may confer protection against enterovirus infection.

**Conclusion:** This is the first population based study using solely molecular methods to verify the prevalence of species HEV-A. Our findings suggest that the prevalence of enterovirus infections in general, and of HEV-A in particular, has been underestimated in epidemiological studies based on virus culture. Furthermore, our results indicate that risk for infection may be modified by diverse factors including breastfeeding and being HLA high risk for T1D. Our collective findings highlight the need for future epidemiological studies of these viruses to elucidate issues relating to viral prevalence, transmissibility and risk factors as well as their association with clinical disease. The worldwide circulation of EV71 and its potential to cause severe disease underscores the need for additional studies to identify the neurovirulent determinants of EV71 in human disease. The present work contributes importantly to the knowledge on enterovirus infections in normal infants.
### 1.3 Selected abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>HEV</td>
<td>human enterovirus</td>
</tr>
<tr>
<td>HEV-A</td>
<td>human enterovirus A</td>
</tr>
<tr>
<td>HEV-B</td>
<td>human enterovirus B</td>
</tr>
<tr>
<td>HEV-C</td>
<td>human enterovirus C</td>
</tr>
<tr>
<td>HEV-D</td>
<td>human enterovirus D</td>
</tr>
<tr>
<td>PV</td>
<td>poliovirus</td>
</tr>
<tr>
<td>CAV</td>
<td>coxsackie A virus</td>
</tr>
<tr>
<td>CBV</td>
<td>coxsackie B virus</td>
</tr>
<tr>
<td>cDNA</td>
<td>complementary DNA</td>
</tr>
<tr>
<td>CI</td>
<td>confidence interval</td>
</tr>
<tr>
<td>CNS</td>
<td>central nervous system</td>
</tr>
<tr>
<td>CPE</td>
<td>cytopathic effect</td>
</tr>
<tr>
<td>CSF</td>
<td>cerebrospinal fluid</td>
</tr>
<tr>
<td>E</td>
<td>echovirus</td>
</tr>
<tr>
<td>EV</td>
<td>enterovirus</td>
</tr>
<tr>
<td>GI</td>
<td>gastrointestinal tract</td>
</tr>
<tr>
<td>HFMD</td>
<td>hand, foot and mouth disease</td>
</tr>
<tr>
<td>HLA</td>
<td>human leucocyte antigen</td>
</tr>
<tr>
<td>IRES</td>
<td>internal ribosome entry site</td>
</tr>
<tr>
<td>NO</td>
<td>Norwegian</td>
</tr>
<tr>
<td>NTR</td>
<td>nontranslated region</td>
</tr>
<tr>
<td>OR</td>
<td>odds ratio</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>RD</td>
<td>rhabdomyosarcoma</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>RT</td>
<td>reverse transcriptase</td>
</tr>
<tr>
<td>TC</td>
<td>tissue culture</td>
</tr>
<tr>
<td>T1D</td>
<td>type 1 diabetes</td>
</tr>
<tr>
<td>VP</td>
<td>viral protein</td>
</tr>
</tbody>
</table>
1.4 List of papers


The papers are referred to in the text by their Roman numbers.
2 Introduction to the enteroviruses

2.1 Background

Enterovirus is a genus of the Picornaviridae family. The genus includes more than 80 antigenically distinct serotypes traditionally divided between polioviruses, coxsackie A viruses, coxsackie B viruses, echoviruses, and newer enteroviruses (1, 2). Human enteroviruses (HEVs) are responsible for a wide variety of clinical syndromes in humans, ranging from sub-clinical or mild upper respiratory illness to more severe diseases, such as aseptic meningitis, encephalitis, acute flaccid paralysis, and neonatal sepsis-like disease (3). HEV may also be implicated in the pathogeneses of severe chronic diseases, including type 1 diabetes (T1D) (4), viral myocarditis (5), and neuromuscular diseases (6).

Most of the HEV serotypes were discovered and described between 1947 and 1963 as a result of the application of cell culture and suckling mouse inoculation for the investigation of cases of paralytic poliomyelitis and other central nervous system diseases (7, 8). Poliovirus and poliomyelitis was the first HEV and enteroviral disease, respectively, to be recognized (and the most important one). The poliovirus eradication program, which include global routine vaccination, has led to the interruption of wild poliovirus (WPVs) transmission worldwide (9). However, at the same time as the incidence of poliomyelitis is reduced, several other HEV serotypes have emerged to cause outbreaks of major public health concern (10-14). Enterovirus 71 (EV71) has been responsible for severe CNS disease in Southeast Asia (10-12, 14). Acute hemorrhagic conjunctivitis was first described in the 1970s, and is associated with the emergence of enterovirus 70 and coxsackievirus A24 (13).

New technologies, such as the polymerase chain reaction (PCR), have provided rapid and sensitive testing methods for diagnosis of HEV infection (15-18), and consequently expanded the list of diseases attributable to this group of pathogens. Comparative phylogeny based on molecular typing methods has been of great help to classify former and new types of HEV (19-21), and to investigate the diversity of HEVs (12, 22-35), as well as the evolutionary mechanisms involved in creating their diversity (36-38). Molecular epidemiology also provides insight as to the surveillance and transmission of HEV infections, yet prospective studies investigating their natural circulation and transmissibility are very limited (28).
2.2 Viral taxonomy

Serotypes of HEV have traditionally been classified into polioviruses (PV), coxsackie A viruses A (CAVs), coxsackie B viruses (CBVs), and echoviruses (Es), based on differences in host range and pathogenic potential, sometimes resulting in overlaps between groups and difficulties with classification. As a result, beginning in the 1960s, newly discovered HEVs receive a numeric designation, beginning with enterovirus 68 (EV68), instead of being assigned to one of the traditional groups (1, 3). The subgenera each contain a number of unique HEV serotypes that are distinguished from one another on the basis of neutralization by specific antisera (see below).

Current taxonomy divides the subgenera into five species by means of VP1 sequence, but keeps traditional names for individual serotypes (1). The five species are denoted: (i) Poliovirus (PV; PV1 to PV3), (ii) Human enterovirus A (HEV-A; CAV2 to CAV8, CAV10, CAV12, CAV14, CAV16 and EV71), (iii) Human enterovirus B (HEV-B; CAV9, CBV1 to CBV6, E1 to E7, E9, E11 to E21, E24 to E27, E29 to E33 and EV69), (iv) Human enterovirus C (HEV-C; CAV1, CAV11, CAV13, CAV15, CAV17 to CAV22, and CAV24), and (v) Human enterovirus D (HEV-D; EV68 and EV70). The distribution of HEV by species only partially corresponds to the groups in the traditional classification. Because molecular techniques of HEV typing are becoming increasingly available, new enteroviruses continue to be identified, and EV71-101 have been recently described (19-21, 39, 40). Recent studies suggest that polioviruses should be reclassified as members of HEV-C (34). E22 and E23 have been reclassified as a new genus (Parecovirus) in Picornaviridae (1, 41).

