
Downloaded from: http://researchonline.lshtm.ac.uk/4651226/

DOI: https://doi.org/10.17037/PUBS.04651226

Usage Guidelines:

Please refer to usage guidelines at https://researchonline.lshtm.ac.uk/policies.html or alternatively contact researchonline@lshtm.ac.uk.

Available under license: http://creativecommons.org/licenses/by-nc-nd/2.5/
Symptomatic dengue and adverse pregnancy outcomes: a population-based record linkage study

Enny da Paixao Cruz
Department of Infectious Disease Epidemiology
Faculty of Epidemiology and Population Heath

Thesis submitted in accordance with the requirements for the degree of Doctor of Philosophy
University of London
April 2018
Funded by National Council for Scientific and Technological Development (CNPq)

Supervisors: Prof. Laura C. Rodrigues  
Faculty of Epidemiology and Population Health,  
Department of Infectious Disease Epidemiology  
London School of Hygiene and Tropical Medicine  

Dr Katie Harron  
Faculty of Public Health and Policy  
Department of Health Services Research and Policy  
London School of Hygiene and Tropical Medicine  

Dr. Elizabeth Brickley  
Faculty of Epidemiology and Population Health,  
Department of Infectious Disease Epidemiology  
London School of Hygiene and Tropical Medicine  

Co-supervisor: Maria Gloria Teixeira  
Institute of Public Health  
Federal University of Bahia  

Advisory Committee members:  

Dr Hannah Blencowe  
Faculty of Epidemiology and Population Heath  
Department of Infectious Disease Epidemiology  
London School of Hygiene and Tropical Medicine  

Collaborating Institutions:  
1. London School of Hygiene & Tropical Medicine  
2. Federal University of Bahia- Brazil
DECLARATION OF WORK

I, Enny da Paixao Cruz, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.
Dengue stands out from other viral mosquito borne diseases because it is the most common; its incidence is growing and it is reaching new geographical areas and spreading worldwide. Indeed, reported cases of negative pregnancy outcomes after dengue infection is not new, however, the evidence of the association between maternal dengue and adverse pregnancy outcomes is limited, controversial and mostly supported by cases reports. The aim of this thesis is to explore the relationship between pregnancy outcome and symptomatic maternal dengue.

I conducted a population-based study using routinely-collected Brazilian data from 2006-2012. These data have the information required to expand the existing knowledge on birth outcomes from women with dengue acquired during pregnancy.

The linkage process imposed complex challenges, and the final linked data showed a low sensitivity. However, it is unlikely that the linkage error introduced bias on the final analysis since it occurred randomly between cases and the comparison group.

Our main findings showed that dengue during pregnancy is associated with adverse maternal and birth outcomes. The effect was higher in the acute disease period (first 10/20 days) and severe disease increased the magnitude of the association. Therefore in areas where dengue virus is circulating, the health of pregnant women should be not only a public health priority, but health professionals attending pregnant women with dengue should more closely observe these patients to be able to intervene in a timely way and avoid adverse outcomes.
# CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>8</td>
</tr>
<tr>
<td>ABBREVIATIONS</td>
<td>9</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>10</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>13</td>
</tr>
<tr>
<td>Section 1- LITERATURE REVIEW</td>
<td>15</td>
</tr>
<tr>
<td>Literature review section</td>
<td>16</td>
</tr>
<tr>
<td>Overview</td>
<td>16</td>
</tr>
<tr>
<td>Chapter 1. Introduction on adverse outcomes of maternal infection during pregnancy</td>
<td>18</td>
</tr>
<tr>
<td>Maternal infection</td>
<td>18</td>
</tr>
<tr>
<td>The effects of infections on birth outcomes</td>
<td>19</td>
</tr>
<tr>
<td>Stillbirth</td>
<td>19</td>
</tr>
<tr>
<td>Prematurity</td>
<td>22</td>
</tr>
<tr>
<td>Intrauterine growth restriction and low birth weight</td>
<td>24</td>
</tr>
<tr>
<td>Summary</td>
<td>26</td>
</tr>
<tr>
<td>Chapter 2. Literature review on dengue in general and dengue during pregnancy together with a systematic review on fetal outcomes. Dengue during pregnancy and adverse fetal outcomes a systematic review and meta-analysis (paper 1)</td>
<td>29</td>
</tr>
<tr>
<td>Overview</td>
<td>29</td>
</tr>
<tr>
<td>Dengue overview</td>
<td>29</td>
</tr>
<tr>
<td>Epidemiology and transmission</td>
<td>31</td>
</tr>
<tr>
<td>Clinical manifestations</td>
<td>33</td>
</tr>
<tr>
<td>General aspects of diagnosis, treatment and prevention</td>
<td>34</td>
</tr>
<tr>
<td>Risk of acquisition and clinical manifestation in pregnant women</td>
<td>37</td>
</tr>
<tr>
<td>Intrauterine transmission rates</td>
<td>38</td>
</tr>
<tr>
<td>Cover sheet</td>
<td>40</td>
</tr>
<tr>
<td>Paper 1 title: Dengue during pregnancy and adverse fetal outcomes: a systematic review and meta-analysis</td>
<td>41</td>
</tr>
<tr>
<td>Thesis rationale</td>
<td>51</td>
</tr>
<tr>
<td>Thesis aims and objective</td>
<td>52</td>
</tr>
<tr>
<td>Thesis overview</td>
<td>53</td>
</tr>
<tr>
<td>Funding</td>
<td>54</td>
</tr>
<tr>
<td>Ethical approval</td>
<td>54</td>
</tr>
<tr>
<td>Section 1 references</td>
<td>56</td>
</tr>
<tr>
<td>Section 2-METHODS</td>
<td>74</td>
</tr>
<tr>
<td>Methods Section</td>
<td>75</td>
</tr>
</tbody>
</table>
Overview ................................................................................................. 75

Chapter 3. Description of study area and design, data source and statistical analysis 77

Study design ............................................................................................ 77
Study Area ............................................................................................... 77
Data source – Brazilian Information System ................................................. 78
Statistical analysis ................................................................................... 86

Chapter 4. Linkage procedures and evaluation- Description of the stillbirth and live
birth linkage process ............................................................................ 88

Cover sheet ............................................................................................... 88

Paper 2 title: Evaluation of record linkage of two large administrative databases in a
middle income country: stillbirth and notifications of dengue during pregnancy in
Brazil ........................................................................................................ 89
Evaluation of records linkage of two large administrative databases: live births and
notifications of dengue ........................................................................... 99
Background ............................................................................................. 99
Results ...................................................................................................... 100
Discussion ............................................................................................... 100

Chapter 5. Linkage procedures and validation – description of the maternal death
linkage process ..................................................................................... 103

Cover sheet ............................................................................................... 104

Paper 3 title: Validating linkage of multiple population-based administrative
databases in Brazil ............................................................................... 106
Section 2 reference .................................................................................. 128

Section 3-RESULTS .................................................................................. 131

Results Section ....................................................................................... 132

Overview ................................................................................................. 132

Chapter 6. Description of the association between symptomatic dengue during
pregnancy and stillbirth .......................................................................... 133

Cover sheet ............................................................................................... 133

Paper 4 title: Symptomatic dengue infection during pregnancy and the risk of
stillbirth in Brazil, 2006–12: a matched case-control study ....................... 134

Chapter 7. Description of the association between symptomatic dengue during
pregnancy and birth outcomes ................................................................ 143

Cover sheet ............................................................................................... 143

Paper 5 title: Dengue during pregnancy and adverse birth outcomes: a cohort analysis
using routine data .................................................................................. 145

Chapter 8. Description of the association between symptomatic dengue during
pregnancy and maternal deaths ................................................................ 165
Cover sheet ........................................................................................................................................... 165
Paper 6 title: Dengue during pregnancy and maternal mortality: a cohort analysis using routine data ........................................................................................................................................... 167

Section 4- DISCUSSION AND CONCLUSION ................................................................................. 186
Discussion section ....................................................................................................................................... 187
Overview..................................................................................................................................................... 187

Chapter 9. Final discussion, conclusion and recommendations ................................................................. 188
Summary of linkage challenges and strength .............................................................................................. 188
Summary of main findings .............................................................................................................................. 189
Interpretation of main findings ...................................................................................................................... 191
Biological Mechanisms................................................................................................................................. 192
Limitations................................................................................................................................................... 195
Implications and recommendations .............................................................................................................. 196
Conclusion ................................................................................................................................................. 199

Section 4 references ................................................................................................................................... 200

Appendices ............................................................................................................................................... 202
Appendix I Ethics approval from Brazil........................................................................................................ 202
Appendix II Ethics approval from LSHTM .................................................................................................... 205
Appendix III Brazilian forms – live birth form (SINASC) ............................................................................ 206
Appendix IV Brazilian forms – death form (SIM) ......................................................................................... 207
Appendix V Brazilian forms – dengue form (SINAN).................................................................................... 208
ACKNOWLEDGEMENTS

Not unto us, O Lord, not unto us, but unto thy name give glory, for thy mercy, and for thy truth’s sake (Psalms115).

I would like to thank my supervisor, Prof. Laura Rodrigues, for her guidance, tolerance and support through this journey. Her kindness and mentorship helped me to achieve my goals and develop my skills as only a true mentor is capable to do. Three years ago I arrived in London as a PhD student without much experience and now I will leave this program ready to give one more step heading to my independence as a researcher and I have to express my gratitude to you Prof Rodrigues. Thank you!

My sincerest thank and gratitude to Prof Maria Gloria Teixeira and Mara Conceicao Costa. This amazing pair has followed me since my Master and has given me amazing opportunities and trusted me without hesitate. Your support and friendship has been very important during my entire career.

An especial thank to Katie Harron and Elizabeth Brickley. Katie supported me during the data linkage process from the beginning and Elizabeth was essencial in the end of my journey. Both of them stepped up to assume duties as my supervisor. You were brilliant and I really appreciate all your effort. Hope to have a long and productive partnership with you.

Thank you to all people that supported me during this process in Brazil and here in London, especially oona Campbell, Hannah Blencowe and my friends from LG22. I had so much fun with you!

I thank you my beloved husband, Tiago, who is my rock and gave up everything to follow me in this crazy dream, without never looking back. Your love, patient and strength made my journey smooth and pleasant, love you. Thank you my little Bean that in the uterus taught me how to be more patient and enjoy little things. Thank you to my amazing family (Pedro, Margo, Pepeu) who have always stood by me along the way, I really miss you.'
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DF</td>
<td>Dengue fever</td>
</tr>
<tr>
<td>DHF</td>
<td>Dengue haemorrhagic fever</td>
</tr>
<tr>
<td>DSS</td>
<td>Dengue shock syndrome</td>
</tr>
<tr>
<td>DENV</td>
<td>Dengue Virus</td>
</tr>
<tr>
<td>LBW</td>
<td>Low birth weight</td>
</tr>
<tr>
<td>IUGR</td>
<td>Intrauterine growth restriction</td>
</tr>
<tr>
<td>SGA</td>
<td>Small for gestational age</td>
</tr>
<tr>
<td>PTB</td>
<td>Preterm birth</td>
</tr>
<tr>
<td>HPV</td>
<td>Human papilloma virus</td>
</tr>
<tr>
<td>HIV</td>
<td>Human immunodeficiency virus</td>
</tr>
<tr>
<td>SINASC</td>
<td>Information System of Birth</td>
</tr>
<tr>
<td>SINAN</td>
<td>Information System for Notifiable Disease</td>
</tr>
<tr>
<td>SIM</td>
<td>Information System of Mortality</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
<tr>
<td>LSHTM</td>
<td>London School of Hygiene and Tropical Medicine</td>
</tr>
<tr>
<td>ICD</td>
<td>International classification of disease</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>NS1</td>
<td>Nonstructural protein</td>
</tr>
<tr>
<td>MM</td>
<td>Maternal mortality</td>
</tr>
<tr>
<td>OR</td>
<td>Odds ratio</td>
</tr>
<tr>
<td>RR</td>
<td>Risk Ratio</td>
</tr>
<tr>
<td>Th</td>
<td>T-helper</td>
</tr>
<tr>
<td>IL</td>
<td>Interleukin</td>
</tr>
<tr>
<td>TNF</td>
<td>Tumour necrosis factor</td>
</tr>
</tbody>
</table>
The tables in this thesis follow two numbering system, those written into the thesis body (which are numbered following the chapter they feature in), and tables in the papers (labelled with P preceding the number) referred in the paper.

