Immune regulation and RNA interference in the aetiology of *Respiratory syncytial virus* disease

Christopher Stephen Inchley

Institute of Clinical Medicine, University of Oslo, Norway

Department of Paediatric and Adolescent Medicine, Akershus University Hospital, Norway
# Table of Contents

1. PREFACE ........................................................................................................................................ 7
   ACKNOWLEDGEMENTS.................................................................................................................... 7
   LIST OF PAPERS................................................................................................................................. 9
   LIST OF ABBREVIATIONS................................................................................................................. 10

2. INTRODUCTION ............................................................................................................................ 11
   THE IMPACT OF RESPIRATORY SYNCYTIAL VIRUS INFECTION ................................................... 11
   CLINICAL DISEASE............................................................................................................................ 12
   TREATMENT AND PREVENTION OF RSV ......................................................................................... 13
   LABORATORY DIAGNOSIS .................................................................................................................. 14
   THE HOST RESPONSE TO RSV ........................................................................................................... 15
      Activation of the innate immune system .......................................................................................... 15
      Cytokine profiles in RSV disease .................................................................................................... 17
      Activation of the adaptive immune system .................................................................................... 19
      RSV-mediated modulation of the immune system .......................................................................... 24
   RNA INTERFERENCE .......................................................................................................................... 25
      Regulation of gene expression ........................................................................................................ 25
      microRNA production and function ............................................................................................... 26
      Dicer is necessary for cell function ............................................................................................... 29
      microRNA regulation of immunological pathways ......................................................................... 30
      microRNAs associated with RSV ................................................................................................. 30
      Direct RNA interference between virus and host ........................................................................... 33
      RNA interference as an anti-viral therapy ..................................................................................... 34
   FACTORS ASSOCIATED WITH RSV DISEASE SEVERITY ............................................................... 35
      Previous medical history .............................................................................................................. 35
      Timing of birth and maternal antibodies ...................................................................................... 35
      Other epidemiological factors – problems with disease definition ............................................... 36
      Genetic polymorphisms ................................................................................................................ 38
      Nasal cavity viral load .................................................................................................................... 38
      Immunological factors .................................................................................................................. 39
   SO WHY STUDY THE IMMUNE SYSTEM IN CHILDREN WITH RSV? .............................................. 43

3. AIMS AND HYPOTHESES ............................................................................................................. 45
   PRIMARY AIM .................................................................................................................................... 45
   SECONDARY AIMS ............................................................................................................................. 45
   HYPOTHESIS 1 – Dicer expression at birth ...................................................................................... 45
   HYPOTHESIS 2 – Cytokine expression at birth ............................................................................... 46
   HYPOTHESIS 3 – miRNA expression during RSV disease ............................................................... 46
   A NOTE ON HYPOTHESIS GENERATION ....................................................................................... 47

4. METHODS ...................................................................................................................................... 49
   BIRTH COHORT 2003 ....................................................................................................................... 49
      Biobank – samples available for analysis ..................................................................................... 49
      Identification of RSV positive infants and controls ..................................................................... 49
      Clinical Data .................................................................................................................................. 50
      Power calculation ........................................................................................................................... 51
   DISEASE COHORT 2011 ................................................................................................................... 53
      Study Population ........................................................................................................................... 53
      Clinical data collection .................................................................................................................. 54
      Confirmation of RSV Infection ........................................................................................................ 56
      Sample collection and processing ................................................................................................ 57
      RNA quality control ...................................................................................................................... 59
Tables

Table 1: Cytokines identified in the immune response to RSV.................................................. 21
Table 2: Clinical cases positive for RSV or other viruses in 2011 .............................................. 56
Table 3: Occurrence of viral co-infection in 2011......................................................................... 56
Table 4: Linear regression analysis of control sample miRNA expression vs.
260/230 ratio, before and after application of a cut-off 260/230 ratio ≥ 0.6
.................................................................................................................................................. 62
Table 5: Linear regression of RSV-positive sample miRNA expression vs.
260/230 ratio, before and after application of a cut-off 260/230 ratio ≥ 0.6
.................................................................................................................................................. 64
Table 6: Linear regression of miRNA expression vs RIN, after application of a cut-off
260/230 ratio ≥ 0.6 .......................................................................................................................... 70
Table 7: Cord blood gene expression in RSV positive infants ...................................................... 91
Table 8: Functions, known targets and disease associations of miRNA
differentially expressed in paper 3 .............................................................................................. 112

Figures

Figure 1: Initiation of the immune response by recognition of RSV-derived
pathogen-associated molecular patterns..................................................................................... 18
Figure 2: The stem-loop structure of human pre-mir-1 ............................................................... 28
Figure 3: RSV positive study participants in the 2003 study cohort ................................. 52
Figure 4: Collection and exclusion of nasal cytology samples in 2011 .............................. 55
Figure 5: miR-23b and miR-27b expression vs 260/230 ratio before and after
application of a cut-off 260/230 ratio ≥ 0.6 ........................................................................... 66
Figure 6: miR-130a and let-7d expression vs 260/230 ratio before and after
application of a cut-off 260/230 ratio ≥ 0.6 ........................................................................... 67
Figure 7: Expression of selected miRNA vs. RIN, after application of a cut-off
260/230 ratio ≥ 0.6 ..................................................................................................................... 72
Figure 8: Hybridization of microarray probes to complementary cDNA targets ............................ 74
Figure 9: Principal Components Analysis of the control group .............................................. 77
Figure 10: The Polymerase Chain Reaction ............................................................................ 79
Figure 11: Normalization of PCR data .................................................................................... 83
Figure 12: Age according to RSV disease group – 2011 cohort ............................................. 90
Figure 13: Study miRNA that target TNFα, TLR4 and NF-κB signalling ................................ 108
1. Preface

Acknowledgements

First, I would like to thank my supervisors Prof. Britt Nakstad and Dr. Hans Olav Fjærli for valuable support, input and drive during this project. I really appreciate the freedom that Britt has given me in choosing the direction of the project and following up all the small details that I have found particularly interesting. She has been patient and encouraged me to be creative in my work, something that has been particularly important to me. Britt has helped me to keep the flow in the project, and not get too bogged down in details and trivialities. Her experience of molecular biology has been a cornerstone of the project and given me a secure foundation to build on. Hans Olav allowed me take over his baby – the biobank of cord blood samples that have been invaluable to this study. He has encouraged me to think freely, and even though he won’t admit it himself, his experience in the field of RSV immunology has been extremely useful when planning which parts of the immune system to follow up. I have appreciated the many spontaneous, informal discussions we have had during the project. They have helped me to land ideas I’ve not been sure about, made me rethink things, or allowed me to take the project one step further.

Our molecular biologist Tonje Sonerud is an invaluable member of our team. She has been of great help in discussing possibilities, planning the studies, explaining to me the basis for lab procedures and interpreting results. I’ve been lucky enough to “help” her in the lab and I’m pleased to say that she is meticulous, thorough and (appropriately) pedantic. I am also grateful to other members of staff at the Department of Clinical Molecular Biology and Laboratory Sciences,
Akershus University Hospital, particularly Torben Lüders, who helped us get started and gave us good advice during the microRNA project.

Dr. Henning Høyte and Dr. Ole Closs identified many additional infants in the birth cohort that had been exposed to RSV. This gave me a great head start when I started the project and increased the power of the statistical analyses.

I am truly indebted to the nurses on the Paediatric Emergency Ward who vigilantly recruited patients, made sure the doctors did their jobs, took nasal tests from angry infants and eagerly tried out the nasal cytology brushings on me first. I also thank all the doctors who filled out clinical data forms despite a busy workload. The project would have looked very different without your help.

I thank the Child health care visitors in Lørenskog, Fjellhamar, Lillestrøm and Skedsmo who showed me great hospitality and allowed me access to their infants.

Statisticians at the Norwegian computing centre have also been of invaluable help in analysing the birth cohort microarray, and in helping me with my analysis of the miRNA microarray.

Pathologist Dr. Katrin Fridrich-Aas kindly assessed nasal mucosa cytology smears for us and reassured us that we were actually getting some useful material from our patients.

Finally I would like to thank the hundreds of patients and their parents who have allowed us to take blood samples, stick brushes into their noses and given us access to their health records. Without your help, we will never come closer to understanding and hopefully defeating *Respiratory syncytial virus.*
List of Papers

Paper 1:
Reduced Dicer expression in the cord blood of infants admitted with severe respiratory syncytial virus disease.

C.S. Inchley, T. Sonerud, H. O. Fjaerli and B. Nakstad

BMC Infectious Diseases 11(1): 59.

Paper 2:
Downregulation of IL7R, CCR7, and TLR4 in the cord blood of children with respiratory syncytial virus disease.

C.S. Inchley, H. C. Osterholt, T. Sonerud, H. O. Fjaerli and B. Nakstad

Journal of Infectious Diseases 208(9): 1431-1435.

Paper 3:
Nasal mucosal microRNA expression in children with respiratory syncytial virus infection

C.S. Inchley, T. Sonerud, H. O. Fjaerli and B. Nakstad

(Paper submitted to BMC Infectious Diseases October 2014)
### List of abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3'UTR</td>
<td>3-prime-untranslated region</td>
</tr>
<tr>
<td>ANOVA</td>
<td>Analysis of variance</td>
</tr>
<tr>
<td>AP1</td>
<td>Activator-protein 1</td>
</tr>
<tr>
<td>BPD</td>
<td>Broncho-pulmonary dysplasia</td>
</tr>
<tr>
<td>CCL</td>
<td>C-C motif chemokine</td>
</tr>
<tr>
<td>CCR</td>
<td>CCL receptor</td>
</tr>
<tr>
<td>CD</td>
<td>Cluster of differentiation</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complimentary DNA</td>
</tr>
<tr>
<td>CHD</td>
<td>Congenital heart disease</td>
</tr>
<tr>
<td>CPAP</td>
<td>Continuous positive airways pressure</td>
</tr>
<tr>
<td>Cq</td>
<td>Quantification cycle</td>
</tr>
<tr>
<td>CX3CL</td>
<td>CX3 motif chemokine</td>
</tr>
<tr>
<td>CX3CR</td>
<td>CX3CL receptor</td>
</tr>
<tr>
<td>CXCL</td>
<td>CXC motif chemokine</td>
</tr>
<tr>
<td>CXCR</td>
<td>CXCL receptor</td>
</tr>
<tr>
<td>DAMP</td>
<td>Damage associated molecular pattern</td>
</tr>
<tr>
<td>DC</td>
<td>Dendritic cell</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>F-protein</td>
<td>Fusion protein</td>
</tr>
<tr>
<td>FDR</td>
<td>False discovery rate</td>
</tr>
<tr>
<td>FWER</td>
<td>Family-wise error rate</td>
</tr>
<tr>
<td>G-protein</td>
<td>Glycoprotein</td>
</tr>
<tr>
<td>HBECs</td>
<td>Normal human bronchial epithelial cells</td>
</tr>
<tr>
<td>HCV</td>
<td>Hepatitis C virus</td>
</tr>
<tr>
<td>ICU</td>
<td>Intensive care unit</td>
</tr>
<tr>
<td>IFN</td>
<td>Interferon</td>
</tr>
<tr>
<td>IKK</td>
<td>Inhibitor of NF-κB kinase</td>
</tr>
<tr>
<td>IL</td>
<td>Interleukin</td>
</tr>
<tr>
<td>IQR</td>
<td>Interquartile range</td>
</tr>
<tr>
<td>IRF</td>
<td>Interferon regulatory factor</td>
</tr>
<tr>
<td>let-7</td>
<td>Lethal-7, a family of miRNAs</td>
</tr>
<tr>
<td>LIMMA</td>
<td>Linear models for microarray</td>
</tr>
<tr>
<td>LPS</td>
<td>Lipopolysaccharide (Endotoxin)</td>
</tr>
<tr>
<td>MDDCs</td>
<td>Monocyte derived dendritic cells</td>
</tr>
<tr>
<td>MHC</td>
<td>Major histocompatibility</td>
</tr>
<tr>
<td>miR-x</td>
<td>Mature miRNA number “X”</td>
</tr>
<tr>
<td>miRNA</td>
<td>microRNA</td>
</tr>
<tr>
<td>mRDAI</td>
<td>Modified respiratory distress assessment instrument</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>MΦ</td>
<td>Macrophage</td>
</tr>
<tr>
<td>NF-κB</td>
<td>Nuclear-Factor-kappa-B</td>
</tr>
<tr>
<td>NK cell</td>
<td>Natural killer cell</td>
</tr>
<tr>
<td>NOD-2</td>
<td>Nucleotide-binding oligomerization domain 2</td>
</tr>
<tr>
<td>NS1</td>
<td>non-structural protein 1</td>
</tr>
<tr>
<td>NS2</td>
<td>non-structural protein 2</td>
</tr>
<tr>
<td>PAMP</td>
<td>Pathogen associated molecular pattern</td>
</tr>
<tr>
<td>PC</td>
<td>Principle component</td>
</tr>
<tr>
<td>PC1</td>
<td>First principle component</td>
</tr>
<tr>
<td>PC2</td>
<td>Second principle component</td>
</tr>
<tr>
<td>PCA</td>
<td>Principal component analysis</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>pre-miRNA</td>
<td>Precursor miRNA</td>
</tr>
<tr>
<td>pri-miRNA</td>
<td>Primary miRNA</td>
</tr>
<tr>
<td>PRR</td>
<td>Pathogen recognition receptor</td>
</tr>
<tr>
<td>qPCR</td>
<td>Quantitative PCR</td>
</tr>
<tr>
<td>RDAI</td>
<td>Respiratory distress assessment instrument</td>
</tr>
<tr>
<td>RIG-1</td>
<td>Retinoic acid inducible gene-1</td>
</tr>
<tr>
<td>RIN</td>
<td>RNA integrity number</td>
</tr>
<tr>
<td>RISC</td>
<td>RNA-induced silencing complex</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>RNase</td>
<td>Ribonuclease</td>
</tr>
<tr>
<td>RSV</td>
<td>Respiratory syncytial virus</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>Reverse transcriptase PCR</td>
</tr>
<tr>
<td>SHIP1</td>
<td>SH2-containing inositol polyphosphate 5-phosphatase 1</td>
</tr>
<tr>
<td>SNP</td>
<td>Single nucleotide polymorphism</td>
</tr>
<tr>
<td>SOCS</td>
<td>Suppressor of cytokine signaling</td>
</tr>
<tr>
<td>ssRNA</td>
<td>Single stranded RNA</td>
</tr>
<tr>
<td>Th1</td>
<td>T-helper 1</td>
</tr>
<tr>
<td>Th2</td>
<td>T-helper 2</td>
</tr>
<tr>
<td>TLR</td>
<td>Toll-like receptor</td>
</tr>
<tr>
<td>TNF-α</td>
<td>Tumor necrosis factor α</td>
</tr>
</tbody>
</table>
2. Introduction

*Respiratory syncytial virus* (RSV) is a significant pathogen in young children, causing bronchiolitis, pneumonia and asthma exacerbations. However, most children infected with RSV have only mild upper respiratory disease. The focus of this project is the immunological response to RSV, including RNA interference as a regulator of the immune system. Of special interest are factors that may predispose infants to more severe disease on infection with RSV. I will briefly introduce clinical aspects of the virus including the issue of prophylaxis, which has been a major motivator internationally for investigating host immune responses to RSV. I will so discuss current knowledge about the immunological mechanisms in RSV disease that I believe to be relevant to this project, and give an overview of gene regulation by RNA interference. Finally, a summary of known predisposing factors for RSV will be provided, with particular attention to those pertaining to the immune system.

The impact of *Respiratory syncytial virus* infection

Bronchiolitis and other lower respiratory tract infections are amongst the most common causes of paediatric admissions [1]. In epidemiological studies the most important viral pathogen causing viral lower respiratory tract infection in young children has consistently been RSV [1-5]. The yearly epidemics of RSV lead to a significant increase in admissions to paediatric wards during the winter and spring months. Infantile bronchiolitis is associated with later development of asthma in childhood [6], and is therefore a major cause of on-going disease burden to patients and significant health costs to society [7].
In an American study from 1986, 69% of children were infected with RSV in the first year of life, and almost all by the age of two years [8]. The majority of children had an asymptomatic illness or only mild symptoms. At our hospital in the county of Akershus, Norway, a retrospective study estimated the annual incidence of infantile bronchiolitis caused by RSV and requiring hospital admission to be 22/1000 newborns [9]. These findings are in accordance with other international studies [1, 10].

Clinical disease

Children admitted with RSV infection are typically under 6 months of age, presenting with respiratory distress, reduced fluid intake and reduced oxygen saturations. Signs of respiratory distress include intercostal, subcostal, supraclavicular and jugular retractions, increased respiratory rate, prolonged expiration, grunting and nasal flaring. On auscultation, generalized coarse inspiratory crepitations may be heard in addition to both inspiratory and expiratory stridor. Finer crepitations are also common, and may indicate localized pulmonary infiltrates. Atelectasis or infiltrates may be found on chest X-rays. There is usually significant upper airways obstruction due to respiratory secretions/secretions, and it is likely that this pathology is common to the entire respiratory tract. The youngest infants are at risk of apnoea due to either exhaustion or mucus plugging of the airways. RSV infection typically takes 4 – 5 days to reach the peak of symptoms, and resolution of infection occurs after 7 – 10 days. Complications including atelectasis and secondary bacterial infection may prolong the course of disease [3, 11-13].
Treatment and prevention of RSV

A number of therapies aimed at reducing the severity or length of infection have been assessed including ribavirin, dexamethasone, inhaled β2-agonists and inhaled racemic adrenalin. Clinical trials have failed to show any significant reduction in length of stay or severity of disease. Inhaled adrenalin and β2-agonists may improve respiratory distress for a short period of time, but do not reduce the duration of disease and are not necessarily better than saline inhalations. Treatment is currently supportive, including aspiration of secretions from the upper airways, nebulized normal saline to ease clearance of respiratory secretions from the lower airways, oxygen administration and fluid administration. In more severe cases, continuous positive airways pressure (CPAP) or mandatory ventilation may be necessary [14-17]. In recent years, oxygen delivery via heated humidified high flow nasal cannula has been introduced as an additional therapy and may reduce use of CPAP [18-20]. Future treatment strategies may include inhaled hypertonic saline, which has had positive results in clinical trials [16, 21].

RSV-vaccine development has been attempted since the 1960s, but so far efforts have been futile. An early trial resulted in vaccine-enhanced disease on natural infection with RSV virus, with 80% hospitalizations on RSV infection, including 2 mortalities, compared to 5% of infants receiving an placebo vaccine [22]. Vaccine development was stopped for many years and focus shifted to understanding the mechanisms underlying vaccine enhanced disease and the host response to RSV [23]. Research has aimed at identifying elements of the immune system that can successfully be targeted by potential vaccines. During the last decade, several
attenuated RSV strains suitable for vaccine-use have been engineered using reverse genetics technology. Clinical trials of several candidate vaccines are currently underway [24].

The anti-RSV monoclonal antibody Palivizumab has been successfully introduced as prophylaxis, and in high-risk groups reduces the risk of admission with RSV by 50% [25]. However, the treatment is costly and must be administered by intra-muscular injection once a month during the RSV season. For this reason, administration has in Norway been restricted to specific high-risk infants who have i) congenital heart disease with significant cyanosis, heart failure, pulmonary hypertension or pulmonary hyperflow; ii) significant chronic pulmonary disease; or iii) been discharged less than three months before the expected RSV season, with birth before 32 weeks gestation and oxygen requirement at 36 weeks post-conception [26].