2.3 Pathogenesis and immunity

HEVs are spread via the fecal-oral route (42). The ingested viruses infect cells of the oropharyngeal mucosa and lymphoid tissue (tonsils) where they are replicated and shed into the alimentary tract (throat) (Fig. 1). From here they may pass further down the gastrointestinal tract. Because of the acid stability of these viruses, they can pass into the intestine and set up further infections in the intestinal mucosa. The virus also infects the lymphoid tissue (Peyer's patches) underlying the intestinal mucosa (Fig. 1). At these sites, the virus replicates and is shed into the faeces, often for months after the primary infection. In the primary viremic
phase, the virus also enters the bloodstream at low levels. At this stage symptoms may occur and the patient may experience fever and malaise.

After multiplication in submucosal lymphatic tissues or in the gastrointestinal (GI) tract enteroviruses may pass to regional lymph nodes and give rise to a secondary viremia, which occurs about 10 days after the initial infection and leads to a humoral and cell-mediated immune response, the latter being of less importance for immunity of HEV infections. In subclinical infections, viremia is at this point limited by the host defence mechanisms.

In the GI tract replication may be sustained for several weeks even though a high titer of neutralizing antibody is present. The cells in which this replication occurs are not known, and it is unclear how replication can continue in the presence of neutralizing antibodies.

**Figure 1. Pathogenesis of HEV infections.** Cox=coxsackievirus, Echo=echovirus, Polio=poliovirus. From: [http://pathmicro.med.sc.edu/virol/picorna.htm](http://pathmicro.med.sc.edu/virol/picorna.htm)

In a minority of cases, further dissemination to susceptible target organs, such as the CNS, heart, and skin, results in major viremia, and may cause severe disease (Fig. 1). The clinical picture depends on the particular HEV subtype or strain, the level of viral replication, as well
as host factors. In these tissues, necrosis and inflammatory lesions are observed, whereas histopathologic lesions are generally not seen in the gut (43).

2.4 Clinical disease

Although the vast majority of non-polio HEV infections is believed to be completely subclinical or accompanied by mild respiratory symptoms, there are estimates of as many as 5 to 10 million symptomatic HEV infections each year in the United States (44). HEV can cause a wide spectrum of disease that involves almost any target organ. The more common syndromes include non-specific febrile illness, aseptic meningitis, herpangina, hand-foot-and-mouth disease, and exanthems. Some syndromes are almost caused exclusively by CAVs (hand-foot-mouth disease), some others by CBVs (epidemic pleurodynia, myocarditis of the newborn) (45). However, the majority of syndromes can be caused by viruses of virtually any HEV serotype. Several serotypes of HEV-A have not definitely been implicated as causative agents of any human disease.

<table>
<thead>
<tr>
<th>Syndrome</th>
<th>PV</th>
<th>CAV</th>
<th>CBV</th>
<th>E</th>
<th>EV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paralytic disease</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Meningitis, encephalitis</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Myocarditis</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Neonatal disease</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Pleurodynia</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Herpangina</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Rash disease</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Haemorrhagic conjunctivitis</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Respiratory infections</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Undifferentiated fever</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Type 1 diabetes/pancreatitis</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Disease in immune-compromised patients</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

PV=poliovirus, CAV=coxsackie A virus, CBV=coxsackie B virus, E=Echovirus, EV=other enterovirus.

Individual serotypes have different temporal patterns of circulation and can be, as already mentioned, associated with different clinical manifestations (1, 45). Although the predominant serotypes change (44), certain HEVs tend to be among those most commonly detected year after year. Disease syndromes considered characteristic of HEVs, such as aseptic meningitis or pericarditis, are in fact unusual manifestations of infection; a 4-year longitudinal family-
based study in New York City detected 291 HEV infections, none with ‘characteristic’ illnesses, and only 6 with exanthems (46).

2.5 Methods for detection and typing of enteroviruses

It is usually difficult to diagnose HEV disease from symptoms alone. The symptoms may be clinically indistinguishable from other, more severe viral and bacterial infections (3). Isolation in virus culture is still being the most common laboratory method prior to neutralization or molecular typing of HEV. The reason for this is that direct detection of HEV type is difficult without prior isolation in virus culture, due to the more than 100 different genotypes of HEV or the great variability in the VP1 region which is applicable for “molecular serotyping” (see below).

Three tissue culture (TC) systems, primary rhesus, cynomolagus or African green monkey kidney tissue culture, and the RD (rhabdomyosarcoma) cell line, allow the isolation of all PV, CBV and EV. Some HEV grow slowly and results are typically only available after the patient has recovered. Rapid techniques such as the shell vial assay can reach a definite diagnosis in 48 hours (47). Traditionally, the primary method for detection of CAV has been isolation from intracerebral inoculation of suckling mouse, because many of the CAV grow poorly in cell culture. However, the use of this technique have declined during the preceding decades (48). After isolation of the virus, identification of its serotype is conventionally done by neutralization, which is an expensive and lengthy process. Moreover, mutations in the capsid region of the genome have resulted in variable neutralization among epidemiologically unrelated isolates of the same serotype or ‘nontypeable’ isolates (49).

Beginning in 1990s, PCR based methods have proven to offer a better diagnostic alternative than TC, at least as to certain HEV related afflictions such as aseptic meningitis (18, 50). However, PCR is not necessarily the most sensitive method for virus detection as it generally involves testing of smaller volumes than can be achieved by tissue culture or viral inoculation. For amplification of HEV RNA by PCR, an initial step of reverse transcription (RT) must be included to convert RNA to cDNA. PCR protocols for the amplification of the conserved 5' 5'NTR allow fast, specific and sensitive detection of HEV directly from the sample; nevertheless, the PCR protocols have drawbacks of their own. DNA amplified by 5'NTR PCR is not appropriate for molecular typing because of low sequence diversity in this region and
frequent recombination events resulting in HEVs with divergent sequences between their 5'NTR and the regions coding for neutralizing epitopes (51). Therefore, samples positive in 5'NTR PCRs need to be examined by less sensitive methods in order to specify the subtype: either by VP1 RT-PCR and sequencing (see next paragraph), or by virus isolation and then subtyping by neutralization with specific antisera or VP1 sequencing.

‘Molecular serotyping’ methods for the identification of HEV isolates are usually based on RT-PCR amplification of a portion of the VP1 capsid gene, which contains the major epitopes associated with neutralization, followed by homology comparison of the amplicon sequence to a database of VP1 sequences (52). Isolates of the same serotype characteristically diverge in the VP1 region by less than 25% and 12%, as to nucleotide and amino acid sequences respectively (52). The epidemiology of HEV has been greatly clarified by these molecular techniques, allowing inferences about strain evolution (15-17, 29, 31, 32, 53, 54).

PCR methodology and traditional antibody techniques for virus detection are measuring very different aspects of a viral infection. PCR applied to blood or serum essentially indicates a viraemia, usually lasting a few days. In contrast, tests for virus antibodies can only indicate a previous virus infection. Depending on the test, the infection could have been recent or at a relatively distant time in the past.