<table>
<thead>
<tr>
<th>No</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Chapter 1</strong></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Maternal infections and adverse pregnancy outcomes</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td><strong>Chapter 2</strong></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Dengue diagnostics and sample characteristics</td>
<td>35</td>
</tr>
<tr>
<td>P1</td>
<td>Study selection</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td><strong>Chapter 3</strong></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Information available on SINASC</td>
<td>79</td>
</tr>
<tr>
<td>2</td>
<td>Brazilian Ministry of Health definition of the terms</td>
<td>80</td>
</tr>
<tr>
<td>3</td>
<td>Information available on SIM</td>
<td>82</td>
</tr>
<tr>
<td>4</td>
<td>International classification of disease-10 chapter XV (pregnancy, childbirth and puerperium)</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td><strong>Chapter 4</strong></td>
<td></td>
</tr>
<tr>
<td>P1</td>
<td>Deterministic rules used to create the gold-standard database</td>
<td>93</td>
</tr>
<tr>
<td>P2</td>
<td>Comparison of linkage strategies for the bespoke algorithm and Reclink III</td>
<td>94</td>
</tr>
<tr>
<td>P3</td>
<td>Performance of linkage algorithms and thresholds</td>
<td>95</td>
</tr>
<tr>
<td>P4</td>
<td>Association between linkage accuracy (using the bespoke algorithm) and characteristics of the cohort</td>
<td>96</td>
</tr>
<tr>
<td>1</td>
<td>Association between linkage accuracy and characteristics of the cohort (live births)</td>
<td>102</td>
</tr>
</tbody>
</table>
## Chapter 5

<table>
<thead>
<tr>
<th>P1</th>
<th>Missing data for the Brazilian Notifiable Diseases Information System, the Brazilian Live Birth Information System, and the Brazilian Mortality Information System, 2007-2012</th>
<th>114</th>
</tr>
</thead>
<tbody>
<tr>
<td>P2</td>
<td>Linkage accuracy</td>
<td>120</td>
</tr>
<tr>
<td>P3</td>
<td>Associations between linkage accuracy and maternal characteristics</td>
<td>121</td>
</tr>
</tbody>
</table>

## Chapter 6

<table>
<thead>
<tr>
<th>P1</th>
<th>Maternal Characteristics for 162188 stillbirths and 1586105 live singleton births in Brazil for 2006-2012</th>
<th>138</th>
</tr>
</thead>
<tbody>
<tr>
<td>P2</td>
<td>Maternal characteristics for 1586105 live singleton births in Brazil by dengue status</td>
<td>139</td>
</tr>
<tr>
<td>P3</td>
<td>Association between symptomatic dengue infection during pregnancy and stillbirth</td>
<td>139</td>
</tr>
</tbody>
</table>

## Chapter 7

<table>
<thead>
<tr>
<th>P1</th>
<th>Maternal characteristics, delivery details, and birth outcomes in relation to dengue status, Brazil, 2006-2012</th>
<th>155</th>
</tr>
</thead>
<tbody>
<tr>
<td>P2</td>
<td>Number of cases notified with dengue during pregnancy (N), risk ratio crude and adjusted for the association among dengue during pregnancy and preterm birth, low birth weight, small for gestational age. Brazil, 2006-2012</td>
<td>157</td>
</tr>
<tr>
<td>P3</td>
<td>Number of cases with dengue during pregnancy (N) and Odds Ratio for the association Dengue during pregnancy by severity of disease and adverse birth outcomes (preterm birth, low birth weight, small for gestational). Brazil, 2006-2012</td>
<td>158</td>
</tr>
<tr>
<td>P4</td>
<td>Number of cases with dengue during pregnancy (N) and risk Ratio for the association Dengue during pregnancy by timing of disease and adverse birth outcomes (preterm birth, low birth weight, small for gestational). Brazil, 2006-2012</td>
<td>159</td>
</tr>
</tbody>
</table>

## Chapter 8

<table>
<thead>
<tr>
<th>P1</th>
<th>Maternal characteristics and delivery details in relation to dengue status, Brazil, 2006-2012</th>
<th>177</th>
</tr>
</thead>
<tbody>
<tr>
<td>P2</td>
<td>Number of cases of dengue during pregnancy and Odds ratio crude and adjusted for the association between dengue during pregnancy and maternal deaths. Brazil, 2007-2012.</td>
<td>178</td>
</tr>
<tr>
<td>P3</td>
<td>Number of cases of dengue during pregnancy (N) and Odds Ratio crude and adjusted for the association Dengue during pregnancy by severity of disease and maternal deaths. Brazil, 2007-2012.</td>
<td>178</td>
</tr>
<tr>
<td>Chapter 9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Summary of main findings</td>
<td>191</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

The figures in this thesis follow two numbering system, those written into the thesis body (which are numbered following the chapter they feature in), and tables in the papers (labelled with P preceding the number) referred in the paper.

<table>
<thead>
<tr>
<th>No</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Chapter 2</strong></td>
<td></td>
</tr>
<tr>
<td>P1</td>
<td>Study selection</td>
<td>44</td>
</tr>
<tr>
<td>P2</td>
<td>Association between dengue infection during pregnancy and miscarriage</td>
<td>46</td>
</tr>
<tr>
<td>P3</td>
<td>Association between dengue infection during pregnancy and preterm birth</td>
<td>47</td>
</tr>
<tr>
<td>P4</td>
<td>Association between dengue during pregnancy and low birth weight or intra uterine growth restriction</td>
<td>47</td>
</tr>
<tr>
<td>1</td>
<td>Countries/areas at risk of dengue transmission, 2008</td>
<td>32</td>
</tr>
<tr>
<td>2</td>
<td>Aim of the thesis with corresponding objectives, sections, chapters and papers</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td><strong>Chapter 4</strong></td>
<td></td>
</tr>
<tr>
<td>P1</td>
<td>Number of records from Brazilian Information System of Notifiable Disease and Brazilian Information System of Mortality</td>
<td>92</td>
</tr>
<tr>
<td></td>
<td><strong>Chapter 5</strong></td>
<td></td>
</tr>
<tr>
<td>P1</td>
<td>Linkage strategy diagram</td>
<td>112</td>
</tr>
<tr>
<td>P2</td>
<td>Linkage description</td>
<td>119</td>
</tr>
<tr>
<td></td>
<td><strong>Chapter 6</strong></td>
<td></td>
</tr>
<tr>
<td>P1</td>
<td>Study profile</td>
<td>137</td>
</tr>
<tr>
<td>P2</td>
<td>Time between the first symptoms of dengue and stillbirth delivery</td>
<td>140</td>
</tr>
<tr>
<td>Chapter 7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>--------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>P1</td>
<td>Flowchart with linkage information Brazil, 2006-2012</td>
<td>151</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 8</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>Flowchart with the linkage information, 2007-2012</td>
</tr>
</tbody>
</table>
Section 1 - LITERATURE REVIEW
Literature review section

Overview

Currently the attention given to adverse pregnancy outcomes of maternal infection is growing and the volume of published studies in this particular field have increased substantially in the last two years. Much of this sudden interest can be attributed to the association between Zika infection during pregnancy and microcephaly.

The discovery of this new association has heated the debate of the harmful effects of maternal infections. Before Zika, only rare cases of adverse birth outcomes had been linked to a virus belonging to the genus flavivirus within the Flaviviridae family, such as West Nile virus\(^1\). In general, studies on the negative outcomes of viral maternal infection is limited, with some exceptions.

It is particularly important to investigate the effects that arboviruses (i.e. virus transmitted by the bite of an infected arthropod, mainly mosquitos and ticks)\(^2\) in the course of gestation produce on the mother and their fetus. This interest is mainly because some of them are already endemic in many countries, particularly in America and South East Asia, such as dengue virus. Others are emerging and spreading to new areas (e.g. Chikungunya and Zika). Control measures are failing and no vaccine is available for most of them.

The first chapter of this thesis is a non-systematic literature review that provides a global perspective of the most discussed maternal infections and their effects on pregnancy outcomes focusing on maternal mortality, stillbirth, preterm birth, low birth weight and small for gestational age and second presents some biological mechanisms to show how infections can affect outcomes (Table 1). There are more adverse fetal outcomes associated with maternal infections than those discussed in this literature.
review, such as early pregnancy loss, congenital anomalies, neonatal deaths and long term disability (cerebral palsy psychiatric disease), however, they will not be addressed here because they were not investigated in this study.

The literature on maternal dengue infection and potential outcomes associated with this disease will be discussed in the next chapter in detail.
Chapter 1. Introduction on adverse outcomes of maternal infection during pregnancy

Maternal infection

Successful pregnancy is an example of immunologic tolerance that women must endure to encompass paternal antigens expressed by the fetus. Although maternal-fetal tolerance is not fully understood, it is known that immunologic competence changes in the mother. There is evidence that pregnancy induces shifts in her systemic immune responses and maternal regulatory T cells. With the progress of the pregnancy, a general shift from T-helper type 1 (Th1) cell-mediated immunity to T-helper type 2 (Th2) occurs⁴. Nonetheless, this expanded tolerance can result in susceptibility to infection. This system is very complex, and resistance against most pathogens is not significantly compromised with pregnancy, except intracellular organisms, for example *Toxoplasma gondii*⁵,⁶. Furthermore, it has been proposed that strong Th1 response may overcome protective T-regulatory and Th2 activity, leading to fetal loss⁵.

When acquired during pregnancy, infections can make women more prone to complications, an example is the increased susceptibility to influenza A during pregnancy and pregnant women have higher rates of hospitalization and mortality compared with non-pregnant peers⁶–⁹. Among all pregnancy related deaths, infections account for over 12% and when the outcome is a miscarriage or stillbirth, the role of maternal infections is even higher, 39% and 16% respectively¹⁰. During pregnancy, those women, with an infectious disease are more likely to suffer maternal deaths than non-infected pregnant women, for example, measles and tuberculosis have been found to increase the risk of maternal deaths⁹¹¹ and 4 fold respectively¹².