Laboratory Diagnosis

At Akershus University Hospital, direct demonstration of RSV in nasopharyngeal swabs or aspirates is performed using a direct antigen test or polymerase chain reaction (PCR). Compliment fixation of anti-RSV antibodies is also available, but is less reliable and has largely been phased out of clinical practice in Norway. At the Department of Paediatric and Adolescent Medicine, RSV-testing in children with suspected infection has been routine practice since before 2003 and facilitates cohort isolation of RSV positive infants. In the last decade, PCR has been the preferred method. PCR is highly sensitive and specific for the presence of RSV in the nasopharynx [27], but must be processed in the lab and results are typically not available until the next day. In cases where a result is desired
immediately, the rapid antigen test may be used. This test is less sensitive, but has a good specificity [28].

The host response to RSV

I will here present an overview of the immune response to RSV. Toll-like receptor 4 mediated NF-κB signalling, RSV-induced cytokine profiles, dendritic cell activation and migration, and antigen presentation to lymphocytes are key areas that are relevant to the current project, and will receive special focus.

Activation of the innate immune system

RSV preferentially infects epithelial cells of both upper and lower airways. On infection, RSV-derived antigens engage with pathogen recognition receptors (PRRs) resulting in activation of various transcription factors and subsequent production of cytokines. Cells of the innate immune system are activated both by these cytokines and by direct contact with RSV antigens.

Pathogen associated molecular patterns (PAMPs) are molecular signatures that are common to many pathogens, and which are recognized by the innate immune system. Archetypal PAMPs include lipopolysaccharide (LPS) present on gram-negative bacteria, and single-stranded viral DNA [29, 30]. The innate immune system has a number of PRRs that recognize PAMPs and activate the immune response. A typical PRR is Toll-like receptor 4 (TLR4), a membrane bound surface protein found on many cells including pulmonary dendritic cells and epithelial cells. On PAMP recognition by TLR4, a signal cascade is initiated culminating in nuclear translocation of several transcription factors, including Nuclear-factor-kappa-B (NF-κB), Interferon regulatory factors (IRF) 3 and 7, and Activator-protein 1 (AP1) [31]. NF-κB is a transcription factor of particular
importance to this project. It transcribes a range of genes relevant to activation of the immune system. The signal cascade leading to NF-κB activation involves a number of proteins with both negative and positive regulatory effects. Fine-tuning of NF-κB activation via feedback mechanisms is necessary to limit the immune response to pathogens [32].

RSV-derived PAMPs that potentially activate the innate immune system are the fusion protein (F-protein), and genomic single stranded RNA (ssRNA).

TLR2 and TLR4 are cell-surface PRRs that initiate an immune response on RSV PAMP recognition. The RSV-PAMP responsible for TLR2 activation has not yet been described [33], but it is well documented that the RSV F-protein shows affinity for TLR4 and its co-factor CD14. This was first demonstrated in human monocytes and mouse peritoneal macrophages, where RSV F-protein - TLR4 interactions stimulated interleukins (IL) 1β, 6 and 8, and Tumour necrosis factor α (TNFα) production [34]. In the same paper, TLR4-deficient mice had persistent pulmonary RSV infection. In a later study, RSV activated NF-κB in a TLR4-dependent manner early in the immune response, but at 24 hours NF-κB activation was less TLR4-dependent, implying other PRRs in the detection of RSV [35]. Other investigators have described similar findings [36, 37], and the association between mutations in the TLR4 gene and increased risk or severity of RSV disease (see the chapter on genetic polymorphisms below) [38-40] further highlights the role of TLR4 in immune responses to RSV. However, in one study TLR4 deletion in mice did not impair immune responses [41], and another study found normal in-vitro responses to RSV and LPS in monocytes from individuals
with TLR4 polymorphisms [42]. The role of TLR4 in the immune response to RSV is therefore currently disputed.

Several intracellular PRRs respond to RSV. On cellular infection, viral replication produces large amounts of RSV-derived RNA, which is detected by the intracellular PRRs TLR3 [43, 44], TLR7 [45], Retinoic acid inducible gene-1 (RIG-1) [46, 47] and nucleotide-binding oligomerization domain 2 (NOD2) [48]. These PRRs activate various transcription factors, including NF-κB and IRF3. Whilst NF-κB transcribes cytokines, IRF3 transcribes type-1 interferon [49, 50]. Figure 1 gives a summary of how RSV activates the immune system.

Cytokine profiles in RSV disease

Cytokines involved in the immune response to RSV include interleukins; CC, CXC and CX3C motif chemokines (CCL, CXCL and CX3CL, respectively); interferons (IFN); and TNFα. Table 1 sums up the major functions of RSV-related cytokines and their receptors, including alternative names previously common in the literature.

Analyses of respiratory secretions of infants with various degrees of RSV infection have identified an abundance of IL1β, 6, 8, 9 and 10; CCL2, 3 and 5, CXCL10 and TNFα in the [51-61]. In-vitro stimulation of respiratory cell lines or alveolar macrophages with RSV results in similar cytokine profiles [62-66]. In addition to these cytokines, a microarray approach identified upregulated mRNA of CCL1, 17, 20 and 22; CXCL1, 2, 3, 5 and 11; and CX3CL1 in RSV-stimulated A549 and primary small airway epithelial cells, although these results were not confirmed by protein analysis [67]. One study assessing anti-viral responses
found upregulation of IFNα and IFNβ but not IFNγ in RSV stimulated A549 cells [68].

Figure 1: Initiation of the immune response by recognition of RSV-derived pathogen-associated molecular patterns

The cytokines produced during RSV infection are associated with chemoattraction and activation of leukocytes. Of note, IL8 is a potent neutrophil chemoattractant, and its levels are associated with neutrophil activity in nasal secretions of RSV-positive infants [51]. CCL2, 3 and 5, and CXCL10 attract
myeloid cells, whilst CCL3 and IFNβ attract natural killer cells. These studies are thus consistent with the infiltration of neutrophils, macrophages, monocytes and natural killer cells described in the initial response to RSV infection [69, 70]. IL1β, IL6 and TNFα are primarily produced by macrophages or lymphocytes and are potent pro-inflammatory cytokines. In contrast, IL10 has anti-inflammatory effects. IFNα and β are powerful cytokines with a wide range of anti-viral effects including apoptosis, restriction of viral replication and leukocyte recruitment [49, 69].

Activation of the adaptive immune system

Dendritic cells (DCs) are myeloid antigen-presenting cells that form a critical link between innate and adaptive immune systems [71]. In the respiratory epithelium they survey epithelial cells for exogenous antigens, expressing PRRs such as TLR4 and RIG-1 in large quantities. RSV-PAMPs [72-75] are recognized by DCs and internalized [76], leading to NF-κB activation. Subsequent maturation of the DC is attributable to NF-κB mediated miR-155 transcription, which targets Suppressor of cytokine signalling 1 (SOCS1), CD155, and SH2-containing inositol polyphosphate 5-phosphatase 1 (SHIP1), factors that maintain a quiescent state [77, 78]. DC maturation results in upregulation of CCR7, a key receptor in DC localization to local lymph nodes. CCR7 binds to CCL19 and CCL21, which are expressed in secondary lymphoid tissue, and it is likely that CCR7 uses these ligands to migrate along afferent lymphatic vessels [79, 80]. On arrival at local lymph nodes, DCs present antigen to T-lymphocytes using MHC class I and class II molecules.
Localization of T-lymphocytes to the lymph nodes is not coincidental, but promoted by several factors including IL7 and CCR7. T-lymphocytes constantly circulate between blood, secondary lymphoid tissue, and efferent lymph channels – a process that allows optimal surveillance for antigen presenting DCs within the body’s lymphoid organs. IL7 exclusively signals via its receptor IL7R,
<table>
<thead>
<tr>
<th>Ligand</th>
<th>Alternative names</th>
<th>Receptor</th>
<th>Major functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCL1</td>
<td>I-309, TCA-3</td>
<td>CCR8</td>
<td>Monocyte, NK cell and dendritic cell chemoattractant</td>
</tr>
<tr>
<td>CCL2</td>
<td>MCP-1</td>
<td>CCR2</td>
<td>Monocyte and dendritic cell chemoattractant</td>
</tr>
<tr>
<td><strong>CCL3</strong></td>
<td>MIP-1α</td>
<td>CCR1</td>
<td>Neutrophil chemoattractant</td>
</tr>
<tr>
<td>CCL4</td>
<td>MIP-1β</td>
<td>CCR1, CCR5</td>
<td>Neutrophil, monocyte, NK cell chemoattractant</td>
</tr>
<tr>
<td>CCL5</td>
<td>RANTES</td>
<td>CCR5</td>
<td>T-cell, eosinophil and basophil chemoattractant, NK cell activation</td>
</tr>
<tr>
<td>CCL7</td>
<td>MARC, MCP3</td>
<td>CCR2</td>
<td>Monocyte chemoattractant; Macrophage modulator</td>
</tr>
<tr>
<td>CCL17</td>
<td>TARC, dendrokine, ABCD-2</td>
<td>CCR4</td>
<td>Lymphocyte chemoattraction to thymus</td>
</tr>
<tr>
<td><strong>CCL19</strong></td>
<td>ELC, Exodus-3, Ckβ11</td>
<td><strong>CCR7</strong></td>
<td>Dendritic cell localization to lymph nodes</td>
</tr>
<tr>
<td>CCL20</td>
<td>LARC, Exodus-1, Ckβ4</td>
<td>CCR6</td>
<td>Lymphocyte and dendritic cell chemoattraction</td>
</tr>
<tr>
<td>CCL21</td>
<td>SLC, 6Ckine, Exodus-2, Ckβ9, TCA-4</td>
<td><strong>CCR7</strong></td>
<td>Dendritic cell localization to lymph nodes</td>
</tr>
<tr>
<td>CCL22</td>
<td>MDC, DC/β-CK</td>
<td>CCR4</td>
<td>Th2 response</td>
</tr>
</tbody>
</table>

*Studied in this project.

Table 1: Cytokines identified in the immune response to RSV.
<table>
<thead>
<tr>
<th>Ligand</th>
<th>Alternative names</th>
<th>Receptor</th>
<th>Major functions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CXCl1</strong></td>
<td>Gro-α, GRO1, NAP-3, KC</td>
<td>CXCR2</td>
<td>Neutrophil chemoattractant</td>
</tr>
<tr>
<td><strong>CXCl2</strong></td>
<td>Gro-β, GRO2, MIP-2α</td>
<td>CXCR2</td>
<td>Neutrophil chemoattractant</td>
</tr>
<tr>
<td><strong>CXCl3</strong></td>
<td>Gro-γ, GRO3, MIP-2β</td>
<td>CXCR2</td>
<td>Monocyte chemoattractant</td>
</tr>
<tr>
<td><strong>CXCl5</strong></td>
<td>ENA-78</td>
<td>CXCR2</td>
<td>Neutrophil chemoattractant</td>
</tr>
<tr>
<td><strong>CXCl7</strong></td>
<td>NAP-2, CTAPIII, β-Ta, PEP</td>
<td>CXCR1, CXCR2</td>
<td>Neutrophil chemoattractant</td>
</tr>
<tr>
<td><strong>CXCl9</strong></td>
<td>MIG, CRG-10</td>
<td>CXCR3*</td>
<td>T-cell chemoattractant</td>
</tr>
<tr>
<td><strong>CXCl10</strong></td>
<td>IP-10, CRG-2</td>
<td>CXCR3*</td>
<td>Chemoattraction of myeloid and lymphoid cells</td>
</tr>
<tr>
<td><strong>CXCl11</strong></td>
<td>I-TAC, β-R1, IP-9</td>
<td>CXCR3*, CXCR7</td>
<td>T-cell chemoattractant</td>
</tr>
</tbody>
</table>

**Interferons**

<p>| IFNα | IFNαR |
| IFNβ | IFNβR |
| <strong>IFNy</strong> | IFNyR1 and 2 | Cellular immunity |</p>
<table>
<thead>
<tr>
<th>Ligand</th>
<th>Alternative names</th>
<th>Receptor</th>
<th>Major functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interleukins</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IL1β</td>
<td>IL1R</td>
<td></td>
<td>Pro-inflammatory</td>
</tr>
<tr>
<td>IL2</td>
<td>IL2R</td>
<td></td>
<td>Proliferation and differentiation of T-cells</td>
</tr>
<tr>
<td>IL4*</td>
<td>IL4R</td>
<td></td>
<td>Humoral immunity/Th2</td>
</tr>
<tr>
<td>IL6</td>
<td>IL6R</td>
<td></td>
<td>Fever, acute phase reaction, Neutrophil production</td>
</tr>
<tr>
<td>IL7</td>
<td>IL7R*</td>
<td></td>
<td>Lymphocyte survival</td>
</tr>
<tr>
<td>IL8</td>
<td>CXCL8</td>
<td>CXC1R, CXC2</td>
<td>Neutrophil chemoattractant; phagocytosis</td>
</tr>
<tr>
<td>IL9</td>
<td>IL9R</td>
<td></td>
<td>Produced by CD4+ T cells; Cell proliferation, anti-apoptotic</td>
</tr>
<tr>
<td>IL10</td>
<td>IL10R</td>
<td></td>
<td>Anti-inflammatory</td>
</tr>
<tr>
<td>IL13</td>
<td>IL13R, IL4R</td>
<td></td>
<td>Th2 response</td>
</tr>
<tr>
<td>IL1R-antagonist</td>
<td>IL1R</td>
<td></td>
<td>Inhibits IL1 signaling</td>
</tr>
<tr>
<td>IL33</td>
<td>IL1RL1 (ST2)</td>
<td></td>
<td>Th2 response</td>
</tr>
<tr>
<td>Others</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TNFα</td>
<td>TNFR 1 and 2</td>
<td></td>
<td>Neutrophil chemoattractant; Phagocytosis; Pyrogen; Induces acute phase reactants; Cortisol response</td>
</tr>
</tbody>
</table>
which is primarily expressed by mature T-lymphocytes and maintains the
peripheral blood T-lymphocyte population. IL7 signalling prevents apoptosis of
mature, naïve peripheral blood T-lymphocytes and can also promote
proliferation of recent thymic emigrants. These are T-cells that have newly
emerged after maturation in the thymus, and which are abundant in neonates.
Naïve T-lymphocytes use CCR7 to cross the endothelium of lymph node
capillaries and migrate to the T-cell zone [81]. It is at this point that interactions
between naïve, RSV-specific T-lymphocytes and dendritic cells can occur.

An appropriate T-lymphocyte response is necessary to clear an RSV infection.
Cellular immune responses, characterized by CD8-positive cytotoxic T-
lymphocytes and Th1-skewed CD4-positive T-lymphocytes, are regarded as
beneficial to the anti-viral host immune response. Cytokines promoting cellular
immunity include IFNγ, IL2 and IL12. The Th2 responses driven by IL4 and IL13
are considered less beneficial, and may promote immunopathology. However,
excessive cellular immune responses may also be damaging, and the optimal
balance of Th1 and Th2 responses to RSV is currently not clear [71, 82].

RSV-mediated modulation of the immune system
There are several mechanisms by which RSV counteracts the immune system,
enhancing conditions for viral replication.

The RSV non-structural proteins 1 and 2 (NS1 and NS2) block transcription
factor IRF3 [83] and inhibit the interferon-induced STAT2 signalling pathway
[84]. This antagonizes type 1 interferon, impairs cytotoxic T-lymphocytes [85],
and can promote a Th2 response. [86].
RSV glycoprotein (G-protein) is produced in large quantities early during RSV infection. It is thought that G-protein may thus flood the immune system and block RSV neutralizing antibodies or their receptors [87]. G-protein also contains a CX3C motif that mimics CX3CL1, a chemoattractant for cells expressing its receptor CX3CR1, such as T-lymphocytes and NK cells that. G-protein antagonizes lymphocyte chemoattraction in vivo [88], which may be a result of CXC3CR1 blockade.

Several studies have assessed dendritic cell (DC) function in RSV disease compared to other viruses. RSV-infected monocyte-derived DCs show inefficient upregulation of CCR7 [75], which is deemed necessary for dendritic cell migration through the lymphatic system. This may explain the increase in the DC population in respiratory epithelium for weeks after infection [76]. RSV infected DCs are also less adept at inducing CD4+ T-lymphocyte proliferation [89].

The sum of these factors may thus be delayed activation of an appropriate adaptive immune response in RSV infection.

RNA interference

Regulation of gene expression

Regulation of gene expression, particularly by RNA interference, is a key concept in this project. Thus far I have discussed the process of receptor activation by a ligand or antigen, leading to a signal cascade, transcription factor activation and transcription of genes relevant to the immune response. There are, however, additional mechanisms that govern the extent to which a gene is translated into a functioning protein. In general, for a gene to be expressed, an activated
transcription factor must translocate to the nucleus, where it binds to the promoter region of target genes and transcribes DNA to mRNA. DNA-methylation and histone modification are molecular processes that make DNA unavailable for binding, inhibiting transcription. After transcription, the mRNA product is cut and spliced to remove introns. Alternative splicing may result in protein products that have differing degrees of function. mRNA is then exported to the cytoplasm where it is translated into protein by ribosomes. RNA interference can inhibit the translation of mRNA into protein.

I will in this section describe the biogenesis and mechanism of action of microRNA (miRNA), the primary effector molecules of RNA interference; the importance of the ribonuclease Dicer to normal cellular function; the effects of miRNA on immunological pathways; and a summary of current knowledge about miRNA in RSV infection.

microRNA production and function

miRNA are small RNA sequences that target specific mRNA sequences to inhibit their expression. Over 1000 miRNA are described in the human genome. miRNA are non-coding (not translated into protein) but act by inhibition of specific mRNA transcripts. A basic concept in miRNA function is that single-stranded RNA readily binds to single-stranded RNA sequences that have near or perfect complementarity to form a double-stranded RNA-helix. Complementarity means the nucleotide sequence on one RNA molecule matches the sequence on the other RNA molecule. The nucleotides Guanine (G) and Cytosine (C) bind to each other, and the nucleotides Adenosine (A) and Uracil (U) bind to each other.
Sequences 1 and 2 below to dot have the same sequence, backwards or forwards, but when aligned the Cs and Gs match, and the As and Us match. The sequences are thus complementary.

Sequence 1: ---CCUGCAACUUAGGCAG---
Sequence 2: ---GGACGUUGAAUCCGUCC---

Complementarity allows the sequences to bind to each other, and the strength of the bond is determined by how perfect the match is. Sequences 1 and 2 have perfect complementarity, and will therefore have strong bond.

miRNA biogenesis and function can be summarized in the following steps:

1. **Transcription**: A number of transcription factors, including NF-κB, are known to transcribe miRNAs. miRNA genes are transcribed to RNA sequences called primary-miRNA (pri-miRNA) that may be several hundred nucleotides long. A single pri-miRNA may incorporate the sequences of between one and six mature miRNA. Each miRNA sequence within the pri-miRNA has RNA segments that are complementary to each other. The result is that the RNA loops back on itself and since the arms are complementary it forms a double-stranded stem. The structure of a pri-miRNA is thus several stem-loops connected by segments of single-stranded RNA. The number of stem-loops corresponds to the number of mature miRNAs encoded by the miRNA gene (between one and six).

2. **Nuclear processing and exportation**: In the nucleus, the ribonuclease Drosha cuts the pri-miRNA at the base of each stem-loop to form precursor-miRNA (pre-miRNA), one for each mature miRNA. These pre-miRNA maintain the stem-loop structure. Figure 2 shows the stem-loop structure of human pre-
mir-1. The mature miR-1 sequence is indicated in red. In addition to areas of perfect complementarity, there are 3 areas of imperfect complementarity on the stem, and the loop that does not show any complementarity. Pre-miRNA are exported from the cell nucleus by exportin-5.