2.6 Characteristics and genome organization

The *Picornaviridae* is a large family of morphologically identical single-stranded RNA viruses that share a common genomic and structural organization (55). Enterovirus virion is icosahedral, 27 nm in diameter (Fig. 2). The viral particle can withstand acidic pH of the human gastrointestinal tract and survive at room temperature for several days. Since there is no envelope, the virus is resistant to lipid solvents. The virion capsid is composed of 60 structural subunits that are formed from four polypeptides, VP1, VP2, VP3 and VP4, initially synthesized as one large polyprotein from a single open reading frame (Fig. 3). (VP1, VP2, VP3)_1 is denoted a protomer, (VP1, VP2, VP3)_5 is denoted a pentamer, and 12 pentamers constitute a virion capsid. The virus-neutralizing epitopes reside mainly in VP1, although the actual sites may cover VP2 and/or VP3. VP4 is internal.
The capsid encloses a linear, single-stranded RNA genome approximately 7.4 kb in length that is divided into 3 regions (Fig. 3): 1) the 5’ nontranslated region (5’NTR), followed by 2) a single long open reading frame encoding about 2,200 amino acids, and 3) the short 3’ NTR ending with a poly(A) tract (25).

**Figure 3.** HEV genome organization, showing the 5’NTR, the open reading frame (ORF), and the 3’NTR region. The ORF is coding for a single polyprotein and is divided into 3 regions; P1 – the viral capsid proteins, and P2 and P3 – proteins involved in protein processing and genome replication.

The four capsid proteins (VP1-VP4) and seven non-structural proteins (2A, 2B, 2C, 3A, 3C, and 3D) (non-structural region), result from the cleaved, long polyprotein that initially is translated from the genomic RNA. The functions of the non-structural proteins have not been
completely defined, but they are predicted to be involved in RNA synthesis and virion formation.

2.7 Enterovirus 71

Clinical disease and molecular epidemiology

Enterovirus 71 (EV71) belongs to species HEV-A (56, 57), which along with some CAV, such as CAV10 and CAV16, have been associated with hand, foot, and mouth disease (HFMD) (58) and herpangina, as occurred in a large outbreak of HFMD in Japan in 1973 and 1978 (59). EV71 usually accounts for less than 3% of the HEVs reported annually in the United States (30, 48). However, a seroepidemiological study conducted in New York State in 1972 showed that EV71 infection is relatively common; 26% of adults had detectable anti-EV71 antibodies in their serum (60).

EV71 infection is sometimes associated with severe neurological diseases, such as brain stem encephalitis and poliomyelitis-like paralysis (11, 35, 59, 61-66). The case severity rate of EV71 in the outbreak was <0.3% (62). This suggests a neuropathogenicity of EV71 similar to that of poliovirus (PV), which causes poliomyelitis in 0.1 to 1.0% of infected individuals (reviewed in reference (67)).

Given the nature of the pathogenicity of this serotype, and the outbreaks of severe disease with which it has been associated, the molecular epidemiology of EV71 has been widely studied (22-24, 30, 35, 68). There are two major EV71 genogroups (B and C) co-circulating worldwide. The EV71 prototype strain BrCr, isolated in 1969, is the only known example of genogroup A. Genogroups B and C have been subdivided into genotypes: B1–B5 and C1–C4, respectively (22, 30, 35, 69). Whereas genotype C1 epidemics in Malaysia, Singapore and Western Australia have been associated primarily with HFMD (22), genotype C2 outbreaks in Malaysia and Taiwan have been associated with severe and fatal neurologic disease (22, 24, 35). The differences in neurovirulence between the C1 and C2 genotypes observed in these outbreaks may provide clues as to determinants of EV71 pathogenicity (61).
2.8 Viral structural elements associated with neurovirulence

The molecular basis of viral determinants of neurovirulence is a topic of great interest, especially in the case of PV (70-76) and more recently EV71 (77, 78), but also in relation to other HEV (79).

A comparative analysis of predicted RNA secondary structures revealed the conservation of common secondary structures for PV and other HEV (80, 81). According to these models, the 5’NTR of PV can fold into several stable, and putatively functionally important, stem-loop structures (schematically presented in Fig. 4). The 5’NTR contains the highly structured internal ribosomal entry site (IRES) which extends from stem-loop II to stem-loop VI. IRES is essential for cap-independent initiation of translation and replication (80).

![Figure 4. RNA secondary structure model of the highly structured internal ribosome entry site (IRES) of the 5’ nontranslated region (5’NTR). Stem-loop structures, or domains, are denoted by roman numbers. A feature known to be highly conserved amongst HEV is the C(NANCCA)G motif (loop in parentheses) in domain V (82). AUG at nucleotide position 743 is the start codon for translation of the mRNA.](image)

Major neurovirulent phenotype determinants have been localized to the 5’NTR of HEV (70-76, 80). Temperature sensitive determinant and also the major determinant of attenuation of the genomes of PV Sabin vaccine strains (at nt 480, 481 and 472 for Sabin 1, 2 and 3, respectively) are located in domain V of IRES (82, 83).

The 5’NTR sequence required for internal initiation of translation has been called the ribosome landing pad (RLP) (82). Data strongly suggests that RLP assumes a highly ordered
tertiary structure that is recognized by trans-acting factor(s) and/or 40S ribosomes. The existence of such a superstructure have not been recognized directly, but the critical elements of RLP has been determined which involve highly conserved sites in HEV and rhinoviruses (82).

In PV, mutations within the RNA-dependent RNA polymerase gene (3D(pol)) (located to the 3' end of the genome) have been shown to affect neurovirulence (84). The highly 'flexible' finger subdomains of the polymerase are involved in modulating substrate recognition and oligomerization of the polymerase for binding to nucleotides (85).

2.9 Epidemiology

Transmission routes
HEV infect humans via direct or indirect contact with virus shed from the gastrointestinal tract or upper respiratory tract (fecal-oral and respiratory routes). The viruses gain entry into the body through the alimentary tract (42). HEV can probably be transmitted in the same way as other viruses causing gastrointestinal or respiratory infection - that is by hand contact with secretions and autoinoculation into the mouth or nose.

HEV has been found in surface water (86-89) and groundwater (90-95) throughout the world. In the case of swimming pools, HEV can be found even after chlorination, and even in the absence of fecal coliforms (96). Although HEV have been found in the environment, and can presumably be transmitted by drinking water, little is known as to the extent of environmental transmission in the developed world (90, 94, 97-99).

Descriptive epidemiology
HEVs are distributed worldwide. Infection rates vary with the season, geography, age and socioeconomic status of the population sampled (3). HEV infections occur throughout the year, but in temperate climates infections are strikingly more prevalent in the summer and autumn months in the Northern Hemisphere (100-102). In tropical climates, HEV infections do not display seasonal variation (103).
Children younger than 5 years of age are the most susceptible to infection (100, 101, 104), due in part to a lack of prior immunity and to poor hygienic habits in this age group. Adults are more likely to be severely affected when compared with other age-groups. In a summary of surveillance of HEV from 1970-2005 from the United States the proportion of reports with fatal outcome had a bimodal distribution by age, with the peaks for ages <1 year and ≥45 years (48). However, severity of specific diseases may be strikingly age-related. In recent outbreaks of HFMD disease caused by EV71, CNS disease was restricted largely to young children (10, 11, 63).

The actual rates of infection in different populations is unknown, as virus is usually only looked for in patients with symptomatic illness, and especially those with ‘characteristic’ symptoms (32, 48, 105-110).