Infections in pregnant women can either resolve themselves spontaneously or require treatment. Some remain localized and have no effect on the fetus, however, the
infecting agent may perhaps invade the bloodstream and subsequently infect the placenta and fetus. Invasion of the maternal bloodstream is common, yet in most cases, neither fetal nor placental infection results. In general the fetus is protected by maternal innate or pre-existing immunity. However, even without direct microbial invasion the fetus can be affected by fever, anoxia, toxins and hematologic impairments in the mother that disturb the pregnancy resulting in abortion, stillbirth, premature delivery, growth retardation.

**The effects of infections on birth outcomes**

The relationship between adverse fetal outcomes and maternal infection with a variety of microorganisms (bacteria, viruses, or protozoa) has been widely debated in the literature. However, the role of maternal infection on negative fetal outcomes has still not been consistently defined and appears to vary according to the pathogen involved, the gestational age and the severity of the maternal disease.

Table 1: maternal infections and adverse pregnancy outcomes

**Stillbirth**

Stillbirth can be defined as the death of a product of conception before the expulsion or complete extraction from the body of the pregnant woman, occurring after 22 week of pregnancy or fetus weighing more than 500g. However, gestational age and birth weight limits that distinguish a spontaneous abortion from a stillbirth vary across studies and regions, ranging from 20-28 weeks. Infection is more clearly associated with early (20-28 weeks) compared with late stillbirth, with an increasing proportion of infected-related stillbirth at lower birth weights. In this literature review, we considered the definition of stillbirth presented by the authors of each study, independent of gestational age or weight limits used. Stillbirth rates differ around the world: modelled
estimates calculate that in high income countries the rate of stillbirth is 4/1000 birth, compared with 26/1000 birth in low-middle income countries 19.

Infection can cause stillbirth by a variety of mechanisms. The intrauterine death may be a result of an overwhelming fetal infection, or the microorganism may interfere with organogenesis to such extent that viability is interrupted 5. Spontaneous termination can occur when the maternal infection leads to systemic illness where the mother is severely ill and the fetus may die without the organism being transmitted to placenta or the fetus, or the placenta may be directly infected resulting in reduced blood flow to the fetus 18. Finally, the infection may precipitate preterm labour, and some of these fetuses are too small to survive birth by caesarean section or vaginal labour and are born dead 18.

Estimates suggest that the worldwide proportion of antepartum stillbirth rate due to maternal infections and chorioamnionitis (colonisation, infection, and inflammation of the placenta and its membranes 20) is approximately 5%. In most cases, however, the cause of stillbirth is unknown 19, so the real role of infections on birth outcomes is likely to be underestimated 16, especially in low-income countries, which normally have higher rates of maternal infection and stillbirth 3. The proportion of infection related to stillbirth in early gestational age tend to be higher than in late stillbirths 21,22. In high income countries, infections accounted for 15% of stillbirths at less than 28 weeks and 6% of stillbirth at more than 28 weeks’ gestation 23.

Ascending bacterial infection is the most common infectious cause of stillbirths 3. These bacteria ascend from the vagina into the uterus, during early pregnancy or even before pregnancy, and enter the amniotic fluid and ultimately infect the fetus 13,24. From all the bacterial infections associated with stillbirth, special emphasis should be given to group B streptococcal (GBS) infection. The rates of GBS-related stillbirths range from
0.04/1000 births in England to 0.9/1000 births in USA and the proportion of GBS among infection related stillbirth could be as high as 50%.\(^{26}\)

In areas where syphilis has a high prevalence, more than 20% of all stillbirths are attributed to this disease.\(^{27,28}\) The most common cause of these deaths appears to be placental infection associated with decreasing blood flow to the fetus. Spirochetes also can cross the placenta and infect the fetus which can either die in utero or be born alive with congenital syphilis.\(^ {18}\)

Malaria is well known as cause of growth restriction, however the risk of stillbirth caused by this disease is difficult to estimate in most populations because, when the prevalence of malaria is high, it does not exist in isolation from other factors.\(^ {18}\) A systematic review showed that placental malaria (primarily in endemic areas) was associated with a higher risk of stillbirth, even after adjusted for parity (OR=2.2, 95% CI=1.5-3.2).\(^ {29}\) In Ethiopia, areas with unstable transmission women with placental parasitemia was associated with 7-fold increased risk of stillbirth, compared with aperasitemic women.\(^ {30}\)

Viral infections are involved in the occurrence of stillbirth, however, the pathological mechanisms and its importance as a cause of stillbirth are unclear, possibly because even in high income countries, virus culture and identification is complex and expensive.\(^ {16,31}\) The proportion of stillbirths due to viral infections have been estimated at 14.5%.\(^ {32}\) However, it is possible that a proportion of unexplained stillbirths could be undiagnosed viral infections.\(^ {33}\) Parvovirus B-19 as a cause of stillbirth has been described since 1984 and it seems to be one of the better understood mechanisms of how a virus can result in a fetal death. This virus can cross the placenta, infecting fetal erythropoietic
tissue, producing fetal anaemia, resulting in hydrops and fetal death. It also can attack cardiac tissue and cause stillbirth.

The magnitude of the association between maternal viral infection and stillbirth varies given the pathogen, the severity of the maternal illness, the gestational age and other factors. For example, maternal influenza requiring hospitalization was found to increase the risk of fetal death odds ratio 4.2 (95% CI 1.4-12.4), whereas the risk of stillbirth in pregnant women with a mild disease was 1.9 (95% CI 1.1-3.4), using women who were pregnant outside the pandemic as reference. The risk ratio of stillbirth among women with HIV infection was 1.67 (95% CI 1.0-2.6) compared to women without the disease. More recently, emerging viral diseases such as Ebola and Zika, have been associated with stillbirth although the magnitude and mechanisms of these associations remain unknown.

**Prematurity**

Preterm birth, defined as birth at less than 37 weeks of gestation, is responsible for a significant proportion of infant mortality and neurological disorders including delay in development. The prevalence of preterm birth varies from about 5% in European countries to 18% in some African countries, although the rates in lesser developed regions may be underestimated due to lack of country-level data. There are three main precursors of preterm birth. The first one is the delivery for maternal or fetal indication, in which the labour is induced or the infant is delivered by caesarean section. These in general account for 30%-35% of all preterm labour. The second one is spontaneous preterm labour with intact membranes, this accounts for 40%-45% of all preterm labour. Finally preterm premature rupture of membranes (PPROM), corresponding to 25%-30% of all preterm births. Preterm can also be divided by
gestational age, less than 28 weeks (extreme prematurity), at 28–31 weeks (severe prematurity), at 32–33 weeks (moderate prematurity), and at 34–36 weeks (late preterm) 43.

There are many maternal factors associated with preterm birth, such as demographic and economic status, nutritional status, pregnancy history, psychological characteristics, behaviours, biological markers and infection 43. Systematic and intrauterine infections are a frequent precursor of preterm birth, predominantly in early gestational ages 15,47.

Bacterial intrauterine infection is the most frequent infection associated with preterm birth 43, accounting for 25-40% of preterm births 15. The mechanism involves the host response to infection, when the pathogen is recognised by receptors increasing the production of pro-inflammatory cytokines and chemokines (e.g., interleukin 8, interleukin 1β, and tumour necrosis factor-TNF α) which can stimulate prostaglandin synthesis by the amnion and decidua and inhibit progesterone synthesis, stimulate uterine contractility, cervical ripening and the rupture of membranes resulting in preterm birth 15,43,48. The most common pathogens reported in the amniotic fluid are *Mycoplasma hominis* and *Ureaplasma urealyticum*, but other organisms have been isolated 3. Other bacterial infections such as periodontitis, bacterial vaginosis, urinary tract infection and systemic infections such as appendicitis have been associated with increasing preterm birth risk, however, the pathological mechanisms are unknown 3,15,49–54.

Compared to bacterial infection, evidence on the association between maternal viral infection and preterm birth is limited. In cases where the mother develops a severe systemic illness a preterm delivery can occur 3. However, in absence of systemic severe disease, the evidence that maternal viral infection leads to preterm birth is contradictory.
and mainly based on case reports \textsuperscript{7,38,55–60}. It seems unlikely that maternal viral infection plays a very important role in preterm birth; however, further research should clarify this.

**Intrauterine growth restriction and low birth weight**

The World Health Organization (WHO) defines low birth weight (LBW) as born alive weighing less than 2,500g \textsuperscript{61}. The prevalence of LBW is monitored through health system surveillance and household surveys. In 2013, approximately 22 million newborns (16\% of all births) had LBW. However, this is probably underestimated since half of the world’s infants are not weighed at birth \textsuperscript{62}. A study estimated that in low and middle-income countries in 2010 36\% of livebirths (43.3 million infants) were born either preterm or small for gestational age (SGA). The LBW definition can include infants with intrauterine growth restriction (IUGR), - birth weight lower than the 10th percentile for sex and gestational age specific \textsuperscript{63} and preterm births \textsuperscript{64}. Small for gestational age is generally used as proxy of IUGR. Overall in low and middle income countries, of 18 million low birth weight infants, 59\% were born term-SGA, whereas 41\% were born preterm, (from those born preterm, 16\% preterm and small for gestational age and 25\% preterm with appropriate size specific for gestational age and sex \textsuperscript{64}).

Low birth weight and growth retardation have been associated with infant mortality, as well as short and long term morbidity. Studies have shown that the earlier the gestational age and the lower the birth weight and the degree of growth restriction, the higher the risk of neurological complications, especially cerebral palsy. The estimated prevalence of term birth but small for gestational age in low and middle income countries ranges from 5.3\% of livebirths in East Asia to 41.5\% in South Asia \textsuperscript{64}.

There are many maternal factors associated with growth retardation, such as low maternal weight, smoking, preeclampsia, toxic exposure and infections \textsuperscript{3}. Between 5-15\%
of all IUGR are of infectious origin\textsuperscript{65}, however, studies into the overall impact of maternal infectious diseases as a cause of LBW are limited, especially in developed countries where growth retardation appears associated with other factors such as below average maternal size, poor nutrition status, hypertension. Therefore, in this settings it is difficult to estimate the portion of growth retardation explained only by infectious aetiology\textsuperscript{3}.

Maternal infections with TORCH (toxoplasmosis, rubella, cytomegalovirus, herpes virus and syphilis) are the main infections associated with growth retardation\textsuperscript{66}. The mechanism proposed for this relationship is fetal cell death caused by direct infection or by changes in placental or fetal blood flow.

Intra uterine growth restriction is the principal fetal outcome related with malaria, which accounts for about 70\% of IUGR in some regions\textsuperscript{67,68}, especially in primigravidae. The risk of low birth weight associated with malaria can be two to seven times higher for primigravid than multigravida women\textsuperscript{67}. A modelling study carried out in Africa in 2010 estimated that without pregnancy-specific protection, 12.4 million pregnant women would have been exposed to malaria, with 92\% of these having placental infection that leads to an estimated 900,000 low birth weight deliveries per year\textsuperscript{69}.