Figure 2: The stem-loop structure of human pre-mir-1

3. **Cleavage:** The cytoplasmic ribonuclease Dicer cuts the loop and tail off pre-miRNA, forming a double-stranded miRNA duplex of about 20-25 base pairs. The dashed lines in figure 1 indicate the sites at which human pre-mir-1 is cut by Dicer. The duplex is handed over to a group of proteins called the RNA-induced silencing complex [90], which discards one of the strands, leaving a mature single-stranded miRNA ready for mRNA-binding, as indicated in red in figure 1.

4. **mRNA-binding and degradation:** RISC binds miRNA to the 3’-untranslated region (3’UTR) of target mRNA that has near or perfect complementarity for the miRNA sequence. On miRNA-binding, RISC may cleave and degrade the targeted mRNA sequence, or simply block mRNA incorporation into the ribosome. The result is prevention of mRNA translation into functional protein [91, 92].

It is not unusual for several mRNAs to have very similar 3’-UTRs, meaning that a single miRNA can target multiple mRNAs. It is therefore possible that one miRNA can regulate the same molecular pathway at multiple points, or several different
pathways simultaneously. The relative abundances of a gene’s mRNA and the miRNAs that target it are thus important determinants of protein levels.

**Dicer is necessary for cell function**

As described above, Dicer is necessary for miRNA biogenesis. The importance of Dicer in miRNA regulation, immunity and general cell function has been described in a number of studies. Experimental knockout of dicer in specific cell-lines significantly alters cell function, often attributable to reduced mRNA targeting by specific miRNA: Developing T-lymphocytes have reduced cytokine expression, increased apoptosis and preference for Th1 phenotype [93]; Mature T-lymphocytes exposed to antigen have prolonged CD69 expression, impairing appropriate T-cell migration [94]; Macrophages fuse excessively to multinucleated giant cells [95]; Natural killer cells have poor survival, turnover and cell function [96]; Differentiation of B-lymphocyte progenitors is blocked, and mature B-cells show impaired antibody production [97, 98]. Knockdown of dicer to 30% activity in cell culture increased viral replication and apoptosis in influenza-infected cells [99]. A viable dicer knockout mouse line had trace Dicer activity in certain tissues. These mice showed increased susceptibility to vesicular stomatitis virus [100] and cytomegalovirus [101], and their Macrophages overexpressed interferon gamma induced genes [101].

In summary, Dicer-dependent production of miRNA is a prerequisite for normal immune cell development and function, indicating a vital role for miRNA-mediated regulation of gene expression.
microRNA regulation of immunological pathways

In the past 2 – 3 years many publications have focused on the role of miRNA in the immune system. miR-146a and miR-155 have well described functions and serve as good examples of how miRNAs can regulate immune responses.

miR-146a modulates activation of NF-κB, a key transcription factor within the innate immune system which is activated by many PAMPs, including RSV antigens. MIR-146a is one of the genes transcribed by NF-κB. Mature miR-146a targets the mRNA of TRAF6 and IRAK1, inhibiting production of these proteins. This is significant because TRAF6 and IRAK1 are proteins that positively regulate NF-κB signalling. The result of NF-κB mediated miR-146a transcription is therefore a negative feedback mechanism on NF-κB signalling [32, 102], regulating the inflammatory response.

miR-155 is also transcribed by NF-κB. In contrast to miR-146a, miR-155 is pro-inflammatory. Known miR-155 targets include SHIP1 in macrophages, which is an inhibitor of differentiation to active phenotypes. NF-κB activation therefore inhibits SHIP1 production via miR-155 transcription, promoting macrophage activation [78]. In dendritic cells, miR-155 targets KPC1 and SOCS1, which are proteins that inhibit dendritic cell maturation. Dendritic cell miR-155 expression increases production of IL12, which is a necessary co-factor for activating cytotoxic T-lymphocytes [77].

microRNAs associated with RSV

Interest in miRNAs as biomarkers for infectious diseases has increased, with recently published clinical studies of sepsis, tuberculosis, influenza, hepatitis B
virus, hepatitis C virus, parvovirus B19, hand foot and mouth disease and human immunodeficiency virus [103, 104].

Whilst there are no published clinical studies of miRNA profiles in RSV positive children, three groups have published studies of miRNA responses in cell-culture on infection with RSV. The pattern of regulation is cell-type dependent and includes:

i) Upregulation of let-7b, let-7c, let-7i, let-7f, miR-24, miR-26b, miR-30b, miR-337-3p and miR-520a-5p.

ii) Downregulation of miR-198, miR-221 and miR-595. Target genes were identified for several of these miRNAs.

These papers are examples of how studying RNA interference can improve understanding of the immune response to RSV, and deserve some attention. The methods used to identify and verify miRNA target genes are typical for other miRNA studies, including those I have cited in table 8 and in the Discussion. I have therefore included a description of these methods.

A data-mining strategy is often used to identify targets genes for miRNAs of interest. In this strategy, the miRNA sequence is compared with all known mRNA sequences in silico to find mRNA targets that are likely to be a good match for the miRNA. Because perfect complementarity is not required for a miRNA to be effective, the algorithms used must be relatively lenient. Thus, putative targets may not be real targets, and data-mining results should always be verified in the lab. The process can also be reversed in order to identify a set of miRNAs that putatively target a particular gene of interest. A number of internet applications
are freely available, and typically 3 – 5 databases are used to generate a list of putative targets.

Bakre et al. infected a human respiratory epithelial cell line (A549) with RSV and used a microarray approach with qPCR confirmation to identify regulated miRNAs. They found upregulation of let-7f, miR-24, miR-337-3p, miR-26b and miR-520a-5p and downregulation of miR-198 and miR-595. Let-7f expression was dependent upon the RSV G-protein, whilst miR-337 and miR-24 were inhibited by G-protein. Using data mining and published literature, the authors generated a set of possible let-7f target genes. Target gene confirmation using let-7f mimics and inhibitors identified CCL7, the immune suppressor SOCS3 and cell cycle genes CCND1, DYRK2 and ELF4 to be let-7f targets [105]. In conclusion, in A549 cells, RSV G-protein stimulates let-7f production, and may thus promote inflammation by suppressing SOCS3, modulate inflammation by suppressing CCL7, and also affect the cell cycle.

Piedemonte’s study group have investigated the relationship between RSV and neurotrophic pathways, which they have found to be dysregulated in RSV infected epithelium, and which may be a mechanism by which RSV modulates pulmonary epithelium resulting in an asthmatic phenotype. Normal human bronchial epithelial cells (HBECs) were infected with RSV and on microarray analysis 104 miRNAs were up or downregulated. Of these 104 miRNAs, data mining identified 6 that might target neurotrophic pathway genes, including miR-221, miR-453 and miR-574 (all downregulated on RSV infection). Cell cultures transfected with mimics and inhibitors of these miRNA were assessed for levels of Nerve Growth Factor and its receptors TrKA and p75NTR. Results
suggest that miR-221 targets NGF and TrKA, and that miR-453 targets p75NTR. miR-221 increased the rate of apoptosis both before and after RSV infection [106]. **In conclusion, RSV infection of HBECs may increase the expression of two neurotrophic genes via downregulation two miRNAs that target them.**

Finally, Thornburg et al. infected monocyte-derived dendritic cells (MDDCs) and HBECs with RSV and analysed miRNA expression using microarray analysis with qPCR validation. MDDCs upregulated let-7b whilst HBECs upregulated let-7c, let-7i and miR-30b, and downregulated miR-221. miRNA upregulation was RSV-dose related, and dependent on viral replication. let-7b levels in MDDCs were associated with RSV G-protein levels. MDDC let-7b expression and HBEC let-7i expression correlated with IFNβ levels. RSV NS1 and NS2 proteins inhibited HBEC let-7i and miR-30b. miR-30b expression was NF-κB dependent. The authors did not attempt to identify target genes for these miRNA [107]. **In conclusion, the authors have identified several miRNA that are induced by RSV and by RSV-derived proteins in two cell lines, and which be associated with type 1 interferon production.**

Direct RNA interference between virus and host

The literature describes three mechanisms of direct interaction between host and viral RNA that may aid viral replication or help the cell defend against attack:

1. Many DNA viruses encode viral miRNAs that directly target host mRNA, altering cell function. It is thought that such interactions are key in permitting long-lasting latent infections with herpes viruses. This mechanism is not thought to be possible for human RNA viruses (such as RSV) because there is a risk of cleavage of the viral genome at the site of encoded miRNA [92].
2. Hepatitis C virus (HCV) uses host miRNA as a protective mechanism. Hepatocellular miR-122 binds directly to HCV-RNA. Surprisingly, this results in impaired degradation of the HCV genome, promoting viral survival and replication [108, 109].

3. Host miRNAs can also directly target HCV-derived mRNA, impairing viral replication [110]. Similar interactions are seen for H1N1 influenza A virus [111] and Vesicular stomatitis virus [100].

RNA interference between RSV and the human hosts has not been identified, but may be a mechanism by which RSV modulates the immune system.

RNA interference as an anti-viral therapy

The principles of RNA interference have informed the development of several anti-viral drugs, including the anti-RSV drug ALN-RSV01, a synthesized small RNA. It functions in a similar manner to cellular-derived miRNA, and targets mRNA coding for the RSV N-protein, inhibiting production of viable virus particles. In clinical studies, nasal treatment with ALN-RSV01 improved symptoms in adults experimentally inoculated with RSV [112], and reduced the risk of serious outcomes in a small group of RSV-positive lung transplant recipients [113]. Drugs based on synthesized small RNA have also been developed to target mRNA of other viruses, including metapneumovirus [114], and Ebola virus [115].

Another RNA interference-based drug is Miravirsen. This is a small RNA sequence that has been developed to combat HCV infection. It is complementary to and has high affinity for human miR-122, which promotes HCV survival, as
described above. Miravirsen competitively binds to miR-122, preventing HCV-miR-122 binding. A phase 2a clinical trial of sub-cutaneous Miravirsen showed prolonged reductions in HCV-RNA levels that were dose-dependent, without significant adverse side effects [116].

Factors associated with RSV disease severity

Most infants do not experience dyspnoea or other severe symptoms on infection with RSV. A number of factors are associated with RSV infection and presentation to hospital, but studies have typically not differentiated between varying degrees of severity. Attempts to adequately explain why some children develop severe lower airways disease, whilst most children only develop upper airways symptoms, are the subject of on-going debate. I will here give an overview of factors associated with RSV infection in clinical studies.

Previous medical history

Patients with certain clinical conditions are at higher risk of severe disease and hospital admission on exposure to RSV. These conditions include infants with low birth weight, congenital heart disease (CHD), bronchopulmonary dysplasia or other lung pathologies, immunodeficiencies and Down syndrome [9, 117, 118]. However, most children admitted with RSV do not have one of these conditions, but have otherwise appeared healthy.

Timing of birth and maternal antibodies

The timing of birth in relation to yearly RSV epidemics impacts the risk of hospital admission. Birth 0 - 3 months before the epidemic is a risk factor for hospitalization with lower respiratory tract infection, whilst birth 0 – 3 months
after the epidemic is a protective factor [119]. Explanations for this phenomenon include young age during the first RSV season – an immature immune system may mean that the response to the virus is weakened or has an inappropriate Th1/Th2 balance [120]; and there is a smaller airways diameter, such that young infants are more prone to blockage by mucous secretions, increasing dyspnoea. However, there may also be other immunological explanations. Maternal RSV antibody titres measured in cord blood increase during RSV epidemics. Infants born immediately after an epidemic will therefore have higher levels of protective antibody transferred from the mother, increasing the chance of neutralizing the virus before severe disease develops. Immediately prior to the RSV season, titres of maternally transferred RSV antibody are at low, reducing the chance of neutralizing the virus [121]. This accumulation of susceptible, young infants prior to the epidemic increases the reproductive number for RSV and may therefore in itself be a factor that contributes to the initiation of an RSV epidemic.

Other epidemiological factors – problems with disease definition

A literature review by Simoes in 2003 concluded that male gender, age < 6 months, birth during the first half of the RSV epidemic, and increased infectious burden (e.g.: older siblings, overcrowding or day care exposure) increase the risk of RSV infection [122]. More recent epidemiological studies have in addition associated RSV infection with birth weight > 4 kg [123]; low cord blood vitamin D levels [124]; Inuit, Maori or Pacific ethnicity [125, 126]; and smoking during pregnancy [126]. Breastfeeding has also been investigated as a risk factor, but with conflicting results. One recent study concludes that breastfeeding does not
protect against hospitalization, but shortens hospital stay and reduces use of oxygen [127]. Another study found that breastfeeding protected against hospitalization [128] in children with acute respiratory infection, whilst breastfeeding in a prospective birth cohort did not protect against LRTI on natural infection with RSV [123]. Simoes [122] concluded that whilst breastfeeding was likely to be protective for lower respiratory tract infection (LRTI), the role in protection specifically against RSV disease was less clear. This was attributed to poor descriptions of the degree of breastfeeding and lack of precise outcomes in the studies reviewed. When corrected for other confounders, the protective effect of breastfeeding disappeared.

A common weakness of epidemiological studies is that RSV disease severity has usually not been defined by factors other than the presence or absence of wheeze, cough, dyspnoea or admission to hospital. Infants with such symptoms may well have a relatively mild course of infection without need for oxygen, fluids or other interventions. Infants may also be admitted for observation because of young age and fear of disease progression, not because of actual symptoms. One study did categorize 195 previously healthy RSV-positive infants according to disease severity (ambulatory, admission to the paediatric ward or admitted to the intensive care unit). The investigators found no association between severity of disease and a range of epidemiological factors including gender, gestational age, birth weight, age at presentation, ethnicity, family history of asthma, exposure to cigarette smoke, breastfeeding, number of siblings, family pets and day-care attendance [129]. As exemplified by this study, future
investigations of epidemiological risk factors should define RSV disease severity more precisely than simple presentation to hospital or the presence of dyspnoea.

Genetic polymorphisms
In the last decade many genetic associations with RSV infection have been described, in particular single nucleotide polymorphisms (SNPs) of genes important to the immune system, such as CXCR1, CX3CL1, CCL5, CCR5, TLR4, surfactant proteins B and D, interleukins 4, 8, 10, 13 and 18, IL1 receptor-like 1 (IL1RL1), the vitamin D receptor, and vitamin D binding-protein [38, 39, 130-144].

These SNPs have most commonly been identified by genetic analysis of RSV-positive infants (irrespective of disease severity) compared to a random selection of controls. Two interesting SNPs are the Asp\textsuperscript{299}Gly and Thr\textsuperscript{399}Ile mutations of TLR4, which have been found in 5% of ambulatory RSV patients, 20% of patients with severe disease [38], and over 85% of prematurely born infants admitted with RSV [145]. In-vitro transfection of human bronchial epithelial cells with these SNPs impaired expression of TLR4 on cell surfaces and reduced NF-κB signalling on RSV-stimulation. [146]. The same study showed impaired RSV-responses in peripheral blood monocytes from TLR4\textsuperscript{299}Asp or \textsuperscript{399}Thr positive children.

Nasal cavity viral load
Studies that quantify the nasal levels of RSV have associated viral load on presentation to hospital with increased clinical severity score on presentation, increased duration of hospitalization, respiratory failure and requirement for intensive care [147-151]. In contrast, one study that also included older children
found that higher RSV viral load correlated with shorter hospital stay and lower respiratory rate [152]. The quantity of RSV in the nasal cavity can conceivably be associated with increased disease severity by several mechanisms:

i) Impaired or inappropriate immunity may allow greater viral replication.

ii) Increased viral load may over-stimulate the immune system, leading to negative effects including excessive mucous secretion and plugging, bronchoconstriction and reduced capacity for gas-transfer.

iii) Greater quantities of RSV in respiratory droplets derived from the nose may increase the risk of viral spread to the lower respiratory tract.

Immunological factors

Analysis of immunological markers can identify immune-phenotypes that predispose to more severe disease. Differences in cytokine levels at the time of disease have been associated with disease severity. However, such results should be interpreted with caution because differing levels of cytokines may be the result of confounding factors that predispose to increased disease severity, for example young age or prematurity [153], and might not necessarily be the cause of this increased severity. Immune-phenotypes identified prior to RSV infection may give more information about predisposing factors associated with disease severity.

*Nasal wash cytokine levels during illness:* Hypoxic bronchiolitis is associated with higher levels of the neutrophil chemoattractants CCL3 and IL8 in nasal secretions [53, 154]. Higher CCL3 and lower IFNγ levels in nasal secretions are associated with longer duration of oxygen treatment [155]. Nasal IFNγ is also significantly less abundant in ventilated compared to non-ventilated infants
[156], whilst nasal IL1 receptor like 1, which promotes Th2 responses, is increased in ventilated infants [142].

In a recently published 10-year longitudinal study of 851 RSV-positive children aged 18 days to 5 years (median age 4 months), 30 cytokines were analysed in nasal washings taken on admission to hospital. 268 were treated as outpatients, 503 were admitted and treated on the ward, and 80 were admitted to intensive care. Increased levels of IL1β, 6, 7 and 8; IL1 receptor antagonist; IL2 receptor; TNFα; IFNα; CCL2, CCL3 and CCL4 were associated with increased disease severity [157].

These studies support each other and the cytokine profiles associated with more severe disease are known to promote a greater influx of neutrophils, dendritic cells, and other innate immune cells, and reduced cellular immune responses.

**Plasma cytokine levels and in-vitro stimulation of peripheral blood leukocytes at the time of illness:** Bont et al. found increased plasma IL8 in ventilated RSV-positive patients, but similar plasma IL4, IL12 and IFNγ levels compared to non-ventilated patients, suggesting a more potent innate immune response, but similar adaptive immune response [158]. However, in a similar experiment that also included a healthy control group, Mella et al. found no difference in plasma concentrations of TNFα, IL6, IL8 or IL10 between disease groups [159].

On stimulation of peripheral blood lymphocytes, Bont et al. found a markedly lower lymphoproliferative response and lower IFNγ and IL4 production in ventilated patients [158]. IL12 production on stimulation of peripheral blood
monocytes was inversely correlated to duration of mechanical ventilation [160]. IFNγ and IL12 promote Th1-mediated cellular immunity, which is physiologically impaired during the first month of life [120]. These studies suggest that persistence of this impairment may predispose to more severe RSV disease [161].

Mella et al. stimulated peripheral whole blood with LPS, and found reduced TNFα and IL8 production capacity in patients admitted to ICU. These findings may indicate poorer innate immune capacity in those with most severe disease [159].

The responses observed on stimulation of peripheral blood cells may be the result of immune exhaustion during the course of a serious disease, but may also represent immunological factors that predispose to more severe disease. If the latter is true, then it should be noted that ICU patients represent a small minority of all those admitted to hospital, and whether observed differences are indicative of a predisposing immune-phenotype common to other patient groups (for example hypoxic patients without need for ICU) is not yet clear.

The difference in peripheral blood leukocyte response in children with more severe disease is also interesting because it confirms that peripheral blood can be used to draw conclusions about an airways infection. This is of particular relevance to papers 1 and 2 in this project.

*Immunological investigations at birth:* Analysis of cord blood in children who later present with RSV is relevant because birth stimulates the immune system [162], and differences in immune regulation at birth may inform how the
immune system responds to later challenges. However, the site of RSV infection is not in the blood, but in the airway mucosa, and this factor should be considered when interpreting results.

Our study group examined mRNA expression in the cord blood of infants subsequently admitted with RSV [163], and found downregulation of TNF receptor superfamily 25, a receptor that is a necessary co-factor for dendritic cell antigen presentation to T-lymphocytes and for effective T-cell function [164, 165]. Interestingly, gene set enrichment analysis showed an almost significant downregulation (adjusted p = 0.08) of the gene set transcribed by NF-κB, although the mechanism behind this downregulation was not assessed.