Transmission of HEV increases with the number of contacts with carriers, especially if the carriers are small children (100-102, 104, 111, 112). Intrafamily transmission is generally greatest in large families of lower socioeconomic status. Transmission also depends on duration of virus excretion. The largest quantity and duration of virus shedding occur upon the first infection with a particular HEV serotype (104, 113). HEV infections are more common among persons living in areas with poor sanitation and in urban areas (62, 63, 104). Co-infections with more than one serotype are common under these circumstances (114).

As there is no specific therapy available against HEV infections, preventive measures are important. Breastfeeding is a unique factor that can provide protection from these infections in infancy (104, 115).

2.10 **Epidemiologic surveillance data**

It is important to note that most widely published epidemiological reports on HEV are passive surveillance data, such as the WHO *Weekly Epidemiologic Record* (48) and similar reports (32, 105-110). These data are subject to several limitations. First, enteroviruses that commonly infect younger patients or that are associated with more severe illnesses, such as poliomyelitis, meningitis, and encephalitis, are overrepresented because clinical specimens from young children and more severely ill patients are submitted for testing more frequently. Second, because the voluntary and passive nature of reporting to many national surveillance
systems, e.g. the National Enterovirus Surveillance System (NESS) (48), with annual numbers of reports varying from year-to-year, these data are not representative for HEV circulation in a population. Passive surveillance data typically describe cases, clusters, outbreaks, and enterovirus surveillance information, and can be useful in identifying targets of diagnostic assay development, for interpreting trends in enteroviral illnesses and for studies of HEV associations with specific disease. For example, the increase in nationwide rates of viral meningitis in the early 1990s occurred when E30 re-emerged after years of relative inactivity (44, 116).

In order to circumvent the selection bias of passive surveillance reports, extended studies of defined populations have been undertaken. These reports represent active surveillance data, which allow interpretations of both infection and disease incidence. One notable example is the Virus Watch programs in various cities in the United States in the early 1960’s, which were population based prospective cohort studies (101, 102, 111).

2.11 Enteroviruses as causative agents of type 1 diabetes

Type 1 diabetes (T1D) is one of the most common chronic diseases with onset in childhood. It is currently not possible to prevent the disease. T1D is characterized by autoimmunity against pancreatic islets, and autoantibodies are present for years before diagnosis (117, 118).

The causes of T1D are poorly understood (118). It is well established that genetic factors play a role in the susceptibility (119). However, the increasing incidence during the past few decades, as well as twin studies showing that the concordance rate for the disease in identical twins is only 40% (120), strongly suggest that non-genetic factors are involved (121).

Viral infections remain one of the strongest candidates for environmental risk factors (121), and in particular HEV infections. It is evident that HEV infection was frequently associated with onset of T1D in many instances. Increased prevalence of HEV RNA in sera of patients with recent-onset T1D compared with healthy controls have been reported (122-126). For example, an Australian group tested 206 newly diagnosed children and 160 control children, and found HEV RNA in 30% of T1D patients and in 4% of the controls (125). An apparently substantial body of serological case-control studies (reviewed in (127)), are also pointing to this conclusion. Nevertheless, there are methodological limitations of several of these studies.
Very different methods were used to test for viral antibodies, and more importantly the selection of controls have not been standardized, especially the differences in age, geographic areas, HLA type, and timing between cases and controls. The importance of proper matching for HLA is underlined by a report demonstrating that children with the diabetes susceptibility HLA-DR3 or DR4 alleles show strong antibody response to CBV4 infection, whereas those with the protective allele HLA-DR2 are weak responders (128). If individuals with diabetes risk alleles have a stronger immune response, a higher proportion of them will make antibodies to HEV, even if they have been equally exposed to virus compared to that of controls.

There is also direct evidence from case reports supporting an association between HEV infection and onset of T1D. A notable case was reported by Yoon et al. (129) in 1979, where a CBV4 isolate had been obtained from the pancreas of a 10-year old boy who died of diabetic ketoacidosis and showed lymphocytic infiltration of the pancreatic islets and necrosis of β-cells. Recently, direct evidence was found when CBV4 was demonstrated in specimens from three of six diabetic patients, compared to none of the 26 control organ donors (130).

Although HEV might in theory precipitate or accompany T1D at onset, implied by several previous epidemiological studies (117, 122-126), we do not know from these studies if HEV is involved in the β-cell damaging longstanding immune process (127). It therefore seems more relevant to look for evidence of viral exposure much earlier in life, and for a role of viruses in the initiation of islet autoimmunity. Some studies suggest that maternal exposure to HEV in pregnancy can increase the risk for T1D in the child (131, 132). However, another study found no evidence of excessive exposure in mothers of T1D children (133).

Long-term cohort studies of infants at high genetic risk, beginning from before birth or early in life, could potentially offer a promising strategy. However, so far these studies have not yielded consistent results. For example, studies from Finland (131, 134-138) have suggested an association between HEV infections and induction of β-cell autoimmunity, while related studies from Germany and Colorado did not find any association (139, 140). However, the studies finding no association had infrequent follow-up and small samples sizes. In addition, in Colorado HEV were rarer, than in studies from Finland, which will affect the statistical
power of the study. More well-designed studies, and careful matching of controls for time, location and HLA type is essential to elucidate the role of HEV in T1D (117).

Based on the limited information available in humans, a wide range of HEV of the echo or coxsackie type could be associated with T1D, with perhaps CBV4 being especially common (117, 126, 141, 142). The mechanism by which HEV may be involved in the pathogenesis of T1D is elusive, but several possibilities have been put forward based on the observations from experimental animal studies and in vitro data. HEV could be involved in non-T-cell mediated β-cell destruction, induction of an enhanced autoantigen-specific T-cell response, molecular mimicry and bystander activation of autoreactive T-cells, as recently reviewed in (4).

There are also other important reports presenting evidence contrary to a role of HEV in T1D (apart from those already presented). For example, studies of pancreases from children who died at or around the time of diagnosis of T1D found no signs of viral infection (143, 144). Animal models, such as the non-obese diabetic (NOD) mouse and the biobreeding (4) rat, develop immune-mediated diseases with features resembling T1D in humans. However, although these animal models have contributed important information regarding the pathogenic mechanisms of T1D, it is questionable whether these models are suitable to investigate the role of viral infections in the aetiology of human T1D (145).
3 Subjects and methods

3.1 Study design

The present studies is part of the Environmental Triggers of Type 1 Diabetes Study (Norwegian acronym ‘MIDIA’, http://www.fhi.no/tema/midia/), a longitudinal cohort study in Norway described in more detail elsewhere (146, 147). Briefly, newborns were recruited from the general population at their first visit at public healthcare centres. Genetic screening identified children at high genetic risk for type 1 diabetes (147), carrying the \( HLA-DQB1*02\)-\( DQA1*05-DRB1*03/DQB1*0302-DQA1*03-DRB1*0401 \) genotype (“HR” children), and children without this genotype (“non-HR” children). Stool samples were taken monthly by the parents from the infants were 3 months old. Data on risk factors of infection were obtained from mailed questionnaires at ages 3, 6, 9 and 12 months. The parents also kept a detailed weekly diary that included information on types and dates of symptoms suggesting an infection (cough and sneezing, diarrhoea or vomiting, and fever above 38°C), as well as breastfeeding and (nutritional) intake of other dietary products, as a help for filling out the questionnaires. The Regional Committee for Medical Research Ethics and the Norwegian Data Inspectorate approved the study, and a written informed consent was obtained from the parents.