In addition to the viral infections contained in the TORCH (rubella, cytomegalovirus and herpes), HIV has been associated with an increased risk of low birth weight and small for gestational age (60\% and 30\% respectively\textsuperscript{38}). More recently, maternal exposure to Zika virus infection has been related with negative pregnancy outcomes. In the first cohort study published on Zika-related adverse outcomes, 9\% of the Zika-exposed babies were small for gestational age compared with 5.3\% of babies in control group\textsuperscript{70}.
Summary

Maternal and fetal adverse outcomes associated with infection, which occurred during pregnancy, have been studied for many years and there is enough evidence to conclude that some infections might lead to stillbirth, preterm birth and growth retardation. For example, there is extensive evidence about the role of maternal bacterial infections on preterm birth. However, the literature on the effects of maternal viral infection is controversial and scarce. Therefore, the occurrence of Zika highlighted the importance of carefully investigating the relationship between pathogen-specific maternal viral infection and pregnancy outcomes. The Zika emergence showed how little we know about the role of arbovirus on adverse fetal and maternal outcomes, even for dengue, that has been endemic in numerous countries for many years, there is not much information. In the next chapter, we will review the general literature on dengue and its association with pregnancy outcomes.
<table>
<thead>
<tr>
<th>Organism</th>
<th>Outcome</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Treponema pallidum</em></td>
<td>Stillbirth 27,28,71-74</td>
<td>In countries where the maternal prevalence is high, syphilis is a significant cause of adverse pregnancy outcomes, especially when untreated.</td>
</tr>
<tr>
<td></td>
<td>Preterm birth 27,28,54,72,75</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low birth weight 27,28,54,72,73,75</td>
<td></td>
</tr>
<tr>
<td>Neisseria gonorrhoeae</td>
<td>Stillbirth 76</td>
<td>Sexually transmitted disease seems to play a role in prematurity and low birth weight although this relationship may be being confounded by other risk factors. Association with stillbirth has just been documented by case reports.</td>
</tr>
<tr>
<td></td>
<td>Preterm birth 54,77-79</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low birth weight 54,80,81</td>
<td></td>
</tr>
<tr>
<td>Chlamydia trachomatis</td>
<td>Stillbirth 79,82,83</td>
<td>This sexually transmitted disease seems to play a role in prematurity. On the other hand, the association between Chlamydia and low birth weight and stillbirth is controversial and should be better investigated.</td>
</tr>
<tr>
<td></td>
<td>Preterm Birth 79,83-87</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low Birth Weight 88</td>
<td></td>
</tr>
<tr>
<td>group-B streptococcus</td>
<td>Stillbirth 24,25,89-92</td>
<td>Group-B <em>streptococcus</em> is considered a common infectious cause of stillbirth, mostly in developed countries.</td>
</tr>
<tr>
<td></td>
<td>Preterm birth 93-96</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low birth Weight 93,94</td>
<td></td>
</tr>
<tr>
<td>Escherichia coli</td>
<td>Stillbirth 24,92,97-99</td>
<td>E. coli is considered a common microorganism found in stillbirth tissue.</td>
</tr>
<tr>
<td></td>
<td>Preterm Birth 100-102</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low birth weight 102</td>
<td></td>
</tr>
<tr>
<td>Plasmodium falciparum, <em>Plasmodium vivax</em></td>
<td>Stillbirth 30,67</td>
<td>Although stillbirth and prematurity, especially in primigravidae, have been associated with malaria in endemic areas, birth weight reduction is the main outcome described.</td>
</tr>
<tr>
<td></td>
<td>Preterm birth 30,67,68,103-106</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low birth weight 30,67-69,104-108</td>
<td></td>
</tr>
<tr>
<td>Toxoplasmosis gondii</td>
<td>Stillbirth 109-112</td>
<td>It is no longer a health problem in developed countries, but the real impact of this disease on negative pregnancy outcomes in low income countries is unknown.</td>
</tr>
<tr>
<td></td>
<td>Preterm Birth 113</td>
<td></td>
</tr>
<tr>
<td>Human papilloma virus</td>
<td>Preterm birth 114-116</td>
<td>HPV infection is likely to result in placental damage leading to preterm birth. Studies that show an association between this infection and other negative outcomes have not been found.</td>
</tr>
<tr>
<td>Cytomegalovirus</td>
<td>Stillbirth 117-119</td>
<td>Exposure to Cytomegalovirus may be associated with stillbirth, preterm birth as well as fetal growth restriction.</td>
</tr>
<tr>
<td></td>
<td>Preterm birth 57,119</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LowBirth weight 56,120</td>
<td></td>
</tr>
<tr>
<td>Human Immunodeficiency Virus</td>
<td>Stillbirth 121,122</td>
<td>Although stillbirth has been associated with HIV, the risk of HIV probably causing fetal death is not due to</td>
</tr>
<tr>
<td></td>
<td>Preterm birth 38,121,123-125</td>
<td></td>
</tr>
<tr>
<td>Virus</td>
<td>Stillbirth</td>
<td>Preterm birth</td>
</tr>
<tr>
<td>-----------------------</td>
<td>------------</td>
<td>---------------</td>
</tr>
<tr>
<td>Parvovirus (B19)</td>
<td>Stillbirth</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Herpes simplex virus</td>
<td>Stillbirth</td>
<td>Preterm birth</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Measles</td>
<td>Stillbirth</td>
<td></td>
</tr>
<tr>
<td>Ebola</td>
<td>Stillbirth</td>
<td></td>
</tr>
<tr>
<td>Zika</td>
<td>Stillbirth</td>
<td></td>
</tr>
</tbody>
</table>
Chapter 2. Literature review on dengue in general and dengue during pregnancy together with a systematic review on fetal outcomes. Dengue during pregnancy and adverse fetal outcomes a systematic review and meta-analysis (paper 1).

Overview

In the first chapter, I discussed the general perspective of adverse outcomes of maternal infection and concluded that especially for viral maternal infection the literature is sparse. This chapter contains a literature review on dengue, the clinical manifestation of dengue when acquired during pregnancy, transmission rates and finally a systematic review on negative fetal outcomes, conducted in 2015 and published in 2016. By the end of this chapter, it is hoped to have established a general idea of the published evidence on adverse outcomes of maternal dengue infection when this study was initiated.

In the following paragraphs, I will inform the importance of dengue in the public health arena. Then, I briefly describe the most relevant aspects of the disease in order to provide enough information on the general aspects to maximize the readers understanding of the importance, mechanisms and potential solutions for dealing with dengue during pregnancy.

Dengue overview

Dengue is the most common viral vector borne disease worldwide. It is estimated that annually 390 million people are infected of which 96 million develop clinical symptoms 136, 500,000 are hospitalized and 2.5% of these die 137. Dengue is endemic in at least 128 countries mostly in South America and Southeast Asia. The disease is spreading to new areas, include Europe, where outbreaks have been reported in more than 10 countries since 2010 137.

There are four serotypes of dengue virus (DENV1, DENV2, DENV3, DENV4), infections with one serotype provides long immunity against that particular serotype;
cross immunity to the others is temporary. Infection by any serotype can cause dengue; co-circulation of four serotypes in the same place for many years generates a complex of antibody-mediated occurrences, such as protection and infection enhancement. The risk of severe dengue increases in subsequent infections, the mechanism that converts mild self-limited dengue to vascular permeability syndrome (DVPS) can be triggered by an immunopathological phenomenon infection of monocytes or macrophages, called antibody-dependent infection enhancement (ADE), which suppresses innate antiviral system, allowing logarithmic intracellular growth of dengue virus. Therefore, when individuals are infected in absence of cross-protective dengue antibodies, the dengue vascular permeability may ensue; this is the underlying pathophysiological mechanism of dengue haemorrhagic fever (DHF) / dengue shock syndrome (DSS).

Dengue infection can be asymptomatic or lead to a range of clinical symptoms. Most cases of dengue are self-limiting and characterised by fever, headache, retroocular pain, myalgia, nausea, vomiting, and a rash (dengue fever – DF): some cases progress to severe illness, DHF or DSS characterized by the rapid onset of capillary leakage accompanied by thrombocytopenia and injury to the liver. The complex pathogenesis of dengue disease is not completely clear, thus it is not known how to predict who will develop the severe form of the disease, but some risk factors such as age, ethnicity, immune system of the host, presence of chronic diseases, viral characteristics, sequential infections, among others, have been described.

There is a live attenuated tetravalent vaccine licensed in 2015 and approved in 19 countries initially recommended only in settings with a high disease burden. After licensed a mass vaccination effort were launched in Philippines and Brazil, targeting 1 million people. However, the reassessment of data from the clinical trials showed an increase risk of severe dengue in vaccinated individuals seronegative before immunization compared with unvaccinated controls, therefore the WHO is no longer
indicating the use of this vaccine. Antiviral drugs are ineffective. Treatment is symptomatic and is designed to treat the clinical manifestations and other conditions with emphasis on fluid replacement therapy and management of bleeding complications.

**Epidemiology and transmission**

During the 19th century, dengue outbreaks occurred sporadically. Only 9 countries had reported cases of the disease in 1950s, however, the average number of cases notified to World Health Organization increased from 908 in 1950-1959 to 514,139 in 1990-1999, reaching 3.2 million in 2015. These numbers are based on the passive detection of symptomatic cases and underestimate the true burden of the disease. Studies calculate that national surveillance systems report 2-28 fold fewer dengue cases than those that actually occurred. Mathematical modelling estimated that in 2013 the incidence of symptomatic cases was more than 58 million resulting in about 10,000 deaths worldwide. The case-fatality rate is about 1% in most endemic areas (with a range of less than 1%-5%).

The geographical areas in which dengue has circulated have expanded in the past decade, and now four serotypes are circulating in the tropical area of Southeast Asia, Latin America and the Pacific and across Africa. Although more than 70% of the population at risk of dengue live in southeast Asia and Western Pacific region, the number of cases and deaths in Latin America have increased considerably and outbreaks have been reported in the USA and southern Europe (autochthonous cases in France, Croatia, Madeira and Italy). Dengue has occurred mainly in urban areas, however, since 2000 rural outbreaks have become more common. Dengue occurs among all age groups, in Southeast Asia there is evidence of an increase in the disease burden towards older age groups and in America the available data on severity by age distribution are inconsistent. Both sexes were equally affected (Figure 1).
Dengue epidemiology is dependent on the ecology of the vector, so transmission has distinct patterns that are seasonal and cyclical reflecting factors that influence the dynamics of transmission.\(^{158}\) The virus infects the mosquito during a blood meal of a viremic individual and after the extrinsic incubation period, the virus can be transmitted to another human via mosquito bites. *Aedes aegypti* and *A. albopictus* are the main vectors of the disease. *A. aegypti* is the most capable vector as it is highly adapted to the urban environment entering homes and breeding in small pools of water, such as plastic cups and bottles. It feeds primarily on humans, frequently biting several times in a single meal with an almost imperceptible bite.\(^{159}\) *A. albopictus* has a wider geographical distribution and can be found in subtropical and temperate climates. Resilient and aggressive, this mosquito can survive in rural and urban areas and is relatively long lived (4-9 weeks) and
more able to survive through winter\textsuperscript{160}, potentially helping to expand dengue infections to new areas, farther from equator\textsuperscript{161}. The most common mode of dengue transmission is vector transmission, but non-vector arboviruses transmission has been reported from vertical transmission\textsuperscript{162}, blood transfusion\textsuperscript{163}, tissue or organ transplantation\textsuperscript{164} and nosocomial\textsuperscript{165}.

**Clinical manifestations**

Dengue infection may be asymptomatic or lead to a spectrum of clinical presentations that can even result in death\textsuperscript{138}. Symptoms usually begin 5-7 days after a bite from an infected mosquito\textsuperscript{166}. Typically, signs and symptoms of a mild dengue presentation are non-specific and include self-limited febrile illness characterised by a headache, retroocular pain, myalgia, nausea, vomiting, and a rash\textsuperscript{167}. Most dengue infections are asymptomatic or mild. Generally one in four infections is symptomatic and commonly result in complete recovery, however, the ratio of asymptomatic to apparent infections varies and depends on age, virus serotype, previous infection with dengue and population background\textsuperscript{168}.