Junnti et al. compared cytokine production in cord blood mononuclear cells, comparing 14 infants who were later admitted with RSV disease to 48 infants who did not have a respiratory tract infection before the age of 6 months [166]. Unstimulated cord blood mononuclear cells showed increased IFNγ and IL4, and reduced IL6 and IL12 compared to the healthy group. On LPS stimulation they found decreased IFNγ and increased IL6 production. Differences in immune regulation already present at birth, in many cases months before RSV challenge, may therefore predispose infants to a dysfunctional immune response and increased disease severity on infection with RSV.

Altogether, these results point towards differences in innate and adaptive immune responses that may result in more severe disease in RSV positive patients.
So why study the immune system in children with RSV?

RSV causes significant morbidity, and treatment is currently largely supportive.

1. Efforts to develop effective strategies for treatment and prevention focus on virus-host interactions. RSV is an interesting virus because the adaptive immune system is not able to create effective long-lasting protection. Reinfection throughout life is common, although not usually as severe as the first infection.

   In this context, improved understanding of the mechanisms governing the host response to RSV infection is necessary, including factors within the immune system associated with different disease outcomes. Deficiencies that allow the virus to avoid effective immune-surveillance or immune-memory in otherwise healthy infants are vital clues that will help the pharmaceutical industry to develop effective vaccines.

2. The availability of monoclonal antibodies as prophylaxis leads to questions about which group of infants should receive prophylaxis. Certain groups of high-risk infants can already be identified according to clinical parameters. If we can in addition use immunological parameters to identify a group of at risk children who otherwise appear healthy at birth, we may be able to further reduce morbidity. While the cost of anti-RSV monoclonal antibodies is currently inhibitive, this may be relevant in the future.

3. Understanding the immune response to RSV will also inform us about the virus-host interaction and immune response to other respiratory viruses such as
Influenza virus, Metapneumovirus, Corona viruses or Rhinoviruses, against which effective treatment strategies have also not yet been developed.
3. Aims and hypotheses

Primary aim

To discover differences in the immune system predisposing infants to RSV disease severe enough that they present to health care providers and are tested for RSV. Specifically, we chose to investigate gene expression at birth and immune regulation by RNA interference during infection.

Secondary aims

1. To discover differences in the immune system that predispose to more severe disease amongst those testing positive for RSV.
2. To identify miRNA that are involved in clinical RSV disease.

Hypothesis 1 – Dicer expression at birth

Infants who present with RSV disease express less Dicer in cord blood, than infants who do not present with RSV disease. Infants with more severe disease have more downregulation than infants with milder disease.

As discussed in the introduction, downregulation of dicer may have effects on immune cell function, expression of immune-associated host or viral genes, and may also have effects on virus-host interactions that are mediated by RNA-interference.

Methods used:

i) qPCR of RNA isolated from cord blood using Dicer specific primers.

ii) Western blot confirmation of qPCR findings.
Hypothesis 2 – Cytokine expression at birth

Compared to other infants, those who present to health care services with RSV disease have up- or downregulation in cord blood of the following PRRs, cytokines and cytokine receptors relevant to the immune response to RSV: CCR3, CCR7, CXCL7, CXCL11, CXCR3, IFNγ, IL4, IL7R and TLR4. Infants with more severe disease have more pronounced up- or downregulation than infants with mild disease.

Birth is a stressful event and leads to changes in expression of immune genes [162]. Differences in immune gene expression at birth may therefore predict immune gene expression on exposure to RSV.

Methods used:

i) qPCR of cord blood RNA using primers for specific cytokines/ receptors.

ii) Western blot confirmation of qPCR findings.

Hypothesis 3 – miRNA expression during RSV disease

Infants infected with RSV have up- or downregulation of immunologically relevant miRNA in their nasal mucosa, compared to healthy uninfected controls. miRNA expression is associated with disease severity.

miRNA expression has not previously been profiled in clinical RSV disease.

Methods used:

i) miRNA microarray of RNA isolated from nasal mucosa.

ii) qPCR confirmation of miRNA differentially expressed in the microarray.
A note on hypothesis generation

The background for hypotheses 1 and 2 was a birth cohort founded in 2003 by our research group [163]. The study included an mRNA microarray analysis of umbilical cord blood in which gene expression of 5 infants who presented to hospital with confirmed RSV disease were compared to five infants who were seropositive for RSV at 12 months of age, but in whom the parents did not report lower respiratory tract disease. Findings from this study are described above in the chapter “Factors associated with disease severity – immunological factors”. The first step in the current project was a re-analysis of the microarray data using a more advanced statistical method, Linear models for microarray (LIMMA). The results were mined for differences in expression of immune regulatory genes (other than those previously published by our study group). There was evidence for differences in expression of *CCR7, CXCL7, CXCR3, Dicer* and *IL7R*. The following genes considered relevant to RSV disease were not included in the microarray results: *CCR3, CXCL11, IFNy, IL4* and *TLR4*. This omission might be due to technical issues, because the gene was not included in the microarray, or because the gene is not expressed in cord blood. Genes for other cytokines and receptors associated with RSV disease (see table 1) were either not differentially expressed in the microarray, or else were at the time not included in our overview of RSV-relevant genes. This comparison of microarray results with the literature led us to hypotheses 1 and 2.

Hypothesis 3 followed confirmation of hypothesis 1. Since *Dicer* was downregulated at birth in infants with severe disease but not in infants with mild disease, we thought that this downregulation might still be present on infection.
If so, this might affect the miRNA profile on exposure to RSV, and indicate ways in which differences in immune regulation might lead to different disease states.
4. Methods

Hypotheses 1 and 2 were tested using cord blood from the birth cohort founded in 2003 by our research group. We tested hypothesis 3 by collecting new clinical samples from infants with current RSV disease assessed at the hospital in 2011. I will first describe data collection methods pertaining to the 2003 birth cohort, and thereafter the 2011 disease cohort. Finally, I will describe analytical methods common to all three hypotheses.

Birth Cohort 2003

Biobank – samples available for analysis

Umbilical cord blood from 2108 healthy neonates was collected between January 2003 and February 2004. There were approximately 3500 births during this period. Samples are stored in the research biobank BarnAhus, Norwegian biobank registration number 314, kept at Akershus University Hospital. Three samples were collected from each infant: an RNA collection tube suitable for mRNA analysis; a serum sample suitable for standard clinical chemistry and protein analysis; and an EDTA sample, which was intended for DNA analysis. At the time, mRNA analysis as a standard research tool was relatively new, and protein confirmation of mRNA expression was not considered necessary. Samples specifically for protein analysis were therefore not gathered. I will discuss this issue further in the chapter “Protein analysis by Western Blot.”

Identification of RSV positive infants and controls

Members of our study group cross-referenced the 2003 cohort with the electronic microbiological database at our hospital to identify which of the 2108
children tested positive for RSV before the age of 3 years. During 2003 and 2004, testing for RSV using nasopharyngeal PCR in clinical practice was relatively new. As a part of the 2003 project, all infants assessed at the hospital with suspected viral lower airways infection January 2003 – February 2005 were cross-referenced with the birth cohort in order to identify study participants. Study participants were tested for RSV, irrespective of the decision to admit. Non-study patients assessed with suspected RSV disease were routinely tested for RSV using a direct antigen test or PCR if they were admitted. This practice limited nosocomial RSV infection by enabling cohort isolation of RSV positive infants. Patients sent home without admission were not routinely tested for RSV. Although popularity was increasing, RSV-testing was not common practice amongst general practitioners in the area.

70 (3.3%) study participants were RSV positive before 3 years of age, in keeping an epidemiological study from our group estimating the incidence of RSV disease requiring hospital admission in our area during the 1990s [9]. In paper 1 we included only children aged < 12 months of age at the time of RSV infection. In paper 2 we also included 10 children who were 12 – 36 months of age at the time of RSV infection.

Controls were chosen at random from the remaining 2038 children who did not deliver a positive RSV test during the first three years of life. Thus the study had a nested case-control design.

Clinical Data

Clinical data was collected by retrospective assessment of the hospital medical records of RSV-positive cohort participants and controls. Exclusion criteria for
controls and RSV-positive infants are described in papers 1 and 2. A summary of RSV-positive study participants is given in figure 3.

We assessed the medical records pertaining to the RSV infection and recorded objective parameters of disease severity such as occurrence of apnoea, use of supplemental oxygen, fluid administration or admission to intensive care. In addition, we assessed the degree of respiratory distress during admission by reading and interpreting the attending physician’s clinical description, which was complete in all cases. Finally we recorded the following objective measures that might be differ between disease severity groups: lowest recorded SpO2, highest recorded respiratory rate and pCO2, pulse on admission, weight, duration of symptoms, length of stay and discharge diagnosis.

Controls were not followed up specifically during the study period. It is likely that they were exposed to RSV during their first two years, but were not sufficiently ill to be referred to hospital. We do not have clinical data about their probable RSV infection, or know at which time it may have occurred.

Power calculation

We used microarray data from the 2003 cohort to estimate the number of individuals required in each group to achieve a power of 80% at a significance level of 5%. Specifically, we used the standard deviation for CXCL7, which had the least statistical significance in the microarray of the differentially regulated genes, and defined biological significance as a 20% up- or downregulation. We assumed that we would be comparing cases to controls using a T-test. The power calculation indicated a need for 11 cases in each group.
Cord blood from 2108 newborns

70 Confirmed RSV disease

64 Included

47 Sufficient RNA

25 severe disease
22 mild disease

37 age < 12 months
10 age > 12 months

6 excluded
Asthma
Cleft palate
Down syndrome
Recurrent Pneumonia

Figure 3: RSV positive study participants in the 2003 study cohort.
Disease Cohort 2011

In the 2011 cohort we wanted to come closer to the site of infection and therefore analysed nasal respiratory epithelial samples. Since children with both mild and severe RSV disease have similar nasal symptoms, with profuse coryza, we considered the nasal epithelium to be a reasonable tissue in which to compare disease severities.

Study Population

Clinical cases: We considered for inclusion infants aged < 12 months referred to the paediatric emergency ward at our hospital with suspected viral upper or lower airways disease January – April 2011. Patients were included if they were RSV-positive on rapid antigen or PCR testing. Exclusion criteria were birth at gestational age < 34 weeks, chronic lung disease, diagnosed asthma or previous bronchial obstruction, previous lower respiratory tract infection, congenital heart disease, Down syndrome, or other neurological diseases including hypotonia or developmental delay. In addition, we excluded patients who had confirmed co-infection with Influenza virus, Parainfluenza virus, Metapneumovirus, Mycoplasma pneumoniae or Chlamydia pneumoniae.

The emergency ward nurses were responsible for inclusion of patients, and recruited a total of 139 infants. 120 were positive for RSV, and of these 110 met the inclusion criteria. Figure 4 summarizes study participants in 2011.

Controls: We recruited healthy controls < 12 months of age from family health clinics in the hospital’s catchment area September – December 2011, with permission from the relevant municipal authorities. Exclusion criteria for controls were the same as for clinical cases: In addition, we excluded controls if
they had symptoms suggestive of an acute upper or lower respiratory tract infection, or if they appeared ill in any other way. We considered sneezing or chronic minor nasal congestion from birth to be a normal variant and included these infants.

In our area, healthy infants have routine appointments with the health visitor and/or doctor at the family health clinic at ages 6 weeks, 3 months, 5 – 6 months, 9 months and approximately 12 months. In addition, infants commonly visit the family health clinics to be weighed, to participate in maternity or educational groups, or to be seen by the physiotherapist. We aimed to match as closely as possible the age distribution of control infants to clinical cases by enrolling infants in the following order of preference: age 0 – 3 months > age 4 – 6 months > age 6 -12 months, although this was not done systematically.

Controls samples were continually assessed for RNA quality during enrolment, and we continued including controls until we considered that we had enough samples of sufficient RNA quality. Altogether we collected control samples from 101 infants. How control samples were handled is further described in the chapter “Sample collection and processing”

Clinical data collection

The attending doctor and nurse filled out a standardised clinical data form on the first assessment of clinical cases, and we completed the form during admission, following up each patient prospectively. The form is reproduced in appendix 1, showing details of which parameters were recorded.
139 with suspected viral respiratory infection

120 RSV positive

10 excluded:
3 Co-infections
4 Premature
1 Previous bronchial obstruction
1 Previous lower respiratory tract infection
1 Chronic liver disease

110 Included

105 RNA samples

5 Only nasal cytology smear taken

40 mild
25 moderate
40 severe

24 mild
11 moderate
25 severe

16 mild
7 moderate
19 severe

19 RSV negative

80 control infants

38 individual samples
21 paired samples

31 control samples

45 RSV positive samples

18 RSV positive samples

45 control samples

19 control samples

28 control samples

19 control samples

31 control samples
Controls and clinical cases sent home without admission were not followed up further in the study, and we assumed that they did not develop any complications or receive any treatments mentioned above. We reviewed the electronic medical record of each clinical case 1–2 weeks after discharge to ensure that patients were not readmitted with complications.

Confirmation of RSV Infection

Clinical cases were tested for the presence of RSV in a nasopharyngeal aspirate by rapid antigen testing on the emergency ward or by PCR at the hospital’s

Table 2: Clinical cases positive for RSV or other viruses in 2011

<table>
<thead>
<tr>
<th></th>
<th>Positive test</th>
<th>Total tested</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any RSV test</td>
<td>120</td>
<td>137</td>
</tr>
<tr>
<td>RSV Rapid Antigen test</td>
<td>73</td>
<td>103</td>
</tr>
<tr>
<td>RSV PCR</td>
<td>114</td>
<td>131</td>
</tr>
<tr>
<td>PCR for other viruses*</td>
<td>11</td>
<td>131</td>
</tr>
</tbody>
</table>

Note: 2 cases were not tested by rapid antigen test or PCR, presumably because samples were lost en route to the Department of Microbiology and Infection Control.

* Other viruses tested for: *Influenza virus, Parainfluenza virus, Metapneumovirus, Mycoplasma pneumoniae and Chlamydophila pneumoniae*

Table 3: Occurrence of viral co-infection in 2011

<table>
<thead>
<tr>
<th></th>
<th>PCR for other viruses</th>
<th>Not PCR tested</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive</td>
<td>Negative</td>
</tr>
<tr>
<td>RSV PCR or Rapid antigen test</td>
<td>Positive</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>8</td>
</tr>
</tbody>
</table>

Note: There were 3 cases of confirmed viral co-infection. 6 cases were only tested by rapid antigen test, and were therefore not tested by PCR for other viruses.

* Other viruses tested for: *Influenza virus, Parainfluenza virus, Metapneumovirus, Mycoplasma pneumoniae and Chlamydophila pneumoniae. We cannot exclude co-infection with viruses not tested for, e.g.: Rhinovirus, Coronavirus, Bocavirus.*
Department of Microbiology and Infection Control. Samples tested by PCR were in addition tested for the presence of Influenza virus, Parainfluenza virus, Metapneumovirus, Mycoplasma pneumoniae and Chlamydophila pneumoniae. Rapid antigen and PCR results for 139 clinical cases are summarised in tables 2 and 3.

Sample collection and processing
Studies of nasal cytokine levels typically analyse nasal wash samples. miRNA can also be found in the extracellular compartment, but in smaller quantities and probably with different functions – for example cell signalling. In addition, Ribonuclease (RNase) present in mucous secretions make reliable RNA quantitation more difficult in nasal wash samples. A study of cytology brushings in 5 children with cystic fibrosis and 10 healthy control children showed that harvested cells were viable and could be grown in cell culture [167], suggesting that it is also possible to assess miRNA profiles in harvested nasal mucosa. We therefore chose to discard nasal wash samples and took cytology specimens instead, using a cytology brush commonly used for cervical examinations.

Cytology sample collection procedure: Cytology specimens were taken from the atrium of the nasal cavity on each side, inserting the brush about 2 cm and gently rotating the brush over the entire circumference of the cavity. A separate brush was used for each nostril. Brushes were immediately immersed in RNALater solution, which prevents degradation or new production of RNA and therefore provides a “freeze-shot” of the transcriptome at the time of sampling. The method was first tried out on myself, and we confirmed isolation of adequate amounts of good quality miRNA. Next we tried the method on 2 infants with
acute viral respiratory disease, also with positive results. We therefore decided that this method would be suitable to analyse miRNA expression in nasal mucosa.

**Cytology smears:** In order to check which cells we were harvesting, we sacrificed some cytology samples. Instead of immersing in RNA later, we smeared the sample onto to a cytology slide and asked the pathologist to describe the slides. Smears were taken from infants with both mild and severe disease. Because RSV status was not known at the time of sampling, some smears were taken from amongst the 19 infants who did not have confirmed RSV disease. The pathologist concluded that there were no significant cytological differences between smears. In all smears, there was an abundance of squamous epithelial cells, ciliary epithelial cells, and granulocytes. It is important to note that the miRNA profiles described in this study can therefore not be ascribed to one cell type alone, but will be affected by at least two epithelial cell types, one immune cell type, and probably several other cell types.

**RNA isolation:** RNA was isolated using a kit bought from Qiagen (Hilden, Germany) specifically designed to capture miRNA. Any mucous was carefully removed prior to RNA isolation to ensure as little RNase activity as possible.

**Pairing of control samples:** A number of the control samples gave a poor RNA yield, which restricted the suitability of the sample for microarray and PCR. We isolated RNA from samples shortly after collection, and were able to identify this as a problem during the second half of the collection period. We therefore paired samples at a stage of RNA isolation that allowed us to increase the RNA concentration in the final output. Criteria for combining samples were that they should be i) of the same gender; ii) no more than 31 days age difference; iii) no
more than 1 kg weight difference. The average weight and age in months was calculated. In the final statistical analysis, 10 control samples were paired samples and 9 control samples were individual samples. Because we did not know the relative contributions of each sample to the combined sample, paired samples were considered as one sample and not weighted prior to statistical analysis.

RNA quality control

We assessed RNA purity and integrity using the 260/280 ratio, the 260/230 ratio and the RNA integrity number (RIN) using the NanoDrop ND-1000 Spectrophotometer (NanoDrop Technologies) and Agilent 2100 Bioanalyzer. These quality measures are commonly utilized in mRNA studies. However, as discussed below, little is known about the relevance of these measures for miRNA expression. I will here summarize the RNA quality measures of the nasal cytology samples that were analyzed by qPCR for the expression of 32 miRNA, because this allows accurate comparison of miRNA expression to RNA quality measures. Of 88 samples analyzed by qPCR, 6 had consistently poor Cq values across a range of miRNAs and were removed prior to statistical analysis.

**i) 260/280 ratio:** This is a measure of sample contamination with protein using spectrophotometric analysis of light absorption at 260 and 280 nm wavelengths. Nucleic acids absorb light at 260 nm and proteins at a wavelength of 280 nm. An ideal ratio is close to 2.0. The 260/280 ratios were satisfactory with a mean > 1.9 for each group, indicating little protein contamination in our samples. We did not pursue further analysis of miRNA expression compared to 260/280 ratio.
ii) 260/230 ratio: This is a measure of sample contamination with organic compounds, which absorb light at 230 nm. A lower ratio therefore indicates greater contamination. Empirical recommendations suggest that ratios should be close to 2.0, ideally > 1.7. Organic compounds are present in all samples prior to RNA isolation, and are also used as part of the isolation process. Steps are taken to remove organic compounds during RNA isolation and purification. Organic compounds may reduce the efficiency of reverse-transcription (RT), and therefore reduce the measured expression of target RNA.