3.2 Subjects

In paper I-III data were based on 113 ”HR” children. 1255 stool samples were submitted between September 2001 and October 2003. Eighty-nine percent of the samples were collected from children residing in the counties of Akershus (southeast Norway) and Hordaland (west coast), two counties separated by more than 400 km.

Paper IV included the same 113 ”HR” children as in paper I-III, except for that stool samples taken after 12 months of age were omitted. Samples were obtained from 133 other “HR” children, as well as from 394 “non-HR” children enrolled between the years 2004 and 2006. In the period from October 2003 until end of Mars 2004 there were no samples analyzed for HEV. More than 70% of the samples included in paper IV were from Akershus and Hordaland County, the remaining from other counties. Of the scheduled stool containers,
approximately 95% were received (paper I-IV). In paper IV, however, 69% of the samples were analyzed (randomly chosen) due to reduced analyzing capacity.

### 3.3 Methods

**PCR protocols for detection and ‘molecular serotyping’ of enteroviruses**

HEV nucleic acid extracted from the stool suspension was detected and quantified using a one-step real-time reverse transcription (RT)-PCR targeting the 5’NTR (paper I-IV). The primer-probe system for HEV detection was first extensively verified, to ensure a high specificity for HEV (paper I).

A partial VP1 gene sequencing method was adopted from Casas et al. (15) to allow for further subtyping of samples positive for HEV (paper II), avoiding the drawbacks of the traditional neutralization assays (see chapter 2.5). This RT nested PCR assay was chosen for the apparent convenience of using only 2 degenerate primers pairs (sense and antisense primers for first and second round PCRs) that allow for direct typing of HEV serotypes. The antisense primers were designed in the 5’ end of the 2A protease region.

In instances where samples could not be typed by this approach (39 of 145 samples), the RNA polymerase region was amplified and sequenced (15). Based on the RNA polymerase sequence, degenerate serotype- or species-specific VP1 primers were designed from all available sequences obtained from the sequence database GenBank (http://www.ncbi.nlm.nih.gov). These were nesting primer sets (localized inside the sequence product of the first amplification). Individual serotypes of the VP1 sequences were compared phylogenetically with all corresponding sequences of the respective serotypes in the GenBank, to distinguish between genotypes of a serotype (paper II).

**Analysis of determinants of neurovirulence in enterovirus 71**

The complete genome sequence of the EV71 strain discovered among Norwegian infants was determined in an attempt to identify neurovirulent determinants using comparative nucleotide analysis of complete genome sequences and secondary structures modelling of the 5’NTR (paper II). The complete genome sequence was obtained after an initial complete genome
sequencing of a representative isolate of the strain. To avoid introduction of mutations by cell culture adaptation of EV71, new primers were designed based on the first sequence, and direct amplification and sequencing was performed to obtain the wild-type EV71 sequence directly from the stool sample (paper III). To minimize introduction of mutations during amplification, a High Fidelity PCR Master Mix (Roche) were utilized. Ambiguous nucleotides were resolved by re-sequencing.

Repeated observations within a child (statistical methods)

Correlated data arise when some of the data are not independent of each other, such as for longitudinal data where observations within the same child are expected to be correlated. It is well known that the correlation must be taken into account in an analysis in order to obtain the correct inferences and standard errors. In this work the dependence in data from repeated measurements for each child was handled in the analysis of associations of viral load with duration of infection, species, sex, age, etc. (paper II), and in the analysis of association of putative predictors with infection (paper IV). In paper II the regression models were fitted by generalized estimating equations (GEE) using the xtgee procedure. In paper IV the model was estimated by maximum likelihood, using the gllamm program for Generalised Linear Latent and Mixed Models (GLLAMMs) (http://www.gllamm.org) which allow for a subject-specific random intercept ($\beta_{0i}$) in the regression model. Both programs are implemented in the newer releases of STATA statistical software (StataCorp. College Station, TX). The difference between the xtgee and glamm programs is that xtgee analysis gives estimates with a population-averaged interpretation, while glamm gives a subject-specific interpretation (148).
4 Aims of the study

The overall aim of the present work was to describe the molecular epidemiology of natural circulating enteroviruses, as well as possible risk factors for sub-clinical infections in early childhood.

The specific aims were:

a. To determine the prevalence of HEV infections in longitudinal stool samples from Norwegian infants at high genetic risk of T1D using highly sensitive molecular detection

b. To characterize serotypes and genotypes of HEV directly by nucleotide sequencing of the VP1 capsid region, and to use this information to estimate their prevalence and for doing molecular epidemiology.

c. To do a complete genome sequencing and molecular characterization of an EV71 strain found in Norwegian infants and to examine genomic determinants of neurovirulence.

d. To assess the effect of breastfeeding, as well as other putative selected predictors, on the risk of HEV infections in the first year of life.
5 Brief summary of papers

Paper I: Longitudinal observation of HEV and adenovirus (ADV) in stool samples from Norwegian infants with the highest genetic risk of T1D

The aim of this study was to estimate the prevalence of HEV and ADV infections in Norwegian infants, to evaluate the duration of infections, to investigate their association with symptoms, and to establish a procedure that will be used to study the association between these viruses and the development of autoimmunity leading to T1D.

Parents of infants submitted monthly stool samples from 113 infants at high genetic risk of T1D, as well as information on symptoms of infection. The samples were analysed for HEV and ADV using quantitative real-time PCR, and HEV-positive samples were sequenced. ADV was studied from the same sample because other viruses of similar routes of spread (e.g. ADV) should be studied to exclude the possibility that a putative association of HEV with pre-diabetes just reflects an increased general propensity towards viral infection in pre-diabetic individuals.

HEV was found in 142/1255 (11.3%), and ADV in 138/1255 (11.0%) of the stool samples. Approximately half of the infants were exposed to these viruses at least once during the first year of observation (period 3–14 months of age). The presence of ADV was associated with fever and with symptoms of cold, but not with diarrhoea and vomiting. HEV positivity was not associated with any symptoms. The prevalences of HEV and ADV were sufficiently high to enable studies of their association with chronic diseases.

In conclusion, the present protocol for evaluating exposure to these viruses is well suited for large-scale efforts aimed at assessing a possible relationship between these viruses and long-term development of autoimmunity leading to T1D.

Paper II: High prevalence of HEV-A infections in natural circulation of HEVs

A total of 1255 longitudinal stool samples were collected from 113 infants at high genetic risk of T1D beginning at age 3 months and continuing to 28 months. A total of 145 HEV positive samples were detected by real-time PCR and typed by nucleotide sequencing of the VP1 region, without prior isolation in cell culture.
HEV-A was detected most frequently, with an overall prevalence of 6.8%. HEV-B was present in 4.8% of the samples, while HEV-C in only 0.2%. No poliovirus or HEV-D group viruses were detected. Twenty-two different serotypes were observed in the study period: the most common were EV71 (14.5%), CAV6 (10.5%), CAV4 (8.9%), E18 (8.9%), and CBV3 (7.3%).