During the febrile phase it is not possible to predict which small proportion of dengue cases patients will progress to more severe cases with symptoms such as systemic vascular leak syndrome, evidenced by growing hemoconcentration, hypoproteinemia, pleural effusions, and ascites, denominated dengue haemorrhagic fever (DHF). Compensatory mechanisms are up-regulated to maintain adequate circulation to crucial organs, narrowing the pulse pressure when loss of plasma becomes critical, dengue shock syndrome (DSS) occurs\textsuperscript{169}. This progression from febrile phase to critical phase normally occurs between 4-7 days of the illness and it is critical for physicians to monitor warning signs such as persistent vomiting, abdominal pain, tender hepatomegaly, a high or increasing haematocrit level which is concurrent with a rapid decrease in the platelet...
count, mucosal bleeding, and lethargy or restlessness. Atypical forms with severe hepatitis myocarditis and encephalopathy, among others, are being seen more often in endemic areas. These states of altered vascular permeability are short and revert to a normal level after about 48-72 hours with recovery of the patient’s symptoms.

**General aspects of diagnosis, treatment and prevention**

The clinical feature of dengue is unspecific and is therefore similar to other non-dengue febrile illnesses making clinical diagnosis difficult without laboratory tests. However, any definitive diagnosis needs sophisticated examinations that are difficult and expensive to implement in many endemic areas.

The appropriate laboratory test for dengue depends on the timing of the sample collection. Virus isolation is the most specific test; however, virus culture is not easily available and maintainable. During the viremic phase, on average 5 days after the onset of symptoms, it is possible to detect the virus nucleic acid in serum by reverse-transcriptase–polymerase-chain-reaction (RT-PCR) or identify the virus non-structural protein (NS1). Because viremia decreases in a very short period of time, a negative RT-PCR does not rule out the suspected diagnosis; serological testing should be performed.

Seroconversion IgM or IgG is frequently used in suspected dengue cases. The presence of IgM or high levels of IgG in acute serum suggests a probable infection. However, due to strong cross reactivity a positive IgM test only indicates a recent flavivirus infection.

Plaque-reduction neutralization tests (PRNT) can be used as an alternative test. It can measure virus-specific neutralizing antibodies and may be able to determine the cause of the primary infection with high specificity. However, in previously infected or vaccinated patients cross-reactions may make it difficult to identify which flavivirus is causing the patient’s current infection. Serological test interpretation is complex and
should be carried out carefully, a positive test could be a result of cross-reaction and PRNT may be able to discriminate only in primary infections. A negative IgM test result obtained 2-12 weeks after travel suggests that infection did not occur\(^{172}\) (Table 1).

<table>
<thead>
<tr>
<th>Clinical sample</th>
<th>Diagnostic method</th>
<th>Methodology</th>
<th>Time to results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Virus detection and its components</td>
<td>Acute serum (1-5 days of fever) and necropsy tissues</td>
<td>Viral isolation</td>
<td>Mosquito or mosquito cell culture inoculation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nucleic acid retection</td>
<td>RT-PCR and real time RT-PCR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Antigen detection</td>
<td>NS1 Ag rapid tests</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NS1 Ag ELISA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Immuno-histochemistry</td>
</tr>
<tr>
<td>Serological response</td>
<td>Paired sera (acute serum from 1–5 days and second serum 15–21 days after)</td>
<td>IgM or IgG seroconversion</td>
<td>ELISA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>HIA (haemagglutination inhibition essay)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Neutralization Test</td>
</tr>
<tr>
<td></td>
<td>Serum after day 5 of fever</td>
<td>IgM detection (recent infection)</td>
<td>ELISA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Rapid Tests</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IgG detection</td>
<td>IgG ELISA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>HIA</td>
</tr>
</tbody>
</table>

*Source: World Health Organization*
There is no effective antiviral drug available to treat dengue patients, so the management of the cases is supportive, with particular emphasis on fluid therapy and careful monitoring of warning symptoms. Patients without complications require oral fluids to compensate for fluid loss. With the earliest threat of severe disease, a parenteral line should be used to provide fluids and avoid haemodynamic instability and hypotension. Early fluids restitution can prevent other complications and restore the plasma volume.

Until very recently, the only approach available to control or prevent dengue transmission was vector control. However, the results of this have been very disappointing and vector-control programs have not prevented the spread of dengue.

Since 2016, a dengue vaccine (CYD-TDV) has been approved in 19 countries. This is a live attenuated tetravalent vaccine that has shown efficacy against virologically confirmed symptomatic dengue of 61% (95% CI 52-68) for any serotype and lower efficacy for DENV 1 and 2 (50% and 39% respectively) when evaluated in 2 parallel Phase 3 randomized clinical trials. There are also safety concerns due to the increased risk of hospitalization and severe dengue identified in the 2-5 years old ages group. As a result, the vaccine is only recommended for children over 9 years old. Another concern about this vaccine is that regardless the age at vaccination, there was an increased risk of severe dengue in vaccinated individuals seronegative before immunization compared with unvaccinated controls. After the release of these new analyses, the WHO changed the recommendation from highly endemic regions to individual with proven past dengue infection. At this moment there are two more vaccines (Takeda TDV and another one from Butantan) that will finish the data collection of the phase 3 trial later this year.
Risk of acquisition and clinical manifestation in pregnant women

Most studies have shown a high seroprevalence rate of dengue among parturient of over 90% in endemic countries such as Brazil and Thailand. In Malaysia the seropositivity rate was much lower, approximately 36%. Advanced maternal age has been positively associated with seropositivity, indicating that younger women are at more risk of acquiring dengue during pregnancy while older pregnant women were more likely to have protective immunity.

It is difficult to know the exact incidence of dengue infections during pregnancy, because physiological changes such as hemodilution that occur during pregnancy can mask thrombocytopenia, leucopenia, or hemoconcentration and common obstetric problems can cause haematological and hepatic issues. This may result in difficulty in differentiating DHF from conditions such as HELLP syndrome leading to misdiagnosis. The incidence of dengue-specific IgM in pregnant women (indication of recent infection within the last 6 months) has been around 2.5% in most studies, ranging from 0.7% in Sudan in Africa to 10% in Northeast Brazil. Variation in the incidence rate of dengue can occur within the same country, for example in Brazil serological evidence of recent dengue was found in 2.8% of parturient in the Midwest and 10% in the Northeast.

There is some evidence that dengue during pregnancy is associated with a more severe presentation when compared with non-pregnant women. A publication using data from the Brazilian surveillance system showed that the risk of severe dengue was threefold higher among pregnant women compared to non-pregnant women. Hospitalization and mortality among pregnant women were more frequent relative to non-pregnant women. In Sri Lanka, among pregnant women, DHF was present in 56% of the patients after primary infection, while in another cohort of non-pregnant
hospitalized adults in the same country, only 24% developed DHF followed a primary infection\textsuperscript{189}. There is preliminary evidence that there is a high risk of severe dengue in the second and third trimester\textsuperscript{187,188,190}, however, it should be better investigated. Atypical presentations, although rare, have been reported such as cardiac\textsuperscript{191} and encephalic\textsuperscript{192} effects.

Dengue during pregnancy has been related with poor maternal outcomes; however, it is difficult to estimate the burden of this illness in this population since most of the studies are case series, so it is difficult to state if the values would be higher or lower than expected in the area where the study was conducted. Some studies have reported maternal deaths in women with dengue, such as Adam et al (2010) who reported maternal clinical complications which occurred among dengue infected patients, 6.4% had vaginal bleeding and 10% delivered by caesarean section. Maternal hemorrhage and difficulties in maintaining hemostasis during caesarian were also observed\textsuperscript{193}. In a study with 53 pregnant women exposed to dengue, the maternal case-fatality was 1.9%, 22% delivered by cesarean and 10% had hemorrhage at birth\textsuperscript{184}. Among 62 pregnant women with dengue in India, 15% developed severe dengue all in the third trimester\textsuperscript{190}. Two cohorts comparing pregnant women exposed and unexposed to dengue in Brazil and Colombia found more maternal deaths among the dengue exposed group\textsuperscript{194,195}. However, both studies had small numbers, only one maternal death occurred in the Colombian cohort and two in the Brazilian.

**Intrauterine transmission rates**

The dengue specific seroprevalence of IgG antibodies transferred at delivery is high, frequently over 90% in endemic regions\textsuperscript{179,181,186}, however, cases of vertical transmission are more rarely described at literature. The placenta is protective to the fetus though it is not always consistent or complete and increased vascular permeability and
endothelial damage in DHF may cause disruption of the placental barrier contributing towards vertical transmission of dengue. Once the virus reaches the placental tissue, pathological changes can be produced such as villous stromal edema, increases in the formation of syncytial knots and chorangiosis, resulting in hypoxia. The hypoxia itself could cause stillbirth, however, in most vertically transmitted dengue cases the newborn survived without sequelae.

The rate of vertical transmission varies, ranging from not observed to 15%. Apparently, vertical transmission is more likely to happen when dengue is clinically apparent and this depends on maternal severity and gestational age. All cases of congenital dengue have occurred in neonates born to mothers infected late in pregnancy, mainly when the symptoms occur at the time of delivery. In Malaysia, a study showed a vertical transmission rate of 1.6%, however, 89% of the women who were dengue IgM-positive did not report symptoms. Studies that only included clinically apparent dengue in pregnant women seem to report a higher number of neonates who are vertically infected (6.8% in Cuba and 15% in French Guiana). However, this is not always true and a study from India of eight pregnancies with symptoms of dengue and in another one among 16 parturient that reported fever in Brazil, no case of vertical transmission was identified.
## RESEARCH PAPER COVER SHEET

**PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED FOR EACH RESEARCH PAPER INCLUDED IN A THESIS.**

### SECTION A – Student Details

<table>
<thead>
<tr>
<th>Student</th>
<th>Enny da Paixao Cruz (Enny S Paixao)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Principal Supervisor</td>
<td>Elizabeth Brickley</td>
</tr>
<tr>
<td>Thesis Title</td>
<td>Symptomatic dengue and adverse pregnancy outcomes: an analysis using routine data</td>
</tr>
</tbody>
</table>

**If the Research Paper has previously been published please complete Section B, if not please move to Section C**

### SECTION B – Paper already published

<table>
<thead>
<tr>
<th>Where was the work published?</th>
<th>Lancet Infectious Disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>When was the work published?</td>
<td>March 2016</td>
</tr>
</tbody>
</table>

If the work was published prior to registration for your research degree, give a brief rationale for its inclusion.

<table>
<thead>
<tr>
<th>Have you retained the copyright for the work?*</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Was the work subject to academic peer review?</td>
<td>Yes</td>
</tr>
</tbody>
</table>

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

### SECTION C – Prepared for publication, but not yet published

**Where is the work intended to be published?**

**Please list the paper’s authors in the intended authorship order:**

**Stage of publication**

Choose an item.

### SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)

**Student Signature:**

**Supervisor Signature:**

**Date:** 8/3/18

Improve health worldwide

www.lshtm.ac.uk
Paper 1 title: Dengue during pregnancy and adverse fetal outcomes: a systematic review and meta-analysis

Authors: Enny S. Paixao, Maria Gloria Teixeira, Maria da Conceicao N Costa, Laura C. Rodrigues

Author contribution:

ESP and LCR designed the study. The literature search was done by ESP and article titles, abstracts, and full-text papers were screened by ESP, MCNC, and MGT. Full texts included in the meta-analyses were scored by ESP and LRC. All authors contributed to the writing of the article and approved the final version.

Permission from copyright holder to include this work:

Title: Dengue during pregnancy and adverse fetal outcomes: a systematic review and meta-analysis
Author: Enny S. Paixao, Maria Gloria Teixeira, Maria da Conceicao N Costa, Laura C. Rodrigues
Publication: The Lancet Infectious Diseases
Publisher: Elsevier
Date: July 2016
© 2016 Elsevier Ltd. All rights reserved.

Please note that, as the author of this Elsevier article, you retain the right to include it in a thesis or dissertation, provided it is not published commercially. Permission is not required, but please ensure that you reference the journal as the original source. For more information on this and on your other retained rights, please visit: https://www.elsevier.com/about/our-business/policies/copyright#Author-rights
Dengue during pregnancy and adverse fetal outcomes: a systematic review and meta-analysis

Emmy S. Pires da Silva, Maria Gloria Teixeira, Maria de Conceição N. Gouveia, Laura Rodrigues

Summary
Background Little is known about the possible adverse effects of dengue infection during pregnancy on fetal outcomes. In this systematic review and meta-analysis we aimed to estimate the increase in risk of four adverse fetal outcomes in women who had dengue infection during pregnancy.

Methods For this systematic review and meta-analysis, we searched Medline, Embase, Global Health Library, and Scopus for articles published before Aug 1, 2015. We included original studies that reported any fetal outcomes for pregnant women who had dengue infection during the gestational period. Case-control, cohort, and cross-sectional studies and unselected case series were eligible for inclusion. We excluded case reports, ecological studies, reviews, in-vitro studies, and studies without data for pregnancy outcomes. We independently screened titles and abstracts to select papers for inclusion and scored the quality of those included in meta-analyses. For each study, we recorded study design, year of publication, study location, period of study, and authors and we extracted data for population characteristics such as the number of pregnancies, dengue diagnostic information, and the frequency of outcomes. We investigated four adverse fetal outcomes: stillbirth, miscarriage, preterm birth, and low birthweight. We estimated the increase in risk of these adverse fetal outcomes by use of Mantel-Haenszel methods. We assessed heterogeneity of odds ratios (OR) with the I² statistic.

Findings We identified 278 non-duplicate records, of which 107 full-text articles were screened for eligibility. 16 studies were eligible for inclusion in the systematic review and eight were eligible for the meta-analyses, which included 6071 pregnant women, 292 of whom were exposed to dengue during pregnancy. For miscarriage, the OR was 3.51 (95% CI 1.15-10.77, P=0.0% p=0.765) for women with dengue infection during pregnancy compared with those without. We did not do a meta-analysis for stillbirth because this outcome was investigated in only one study with a comparison group, we calculated the crude relative risk to be 6.7 (95% CI 2.37-23.1) in women with symptomatic dengue compared with women without dengue. Preterm birth and low birthweight were the most common adverse pregnancy outcomes. The OR for the association with dengue was 1.71 (95% CI 1.06-2.76, P=5% p=0.055) for preterm birth and 1.41 (95% CI 0.98-2.12, P=0.0% p=0.343) for low birthweight.

Interpretation Evidence suggests that symptomatic dengue during pregnancy might be associated with fetal adverse outcomes. If confirmed, it would be important to monitor pregnancies during which dengue is diagnosed and to consider pregnant women in dengue control policies.

Funding National Council for Scientific and Technological Development (CNPq).

Introduction Dengue, a mosquito-borne viral disease, is endemic in more than 100 countries (mainly in South America and southeast Asia) and is spreading to new areas, with outbreaks of increasing magnitude and severity.1 It is estimated that each year, 390 million people are infected with dengue and 96 million develop clinical symptoms.2 A study of dengue seroprevalence in pregnant women in Brazil showed that recent infection (IgM positive) had occurred in 2.8% of participants.3 Most people with dengue infection either have no symptoms or have mild self-limited disease (including fever, headache, retro-ocular pain, muscle and joint pain, nausea, vomiting, and rash); a small proportion of infections progress to severe illness, with rapid onset of capillary leakage accompanied by bleeding, thrombocytopenia, and liver injury.4

There are four serotypes of dengue virus: DENV 1, DENV 2, DENV 3, and DENV 4. Infection with one serotype provides long-lasting immunity against that particular serotype, whereas cross-immunity to the other serotypes is temporary.6 The risk of severe dengue increases with subsequent infections.7 The complex pathogenesis of dengue disease is not completely understood, and accurate prediction of which patients will develop severe disease is not possible, although some risk factors for progression to severe disease have been identified, including age (mainly children),4 presence of chronic diseases8,9 sequential infections,10 and ethnic origin (African ancestry is protective against the severe form in admixed populations).11 No licensed vaccine exists for dengue and antiviral drugs are not effective. Treatment is symptomatic and targeted at clinical manifestations, mostly consisting of fluid replacement therapy and management of bleeding.12 Since women of reproductive age in endemic areas are at risk of dengue infection, whether dengue infection...
Research in context

Evidence before this study
Prenatal and low birthweight are among the main causes of neonatal and infant death and although awareness of the importance of stillbirth is increasing, the cause remains unknown. Evidence suggests a role in perinatal birth, low birthweight, and stillbirth for some infections such as syphilis, toxoplasmosis,rubella, cytomegalovirus, and herpes. Dengue is epidemic in many regions of the world, including Asia, Africa, and Oceania, but the effect of dengue infection during pregnancy on fetal outcomes is not well understood or documented. We searched Medline, Embase, Global Health, and Cochrane for articles published before Aug 1, 2015, using the terms “dengue”, “dengue haemorrhagic fever” AND “pregnancy outcomes”, “pregnancy complications”, “low birth weight”, “small for gestational age”, “intrauterine growth restriction”, “stillbirth”, “fetal death”, “preterm birth”, “preterm delivery”, “preterm labour”, “abortion”, or “miscarriage". We independently screened titles and abstracts to select the papers for inclusion and scored the quality of those included in meta-analysis. We identified 16 published studies that met our inclusion criteria. The research base is sparse and has many limitations, with only seven studies investigating the evidence in a comparative way. Some preliminary evidence suggests that dengue infection during pregnancy is associated with adverse fetal outcomes needs to be established. In 2010, a systematic review of dengue infection during pregnancy and fetal outcomes was published. The investigators reviewed 19 case reports, nine case series, and two cohorts, and concluded that vertical transmission is possible; however, the evidence was not sufficient to confirm whether dengue infection during pregnancy increases the risk of adverse outcomes. The effects of infection during pregnancy on fetal outcomes remain unclear.

Methods

Search strategy and eligibility criteria
We report this systematic review and meta-analysis in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement.6 We searched Medline, Embase, Scopus, and Global Library to identify studies reporting fetal outcomes in women who had dengue infection during the pregnancy. We searched databases up to Aug 1, 2015, with the terms “dengue” or “dengue haemorrhagic fever” AND “pregnancy outcomes”, “pregnancy complications”, “low birth weight”, “small for gestational age”, “intrauterine growth restriction”, “stillbirth”, “fetal death”, “preterm birth”, “preterm delivery”, “preterm labour”, “abortion”, or “miscarriage”. We supplemented our database searches by screening the bibliographies of the previous systematic review article.6 We used the explode function for dengue, pregnancy outcomes, and pregnancy complications. This function automatically includes all narrower terms in the hierarchical list during the search to retrieve citations that carry the specified MeSH heading (or subheading). We used no language restrictions. We reviewed all titles and abstracts of publications identified in the primary search for relevance and eligibility, after duplicates had been removed.

Eligible publications were original studies that reported any fetal outcome for pregnant women who had dengue infection during the gestational period. To avoid overlapping populations, if participants were included in more than one report, the study with the largest sample size was included. Eligible study designs were case-control, cohort, and cross-sectional studies and unselected case series (ie, those in which participants were selected independently of outcome). Case reports, ecological studies, reviews, in-vitro studies, and studies without information about pregnancy outcomes were excluded. We sought to extract patient-level data.

Article titles and abstracts were screened independently by two reviewers (ESPN and MHCNC) to select papers for full-text screening. Full texts were independently assessed by three reviewers. In case of disagreements, a third reviewer (MCJ) was consulted and a decision was agreed by consensus.
Data analysis
We used a spreadsheet to record information from eligible articles about study design, year of publication, study location, study period, and authors. We also recorded population characteristics such as the number of pregnancies, dengue diagnostic information, and frequency of each outcome.

We studied four adverse fetal outcomes for which data were available from more than one study: miscarriage, stillbirth, defined as fetal death in utero at or after 22 weeks of gestation or at a weight of more than 500 g; low birthweight, defined as birthweight less than 2500 g or intranatal growth restriction, defined as birthweight less than the tenth percentile for gestational age. We defined dengue infection during pregnancy by use of clinical criteria (symptoms of dengue), laboratory criteria (positive test from one of IGM detection by ELISA, viral RNA detection via PCR, NS1 viral antigen detection, or positive viral culture), or both.

Two authors (ESP, LCR) independently scored the quality of the studies included in meta-analyses in accordance with the Newcastle-Ottawa scale (NOS). This scale was used for cohort and case-control studies and a modified version was used for case series studies and cross-sectional studies. In the NOS, cohort and case-control studies are scored between zero and nine stars for nine questions that cover three items (selection, comparability, and outcome); cross-sectional studies are scored between zero and eight, and case series studies are scored between zero and six. We deemed nine stars to be 100%, so a perfect case series (six stars) would be assessed as scoring 67%. The final score was agreed between the two reviewers.

When effect estimates were not presented in the papers, we calculated the proportion of participants with outcome events on the basis of the data within each study. For stillbirths, preterm births, and low birthweight the denominator was the total number of pregnant women beyond week 22 of gestation, and for miscarriages the denominator was the total number of all pregnancies. We then estimated the 95% CIs for each outcome and study with the Poisson distribution, because of the small numbers of cases.

In cohort, case-control, and cross-sectional studies, we estimated odds ratios (OR) of fetal outcomes in pregnancies with and without dengue infection during pregnancy. We did meta-analysis for miscarriages, preterm birth, and low birthweight or intranatal growth restriction; we did not do a meta-analysis for stillbirths because this outcome was investigated in only one study with a comparison group. It was possible to estimate the OR for preterm birth in one of the case series studies because it provided the required comparative data for the same year. This allowed us to include this study in the meta-analysis with the other studies that had a comparison group. To estimate the increase in risk of adverse fetal outcomes, we used the Mantel-Haenszel test, because the data are sparse in terms of events and study size. This test makes an adjustment to the study weights according to the variation or heterogeneity, among the varying effects. We assessed heterogeneity of OR with the B statistic. We analysed the data with Stata version 14.0.