260/230 ratios in the remaining 82 samples were low indicating organic contamination (Control mean 1.1; 95% CI 0.71 – 1.49. RSV-positive mean 1.18; 95% CI 0.98 – 1.38. p = 0.31, Student’s t-test). The contamination in our samples is most likely to be due to trace amounts of guanidine thiocyanate, which is a constituent of the TRIzol reagent used to isolate RNA. Normalization of samples using a reference gene can help correct this error since RT of target genes and reference genes within a sample are likely to be equally affected by organic compound contamination. To our knowledge, few investigators have studied the relationship between contamination with organic compounds and RT reaction inhibition, and there is no consensus on an acceptable lower limit for the 260/230 ratio. Cicinatti et al. investigated the suitability of various genes as reference genes in clinical samples and cell lines of hepatocellular carcinoma. The 260/230 ratio ranged from 0.43 – 2.23. They found no correlation between 260/230 ratio and Cq values [168]. von Ahlfen and Schlumpberger intentionally contaminated RNA samples with guanidine thiocyanate, and showed that the 260/230 ratio fell even with trace amounts of contaminant. A one step RT-qPCR
reaction was robust despite considerable guanidine thiocyanate concentrations. RT-qPCR reaction inhibition first occurred at a cut-off contamination level that was over 100 times that required for a 260/230 ratio < 0.5 [169].

Given this disparity between empirical recommendations and published literature, we chose to investigate the relationship between normalized miRNA expression data and 260/230 ratio in both control and RSV-positive samples, seeking to identify a cut-off 260/230 ratio below which the RT reaction is inhibited. Results are shown in tables 4 (control samples) and 5 (RSV positive samples) and exemplified by 4 miRNAs shown in figures 5 and 6. We found a positive correlation between miRNA expression and 260/230 ratio for most miRNAs in the RSV positive samples, but not in control samples. Examination of scatter plots (figures 5 and 6, upper panels) indicated that this correlation was in all cases likely due to markedly reduced miRNA expression at a 260/230 ratio < 0.6. We therefore introduced a cut-off minimum 260/230 ratio of 0.6 (figure 5 and 6, lower panels). A total of 21 samples were removed from the analysis, including 9 control and 12 RSV positive samples. On re-analysis, the correlation between miRNA expression and 260/230 ratio disappeared in all cases (table 5), supporting the introduction of this cut-off in our samples.

There were still many samples with a 260/230 ratio lower than recommended. However, since the contamination was evenly distributed between groups, we considered that it would not affect comparisons between groups other than a possible increase in confidence interval. This could increase the rate of type 2 errors (failure to reject a null hypothesis).
Table 4: Linear regression analysis of control sample miRNA expression vs. 260/230 ratio, before and after application of a cut-off 260/230 ratio ≥ 0.6

<table>
<thead>
<tr>
<th>miRNA</th>
<th>All control samples</th>
<th>Samples with a 260/230 ratio ≥ 0.6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Slope (95% CI)</td>
<td>Significance</td>
</tr>
<tr>
<td>miR-16</td>
<td>0.6 (-0.2 - 1.5)</td>
<td>p = 0.1</td>
</tr>
<tr>
<td>miR-19</td>
<td>0.7 (-0.2 - 1.6)</td>
<td>p = 0.1</td>
</tr>
<tr>
<td>miR-21</td>
<td>0.7 (-0.1 - 1.5)</td>
<td>p = 0.1</td>
</tr>
<tr>
<td>miR-23b</td>
<td>0.5 (-0.6 - 1.7)</td>
<td>p = 0.3</td>
</tr>
<tr>
<td>miR-26b</td>
<td>0.8 (0.0 - 1.6)</td>
<td>p = 0.05</td>
</tr>
<tr>
<td>miR-27b</td>
<td>0.4 (-0.7 - 1.5)</td>
<td>p = 0.4</td>
</tr>
<tr>
<td>miR-29c</td>
<td>0.1 (-0.5 - 0.8)</td>
<td>p = 0.6</td>
</tr>
<tr>
<td>miR-30b</td>
<td>0.4 (-0.3 - 1.2)</td>
<td>p = 0.2</td>
</tr>
<tr>
<td>miR-30d</td>
<td>0.2 (-0.5 - 0.9)</td>
<td>p = 0.5</td>
</tr>
<tr>
<td>miR-31</td>
<td>0.2 (-0.4 - 0.8)</td>
<td>p = 0.5</td>
</tr>
<tr>
<td>miR-34b</td>
<td>-0.2 (-1.1 - 0.8)</td>
<td>p = 0.7</td>
</tr>
<tr>
<td>miR-34c</td>
<td>-0.1 (-1.1 - 0.9)</td>
<td>p = 0.9</td>
</tr>
<tr>
<td>miR-96</td>
<td>0.6 (-0.7 - 1.9)</td>
<td>p = 0.3</td>
</tr>
<tr>
<td>miR-125a</td>
<td>-0.01 (-0.4 - 0.4)</td>
<td>p = 1.0</td>
</tr>
<tr>
<td>miR-125b</td>
<td>0.1 (-0.9 - 1.0)</td>
<td>p = 0.9</td>
</tr>
<tr>
<td>miR-130a</td>
<td>0.7 (-0.3 - 1.7)</td>
<td>p = 0.1</td>
</tr>
<tr>
<td>miR-146a</td>
<td>-0.2 (-1.1 - 0.7)</td>
<td>p = 0.7</td>
</tr>
<tr>
<td>miR-148a</td>
<td>0.2 (-0.7 - 1.1)</td>
<td>p = 0.6</td>
</tr>
<tr>
<td>miR-155</td>
<td>0.6 (-0.3 - 1.5)</td>
<td>p = 0.2</td>
</tr>
<tr>
<td>miR-183</td>
<td>-0.1 (-1.0 - 0.9)</td>
<td>p = 0.9</td>
</tr>
<tr>
<td>miR-200b</td>
<td>0.2 (-0.4 - 0.8)</td>
<td>p = 0.5</td>
</tr>
</tbody>
</table>
Table 4 (continued)

<table>
<thead>
<tr>
<th>miRNA</th>
<th>All control samples</th>
<th>Samples with a 260/230 ratio ≥ 0.6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Slope</td>
<td>(95% CI)</td>
</tr>
<tr>
<td>miR-203a</td>
<td>0.3</td>
<td>(-0.4 - 0.9)</td>
</tr>
<tr>
<td>miR-205</td>
<td>-0.1</td>
<td>(-0.9 - 0.7)</td>
</tr>
<tr>
<td>miR-223</td>
<td>0.3</td>
<td>(-1.2 - 1.8)</td>
</tr>
<tr>
<td>miR-324</td>
<td>-0.6</td>
<td>(-1.2 - 0.1)</td>
</tr>
<tr>
<td>miR-331</td>
<td>-0.02</td>
<td>(-0.4 - 0.4)</td>
</tr>
<tr>
<td>miR-375</td>
<td>0.04</td>
<td>(-0.8 - 0.9)</td>
</tr>
<tr>
<td>miR-429</td>
<td>0.7</td>
<td>(0.0 - 1.4)</td>
</tr>
<tr>
<td>let-7d</td>
<td>0.5</td>
<td>(-0.4 - 1.4)</td>
</tr>
<tr>
<td>let-7f</td>
<td>1.0</td>
<td>(0.0 - 1.9)</td>
</tr>
<tr>
<td>let-7g</td>
<td>0.4</td>
<td>(-0.3 - 1.1)</td>
</tr>
<tr>
<td>let-7g</td>
<td>0.4</td>
<td>(-0.3 - 1.1)</td>
</tr>
<tr>
<td>let-7i</td>
<td>0.3</td>
<td>(-0.5 - 1.2)</td>
</tr>
</tbody>
</table>

Note: There is no significant correlation between miRNA expression and 260/230 ratio neither before or after application of the 260/230 ratio ≥ 0.6 cut-off. miR-26b, miR-324, miR-429 and let-7f have p < 0.1 before cut-off application. For miR-324, this finding remains after cut-off application. However, the slope is negative, indicating decreasing expression with increasing 260/230 ratio, which is the opposite of what is expected if organic contamination inhibits the qPCR. This phenomenon may be explained by the fact that significance values are not adjusted for multiple testing. Fold change was used for statistical analysis because it is normally distributed.
Table 5: Linear regression of RSV-positive sample miRNA expression vs. 260/230 ratio, before and after application of a cut-off 260/230 ratio ≥ 0.6

<table>
<thead>
<tr>
<th>miRNA</th>
<th>Slope (95% CI)</th>
<th>Significance</th>
<th>Slope (95% CI)</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>miR-16</td>
<td>1.2 (0.7 - 1.8)</td>
<td>p &lt; 0.0001</td>
<td>0.2 (-0.6 - 1.1)</td>
<td>p = 0.6</td>
</tr>
<tr>
<td>miR-19</td>
<td>0.3 (-0.3 - 0.9)</td>
<td>p = 0.3</td>
<td>0.0 (-1.1 - 1.0)</td>
<td>p = 1.0</td>
</tr>
<tr>
<td>miR-21</td>
<td>1.7 (1.1 - 2.2)</td>
<td>p &lt; 0.0001</td>
<td>0.8 (-0.1 - 1.7)</td>
<td>p = 0.08</td>
</tr>
<tr>
<td>miR-23b</td>
<td>1.9 (1.1 - 2.7)</td>
<td>p &lt; 0.0001</td>
<td>-0.4 (-1.5 - 0.8)</td>
<td>p = 0.5</td>
</tr>
<tr>
<td>miR-26b</td>
<td>1.5 (1.0 - 2.1)</td>
<td>p &lt; 0.0001</td>
<td>0.5 (-0.3 - 1.4)</td>
<td>p = 0.2</td>
</tr>
<tr>
<td>miR-27b</td>
<td>1.5 (0.9 - 2.1)</td>
<td>p &lt; 0.0001</td>
<td>0.2 (-0.8 - 1.1)</td>
<td>p = 0.7</td>
</tr>
<tr>
<td>miR-29c</td>
<td>0.5 (0.1 - 0.9)</td>
<td>p = 0.014</td>
<td>0.2 (-0.6 - 0.9)</td>
<td>p = 0.6</td>
</tr>
<tr>
<td>miR-30b</td>
<td>1.2 (0.7 - 1.6)</td>
<td>p &lt; 0.0001</td>
<td>0.2 (-0.6 - 0.9)</td>
<td>p = 0.7</td>
</tr>
<tr>
<td>miR-30d</td>
<td>0.7 (0.3 - 1.0)</td>
<td>p = 0.0005</td>
<td>0.0 (-0.6 - 0.6)</td>
<td>p = 1.0</td>
</tr>
<tr>
<td>miR-31</td>
<td>0.4 (-0.1 - 1.0)</td>
<td>p = 0.1</td>
<td>0.6 (-0.3 - 1.6)</td>
<td>p = 0.2</td>
</tr>
<tr>
<td>miR-34b</td>
<td>0.1 (-0.8 - 0.9)</td>
<td>p = 0.9</td>
<td>-0.3 (-2.0 - 1.3)</td>
<td>p = 0.7</td>
</tr>
<tr>
<td>miR-34c</td>
<td>0.5 (-0.5 - 1.5)</td>
<td>p = 0.3</td>
<td>0.4 (-1.6 - 2.3)</td>
<td>p = 0.7</td>
</tr>
<tr>
<td>miR-96</td>
<td>1.2 (0.2 - 2.2)</td>
<td>p = 0.02</td>
<td>0.7 (-1.0 - 2.5)</td>
<td>p = 0.4</td>
</tr>
<tr>
<td>miR-125a</td>
<td>0.4 (-0.1 - 0.9)</td>
<td>p = 0.1</td>
<td>0.4 (-0.6 - 1.3)</td>
<td>p = 0.4</td>
</tr>
<tr>
<td>miR-125b</td>
<td>1.1 (0.7 - 1.5)</td>
<td>p &lt; 0.0001</td>
<td>0.3 (-0.4 - 1.0)</td>
<td>p = 0.4</td>
</tr>
<tr>
<td>miR-130a</td>
<td>1.7 (1.3 - 2.2)</td>
<td>p &lt; 0.0001</td>
<td>0.5 (-0.2 - 1.2)</td>
<td>p = 0.2</td>
</tr>
<tr>
<td>miR-146a</td>
<td>0.8 (0.2 - 1.4)</td>
<td>p = 0.007</td>
<td>1.0 (-0.2 - 2.1)</td>
<td>p = 0.09</td>
</tr>
<tr>
<td>miR-148a</td>
<td>0.7 (0.2 - 1.2)</td>
<td>p = 0.01</td>
<td>-0.2 (-1.0 - 0.7)</td>
<td>p = 0.6</td>
</tr>
<tr>
<td>miR-155</td>
<td>0.9 (0.4 - 1.4)</td>
<td>p = 0.001</td>
<td>0.4 (-0.5 - 1.4)</td>
<td>p = 0.4</td>
</tr>
<tr>
<td>miR-183</td>
<td>1.2 (0.5 - 1.8)</td>
<td>p = 0.001</td>
<td>0.6 (-0.8 - 2.0)</td>
<td>p = 0.4</td>
</tr>
<tr>
<td>miR-200b</td>
<td>0.8 (0.3 - 1.3)</td>
<td>p = 0.004</td>
<td>0.4 (-0.5 - 1.4)</td>
<td>p = 0.4</td>
</tr>
</tbody>
</table>
Table 5 (continued)

<table>
<thead>
<tr>
<th>miRNA</th>
<th>All RSV positive samples</th>
<th>Samples with a 260/230 ratio ≥ 0.6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Slope (95% CI)</td>
<td>Significance</td>
</tr>
<tr>
<td>miR-203a</td>
<td>0.9 (0.4 - 1.5)</td>
<td>p = 0.001</td>
</tr>
<tr>
<td>miR-205</td>
<td>0.9 (0.5 - 1.4)</td>
<td>p = 0.0002</td>
</tr>
<tr>
<td>miR-223</td>
<td>0.7 (-0.3 - 1.6)</td>
<td>p = 0.2</td>
</tr>
<tr>
<td>miR-324</td>
<td>-0.1 (-0.5 - 0.2)</td>
<td>p = 0.4</td>
</tr>
<tr>
<td>miR-331</td>
<td>0.1 (-0.3 - 0.5)</td>
<td>p = 0.6</td>
</tr>
<tr>
<td>miR-375</td>
<td>1.2 (0.5 - 1.9)</td>
<td>p = 0.001</td>
</tr>
<tr>
<td>miR-429</td>
<td>1.5 (0.9 - 2.1)</td>
<td>p &lt; 0.0001</td>
</tr>
<tr>
<td>let-7d</td>
<td>1.5 (0.9 - 2.0)</td>
<td>p &lt; 0.0001</td>
</tr>
<tr>
<td>let-7f</td>
<td>2.8 (1.9 - 3.7)</td>
<td>p &lt; 0.0001</td>
</tr>
<tr>
<td>let-7g</td>
<td>1.3 (0.8 - 1.8)</td>
<td>p &lt; 0.0001</td>
</tr>
<tr>
<td>let-7i</td>
<td>1.3 (0.8 - 1.8)</td>
<td>p &lt; 0.0001</td>
</tr>
<tr>
<td>let-7i</td>
<td>1.4 (0.9 - 2.0)</td>
<td>p &lt; 0.0001</td>
</tr>
</tbody>
</table>

Note: Before application of the 260/230 ratio ≥ 0.6 cut-off, there was a highly significant correlation between miRNA expression and 260/230 ratio for most miRNAs. After application of the cut-off, this correlation disappeared in all cases, supporting the validity of a cut-off 260/230 ratio ≥ 0.6. Fold change was used for statistical analysis because it is normally distributed.
Figure 5: miR-23b and miR-27b expression vs 260/230 ratio before and after application of a cut-off 260/230 ratio ≥ 0.6

Note: Upper panels show regression lines before cut-off application; lower panels after cut-off application. Cut-off is indicated by the vertical dotted line. p-values are taken from tables 4 and 5. The plots show a clear reduction in miRNA expression in the RSV-positive group below the cut-off. After cut-off application, there is no longer a significant correlation between miRNA expression and 260/230 ratio.
Figure 6: miR-130a and let-7d expression vs 260/230 ratio before and after application of a cut-off 260/230 ratio ≥ 0.6

Note: Upper panels show regression lines before cut-off application; lower panels after cut-off application. Cut-off is indicated by the vertical dotted line. p-values are taken from tables 4 and 5. The plots show a clear reduction in miRNA expression in the RSV-positive group below the cut-off. After cut-off application, there is no longer a significant correlation between miRNA expression and 260/230 ratio.

iii) RNA integrity number: The RNA integrity number (RIN) is a measure of the integrity of large RNA sequences such as mRNA and rRNA. RNA integrity is important because RNA degradation by RNases can significantly affect RNA quantification. In brief, an algorithm compares various measures of 18S and 28S ribosomal RNA quantity. 18S is shorter than 28S, and their ratios are fairly constant. Therefore, if there is less 28S than expected compared to 18S, there is
likely to be some RNA degradation [170]. The algorithm returns a number between 1 and 10, where 10 indicates intact RNA. RIN > 7 is recommended empirically for studies of mRNA because mRNA has long sequences and is prone to degradation. mRNA cleavage at the complementary region for a qPCR probe inhibits PCR-amplification, and the measured abundance of that mRNA transcript is therefore falsely low.

miRNA, however, are more stable than mRNA in a range of biological systems, are less prone to degradation and the biological mechanisms by which miRNA and mRNA are degraded are likely to differ [171]. To our knowledge there are no studies that adequately support the use of RIN as a measure of miRNA integrity in clinical samples. In RNA samples degraded by UV light for 90 minutes there was a positive correlation between miRNA expression and RIN [172]. This suggests that UV light affects large and small RNA indiscriminately but does not seem relevant to our study since samples were not exposed to UV light. In another study, heat degradation resulting in lower RIN values had no effect on measured miRNA expression, but did affect mRNA expression [173], suggesting that miRNA tolerates heat more robustly than mRNA. A number of investigators have studied the expression of miRNA in formalin-fixed paraffin-embedded tissue samples, in which RNAses are expected to degrade larger RNA. mRNA expression is strongly associated with RIN, whilst miRNA expression is not [174, 175], identifying miRNAs as robust biomarkers that are also reliable in degraded samples. RIN therefore seems to be less relevant for miRNA expression studies than mRNA studies.
RIN was significantly lower in control samples (mean 3.1; 95% CI 2.1 – 4.0) compared to RSV-positive samples (mean 5.7; 95% CI 4.8 – 6.7); p < 0.0001, Student’s t-test. A positive correlation between miRNA expression and RIN would therefore adversely affect the statistical comparison of control and RSV positive samples, and we could expect miRNAs to falsely appear upregulated in RSV disease in our samples. We therefore compared miRNA expression to RIN in using a linear regression analysis in both control and RSV positive samples (table 6 and figure 7). We found no correlation between miRNA expression and RIN in 31 of 32 miRNAs. miR-324 expression had an unexpected negative correlation with RIN. We conclude that RIN does not appear to be an important measure of miRNA integrity in our samples. Nevertheless, to be certain that any possible effects of low RIN were eliminated we chose to adjust for RIN statistically using a multiple regression analysis in the comparison of miRNA expression between control and RSV positive groups.

Whilst we have taken measures to account for RNA degradation and contamination in our samples, we cannot exclude an increased intra-group variation in miRNA expression that is due to sample quality, not simply biological variation. Variation in sample quality may thus reduce the power of our study to identify differences in miRNA expression. We consider that differences in miRNA expression identified by this study are true differences that have become apparent despite possible effects of RNA degradation and contamination.

Finally, the RIN values indicated that our samples were not good enough for mRNA analysis so we did not attempt to compare miRNA expression with mRNA
expression, a technique that some investigators have used to identify likely
miRNA-mRNA interactions in the biological system studied.