These findings suggest that the prevalence of HEV infections in general, and HEV-A infections in particular, has been underestimated in epidemiological studies based on virus culture.

**Paper III: Sub-clinical circulation of EV71 in Norway**

Molecular characterization of HEV in monthly stool samples from infants recruited to a prospective cohort study discovered widespread circulation of EV71 in healthy infants. Phylogenetic analysis of VP1 sequences revealed the presence of a single strain belonging to genotype C1. Clinical records of diseases associated with EV71, such as HFMD, herpangina and encephalitis, obtained from the Norwegian Surveillance System for Communicable Diseases and the Norwegian Patient Register, showed no increase in Norwegian infants during the period of study.

The genomic sequence of the Norwegian strain was compared to C2 strains from fatal and non-fatal cases in an attempt to find determinants associated with neurovirulence. Most of the amino acid differences were at non-structural sites, with the highest number (eleven amino acid changes) in the RNA-dependent RNA polymerase (3D(pol)) gene. Structure prediction algorithms did not suggest that these changes modified the tertiary folding of the protein, but functional effects cannot be excluded.

A secondary structure of the 5’NTR was predicted in an attempt to elucidate structural elements and motifs associated with neurovirulence of EV71. Possible relevant mutations in conserved motifs of the secondary structure of domain V of the highly structured IRES were observed in the Norwegian C1 strain, differentiating this genotype from all other known genotypes.
In conclusion, whether the absence of disease in Norwegian children reflects the genomic differences found in the 5’NTR and the polymerase, other intrinsic viral properties, or even host factors remains to be determined.

**Paper IV: Breastfeeding and other predictors of sub-clinical HEV infections in infants: A prospective cohort study**

The purpose of this study was to evaluate to what extent breastfeeding and other predictors were related to sub-clinical HEV infections in infants aged 3-12 months. The analysis was based on data from the same prospective cohort study as the other papers (paper I-III) (see chap. 3.1, “Study design”), however the data here included far more children (639 children), where 62% (394/639) did not carry the high risk genotype. 4279 monthly stool samples were analyzed for HEV by real-time PCR. Predictors, such as the average number of breastfeeds per day, were based on parental reports submitted at 3-months intervals.

Logistic regression, taking all predictors into account simultaneously, indicated that the more intensive breastfeeding, the more protection from HEV infections. The most pronounced effect was at 3 months of age (odds ratio=0.20, 95% confidence interval: 0.11, 0.35, for a child breastfeeding ten times per day), while the protective effect decreased to zero at 12 months. Factors associated with an increased risk of infection included: season ($p<0.001$), year of sampling ($p=0.03$), number of siblings ($p=0.01$), siblings attending day-care ($p=0.001$), and day-time company of other children ($p=0.05$). Socioeconomic status, gender, county of residence, household drinking-water and domestic animals were not associated with infection.

In conclusion, this study identified several predictors of HEV infections, the more important being breastfeeding.
6 Discussion of main results

The primary aim of the work included in this thesis was to estimate the prevalence of HEV infections in Norwegian children, and to gain insight into the molecular epidemiology of natural circulating HEV. Other aims included looking at associations of infection with disease, and to investigate possible protective factors, as well as factors associated with risk. HEV infections were prospectively assessed in stools of normal children by using molecular methods for detection and typing, and it was found that HEV infections were more common than expected. In particular, the prevalence of HEV-A viruses has been underestimated in epidemiological studies based on virus culture. Based on parental reports it was found that the presence of HEV was not significantly associated with symptoms of disease. Widespread circulation of a single strain of EV71 was discovered among Norwegian children, but was not found to be associated with CNS disease, HFMD or herpangina. It was found that the risk of HEV decreases with more frequent breastfeeding in the first year of life. The risk of infection increased significantly with the number of siblings and other children an infant was together with, such as in a day-care setting.

6.1 Methodological considerations

Study design

The MIDIA study was feasible in the context of assessing the natural circulation of HEV because healthy children, or children with no known specific or serious disease, are randomly selected from the general population and then repeatedly observed for some time. There were a few limitations of this approach, however, as discussed below.

Measurement error

Measurement error, both random and systematic, may influence the study’s results. A reduction in random errors will improve the precision while reduction of systematic errors will improve the validity of the results. Random errors in epidemiological studies are primarily reduced by enlarging sample size (149). Three of the papers included in this thesis (paper I-III) were based on 1255 observations from 113 infants, and one paper (paper IV) on 4279 observations from 639 infants.
Systematic error (bias)

Selection bias Procedures used to select subjects and factors that influence study participation may introduce systematic error in the study. As a result, the observed occurrence of disease (e.g. infection) and the relation between exposure and disease will differ from the true values in the population. In our study (paper IV), such bias may occur if children who are more frequently infected by HEV tend to participate more or less often than children who are less frequently infected. For example, children becoming ill from infection could be systematically not selected to the cohort. However, since HEV infections are usually mild or self-limiting and very common in childhood, we do not expect that infection status, or future infection status, influences the selection of subjects to the study of individuals different for exposed and non-exposed individuals, e.g. among breastfeed and not breastfeed infants (paper IV). Selection bias is therefore not regarded as a problem in this study (paper IV).

Can we generalize our results to the general population?

We did not have a random selection of children from the general Norwegian population. For one, the children were included based on their HLA genotype; and two, they all came from a few selected counties and within a relative short period of time. Previous studies have reported an association between HLA alleles with high risk of T1D, and susceptibility for HEV infections (125, 128, 150, 151). According to the results of paper IV, this HLA genotype actually appeared to protect against HEV infection, although the association was only borderline significant (p=0.050). However, the possible bias as to the estimated HEV prevalence (paper I and II) is most likely minimal.

It cannot be ruled out that selection of HLA genes also could have changed the distribution of serotypes detected (paper II). HLA could modify the susceptibility for certain serotypes compared to children representing the general population for HLA-DQ and DR allele frequencies, but any significant change seems unlikely.

The selection of counties and years of sampling could have caused a bias, i.e. the estimated frequencies of HEV-A and HEV-B may not be correct as to a national average, or other years. In the short time frame of this work, certain HEV-A types might have dominated as compared to other periods, but the effect is probably minor. Future studies might focus on longer time periods as well as other geographical regions.
Within the material obtained, county of sampling was not, however, significantly associated with circulation of HEV serotypes (paper II). Moreover, the two counties providing the most samples are located far away from each other, one outside Oslo (Akershus), one at the Western coast of Norway (Hordaland), and might be somewhat representative for main regions of Norway. The selection of county and year of sampling could not lead to any bias in the effect of breastfeeding and other predictors, as a regression model was applied that adjusts for these variables (paper IV).

Since HEV prevalence varies by year and by season, the short period of observation can not suggest patterns of circulation of HEV types or predominate types from year to year. A larger study population is necessary to assure generalizability, and lack of throat swabs may have decreased the prevalence of HEV and certain serotypes since respiratory excretions could have been missed. It has been reported that respiratory shedding is more common for coxsackieviruses than with echoviruses (111).