Role of the funding source
The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results
Our initial search identified 665 papers, 387 of which were duplicates (figure 1). After screening, 107 articles were assessed for eligibility and 91 were excluded because they did not meet the inclusion criteria. We included 16 articles14–31 in the systematic review (five cohorts, one case-control, one cross-sectional study, and nine case series), and eight32–39 in the meta-analysis (table). The studies were published from 1994 to 2014, and were done in ten countries (Brazil,33,34 Colombia,35 Cuba,36 France [French Guiana]37 India,38,39 Malaysia,40 Mexico,41 Sri Lanka,42,43 Sudan,44 and Venezuela).
<table>
<thead>
<tr>
<th>Country</th>
<th>Study design</th>
<th>Number of pregnant women admitted to birth cohort</th>
<th>Prevalence (%) of fetal outcomes in pregnant women with dengue infection</th>
<th>Other findings and comments</th>
<th>NOS score</th>
<th>Included in meta-analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Friedman et al., 2004</td>
<td>France, Gerays</td>
<td>Retrospective cohort</td>
<td>Pattern birth 17% (95% CI: 0.17-0.30) Low birth weight 10% (95% CI: 0.03-0.27)</td>
<td>Study used medical records from a maternity ward.</td>
<td>9</td>
<td>Yes</td>
</tr>
<tr>
<td>Brazil, 2003</td>
<td>Ref.</td>
<td>Prospective cohort</td>
<td>Low birth weight: none</td>
<td>Study was conducted in women admitted at the time of delivery.</td>
<td>7</td>
<td>Yes</td>
</tr>
<tr>
<td>Malaysia, 2007</td>
<td>Ref.</td>
<td>Prospective cohort</td>
<td>Term birth 31% (95% CI: 0.03-0.27) Low birth weight 9% (95% CI: 0.03-0.25)</td>
<td>Study was conducted in women admitted at the time of delivery.</td>
<td>7</td>
<td>Yes</td>
</tr>
<tr>
<td>Colombia, 2004</td>
<td>Colombia</td>
<td>Prospective cohort</td>
<td>Pattern birth 18% (95% CI: 0.17-0.25) Low birth weight: 10% (95% CI: 0.03-0.27)</td>
<td>No differences in fetal outcomes between dengue group and non-infected group for gestational age.</td>
<td>6</td>
<td>Yes</td>
</tr>
<tr>
<td>Cuba, 2009</td>
<td>Ref.</td>
<td>Prospective cohort</td>
<td>Intracranial perfusion infection: 3% (95% CI: 0.3-0.27) Non-perinatal occurred 6% (95% CI: 0.3-0.27)</td>
<td>Study was conducted in women admitted at the time of delivery.</td>
<td>6</td>
<td>Yes</td>
</tr>
<tr>
<td>Malaysia, 2007</td>
<td>Ref.</td>
<td>Case-control</td>
<td>6% dengue cases among asymptomatic controls</td>
<td>Study was conducted in women admitted at the time of delivery.</td>
<td>8</td>
<td>Yes</td>
</tr>
<tr>
<td>Venezuela, 2013</td>
<td>Venezuela</td>
<td>Cross-sectional</td>
<td>Pattern birth 38% (95% CI: 0.18-0.72)</td>
<td>Study was conducted in women admitted at the time of delivery.</td>
<td>6</td>
<td>Yes</td>
</tr>
<tr>
<td>Brazil, 2009</td>
<td>Ref.</td>
<td>Cases only</td>
<td>Pattern birth 10% (95% CI: 0.03-0.27) Low birth weight: 5% (95% CI: 0.03-0.27)</td>
<td>Study was conducted in women admitted at the time of delivery.</td>
<td>5</td>
<td>Yes</td>
</tr>
<tr>
<td>Sri Lanka, 2010</td>
<td>Sri Lanka</td>
<td>Cases only</td>
<td>Pattern birth 6.6% (95% CI: 0.03-0.27) Low birth weight: 6% (95% CI: 0.03-0.27)</td>
<td>Study was conducted in women admitted at the time of delivery.</td>
<td>4</td>
<td>Yes</td>
</tr>
<tr>
<td>Sudan, 2010</td>
<td>Sudan</td>
<td>Cases only</td>
<td>Pattern birth 10% (95% CI: 0.11-0.19) Low birth weight: 26% (95% CI: 0.14-0.31)</td>
<td>Study was conducted in women admitted at the time of delivery.</td>
<td>7</td>
<td>Yes</td>
</tr>
</tbody>
</table>

(Continued from previous page)
<table>
<thead>
<tr>
<th>Country</th>
<th>Study design</th>
<th>% of female outcomes</th>
<th>Other findings and comments</th>
<th>NOS score included in meta-analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brazil, 2016</td>
<td>Case series</td>
<td>36</td>
<td>Puerperal death: 25.4% (95% CI 17.3-33.4)</td>
<td>- No</td>
</tr>
<tr>
<td>Sri Lanka</td>
<td>Case series</td>
<td>28.6</td>
<td>Odds ratio: 1.4 (95% CI 1.13-1.78)</td>
<td>- No</td>
</tr>
<tr>
<td>Maldives, 2005</td>
<td>Case series</td>
<td>83.3</td>
<td>None of the women had a previous miscarriage</td>
<td>- No</td>
</tr>
<tr>
<td>China and Pakistan, 2017</td>
<td>Case series</td>
<td>34.0</td>
<td>Puerperal death: 18.9% (95% CI 12.6-26.5)</td>
<td>- No</td>
</tr>
<tr>
<td>India, 2014</td>
<td>Case series</td>
<td>35.0</td>
<td>Puerperal death: 18.6% (95% CI 12.9-26.4)</td>
<td>- No</td>
</tr>
<tr>
<td>Mexico</td>
<td>Case series</td>
<td>26.0</td>
<td>Puerperal death: 17.8% (95% CI 11.1-26.3)</td>
<td>- No</td>
</tr>
</tbody>
</table>

For stillbirths, puerperal deaths, and low birthweight the denominators were the total number of pregnant women exposed to dengue infection, and for miscarriage the denominators were the total number of pregnant women exposed to dengue infection and stillbirths or afterbirths and gestational age. The miscarriage rate is defined as the total number of miscarriages in the study period divided by the total number of pregnant women with confirmed dengue infection admitted to a maternity ward during the study period.

For stillbirths, puerperal deaths, and low birthweight the denominators were the total number of pregnant women exposed to dengue infection and stillbirths or afterbirths and gestational age. The miscarriage rate is defined as the total number of miscarriages in the study period divided by the total number of pregnant women with confirmed dengue infection admitted to a maternity ward during the study period.

Miscarriage as a potential adverse outcome associated with dengue infection during pregnancy was described in six case series,\textsuperscript{25,26,27,28,29,30} one case-control study,\textsuperscript{31} and one cohort study.\textsuperscript{32} Prevalence of miscarriage associated with dengue infection during pregnancy ranged from 3.8% (95% CI 0.0-6.1) in Sri Lanka\textsuperscript{33} to 13.8% (4.3-24.1) in India.\textsuperscript{34} In the single study that controlled for confounding,\textsuperscript{35} which was done in Malaysia, the OR for recent dengue infection was 4.2 (95% CI 1.2-14.4) for cases of miscarriage versus controls after adjustment for maternal age, gestational age, parity, and ethnic origin.\textsuperscript{36,37} We used two studies (a case-control and a cohort study)\textsuperscript{32} to do the meta-analysis for miscarriage as a pregnancy outcome potentially associated with dengue pregnancy: the crude overall OR was 3.5 (95% CI 1.15-10.77, P=0.0%, 5=0.765, Figure 2).

Stillbirths were investigated in four case series,\textsuperscript{23,24,38,39} and one cohort study.\textsuperscript{40} Prevalence of stillbirth in pregnant women who had dengue infection during pregnancy varied between 4.7% (95% CI 0.1-26.9) in India\textsuperscript{34} and 13.8% (1.6-48.0) in Sri Lanka.\textsuperscript{33} In three of the four case series, stillbirths occurred only in women who had severe dengue infection.\textsuperscript{23,26,38} We did not do a meta-analysis for stillbirths because this outcome was investigated in only one study with a comparison group. In the cohort study done in France Guiana,\textsuperscript{41} the crude relative risk (calculated from data presented in the Table) was 3.0 (95% CI 0.5-16.0) for stillbirths.

For stillbirths, puerperal deaths, and low birthweight the denominators were the total number of pregnant women exposed to dengue infection, and for miscarriage the denominators were the total number of pregnant women exposed to dengue infection and stillbirths or afterbirths and gestational age. The miscarriage rate is defined as the total number of miscarriages in the study period divided by the total number of pregnant women with confirmed dengue infection admitted to a maternity ward during the study period.
study) for stillbirth was 6.7 (95% CI 2.1-21.3) for women with symptomatic dengue compared with women without dengue.

Preterm birth and low birthweight (or intrauterine growth restriction) were the most common adverse pregnancy outcomes for women with dengue infection during pregnancy. There was substantial between-study variation in the prevalence of preterm birth and low birthweight, even in studies done in the same country. In studies in Malaysia, for example, the proportion of preterm births among women with dengue infection during pregnancy ranged from 3.8% (95% CI 0.3-11.0)% to 26.5% (7.2-68.0)%.

Figure 3: Association between dengue infection during pregnancy and preterm birth
All studies (A). Only studies including women with clinically diagnosed dengue symptoms in B. OR: odds ratio.

We assessed the quality of the studies included in the meta-analysis using the modified Newcastle-Ottawa scale. The cohort in French Guiana was awarded nine stars and the case-control study in Malaysia was awarded eight stars. Two prospective cohort studies that used only laboratory criteria to define dengue during pregnancy were awarded seven stars because they did not control for confounding factors. The studies by Restrepo and colleagues and Barros and colleagues (two small prospective cohort studies) and a cross-sectional study were awarded six stars. The case series with a comparison group was awarded five stars. In most of the studies included in the meta-analysis, the outcomes of interest were shown not to be present at the start of the study; only two studies were not assessed with criteria 4 of NOS (demonstration that outcome of interest was not present at start of study). All studies included in a meta-analysis received a star in two criteria: representativeness of the exposed cohort and adequacy of follow-up.

We did post-hoc sensitivity analyses for preterm birth and low birthweight because we found that the studies that included women with infection but without clinical symptoms had ORs less than 1. Among studies with preterm birth as an outcome, only Tan and colleagues defined dengue exposure based only on serology for recent dengue infection. In their study, some of the women classified as having dengue infection during pregnancy had no clinical symptoms. The OR for preterm birth in this study was 0.42 (95% CI 0.10-1.73) (figure 3). Sensitivity analysis excluding Tan and colleagues gave an OR of 2.50 (95% CI 1.43-4.34). (4.25-5.9), p=0.259 (figure 3). In a sensitivity analysis that excluded studies with quality assessments less than six stars, the OR was 2.46 (95% CI 1.24-4.5).
Consistent with the previous systematic review, the most common fetal outcomes were preterm birth and low birthweight. The increase in risk of adverse fetal outcomes associated with dengue infection during pregnancy was highly variable across studies, which we expected given the heterogeneity of studies in terms of site, study design, and control of confounding. We pooled intrauterine growth restriction with low birthweight because low birthweight includes infants born preterm and infants with intrauterine growth restriction. Because of the overlap between low birthweight and preterm birth, it would have been better to use intrauterine growth restriction or small for gestational age instead, but data were not available; only Barros and colleagues provided data about intrauterine growth restriction defined by Dueto's curves. All other studies classified neonates only as either low birthweight or normal birthweight. Since the publication of the systematic review by Pouillot and colleagues, more studies have included comparison groups, enabling us to do our meta-analyses.