Table 6: Linear regression of miRNA expression vs RIN, after application of
a cut-off 260/230 ratio ≥ 0.6

<table>
<thead>
<tr>
<th>miRNA</th>
<th>Control samples</th>
<th>RSV-positive samples</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Slope (95% CI)</td>
<td>Sig</td>
</tr>
<tr>
<td>miR-16</td>
<td>-0.1 (-0.3 - 0.2)</td>
<td>p = 0.5</td>
</tr>
<tr>
<td>miR-19</td>
<td>-0.3 (-0.7 - 0.1)</td>
<td>p = 0.1</td>
</tr>
<tr>
<td>miR-21</td>
<td>0.1 (-0.2 - 0.5)</td>
<td>p = 0.4</td>
</tr>
<tr>
<td>miR-23b</td>
<td>-0.1 (-0.6 - 0.5)</td>
<td>p = 0.8</td>
</tr>
<tr>
<td>miR-26b</td>
<td>0.1 (-0.2 - 0.5)</td>
<td>p = 0.4</td>
</tr>
<tr>
<td>miR-27b</td>
<td>-0.2 (-0.7 - 0.4)</td>
<td>p = 0.5</td>
</tr>
<tr>
<td>miR-29c</td>
<td>-0.3 (-0.5 - 0.01)</td>
<td>p = 0.1</td>
</tr>
<tr>
<td>miR-30b</td>
<td>0.2 (-0.2 - 0.6)</td>
<td>p = 0.3</td>
</tr>
<tr>
<td>miR-30d</td>
<td>-0.1 (-0.3 - 0.1)</td>
<td>p = 0.2</td>
</tr>
<tr>
<td>miR-31</td>
<td>0.1 (-0.2 - 0.4)</td>
<td>p = 0.6</td>
</tr>
<tr>
<td>miR-34b</td>
<td>-0.4 (-1.0 - 0.1)</td>
<td>p = 0.1</td>
</tr>
<tr>
<td>miR-34c</td>
<td>-0.2 (-0.8 - 0.4)</td>
<td>p = 0.5</td>
</tr>
<tr>
<td>miR-96</td>
<td>0.02 (-0.6 - 0.6)</td>
<td>p = 0.9</td>
</tr>
<tr>
<td>miR-125a</td>
<td>0.1 (-0.2 - 0.3)</td>
<td>p = 0.6</td>
</tr>
<tr>
<td>miR-125b</td>
<td>-0.1 (-0.6 - 0.3)</td>
<td>p = 0.5</td>
</tr>
<tr>
<td>miR-130a</td>
<td>-0.2 (-0.6 - 0.1)</td>
<td>p = 0.2</td>
</tr>
<tr>
<td>miR-146a</td>
<td>0.1 (-0.4 - 0.5)</td>
<td>p = 0.8</td>
</tr>
<tr>
<td>miR-148a</td>
<td>-0.1 (-0.5 - 0.3)</td>
<td>p = 0.6</td>
</tr>
<tr>
<td>miR-155</td>
<td>0.3 (-0.01 - 0.5)</td>
<td>p = 0.1</td>
</tr>
<tr>
<td>miR-183</td>
<td>0.1 (-0.5 - 0.6)</td>
<td>p = 0.8</td>
</tr>
<tr>
<td>miR-200b</td>
<td>0.1 (-0.2 - 0.4)</td>
<td>p = 0.5</td>
</tr>
<tr>
<td>miR-203a</td>
<td>0.01 (-0.2 - 0.3)</td>
<td>p = 0.9</td>
</tr>
</tbody>
</table>
Table 6 (continued)

<table>
<thead>
<tr>
<th>miRNA</th>
<th>Control samples</th>
<th>RSV-positive samples</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Slope</td>
<td>(95% CI)</td>
</tr>
<tr>
<td>miR-205</td>
<td>-0.3</td>
<td>(-0.7 - 0.1)</td>
</tr>
<tr>
<td>miR-223</td>
<td>-0.01</td>
<td>(-0.8 - 0.8)</td>
</tr>
<tr>
<td>miR-324</td>
<td>-0.3</td>
<td>(-0.6 - -0.1)</td>
</tr>
<tr>
<td>miR-331</td>
<td>0.1</td>
<td>(-0.1 - 0.3)</td>
</tr>
<tr>
<td>miR-375</td>
<td>0.1</td>
<td>(-0.4 - 0.6)</td>
</tr>
<tr>
<td>miR-429</td>
<td>0.03</td>
<td>(-0.3 - 0.4)</td>
</tr>
<tr>
<td>let-7d</td>
<td>0.1</td>
<td>(-0.3 - 0.5)</td>
</tr>
<tr>
<td>let-7f</td>
<td>0.1</td>
<td>(-0.3 - 0.5)</td>
</tr>
<tr>
<td>let-7g</td>
<td>0.0</td>
<td>(-0.3 - 0.3)</td>
</tr>
<tr>
<td>let-7i</td>
<td>-0.01</td>
<td>(-0.3 - 0.3)</td>
</tr>
<tr>
<td></td>
<td>-0.1</td>
<td>(-0.5 - 0.3)</td>
</tr>
</tbody>
</table>

Note: There is no correlation between miRNA expression and RIN for any of the 32 miRNA analyzed, except for miR-324, which shows a negative correlation (expression decreases with increasing RIN), the opposite of what we would expect. This may be explained by the fact that we have not adjusted for multiple testing.
Note: There is no correlation between normalized expression data and RIN for either control or RSV positive samples. p-values are taken from table 6. These figures are typical for all miRNAs analyzed.
Disease Severity

Subgroup analysis according to disease severity was central to our hypotheses, and we therefore devised definitions of disease severity that could be applied to retrospectively or prospectively collected data.

Disease severity for the 2003 cohort is defined in paper 1 (table 1), and for the 2011 cohort in paper 3 (methods). Weaknesses and strengths of our categorization are addressed in the chapter “Discussion of methods.”

Microarray Analysis

Microarray analysis is a technique that tests a sample for many different variables at the same time. In this project, we used mRNA and miRNA microarrays to analyse expression of approximately 25000 mRNAs and 900 miRNAs per sample. Each microarray chip contains thousands of spots, each imprinted with probes specific for an mRNA or miRNA transcript (see figure 8). First, the enzyme reverse transcriptase (RT) is used to convert sample RNA to complementary DNA (cDNA). There may also be a pre-amplification stage in which cDNA is PCR - amplified to increase the sensitivity of the microarray. cDNA is labelled with a colour dye then introduced onto the chip. Hybridization is a process in which cDNA complementary to a probe on the chip is bound to that probe. The result is activation of the colour label. After hybridization, the chip is placed in a scanner, which uses a laser to excite the colour label. The amount of light emitted correlates with the amount of probe-specific cDNA loaded onto the chip. This output can then be compared between samples.
Sources of error in microarray experiments: There may be sample errors and technical errors in a microarray experiment. Sample errors are the same as for PCR, and are discussed in the chapter “Quantitative RT-PCR – Sources of error in RT-qPCR experiments.” Technical errors specific to microarray include inefficient pre-amplification, poor sample loading, inefficient sample colour labelling and inefficient probe hybridization. Such differences may occur both within a sample, within a chip and between chips. For the miRNA microarray, we used a total of 6 chips, with 8 samples on each chip. Small differences in reaction conditions between each chip (e.g.: room temperature) may result in differences in reaction efficiencies. For the miRNA microarray we chose chips that included spots with the same probe at a number of different places on the microarray, and used a spike-in. Multiple spots per probe are essentially technical replicates, and can account for differences in reaction efficiency within the chip. The spike-in is a standard RNA transcript provided by the chip manufacturer that hybridises to
a specific probe on the chip. If the same amount is added to each sample, then the spike-in output should be the same for each sample. This can therefore identify differences in reaction efficiencies between samples and between chips. Finally, we applied a normalization procedure (briefly described in the next section) to further account for differences between chips.

**miRNA microarray data processing**

In paper 3, the microarray data was analysed using the R statistical environment with the extension packages “AgiMiRNA” [176] and “Linear Models for Microarray” (LIMMA) [177], available from Bioconductor. R-statistics and Bioconductor are open source projects providing tools for the analysis and understanding of high throughput genomic data, such as microarray analysis [178]. I attended a necessary course in processing, statistical analysis and interpretation of microarray data. Standard scripts for normalization, quality control and statistical analysis were used, published on the Bioconductor website specifically for use with the LIMMA and AgiMiRNA packages.

We processed the raw microarray data using Agilent’s Feature Extraction Software, producing a file for each sample that includes the signal strength of each microarray spot. The AgiMiRNA package allows automated extraction of microarray data from the Feature Extraction Software files, normalization of the data and principal component analysis.

*Normalization:* Normalization of microarray data reduces differences between samples due to technical error. We applied quantile normalization, which assumes that the variation in the data is the same for each sample [179], and is a
robust method to normalize miRNA microarray data [180]. Instead of using a reference gene, information from the entire data set is used to adjust each value.

*Principal component analysis and other quality measures:* LIMMA analysis is vulnerable to samples in which gene expression is very different to other samples in the same group (outliers). Such samples can have great leverage on the data and can result in type 1 and type 2 errors. Several methods are therefore used to examine the quality of the data and assess the suitability of each sample for inclusion in the LIMMA analysis. I will here briefly describe Principal component analysis (PCA), which had implications for the analysis of the data. PCA transforms the entire gene expression data set into a new set of variables (principal components, PC) that replace the individual miRNA expression levels. The number of PCs is equal to or less than the number of miRNA analysed. Each principal component [181] is linearly independent of the others (i.e.: a linear regression comparing one PC to another would not be significant). The first PC (PC1) has the greatest variability, and the second PC (PC2) has the second greatest variability and so on. The components with greatest variability can be used to identify outlier samples. On a plot of PC1 vs PC2 biological samples within the same group should have similar PC values and should therefore be clustered. PCA of the miRNA microarray data revealed 4 control samples and 3 severe disease samples that were outliers (see figure 9). In addition, one mild disease sample did not load appropriately onto the microarray chip and had to be excluded. This left us with 40 samples from the original 48. Defining 20 – 25% of samples in the control and severe groups as outliers seems inappropriate. However, LIMMA analysis prior to removal of
these samples showed only two differentially expressed miRNA, which does not seem likely. After removing the samples there were 26 differentially expressed miRNA, which is a more likely scenario. Confirmation by qPCR of differential expression of 11 of these miRNA confirms that removal of the outliers was a

Figure 9: Principal components analysis of the control group

Note: Principal component 2 plotted against principal component 1. Only the control group is included in the plot (Samples names do not reflect RSV positivity or not). The majority of control samples are clustered at the top right of the plot. 4 samples, deemed outliers, are found at the top left and bottom left.

Generating hypotheses using microarray results

An mRNA microarray analysis had already been done on the 2003 cohort before the start of this project. How we used this analysis to generate hypotheses for
papers 1 and 2 is described in the Chapter “Aims and hypotheses – a note on hypothesis generation.”

In 2011, we conducted a miRNA microarray on nasal epithelial samples. We chose the samples that had the best RNA quality, assessing both 260/230, 260/280 and RIN values. Moderate and severe disease groups were grouped together to increase power. Of 877 miRNA transcripts tested, 199 were expressed. 26 differentially expressed miRNA and 3 miRNA with a tendency to significance were selected for further testing. In addition, 3 miRNA that were not significantly expressed were selected because they have well described functions in the immune system and we considered the lack of regulation as interesting in itself.

Quantitative RT-PCR

The polymerase chain reaction allows us to measure the quantity of nucleic acid sequences with great precision and specificity. Since polymerase only amplifies DNA, RNA samples are first reverse transcribed to cDNA. PCR analyses measuring RNA are therefore called RT-PCR. The amount of cDNA produced is directly related to the amount of RNA present, allowing us to determine how much RNA was present at the time the sample was collected, and therefore the extent to which a gene is expressed. The mechanism of the PCR reaction is shown in figure 10. In this project we conducted a series of 40 cycles, allowing amplification of cDNA up to $2^{40}$ times, depending upon reaction efficiency and the availability of reagents (primers and nucleic acids).
Figure 10: The Polymerase Chain Reaction

Note: The different steps of the PCR reaction are accomplished by varying the temperature. 1. Double stranded cDNA is melted into single strands. 2. Short DNA molecules called primers bind to complementary cDNA sequences. The primer is chosen to be specific for cDNA produced from the RNA transcript of interest, allowing accurate measurement of that RNA transcript, and no others. 3. The enzyme polymerase extends the primer, copying the cDNA sequence. The quantity of the primer-specific cDNA sequence is thus doubled.

Picture: "Polymerase chain reaction" by Enzoklop - Own work. Licensed under Creative Commons Attribution-Share Alike 3.0 via Wikimedia Commons.

Primers are bound to a fluorescent reporter probe that emits light of specific wavelengths when the primer binds to its complementary DNA and is extended. The amount of light produced by the reaction is measured at the end of each cycle, allowing us to quantify the amount of primer-specific DNA present at the end of each cycle. This is called therefore called quantitative RT-PCR (RT-qPCR).

Sources of error in RT-qPCR experiments

RT-qPCR is widely regarded as both sensitive and accurate, but there are important sources of error that should be considered when planning, interpreting and reporting an RT-qPCR experiment [182]. In research
applications, sources of error are usually distributed evenly between the groups being compared, meaning that the estimated difference in means is little affected, but the confidence interval is increased, reducing the power of the experiment. With small group sizes the estimated mean may also be affected. Sources of error associated with RNA quality and purity are discussed above in the chapter

*Disease cohort 2011: RNA quality control.*

**Sample errors:**

Sample errors are intrinsic to the biological samples themselves and cannot be improved by laboratory procedures. Sample errors include:

i) Sample contamination. Exogenous biological material can alter the quantity of target RNA in the sample. To avoid this, cord blood samples were collected directly into a vacuum receptacle, and nasal epithelial sampling was done using rubber gloves and masks.

ii) Tissue heterogeneity. Biological samples are not homogenous materials, and the amount of RNA harvested from a sample can vary considerably. In clinical specimens there is usually less control over the sampling procedure than, for example, in cell culture specimens.

iii) Presence of RNase in mucous secretions and body fluids. RNase causes rapid RNA degradation that can severely affect RNA quality. Aspiration of nasal secretions prior to sampling and rapid immersion of cord blood and nasal samples in an RNA stabilization solution was critical to the integrity of samples.
iv) Poor storage. RNA degradation increases with temperature. Samples were stored at -80 °C to reduce the risk of degradation. Samples were thawed as few times as possible.

**Technical errors:**

Differences in the efficiency of procedures and reactions used to process samples and produce qPCR results can result in technical errors. Sources of technical error include variations in:

i) RNA-yield during isolation procedures.

ii) The amount of sample RNA included in the reaction

iii) The efficiency of the reverse transcription reaction.

iv) The efficiency of the PCR reaction.

Sampling and technical errors can be reduced by stringent sampling procedures, strict adherence to laboratory protocols and use of standardised solutions, but can never be excluded. Two methods can improve experimental results:

i) Increase the number of biological replicates. This may be done by increasing the group size, or by increasing the number of samples taken from each tissue. In paper 3, we were compelled to combine biological replicates in order to increase the RNA yield, which reduced the power of the experiment.

ii) Include technical replicates in the qPCR analysis. Technical replicates are created when a sample is divided into several samples at some point during the laboratory process (eg: prior to or during isolation; prior to reverse transcription; prior to qPCR), and each sample is thereafter
treated individually. Technical replicates may be created at several points during processing. Prior to statistical analysis, the mean result of the technical replicates is calculated for each sample. This combined value is then included in the statistical analysis. In papers 1 and 2, technical replicates were not included. For the qPCR experiment in paper 3, two technical replicates per sample were created prior to reverse transcription. We were therefore able to control for variations in reverse transcription and qPCR efficiency.

Comparing RNA quantity between samples
Before samples can be compared to each other a number of steps must be performed. We used the ΔΔCq method, normalising target gene results against a reference gene (see figure 11), and then calibrating to an arbitrary Cq value.

i) Establish the number of PCR cycles it took for each sample to be amplified to a pre-set quantity threshold. If a sample has twice as much RNA to begin with, then it will need one cycle less to achieve that threshold. The cycle at which a sample reaches the threshold is called the quantification cycle (Cq), and may theoretically include a fraction of a cycle (e.g.: Cq = 26.2534).

ii) Normalization adjusts for technical errors such as differences in reaction efficiencies and the amount of RNA included in the reaction, and is described in figure 11. It is assumed that technical errors will affect similar RNA sequences within a sample in the same way (i.e.: all short RNA sequences will be affected similarly; all long RNA sequences will be affected similarly). The ratio of target RNA to reference RNA should therefore be constant regardless of technical
errors. The mean value of technical replicates for each sample can be calculated either before or after normalization.

Figure 11: Normalization of PCR data

A

B

Note: Plots A and B show the light intensity emitted at the end of each PCR cycle plotted against the number of PCR cycles. Note the logarithmic y-axis. Plot A: Theoretical example showing real time qPCR plots for the target gene and reference gene of an individual sample. There is an initial non-specific phase followed by an exponential amplification phase. After the reagents are used up, the plot enters a plateau phase. The horizontal dotted lines indicate an arbitrary threshold level set to intersect all samples during the exponential phase. The threshold is different for each gene investigated. The cycle quantity (Cq) is the x-value at which the plot crosses the threshold. The reference Cq is subtracted from the target Cq, giving the ΔCq. The ΔCq is used to compare samples. A difference in ΔCq of three confers a $2^3 = 8$ times difference in quantity. Plot B: Real time PCR plot for a number of samples using a primer specific for TLR4 mRNA, taken from the raw data for paper 2. The threshold for TLR4 is indicated by the horizontal line, and is chosen to intersect all samples during the logarithmic phase. The TLR4 Cq for each sample is determined and normalized against each sample’s reference Cq.
iii) Calibration. An arbitrary ΔCq value (the calibrator) is chosen, for example the mean or median of the control group, which is subtracted from sample ΔCq values. The product is the ΔΔCq, indicating the relative abundance of gene-specific mRNA or miRNA in that sample compared to the calibrator, on a log2 scale. Positive values indicate less initial RNA than the calibrator, negative values more RNA. Calibration is not strictly necessary since it does not affect the results of most statistical tests.

ΔΔCq values can now be compared between groups using Student’s t-test, one-way ANOVA, non-parametric tests or more complicated statistics as appropriate.

Protein analysis by Western Blot

In an mRNA experiment it is necessary to provide some evidence that protein levels or function correlate with mRNA levels. This is because post-transcriptional processes such as mRNA splicing and RNA interference may reduce the extent to which mRNA is translated into functional protein. Differences in mRNA levels may therefore not have any affect on cell function.

Western blot was used to analyse protein levels. Briefly, gel electrophoresis separate proteins in a sample according to their molecular weight because smaller proteins migrate further during electrophoresis than larger proteins. A reference sample incorporating a series of proteins of known molecular weight is included in order to easier identify target proteins. After gel electrophoresis, the proteins are transferred onto a membrane by immunoblotting, making the proteins available for antibody binding.
Primary antibodies targeting amino acid sequences specific to the protein of interest are added. Secondary antibodies targeting the primary antibodies are added, and contain a light-emitting enzyme. The membrane is then placed in a booth and scanned by a camera for a set period of time. The intensity of light emitted is related to the amount of primary antibody bound to target protein, and therefore related to the amount of protein present in the sample.

As with PCR, there are both sample errors and technical errors associated with Western blot.

**Sampling errors:** In our project, the main source of error was the contamination of cord blood white-cell samples with an unknown amount of plasma and red cells. This came about because in 2003 the EDTA-samples were primarily intended for DNA analysis. After centrifugation the leukocyte layer was taken and in order to harvest as much DNA as possible, neighbouring plasma and erythrocytes were also taken. In order to correct for this error we normalized target protein levels against a reference gene, CD45, which is ubiquitously expressed on all leukocytes. The target gene/CD45 ratio therefore indicated the level of target protein per leukocyte, assuming that CD45 is evenly expressed in all samples. We tested for CD45 specificity in new cord blood samples taken in 2011, in which the leukocyte layer was carefully separated (see paper 2, supplementary material). CD45 was identifiable in leukocyte layers. There was some non-specific binding of CD45 antibodies to plasma, but not to erythrocytes. We could therefore use CD45 to adjust for contamination with erythrocytes.