**Confounding**

A number of putative confounders for the effect of explanatory variables (breastfeeding, siblings, etc) on risk for HEV infections were included and adjusted for in the regression model, such as the child’s age, year, season, HLA high risk genotype and county of sampling (paper IV).

**Molecular typing of enteroviruses**

Positive samples was subtyped directly (without prior isolation in cell-culture) using VP1 sequencing to avoid bias in the selection of HEV types (paper II), since several CAV serotypes are difficult to grow in cell cultures and require inoculation in suckling mice for virus isolation. These are not routinely performed in most laboratories. The RD cell line support replication of a number of CAV serotypes, but some serotypes and strains even fail to replicate on human RD cells. RD are recommended by the World Health Organization for poliovirus surveillance (152, 153). Thus poliovirus surveillance efforts may produce some data on CAV circulation but some CAV types are still overlooked by this approach, leaving the picture of HEV surveillance somewhat incomplete (154).

Isolation in virus culture is still being the most common laboratory method prior to neutralization or molecular typing of HEV. The reason for this is that direct detection of HEV
type is difficult without prior isolation in culture, due to the more than 100 different genotypes of HEV or the great variability in the VP1 region. Typing of HEV therefore requires a huge effort and a numerous of primers to achieve a high success rate, as was demonstrated in paper II. Fourteen degenerate primers pairs had to be designed in addition to the typing of positive samples with the protocol by Casas et al. (15) in the first round. Ninety-three % of the infection episodes could be assigned to the serotype lever, and 98% to the species level. The typing approach could also detect mixed infections, although some could have been missed (paper II). Phylogenetic analysis of the C-terminal domain of the VP1 gene has previously facilitated the differentiation of circulating serotypes and the association of specific HEV genotypes with disease (27, 30, 31, 155).

**Breastfeeding and protection against enterovirus infections**

Previous studies investigating whether breastfeeding can protect against HEV infections (115), found that exclusive breastfeeding can protect against these infections. Exclusive breastfeeding is according to the WHO classification system (156), meaning those infants who receives no other food or drink besides breast milk. In paper IV, however, a continuous measure of breastfeeding was applied as the main measure of breastfeeding, i.e., the parents reported the number of times the infant was breastfeed each day. This was adopted in order to investigate a dose effect of breastfeeding as to the risk of HEV infections. The measure used is only an indirect measure for the amount of mother’s milk ingested by the infant. However, the inverse correlation between the frequencies of breastfeeds and supplementary feeding strongly suggests that the former measure is relevant as a proxy for the amount of mother’s milk ingested. In a study from Finland (115), the protective effect of breastfeeding was related to exclusive breastfeeding rather than total breastfeeding (exclusive or non-exclusive), which only can suggest that there might be a dose-effect related to the greater amount of mother’s milk ingested by exclusively breastfeed infants.

**Reliable diagnosis of enterovirus infections in type 1 diabetes**

Clearly, to elucidate the role of HEV in T1D pathogenesis, further adequately designed and powered studies must be performed (see chap. 2.11). Indeed, we have shown that a highly sensitive virological test for detection and short sample intervals (stools samples collected each month the first 3 years of life) are essential for reliable diagnosis of HEV infections, and for the possibility to detect as many infections as possible. We have also developed a reliable typing strategy which allows for minimal bias for identification of potential “diabetogenic”
strains. This work will therefore benefit the MIDIA study, making it feasible to study the association between HEV and T1D in the future, when disease has been developed in a sufficient number of children conferring HLA high risk for T1D.

6.2 Theoretical considerations

Taking the methodological limitations into consideration, the present work adds a valuable contribution to the knowledge on the molecular epidemiology of natural circulating HEV and predictors of infection.

Natural circulation of enteroviruses

An important aim of the present work was to examine the natural or sub-clinical circulation of HEV infections in normal infants (paper I-IV). Most observations on the epidemiology of HEV infections and etiological associations with disease arise from studies of patients and sometimes controls during an epidemic outbreaks (106-109, 157-161). This has resulted in a biased picture of the epidemiology of HEV infections, as they are often sub-clinical. In addition, the more limited previous epidemiological observations of healthy or normal children are in general old (28, 100-102, 111, 162-164), which make comparison difficult due to methodological differences.

The prevalence of HEV RNA in stool samples from healthy Norwegian infants was 11.3% (paper I), which is higher than was reported in a similar study from Finland (8.2%) (138). The prevalence of HEV in serum is typically lower (115). Indeed, these estimates are considerably higher than frequencies in populations in the Unites States obtained using cell culture techniques (101, 102, 163), presumably due to the differences in detection methods.

Paper II suggests that HEV-A infections are as least as common as HEV-B infections in healthy children (6.8% and 4.8%, respectively). Thus, not only the prevalence of HEV infections in general (paper I), but in particular the prevalence of HEV-A infections, has been greatly underestimated, presumably due to the differences in detection methods. Previous epidemiological studies based on isolation in virus culture followed by serology or PCR typing report only infrequent HEV-A infections in symptomatic (106-109, 157, 158, 160, 161) or healthy children (28, 102, 111, 162, 163). However, the present findings are
consistent with earlier studies where suckling mice were used for in vivo assays and HEV-A infections were observed to be frequent in healthy children (14, 100, 164). There seems to be no previous data on the prevalence of HEV-A in clinical samples using solely molecular methods, which should have more equal sensitivities for both HEV-A and HEV-B (paper II). Thus, the present findings probably represent accurate prevalences of circulating HEV types (paper II).

A total of 22 different serotypes were detected among the 113 infants observed (paper II). The most prevalent serotypes were EV71 (14.5%), CAV6 (10.5%), CAV4 (8.9%), E18 (8.9%), and CBV3 (7.3%). None of these were among the 15 most common serotypes detected since 1993 in the United States (159, 165, 166). However, during the period from 2001 to 2003, most of the infections causing outbreaks of aseptic meningitis in the United States were E13 and E18 (157, 159), where most sequences (8.9% versus 1.6%) were E18. EV71, CAV6 and CAV4 have been rarely reported in the United States since 1970, contributing 0.5%, 0.1%, and 0.6% of all reports, respectively (48).

Many clinical syndromes have been attributed to HEV infections, including serious illness (43). The presence of HEV was not significantly associated with minor symptoms of disease in this study (paper I). This finding is consistent with a report of coxsackieviruses isolated in normal children (164), where no serious illness was noticed among the children. Minor respiratory illness was related to HEV in the Virus Watch experience, but the more serious syndromes were not encountered (111).

The prevalence was significantly higher from June to December (paper IV), confirming the seasonal distribution of HEV infections in the Northern Hemisphere (100-102). This observation held for both HEV-A and HEV-B (paper II).

We found no association between sex and either the frequency or duration of infection (paper II), which is in contrast to the common observation that HEV disease occur more frequently in males than in females among persons aged <20 years, and in particularly for more severe diseases, such as meningitis and carditis (3, 48, 167). Higher isolation rates among males than in females in younger age groups of healthy individuals have also been reported (101, 102, 114).
Analysis of age-specific frequencies of HEV-A and HEV-B showed lower frequencies in the first year of life, with a distinct increase after 12 months of age (paper II). These findings are consistent with those of Gamble et al. (100), in which the frequencies of both species were lower in the first year of life, and higher in the second and third year of life. A lower prevalence of HEV, particularly around the age of 10 months, was found in paper I with 113 infants included. However, in paper IV, with more data or 639 infants included, it was no clear indication for this pattern (data not shown).

Prolonged duration of infection in stool was observed at the most for 4 consecutive months (paper II). Excretion up to 70 days was not uncommon for coxsackievirus isolated in normal children (111). Prolonged duration of infection was associated with a higher viral load in the first positive sample, but not with sex, viral species (HEV-A, HEV-B), or season. No persistent infection was found (paper II).

Co-infections or mixed infections of HEV were observed in 10 of 113 infants (paper II). Co-infections seem to be common in poorer countries (28, 114), but they are only infrequently reported in children in industrialized or temperate areas (163).

**Sub-clinical circulation of enterovirus 71 in Norway and determinants of neurovirulence**

The circulation of the EV71 genotype C1 in Norway was not associated with severe disease (paper III), in line with other outbreaks of C1 associated primarily with HFMD or other mild symptoms (22). Genotype C2, on the other hand, have been associated with severe and fatal neurologic disease (22, 24, 35). This might suggest differences in the neurovirulent phenotypes of these two genotypes. However, there are exceptions to this generalization. Recent genotype C2 outbreaks in Japan were associated only with HFMD (69).

RNA secondary structures of the 5’NTR (domain V, VI and VII) of the Norwegian strain, as well as representatives of the other genotypes of EV71 (B1-B4, and C2), were predicted in an attempt to elucidate structural elements and motifs associated with neurovirulence (paper III). The consensus structure for each of the genotypes was conserved in all available sequences for each genotype and resembled that of EV71-BrCr (168). A series of conserved covariant mutations in the “central loop” in domain V of the IRES were observed in the Norwegian strain that were not observed in the other genotypes. A covariant mutation is defined as paired
mutations in the nucleotide sequence which leads to the predicted secondary structure being maintained of the 5’NTR. Mutational studies have localized functional determinants of PV transcription and translation in the same region (82, 83).

Mutations in the motif C(NANCCA)G in domain V of IRES, which is highly conserved for HEV (79), were also observed in the Norwegian strain (paper III). Mutations here could have impact on the tertiary folding structure of the IRES and its interaction with the 18S ribosomal RNA (82).

A K to R substitution at position 75 in the characteristic “palm” subdomain of the 3Dpol gene was found for other C2 strains compared to the Norwegian strain (paper III), which could affect polymerase activity (for review see (169)). Substitutions here have been reported to affect pathogenesis or viral replication of PV vaccine strains (84), as well as resulting in an attenuated neurological phenotype of the prototype EV71-BrCr in a monkey model of infection (78).

The comparison of genomic sequences of EV71 strains derived from fatal and non-fatal cases, including the non-fatal Norwegian strain, was performed to provide clues about neurovirulent determinants (paper III). Studies using a similar approach have not been able to confirm specific sequences or mutations responsible for neurovirulence (12, 22, 58, 79, 170). This suggests that severity of illness from an EV71 infection may not be solely due to strain virulence, but could also reflect population differences such as host-factors (genetic susceptibility, cross-reactive immunity) and environmental factors (e.g. nutritional status).

**Predictors of sub-clinical enterovirus infections**

A protective effect of breastfeeding on the risk of sub-clinical HEV infections has been suggested previously (104, 115), but these studies were based on smaller sample sizes (paper IV). A dose-response relationship as to the effect on HEV infection was found (paper IV). A dose-effect of breastfeeding is supported by data indicating that the magnitude of PV replication in the intestine depends on the level of secretory IgA antibodies in breastmilk (171). The observation that the protective effect of breastfeeding gradually decreased with age (paper IV) could be related to a decline in the concentration of immune components in human milk (172), or to the maturation of the infant’s own immune system. Clinical studies in other
populations have pointed towards predominant breastfeeding for at least six months (173, 174), and partial breastfeeding for up to one year (173), as important protective measures for respiratory infections and illness in infancy.

Approximately 40% of the infants had HLA-conferred susceptibility to T1D, while the remaining (60%) did not carry this particular genotype (paper IV). As HLA genes might modulate or increase the immune responses against enteroviruses (128, 150, 151), HLA genotype was included as a putative confounder in the regression model (paper IV). Interestingly, the high risk genotype was associated with a reduced frequency of HEV RNA in stools, but the effect was only borderline significant (p=0.050). More data is required to confirm if HLA is associated with susceptibility for HEV infections.
7 Conclusions and future research

We found that HEV infections are common in infancy, and that strains of species HEV-A and HEV-B are more or less equally represented, while strains of HEV-C are rare. Poliovirus and HEV-D were not detected. The presence of HEV in stool was not associated with fever and symptoms of cold, diarrhoea and vomiting. These viruses can cause severe disease. This thesis, however, suggest that one should be careful when implicating HEV in the aetiology based solely on the presence of virus in stool. Widespread circulation of a single EV71 strain belonging to genotype C1 was discovered in a restricted period of time, but was not associated with CNS disease. Complete genome sequencing of this strain revealed differences in the 5’NTR and the non-structural polymerase gene (3D(pol)), with respect to genotypes associated with outbreaks of CNS disease. We also found that the more intensive breastfeeding, the more protection from HEV infections for infants aged 3-12 months.

The present work adds important information to the current knowledge on HEV infections in normal infants. For example, this is the first study using molecular methods to report an unbiased estimate of the HEV-A prevalence. The lack of molecular studies on HEV-A is reflected by the limited number of GenBank sequences of HEV-A compared to other types, despite that they are very common. Accordingly, the prevalence, transmissibility and risk factors of these viruses, as well as their association with clinical disease, need more attention in future epidemiological studies. Molecular methods of typing should be implemented in public health laboratories to allow for more accurate surveillance of HEV-A infections.

The worldwide circulation of EV71 and its potential to cause severe disease underscore the need for additional studies for the identification of the neurovirulent determinants of EV71 in human disease.

Autoimmune T1D seem to result from dysfunctional immune responses, and amounting candidate genetic loci for the disease are involved in immune regulation (119). To elucidate the role of HEV in T1D, future studies might focus on the genetic susceptibility for HEV infections, as to identify if the apparent association between HEV and risk for T1D to a certain degree could be explained by a common genetic predisposition to both conditions.
8 Reference list


125. Craig ME, Howard NJ, Silink M, Rawlinson WD. Reduced frequency of HLA DRB1*03-DQB1*02 in children with type 1 diabetes associated with enterovirus RNA. J Infect Dis 2003;187:1562-70.


128. Sadeharju K, Knip M, Hiltunen M, Åkerblom HK, Hyöty H. The HLA-DR phenotype modulates the humoral immune response to enterovirus antigens. Diabetologia 2003;46:1100-5.


147. Stene LC, Witsø E, Torjesen PA, Rasmussen T, Magnus P, Cinek O, et al. Islet autoantibody development during follow-up of high-risk children from the general