**Technical errors:** There are a number of technical errors associated with Western blot, most of which can be overcome by conducting a series of experiments to
optimize conditions for Western blot before running the samples of interest.
These experiments allowed us to determine the best concentrations of reagents,
amount of antibodies, and duration of shutter opening for each protein. A source
of error of particular relevance to this project was inefficient or non-specific
binding of primary or secondary antibodies. In paper 1 we were unable to
identify Dicer in our samples despite trying 2 different primary antibodies. We
were, however able to identify Dicer in colorectal cancer samples. The reason for
this discrepancy is unknown. Alternative splicing may have produced an mRNA
recognised by the PCR analysis, but a protein not recognised by our antibodies.
This is further discussed in paper 1. In paper 2, we had initial problems with
non-specific binding of primary antibodies. Positive (available for CCR7 and
TLR4) and negative controls, allowed us to identify which bands on the
membrane corresponded to target protein and which indicated non-specific
binding.

Statistical analysis

Minitab 15 (papers 1 and 2), the R computing environment (paper 3) and for all
papers IBM SPSS software and Prism 5, were used to analyze the data.

LIMMA analysis was used to analyse microarrays in both cohorts. LIMMA uses an
empirical Bayes approach to increase the statistical power when analysing each
gene. Because genes are not independent of each other, it is reasonable to use
information from the entire data set to inform the probability that a gene is
differentially expressed or not. This approach produces a set of hypervariables
for each sample that can be used in a general linear model for the contrasts of
interest. The t-distribution for pair-wise comparisons and F-distribution for
multiple comparisons is then used to compare groups in a similar way to Student’s t-test or one-way ANOVA.

After normalization and calibration of qPCR data, the distribution of each gene/miRNA was assessed for normality. Genes/miRNAs with visual evidence of normal distribution were assessed using parametric methods (t-test and one-way ANOVA). Genes with statistical evidence of non-normality, or in which we were in doubt about the distribution were assessed using non-parametric methods (Mann-Whitney test, Kruskal-Wallis test).

**Adjustment for multiple testing:** The familywise error rate (FWER) is the risk of making a single type 1 error when testing multiple hypotheses. If we have set the level of significance to \( p = 0.05 \), then there is a FWER of 9.75\% when testing two hypotheses, 14\% when testing three hypotheses and 81\% when testing 32 hypotheses. Testing multiple hypotheses therefore demands correction for the FWER.

Bonferonni’s correction is strict and reduces the FWER to a minimum by dividing the significance level by the number of hypotheses. The power of the study is thus reduced, but the risk of type 1 error is minimised, increasing our confidence in the results.

The False discovery rate (FDR) is the proportion of “discoveries” (null hypothesis rejected) that are false. Benjamini-Hochberg’s method adjusts for a pre-set false discovery rate. If we set the FDR to \( q = 0.05 \) then we accept that up to 5\% of reported discoveries may be false. This method is less stringent than
Bonferroni’s correction. It increases the power of the study at the risk of increasing type 1 error.

In papers 1 and 2, a total of 10 genes were tested. Reported significance values are unadjusted, but the level at which a value is considered significant is reduced using Bonferroni’s method. In paper 3, 32 miRNAs were tested by qPCR. Reported significance values are adjusted for a false discovery rate of 5%, meaning that an adjusted p-value < 0.05 is considered significant.

Ethics

Informed consent was obtained from the guardians of all patients prior to inclusion in the study. Study protocols were approved by the regional ethics committee and by the local data protection officer.

Other issues

Paper 2 was also included in Helene Østerholt’s PhD thesis “Hyaluronan in the neonatal period.” The focus of that thesis was Hyaluron metabolism. TLR4 is a hyaluronan receptor, and this was the basis for a collaboration on paper 2. The genes RHAMM and CD44 were included in paper 2 but are not the subject of this thesis, and will not be discussed further. Correction for multiple testing in paper 2 also considered these genes.
5. Summary of Results

Papers 1 and 2 - Cord blood gene expression:

In paper 1 we included children aged < 12 months of age at the time of their RSV infection. There were 16 in the control group, 17 in the mild disease group, and 20 in the severe disease group. Dicer expression was not normally distributed, so groups were compared using the Mann-Whitney test.

In paper 2 we included children aged < 36 months of age at the time of their RSV infection in order to increase the power of the study. There were 17 in the control group, 22 in the mild disease group, and 25 in the severe disease group.

There was downregulation of Dicer, IL7R, CCR7, and TLR4 at birth in infants who later test positive for RSV. For Dicer and TLR4 these results were specific to the severe disease subgroup. Details are summarized in table 7.

Dicer protein was not detectable in our samples or in adult peripheral blood samples, and we were therefore not able to confirm mRNA findings. Western blot for IL7R, CCR7 and TLR4 did not have sufficient power to definitively confirm mRNA findings, but did show a similar trend.

Paper 3 – Nasal epithelial miRNA expression in RSV positive infants

In the final statistical analysis of qPCR data there were 19 control samples, 16 mild disease, 7 moderate disease and 19 severe disease samples. RSV positive infants had a similar median age and interquartile range (IQR) to control infants (see figure 12). The control and mild disease samples were taken from infants who were a little older than the moderate and severe infants, but this difference
was not statistically significant. We conducted a linear regression of miRNA expression against age for all miRNAs included in the qPCR and found no correlations (data not shown). We conclude that it is unlikely that differences in age between disease subgroups affect our results.

Figure 12: Age according to RSV disease group – 2011 cohort

*miRNA expression:* In the final qPCR analysis RSV positive infants had downregulation of miR-34b, miR-34c, miR-125b, miR-29c, mir125a, miR-429 and miR-27b and upregulation of miR-155, miR-31, miR-203a, miR-16 and let-7d in nasal mucosa. On disease subgroup analysis, miR-125a and miR-429 were downregulated in mild disease, but not in severe disease. A summary of the functions and disease associations of these miRNA is given in the discussion of results and in table 8.
### Table 7: Cord blood gene expression in RSV positive infants

<table>
<thead>
<tr>
<th>Downregulated</th>
<th>Not regulated</th>
</tr>
</thead>
<tbody>
<tr>
<td>All RSV positive</td>
<td>Severe disease only</td>
</tr>
</tbody>
</table>

**Paper 1**

- *Dicer*  

**Paper 2**

- *IL7R*  
- *CCR7*  
- *TLR4*  
- *CXCR3*  
- *IFNy*  
- *IL4*  
- *CXCL7*  
- *CXCL11*  
- *CCL3*
6. Discussion

Discussion of methods

Patient selection and categorization
We relied on the microbiology database to identify children from the 2003 birth cohort with RSV disease. The underlying assumption was that RSV-infected children with severe disease would be referred to hospital and therefore RSV-tested. We believe that this assumption is valid, but were careful to ensure that controls were not admitted to other hospitals in the area. It is highly likely that members of the control group also had an RSV infection, but we assume that this infection was not severe enough to warrant hospital referral with subsequent RSV-testing. Given the spectrum of RSV disease there is likely to be an overlap between children with mild disease and children in the control group, although we believe it a reasonable assumption that the control group would in general have an even milder disease than the mild group.

The 2011 cohort was enrolled on assessment in the emergency ward, with subsequent removal of those who were not RSV positive. We did not test the control group for RSV, but since they appeared healthy, without evidence of respiratory infection, and were enrolled outside of the viral airways season, we consider the risk of RSV disease in this group at the time of sampling to be small.

Objective and subjective indicators of disease severity
A secondary aim of this study was to find differences in the immune system that predispose to more severe RSV disease. In order to achieve this, we needed to define different degrees of disease severity. RSV disease phenotypes are
heterogeneous, meaning that it is not suitable to define disease severity as a single continuous variable. It was therefore necessary to construct categorical data. Several considerations were important:

i) In many studies the severe disease group includes only children admitted to the paediatric intensive care unit, especially those who are intubated. We did not have enough children with disease of this severity, and so also included children receiving supplementary oxygen or fluids in the severe group.

ii) In many studies children who are not admitted are categorized with mild disease, whilst admitted patients are categorized with more severe disease. However, in our experience children with mild disease are often admitted for observation, for example due to young age and reasonable caution, uncertainty on the part of the treating physician, or parental anxiety.

iii) Division of patients into too many disease subgroups reduces the power of statistical analysis.

*Objective measures of respiratory distress:* Objective end-points are more reliable than subjective ones, and we therefore based disease severity categorization on the following foundation: i) Use of supplemental oxygen, supplemental fluids, CPAP or a ventilator always gave a severe disease categorization; ii) Discharge from the emergency ward without admission to hospital always gave a mild disease categorization. In our experience, these measurements are clearly defined and reliably recorded in medical records. The chance of error during retrospective data collection was therefore reduced, and the study is more reproducible.
Subjective measures of respiratory distress: We chose the degree of respiratory distress on admission as the final determinant for disease severity categorization, applying it to children that were admitted but did not receive treatments conferring severe disease.

Retrospective data collection for the 2003 cohort meant that we had to rely on a clinical description of the degree of respiratory distress. This was by no means precise, but the description in general seemed to be quite clear if there was little respiratory effort, and was often confirmed by subsequent descriptions, particularly if the infant was discharged the next day. There was too much uncertainty to define a moderate disease group, and so we included children in the severe disease group in which there was a description of moderate or severe dyspnoea, including a history of apnoea as an indicator of severe respiratory effort. These children represented about half of the severe group.

For the 2011 cohort, prospective data collection included use of the modified respiratory distress assessment instrument (mRDAI). This gave a more precise assessment of respiratory distress, allowing us to divide the cohort into 3 disease severity subgroups. The mRDAI was developed for this study by modifying Lowell et al.’s Respiratory Distress Assessment Instrument (RDAI) [183]. This instrument has been used in a number of studies of asthma and bronchiolitis, and has been validated for inter-rater reliability [184-186]. The RDAI is designed for repeat assessment of respiratory effort, where the variable produced reflects an increase or reduction in respiratory effort between assessments. The objective difference in respiratory rate between assessments is part of the RDAI algorithm, in addition to semi-objective measures of respiratory distress (degree
of retractions and degree of wheeze) taken at each assessment. To allow a single assessment, we modified the RDAI by scoring the respiratory rate, using age-appropriate upper limits of respiratory rate as a baseline. See paper 3, table 1 for details. The distinction between mild and moderate disease was arbitrary – an mRDAI respiratory distress score ≥ 9 amongst children who were admitted gave a moderate disease categorization. mRDAI intra-rater and inter-rater reliability were not assessed in our study, conferring a risk of error when differentiating between mild and moderate disease.

Timing of assessment: Assessment on admission gave a consistent time-point during the course of the illness, prior to use of inhaled medications or other treatments that might improve respiratory distress. The median duration of illness on assessment was approximately 4 days for all disease severity subgroups in both cohorts (paper 1, table 3; paper 2, table 1; paper 3, table 3). This is particularly important for paper 3, in which a similar duration of illness at the time of sampling increases the chance that we are assessing the same immunological process in each group.

Repeat respiratory assessment of children during admission in the 2003 cohort was not feasible because of retrospective clinical data collection. Repeat examination of the 2011 cohort to ensure respiratory distress assessment during the period of maximal respiratory distress was possible but would require considerable resources in order to be practically achievable. We would either have to train a larger group of nurses and doctors in the study protocol for respiratory distress assessment, reducing inter-rater reliability, or else the study team would have to follow up each child individually several times daily.
Furthermore, such follow up would not eliminate the risk of failing to assess the child during periods of maximal respiratory distress. We therefore considered that the amount of effort required would not be proportional to the benefit.

Retrospective clinical data collection – implications for interpretation

The clinical data for the 2003 cohort was collected retrospectively. This conferred a risk of bias when defining disease severity. As discussed above, medical records were not written in a structured manner, and although we aimed to be as objective as possible there may have been an interpretation bias when reading the physician’s assessment of respiratory effort. Use of interventions such as oxygen or fluid use may not have been adequately recorded. Retrospective data collection may also result in an overlap between mild and severe disease groups, if children classified with mild disease had more respiratory distress than was evident from the medical record, or if physicians exaggerated the degree of respiratory distress.

Prospective clinical data collection – 2011 cohort

The 2011 data was collected prospectively, and is therefore more reliable. In particular, we were careful to follow up each case individually during admission. Since the requirements for allocation to the severe disease groups were purely objective we are confident about the distinction between severe RSV disease and the other categories. However, it is interesting to note that infants with moderate disease had more respiratory effort than children with severe disease (see paper 3, table 3). This cannot be explained by differences in age or weight. The moderate disease group had been ill for a median 5 days on admission (IQR 4 – 5); the severe disease group 4 days (IQR 3 – 5), which may explain some of the
difference in respiratory effort. Alternatively, exhaustion of severe disease patients may reduce observed respiratory effort. We also consider that differences in the physiological or immunological response to RSV may result in a smaller increase in respiratory effort amongst severe disease cases. This lack of response may reduce the ability to clear respiratory secretions and to ventilate adequately, increasing the risk of supplementary oxygen or fluids, or the need for mechanical respiratory support.

RSV gives a variety of clinical phenotypes, and distinguishing between mild and moderate disease in clinical practice is not likely to be so easy. We do not know if infants with mild disease developed respiratory distress consistent with moderate disease after assessment. However, this question becomes purely hypothetical because in the final qPCR analysis, all infants with admitted with mild disease had poor RNA quality and were removed, leaving only mild disease infants discharged from the emergency ward without admission. Furthermore, the majority of miRNAs had equal expression across disease subgroups. For the 2 miRNAs that showed differential expression between disease groups, it was the severe disease group that was different, not the mild or moderate group (see figure 2, paper 3). Poor distinction between mild and moderate disease group will therefore not affect our results.

Reliability of Gene Expression Analysis

RNA quality

The quality of isolated RNA has significant consequences for the utility of biological specimens and interpretation of results in paper 3, as discussed in the chapter “Methods – Disease cohort 2011 – RNA quality control.”
The RNA quality in the 2003 cohort was considered good enough for mRNA analysis (see paper 1 – methods section – Preparation of RNA samples for analysis and RNA quality).

We removed a number of 2011 samples due to poor RNA yield or quality, and adjusted for RIN in the statistical analysis. Removing samples meant that only 40% of disease samples and 32% of control samples were included in the final analysis (see figure 4). Sample removal and adjustment for RIN reduced the power of statistical analyses. Of greater consequence, the low proportion of samples analysed has clear implications for whether the results are applicable to the entire cohort. We conducted our microarrays in a sub-cohort and expanded the size of the group in the qPCR. This increases the probability that results are applicable to a wider group. Ideally, a microarray should be conducted in one cohort and the qPCR confirmation in a completely different cohort. Unfortunately, this was not feasible due to the limited number of samples available.

Finally, RNA degradation in the 2011 sample meant that samples could not be reliably tested for mRNA expression. mRNA data would have increased the value of the study by allowing comparison of miRNA and mRNA expression data. Such an analysis could have provided valuable information about miRNA-mRNA interactions and the role of miRNA in biological pathways.

**Microarray analysis**

Studies that include microarray analysis are often referred to as hypothesis generating research. This is because microarrays are typically used when there is a gap in the knowledge of what makes up a biological system, and what the
function or relevance of these constituents are. For example, we are aware of over a thousand miRNA encoded in the human genome, but there is a paucity of previously generated knowledge about what the majority of these miRNA do, or if they are relevant for RSV disease. In classical research, previous knowledge has either been generated by decades of research or by empirical experience. This knowledge leads to specific hypotheses that can then be tested (e.g.: breastfeeding reduces the risk of severe RSV disease). Detailed knowledge about the genome and transcriptome has been available for a relatively short time meaning that it is difficult to guess from previous experience which genes or miRNAs are relevant. In addition, focusing only on genes or miRNAs that are previously well described results in a bias in the knowledge – we increase our knowledge about elements that we already are well acquainted with. The complexity of biological systems demands a high-throughput approach in order to increase knowledge within a reasonable timeframe.

The microarrays used in this project test a sample for many different transcripts at the same time. In the 2003 cohort a microarray measured the abundance of about 25 000 mRNA transcripts per sample. In the 2011 cohort, about 900 miRNA transcripts per sample were tested. Testing so many hypotheses concurrently leads to several issues:

i) There is a likelihood of type 1 error due to multiple testing. Adjusting for this using a stringent method may exclude results relevant to the system being tested, whilst less stringent adjustment increases the risk of false positives. Our microarray results were adjusted using the false discovery rate, which is less stringent than Bonferroni’s correction (see Methods - Statistical analyses). It is
therefore not surprising that a number of miRNAs differentially expressed in the microarray were not differentially expressed in the qPCR. We considered this necessary in order to maximise the chance of discovering differentially expressed genes and miRNAs.

ii) The risk of technical error increases, and positive results may not be positive when re-tested. Microarray is also less precise than qPCR. It is therefore necessary to verify microarray results using qPCR.

Testing many elements of a biological system identifies those elements that are worth investigating further. Thus, microarray analyses are hypothesis generating, but results should always be verified. Because microarrays are expensive, it is usual to test a limited number of samples and then expand the group size in subsequent analyses.

**Gene selection after microarray analysis**

The selection of genes for PCR verification in paper 1 and 2 was based upon a microarray in which the cord blood of 5 RSV positive infants who presented with lower respiratory tract infection was compared to cord blood of 5 RSV positive infants who did not present with lower respiratory tract infection. A number of genes were excluded from further analysis because they were not significantly different in this analysis. The literature was also consulted, and relevant genes not included in the microarray analysis (such as TLR4) were included. There are two drawbacks to this method: i) the sample size was small, so the microarray may have been underpowered to detect relevant changes in gene expression. Genes that are differentially expressed may therefore have been excluded from further analysis.; ii) Using the literature to identify relevant genes introduces a
bias because we may exclude genes that are relevant but which are not sufficiently studied. Many genes were excluded from further analysis because we did not consider them sufficiently relevant.

The sample size in the miRNA microarray was bigger than in the birth cohort, but still relatively small for a clinical investigation. Some relevant miRNAs may therefore have been excluded from further analysis. When selecting miRNA for the qPCR, we included all miRNA that were differentially expressed in the microarray. This excludes a bias caused by consultation of the literature, even though we did include 3 miRNA in the qPCR that were not differentially expressed in the microarray. Interestingly, one of these three was significantly upregulated, whilst the other two lost statistical significance after correction for multiple testing.

Quantitative RT-PCR

qPCR is considered highly precise given that the biological samples are of good quality and that technical errors are kept to a minimum and adequately controlled for. Any remaining errors can be dealt with statistically by increasing the number of samples, thereby reducing the confidence interval of the estimated mean for each group and increasing the power of the study. We removed miRNA samples with a low 260/230 ratio because low ratios seemed to favour one group over the other. We have not identified any other sources of bias in our experiments, and are therefore confident that the differences we have found are true differences. We are less confident about the hypotheses we have rejected. In papers 2 and 3, there were genes/ miRNAs that tended to significance, even after correction for multiple testing. Of note are CXCR3, miR-
21, let-7i, miR-223, miR-375 and miR-331 (paper 2, figure 1A; paper 3, figure 1). The subgroup analysis in paper 3 was particularly hampered by a reduced group size, and we believe that this may have affected results for miR-31, miR-223 and miR-429 (see paper 3, figure 2). Increasing the group size, particularly those with moderate disease, would have increased the power of the study.

**Western blot**

In papers 1 and 2 we used Western blot to confirm that protein expression mirrors mRNA expression. Western blot is far less precise than qPCR, but a necessary step to ensure that mRNA expression really does reflect protein levels. Western blot is also far more time consuming than qPCR, which is why only genes identified by qPCR were tested. In paper 2, we restricted Western blot testing to samples that represented the interquartile range of each group. In hindsight, we should have included all the samples because the Western blot results suffered from a lack of power. Visually, however, it seems reasonable to conclude that the protein quantification follows the same pattern as mRNA expression.

In paper 1 we were unable to detect Dicer protein, and as discussed in the chapter “Methods – Protein analysis by Western blot,” we do not know the reason for this. It is possible that the antibodies we used did not match the protein produced from the mRNA splice variant identified by the qPCR probe. There is therefore a greater uncertainty about the interpretation of results in paper 1. Further investigations of the role of Dicer in RSV disease may clarify this issue.
Discussion of main results

Gene expression at birth as a predictive factor for RSV infection

We found downregulation at birth of \textit{IL7R} and \textit{CCR7} in infants who later tested positive for RSV. In infants who had severe RSV disease, we also found downregulation of \textit{Dicer} and \textit{TLR4}. These four genes have important functions in the immune response to RSV, as described in the introduction. A number of other studies have identified gene mutations and epidemiological factors associated with RSV (see Introduction – Factors associated with RSV disease severity). We did not test for these factors, but may they be the cause of the gene dysregulation seen in our project.

Finding gene dysregulation at birth is interesting because it occurs prior to the RSV infection, during a process that all infants in our cohort underwent, and which stimulates the immune system. The statistical association therefore also seems to be biologically relevant. Differences in the way children express immune genes at birth may mirror differences in the expression of those same genes immediately prior to or during RSV infection, resulting in different disease phenotypes. We measured gene expression in blood, which is not the primary site of RSV infection. We should therefore consider the relevance of these results to respiratory disease with caution. However, the mRNA analysed in blood is primarily derived from leukocytes and our results can give us insight into how the immune system responds to stressful stimuli.

\textit{Dicer} is a key enzyme in the production of miRNA, which have been shown to regulate a variety of immunological pathways and cell functions. \textit{Dicer} is necessary for adequate in-vitro responses to influenza virus [99], and for
development and function of T-cells, B-cells, macrophages and natural killer cells [93-98]. A Dicer knockout mouse line has increased susceptibility to viral infection [100, 101]. Dicer downregulation does not necessarily mean a global reduction in miRNA and a subsequent global increase in mRNA translation to protein - biological systems are more complicated than that, with multiple regulatory mechanisms and feedback loops. In one experiment, reduced Dicer expression was associated with increased levels of a subgroup of miRNAs [187]. The net effect of Dicer downregulation on the immune system of the infants in our cohort is therefore not known.

TLR4 is a pathogen recognition receptor that shows affinity for RSV F-protein, and is expressed on cells that survey tissues for foreign antigen, e.g. pulmonary dendrocytes. It is likely that TLR4 is involved in the initial recognition of RSV, driving the innate immune response, and promoting maturation of pulmonary dendrocytes. TLR4 downregulation at the time of RSV infection may thus make the immune system less responsive to RSV.

CCR7 is a cell surface receptor expressed by dendrocytes and T-lymphocytes. CCR7 promotes activated dendrocyte and naïve T-cell migration into the T-cell zone of lymph nodes. CCR7 is therefore necessary for antigen presentation to lymphocytes and recruitment of the adaptive immune system.

IL7R promotes survival and proliferation of naïve T-cells and recent thymic immigrants. Individuals with CCR7 or IL7R deficiency have a severe combined immunodeficiency. It is therefore clear that these proteins are necessary for adequate immune function.
When interpreting our results, it should be noted that research on these proteins often utilises knockout studies – the protein in question is taken out of the biological system. In our birth cohort we did not find lack of expression, but downregulation. We cannot make definitive conclusions about the effects of knockdown to the mRNA levels seen in our results, but assume that these have a deleterious effect on immune function. It is clear, however, that the children in our cohort did not have a severe immunodeficiency and had otherwise appeared healthy. We did not follow up the children to identify if they had an increased frequency of respiratory diseases, but both empirical experience and a range of studies indicate that infants admitted with RSV disease have a greater risk of repeat admission with other infections, and asthma [188]. Taken together, we suggest that our results support the theory that a delayed or muted response to RSV, within both the innate and adaptive immune systems, predisposes to more severe disease. Increased nasal viral load is associated with more severe RSV disease [147-151], and this may reflect a delayed immune response that allows greater replication of the virus before effective defensive mechanisms are in place. Increased viral replication may increase spread of the virus within the tissue, and increase the concentration of virus particles in respiratory secretions. This may in turn increase the likelihood of aspiration of viral particles to the lower respiratory tract and subsequent lower respiratory tract infection. When the immune system is finally activated, the increased tissue viral load may result in an excessive response with increased oedema and respiratory secretions culminating in more severe disease.
miRNA regulation of the immune response to RSV

We profiled miRNA expression in the nasal mucosa of RSV positive infants on assessment in the emergency ward, comparing them to healthy controls. We found upregulation of 5 miRNA and downregulation of 7 miRNA in RSV disease.

In a literature search we found that 11 of these miRNA have functions described in the immune system, with specific mRNA targets identified. For statistically significant study miRNA, a brief summary of these functions is given in table 8 miRNAs that target innate immune genes are particularly represented, including direct targeting of TNFα and modulators of NF-κB signaling. Figure 13 shows a simplified NF-κB signalling cascade detailing the targets of 8 miRNA differentially regulated in our study and the presumed effect of our study results on NF-κB activation.

Several of our significant miRNA also have other functions related to Dendritic cells (DCs), Macrophages (MΦs) and T-lymphocytes. Neutrophils were particularly abundant in our nasal samples, but there are few investigations of the role of miRNA in neutrophil populations. After antigen challenge miR-125b promotes DC differentiation, and miR-155 promotes DC maturation. miR-125b also activates MΦs, whilst both miR-125a and miR-125b have specific effects on MΦ polarization to different MΦ functional subsets. miR-155 also promotes proliferation of all myeloid cells.

In T-lymphocytes, miR-125a reduces CCL5 production. miR-125b contributes to maintaining the naïve state. Let-7d contributes to cellular immunity by promoting Th1 differentiation.
Figure 13: Study miRNA that target TNFα, TLR4 and NF-κB signalling

Note: Both the TNFα receptor (TNFR) and TLR4 can activate NF-κB, although this is not the only signalling pathway activated by these receptors. TLR4 activation is dependent upon the adapter protein MyD88. Signal cascades activated by TNFR or TLR4 culminate in activation of Inhibitor of NF-κB Kinase (IKK) which is made up of sub-units IKKα, IKKβ and IKKγ. Activated IKK removes inhibitory proteins from cytoplasmic NF-κB, allowing nuclear translocation, DNA-binding and gene transcription. Significant study miRNA for which there is evidence of transcription by NF-κB are also listed. SOCS1, TNFα-induced protein 3 (TNFAIP3) and Silencing mediator for retinoid or thyroid-hormone receptors (SMRT) each negatively regulate the signal cascade by inhibition of specific proteins. Their mRNA is targeted by study miRNA. miR-16 appears three times; miR-125b and miRNA-203a appear twice.

miRNAs are coded red or blue according to the overall effect their up- or downregulation will have on NF-κB signaling. Blue capped lines: inhibition via RNA interference or protein-protein interactions. Thin black arrows: a biological pathway or activation. Short black arrows: up- or downregulation in our study. Yellow boxes: proteins whose mRNA is targeted by study miRNA. Only statistically significant miRNA are included.
miR-34b and miR-34c were strongly downregulated in disease, but their role in the immune system has been little studied. Damage Associated Molecular Pattern molecules (DAMPs), cellular molecules exposed upon tissue damage, induce miR-34c expression in peripheral blood monocytes [189]. In umbilical cord blood monocytes, IFN-γ induced miR-34c expression, but not the TLR4 ligand LPS [190]. miR-34b is upregulated in mouse lung tissue 6 hours after injection with LPS [191]. Specific immune functions of miR-34b have not been identified, but in peripheral blood monocytes, miR-34c negatively regulates NF-kB signaling by targeting IKKγ (see figure 13).

In summary, it is clear that the miRNA differentially regulated in this study have a number of specific functions in the immune system, especially the innate immune system. There are also likely to be number of functions that are not yet described, and which are relevant to clinical RSV disease. Interpreting the overall effects of the regulation seen in this study is difficult. Some miRNAs have similar effects, but are regulated in opposite directions. Other miRNAs have opposite effects, but are regulated in the same direction. It is also likely that the miRNA we have investigated are derived from several cell lines, particularly nasal epithelial cells and neutrophils, and this further complicates efforts to draw conclusions. We have, however, identified a set of miRNA that are regulated in RSV disease, in line with our aims and the motivation for doing this study.

Differences in miRNA expression according to disease severity
Two of our miRNAs, miR-125a and miR-429 showed evidence of differential expression between mild, moderate and severe disease. In addition, miR-223 showed a tendency to differential expression in the moderate group. This group
was particularly small (only 7 samples in the final analysis), so data should be interpreted with caution. There may be a type 1 error due to lack of power, alternatively the moderate group may not be representative of the population from which study participants were drawn.

Activated MΦs express miR-125a, and miR-125a can directly effect MΦ polarization. miR-125a positively regulates NF-κB signaling by targeting the inhibitor protein TNFAIP3, and negatively regulates CCL5, an important chemoattractant for both myeloid and lymphoid leukocytes. Studies of miR-223 show that it negatively regulates innate immunity via several confirmed targets. In addition to negative modulation of NF-κB by targeting IKKα during monocyte-to-macrophage differentiation [192] (see figure 13), miR-223 reduces IL6 and IL-1β production in MΦs by targeting the Inflammasome protein NOD-like receptor family, pyrin domain containing 3 (NLRP3) [193] and the cytokine induced transcription factor Signal transducer and activator of transcription 3 (STAT3) [194]. miR-223 is also expressed in NK-cells, where it has been shown to target Granzyme B, an apoptosis inducing protease found in the granules of NK cells.

In our study, the severe disease group showed a lack of miR-125a downregulation, and a possible lack of miR-223 upregulation. Given the functions described above, this corresponds to increased innate immune activation in the severe disease group. As postulated above, reduced Dicer, TLR4, CCR7 and IL7R expression prior to RSV exposure may result in delayed innate and adaptive immune responses, greater viral replication and viral load, and therefore greater pathology when the immune system does respond. The
differential expression of miR-125a, mir-429 and miR-223 may therefore be the result of this increased response.
<table>
<thead>
<tr>
<th>miRNA</th>
<th>Expression in this study</th>
<th>Functions in the immune system†</th>
<th>Specific targets†</th>
<th>Disease associations (Human studies)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>miR-29c</strong> Downregulation</td>
<td>Positive regulation of NFκB signaling [198]</td>
<td>TNFAIP3 [198]</td>
<td>Tuberculosis – upregulated [195]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Influenza→Upregulation. †† Promotes apoptosis</td>
<td>BCL2L2 [199]</td>
<td>Asthma – downregulated [196]</td>
</tr>
<tr>
<td><strong>miR-31</strong> Upregulation</td>
<td>Proliferation of myeloid cells [78]</td>
<td>E-selectin [200]</td>
<td>Influenza – downregulated [104]</td>
<td></td>
</tr>
<tr>
<td><strong>miR-155</strong> Upregulation</td>
<td>Proliferation of myeloid cells [78]</td>
<td>SHIP1 [78]</td>
<td>Influenza – downregulated [104]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dendritic cell maturation [77, 201]</td>
<td>PU.1 [201], SOCS1, CD115, KPC1 [77]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Positive regulation of TLR signaling [202]</td>
<td>SOCS1 [202]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Required for dendritic, B and T-cell function [204]</td>
<td></td>
</tr>
<tr>
<td><strong>let-7d</strong> Upregulation</td>
<td>Promotes T&lt;sub&gt;H&lt;/sub&gt;1 polarization [203]</td>
<td>IL13 [203]</td>
<td>Asthma – downregulated [197]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Endotoxin→Upregulation (TLR4 dep.) †† [191]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 8: Functions, known targets and disease associations of miRNA differentially expressed in paper 3

---

Both innate and adaptive immune functions
<table>
<thead>
<tr>
<th>miRNA</th>
<th>Expression in this study</th>
<th>Functions in the immune system</th>
<th>Specific targets</th>
<th>Disease associations (Human studies)</th>
</tr>
</thead>
<tbody>
<tr>
<td>miR-16</td>
<td>Upregulation</td>
<td>Negative regulation of inflammation [206]</td>
<td>TNFα [206]</td>
<td>Sepsis – improved survival [205]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Positive regulation of NFκB signaling [207]</td>
<td>SMRT [207]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Negative regulation of NFκB signaling [192]</td>
<td>IKKα [192]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Endotoxin→Upregulation (TLR4 dep) †† [191]</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Endotoxin→Upregulation (NFκB dep) †† [209]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-34b</td>
<td>Downregulation</td>
<td>Endotoxin→Upregulation (TLR4 dep) †† [191]</td>
<td></td>
<td>Asthma – downregulated [196]</td>
</tr>
<tr>
<td>miR-34c</td>
<td>Downregulation</td>
<td>Negative regulation of NFκB signaling; DAMPs*→Upregulation [189]</td>
<td>IKKγ [189]</td>
<td>Asthma – downregulated [196]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>In cord blood monocytes: IFNγ→Upregulation ††; Endotoxin→no regulation †† [190]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>miRNA</td>
<td>Expression in this study</td>
<td>Functions in the immune system</td>
<td>Specific targets†</td>
<td>Disease associations (Human studies)</td>
</tr>
<tr>
<td>--------</td>
<td>--------------------------</td>
<td>--------------------------------</td>
<td>-------------------</td>
<td>-------------------------------------</td>
</tr>
<tr>
<td>miR-125a</td>
<td>Downregulation</td>
<td>Positive regulation of NFκB signaling [212]</td>
<td>TNFAIP3 [212]</td>
<td>CERS** – upregulated [210]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inhibits CCL5 production [213]</td>
<td>KLF13 [213]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Macrophage polarization [220, 221]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-125b</td>
<td>Downregulation</td>
<td>Positive regulation of NFκB signaling [212]</td>
<td>TNFAIP3 [212]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Negative regulation of NFκB signaling [202]</td>
<td>MyD88 [202]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Activates Macrophages [214]</td>
<td>IRF4 [214]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Modulates dendritic cell differentiation [215]</td>
<td>PRDM1 [215]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Maintains naïve state in CD4+ T-cells [216]</td>
<td>IFNγ, IL2RB, IL10RA, PRDM1 [216]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Negative regulation of inflammation [217]</td>
<td>TNFα [217]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Negative regulation of inflammation [219]</td>
<td>TNFα, IL24 [219]</td>
<td></td>
</tr>
<tr>
<td>miRNA</td>
<td>Expression in this study</td>
<td>Functions in the immune system†</td>
<td>Specific targets†</td>
<td>Disease associations (Human studies)</td>
</tr>
<tr>
<td>----------</td>
<td>--------------------------</td>
<td>---------------------------------</td>
<td>-------------------</td>
<td>-------------------------------------</td>
</tr>
<tr>
<td>No known immune functions</td>
<td>Downregulation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-429</td>
<td>Downregulation</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

†: The functions described are extracted from a literature search April 2014, excluding studies of cancer. This table is not an exhaustive description of all known functions for each miRNA. All studies were carried out in vitro in specific cell lines, and caution should be taken when extrapolating the data to other cells or biological systems, including clinical disease. Specific miRNA functions and mRNA-target verification have in most cases been determined both by removing the miRNA from the biological system and by stimulating the system with the miRNA. miRNA upregulation after cell stimulation has usually been verified by simple qPCR. miRNA responses to cell stimulation have been included in the table if deemed relevant to virus infection or TLR/NFκB signaling.

††: X→Upregulation (Y dep): indicates that in-vitro stimulation of a cell line with molecule or pathogen X causes upregulation of the miRNA, and that this is dependent upon protein Y.

*DAMPs: Damage-associated molecular pattern molecules

**CERS - Chronic Eosinophilic Rhinosinusitis
7. Summary and Conclusions

In cord blood, we found downregulation of *IL7R* and *CCR7* in infants who later tested positive for RSV, and amongst those with severe disease we found downregulation of *TLR4* and *Dicer*. These genes are important for gene regulation, activation of the immune response to RSV and communication between innate and adaptive immune systems. We have also described twelve miRNA up- or downregulated in the nasal epithelium of children during RSV disease, many of which have functions directly related to modulation of the innate immune response. Two miRNA showed differential regulation between disease groups, whilst a third tended to this. Two these three disease severity-associated miRNA have specific targets involved in the modulation of macrophage activation, polarization and function.

We propose that downregulation of immune genes in infants prior to RSV exposure leads to delayed innate and adaptive immune activation on infection, allowing greater viral load, spread of virus to the lower airways and an increased inflammatory response when the immune system is finally activated.
8. Future Perspectives

This project has identified genes and miRNAs that were previously not associated with risk for RSV infection. We have also taken a number of miRNAs for which immune functions were described in vitro and associated them with a clinical disease, supporting their role in the immune system in vivo. We suggest the following avenues of investigation to further expand the knowledge generated by this project:

1. IL7R promotes naïve T-lymphocyte homeostasis. The T-lymphocyte population may therefore be reduced in children with severe RSV disease. We are collaborating with Dr Tore Abrahamsen and Dr Kiran Aftab Gul at the Department of Pediatric Research, Rikshospitalet, Oslo to investigate this further.

2. TLR4, CCR7, IL7R and Dicer expression in the nasal mucosa and peripheral blood of infants at the time of RSV infection. Expression patterns may be similar to those found at birth, but may also be ablated by the active infection. Unfortunately RNA quality did not allow us to do this in the current project, but may be improved in future studies. Protein quantification alone may also be an option.

3. Gene methylation studies have not previously been done in children with RSV. Such studies are expensive and require a large number of patients. Gene methylation of TLR4, IL7R, CCR7 and Dicer may explain our results.

4. Single nucleotide polymorphisms may also explain our results. Some TLR4 SNPs are already associated with RSV, but CCR7, IL7R and Dicer have not been investigated.
5. Nasal mucosal gene and miRNA expression after clearing RSV infection may also give insights into their relevance for predisposition to disease severity.

6. Intriguingly, there is a similarity between our findings and the miRNA profile described in a study of bronchial epithelial brushings from adult patients with asthma [196]. Given the known epidemiological association between RSV and later development of asthma, and the airways obstruction that can occur in bronchiolitis, we consider that this finding deserves closer attention in later investigations addressing a possible causative role of RSV in asthma development.

7. The miRNA identified in our study deserve more attention in vitro. As an example, immune cells (e.g.: umbilical cord blood monocytes) could be stimulated with RSV. Which miRNA are induced? What is the effect on cell function of inhibiting selected miRNA? What are the targets of these miRNA? We are currently developing a flow-cytometry based technique which we can use to investigate these questions.

8. Finally, we have identified gene expression patterns at birth that may be predictive for later RSV disease. If we can predict which children are at risk of disease then this may influence policies for RSV prophylaxis. In a new birth cohort we could measure expression of selected genes and thus categorize children according to risk. Prospective follow up could determine the validity of this approach for predicting RSV disease. We would, however, require several thousand neonates in order to include enough children with RSV disease severe enough to require hospital admission.
9. References

100. Otsuka M, Jing Q, Georgel P, et al. Hypersusceptibility to vesicular stomatitis virus infection in Dicer1-deficient mice is due to impaired miR24 and miR93 expression. Immunity 2007; 27:123-34.


208. Zhou R, Gong AY, Eischeid AN, Chen XM. miR-27b targets KSRP to coordinate TLR4-mediated epithelial defense against Cryptosporidium parvum infection. PLoS pathogens 2012; 8:e1002702.