The overall results for the meta-analyses of the association between dengue infection during pregnancy and preterm birth and low birthweight were not significant. However, two studies classified women as having dengue infection during pregnancy if they had positive IgM serology, even if they had no clinical symptoms. Neither of these two studies showed significant associations between adverse fetal outcomes and dengue infection during pregnancy. In one of these studies, 89% of the IgM-positive women did not have dengue-clinical symptoms; in the other study, this proportion was 63%. When we did a sensitivity analysis excluding these studies, the association with symptomatic dengue was significant for both preterm birth and low birthweight, suggesting that the presence of clinical symptoms and severity of disease are related to the risk of adverse fetal outcomes. The occurrence of stillbirth mainly in women who had severe dengue also supports this hypothesis. Consequently, we think that the presence of clinical symptoms is an important factor in the increased risk of adverse birth outcomes for women who have dengue infection during pregnancy. Further exploration of the role of dengue severity was not possible in the meta-analysis because studies did not report this information. It is, however, likely that the severity of disease varies between women in different studies and this might have contributed to estimated variations in ORs.

This literature review has some limitations. Of the 16 studies reviewed, 56% were case series, and, in the studies with comparison groups, sample sizes were small, which might have led to publication bias and consequently outcomes might have been overestimated. To our knowledge, our study is the first to include meta-analyses of fetal outcomes and dengue infection during pregnancy, and these meta-analyses also have some limitations. First, few studies had been done and many had imperfect...
methodology. Most of these studies were assessed as having less than seven stars in the NOS. Second, some studies have higher weight in the meta-analyses, so the results are mainly based on these studies. Third, most studies did not control for confounding or stratify by gestational age. In the two studies that controlled for confounding, a risk became higher when adjusted by confounders, suggesting that, at least in those settings, any confounding was negative confounding. Despite these limitations, this systematic review and meta-analysis consistently suggest an association between symptomatic dengue in pregnancy and each of the four adverse fetal outcomes.

We recommend that further epidemiological studies be done, with larger sample sizes, adequate comparison groups, and control for confounding. Many opportunities exist for such studies to be done, mainly in South America and Southeast Asia, where dengue incidence has been increasing and large outbreaks occur frequently. If an association between dengue infection during pregnancy and adverse fetal outcomes can be confirmed, recommendations should be made for the close monitoring of pregnancies during which dengue is diagnosed and for strategies for dengue control to include pregnant women as an at-risk population.

Declaration of interests
We declare no competing interests.

References
Thesis rationale

Dengue stands out from others viral mosquito borne disease because it is the most common; its incidence is growing and it is reaching new geographical areas and spreading worldwide. Indeed, reported cases of negative pregnancy outcomes after dengue infection are not new; however, the evidence of the association between maternal dengue and adverse pregnancy outcomes is limited, controversial and mostly supported by cases reports. The studies published until I start this thesis were: 1) mostly cases series, without a comparison group; 2) those with comparison group were hospital based, therefore only capable of capturing the most severe cases; 3) they did not have enough number to study less frequent outcomes (such as stillbirth or maternal deaths) or categorize the exposure by severity; 4) most of them did not control for confounders. There is a lack of population based studies that could estimate the magnitude of the risk (if proved), appropriately controlled by confounding and with enough number of observation to study rare outcomes and categorize the exposure (dengue) by severity of the maternal disease.

To be able to fill this gap in the literature I will use routine collected data from Brazil (a country where detailed, high quality data are captured and dengue is endemic). These population-level longitudinal data have the information required to expand the existing knowledge on birth outcomes from women with dengue acquired during pregnancy and they have sufficient number and less attrition than set up an ordinary community based cohort.

In this thesis, I will contribute to the debate on 1) infectious disease occurring during pregnancy: this information would enable physicians to be able to properly conduct a risk assessment in pregnant patients and for policy maker to use these information to support the inclusion of this group in at risk population. 2) Use of secondary data in epidemiological studies: the recent availability of administrative data
for research provides new opportunities for creating informative datasets to help us to answer research questions that require large numbers and it is difficult to set up an ordinary cohort due to expensive cost and potential loss of follow up.

**Thesis aims and objective**

The aim of this thesis is to explore the relationship between pregnancy outcome and symptomatic maternal dengue.

The aim of the thesis was achieved by completing the following objectives:

1. To develop a specific linkage algorithm and apply in a large administrative databases from low-middle income country:
   a. I extracted data from Brazilian databases on pregnancy outcomes and dengue notification;
   b. I used common variable between the datasets (name of the mother, age and place of residency) and calculated match weight;
   c. I applied the specific created algorithm to link the dengue and birth and maternal deaths records.

2. To describe and evaluate the linkage and discuss potential bias associated with the linkage process. I conducted this step to ensure valid results, showing that the linkage did not affect the measure of association in the next step
   a. I created a gold standard to compare its results with the results obtained by our algorithm;
   b. I compared the accuracy of our algorithm with a widely used software for linking data in Brazil.

3. To estimate the association between dengue and adverse pregnancy outcomes:
a. I analysed the data obtained after the linkage process; focus on three main outcomes: stillbirth, births outcomes (preterm birth, low birth weight and small for gestational age) and maternal deaths.

**Thesis overview**

There are four section, 9 chapters and 6 papers in this thesis (Figure 2). The first section contains the chapter one and two. Together they provide a background to the thesis. They discuss issues arising the nature of the current knowledge about adverse outcomes of maternal infections, such as the lack of evidence on viral maternal infection when compared with bacterial infections. I then contrast with the importance of dengue worldwide and the lack of evidence about its effects during pregnancy. The final section of chapter 2 is a systematic review on adverse fetal outcomes of maternal dengue infection.

The second section contains chapter three to five, these chapters describe the data source, analyses and the linkage process, providing detailed information on Brazilian Information System (data source), and the evaluation and validation of the linkage process.

The third section comprises the results chapters formatted as three papers. Each of them includes an introduction, methods, results and discussion, as well as tables and figures which is referred thought the paper. Any other relevant information is included in the chapter introduction. Then the thesis is concluded with a final discussion and recommendations for public health policies and future areas of research.

It is important to highlight that each paper was written as a separate article, and as such, there is some repetition of information. I have endeavoured to keep terminology
consistent throughout the papers, but due to variation in journal specifications and in the published literature, differences in terminology may exist.

Funding

This PhD project was funded through a studentship granted by National Council for Scientific and Technological Development (CNPq)-Brazil.

Ethical approval

This thesis was approved by the Ethics Committee of the Public Health Institute of the Federal University of Bahia (reference number 26797814.70000.5030) and by London School of Hygiene & Tropical Medicine (reference number 10269). Ethical approval letters can be found in Appendix I and II.
Figure 2: Aim of the thesis with corresponding objectives, sections, chapters and papers.
Section 1 references


69. Walker, P. G. T., ter Kuile, F. O., Garske, T., Menendez, C. & Ghani, A. C.


80. Elliott, B. *et al.* Maternal gonococcal infection as a preventable risk factor for low


123. Burnett, E., Loucks, T. L. & Lindsay, M. Perinatal outcomes in HIV positive


134. Baggi, F. M. *et al.* Management of pregnant women infected with Ebola virus in a


147. Lancet Infectious Diseases, T. The dengue vaccine dilemma.


172. CDC. Revised diagnostic testing for Zika chikungunya and dengue viruses. (2016).


Section 2-METHODS
Methods Section

Overview

The overall aim of this thesis is to examine the relationship between adverse pregnancy outcomes and symptomatic maternal dengue. To accomplish this, I used administrative collected data. Firstly, this was mainly because of its availability in the country where the research was conducted and secondly to answer this specific research questions, due to the rarity of the outcomes and exposure, would require establishing a cohort with large numbers of pregnant women. To set up an ordinary cohort would be expensive and they have been proven inefficient in other research to achieve sufficient sample size to analyse rare outcomes (such as stillbirth) or to do sensitivity analyses, for disease severity for example. Therefore, the use of administrative collected data is a cheaper yet efficient alternative.

The data available in Brazil have the information on exposure (dengue) and outcome (pregnancy outcomes), however they are recorded in different databases and not routinely linked together or have a unique identifier available. Therefore, in order to assess the exposure (dengue during pregnancy) the method of choice for linkage is to calculate the probabilistic match weight that measures the similarity between records from different datasets, taking into account errors in the identifier and missing data. Linkage between data without unique identifier relies on calculate probabilistic match weights that measure the similarity between records from different sources, with considerable potential for misclassification (errors in the linkage- records not correct match or a match that did not exceeds the used threshold). Even a slight linkage error can produce substantial bias in the final analyses, especially if the data belonged to a particular group, e.g. less educated women or the youngest ones. As a result, providing information on data quality and evaluating the linkage is a crucial step in this thesis as well as
understanding the extent of linkage error and potential impact on the analysis of the results.

The methods section contains three chapters and two papers, altogether this should inform the readers of the quality of the data used and the quality of the performed linkage process. In the third thesis chapter (first chapter of the methods section), I describe the country where the study was conducted and the details of the administrative data. Information on the data collection and processing are included for each of the datasets and finally I summarize the analytical methods performed to calculate the association between dengue during pregnancy and adverse maternal and fetal outcomes.

The following chapters are two papers that describe in detail the complete linkage process and its evaluation. The fourth chapter (second paper of the thesis) describes the linkage between stillbirths and dengue cases, evaluates the linkage accuracy, discusses potential bias and its potential effects on the final analyses. At the end of this chapter, I include supplementary information on live births linkage.

The fifth chapter (paper number three) evaluates the linkage between records of maternal deaths and dengue cases. Although this paper describes a linkage process using similar techniques as applied in the paper described in chapter four, some unique factors are applied and will be described in detail in this thesis chapter. This chapter also includes information on linkage accuracy, linkage error and its effects on the final analyses.
Chapter 3. Description of study area and design, data source and statistical analysis

Study design

This thesis used large population-based health datasets by linking routine records of birth outcomes and maternal deaths with records of women notified and confirmed with dengue disease in Brazil from January 1, 2006 until December 31, 2012. I used this period because Chikungunya and Zika arrived in Brazil in 2014/2015, so dengue had been the most important vector borne disease circulating in the country during the study period.

It is important to highlight (as a paper style thesis) that each paper was an independent study; as a result, there was some variation in study design and study period. Specific information is included in the methods section of each paper.

Study Area

The study area is Brazil, a South American country with more than 200 million inhabitants. According to national data, almost 3 million births and more than 30 thousand stillbirth are registered per year, prematurity and low birth weight rates were 10.8% and 8.4% in 2015.2

Dengue control in Brazil is considered a failure because since 1986 the incidence and the number of severe cases have increased 3. All four serotypes are circulating throughout the country. From 2001 to 2012 more than 5.5 million cases were notified to the Brazilian Ministry of Health, 55% of them in women 4.

Brazil has a long tradition of collecting administrative data and its quality is constantly improving. There is a large body of research evaluating the quality of routinely produced electronic data as well as substantive health research conducted using these data 5–7.
Data source – Brazilian Information System

This thesis used routine data from the Information System of Brazil. This system records health data in the country, including births, dengue cases and deaths. I will describe each of them in detail below:

a) SINASC (Sistema de Informação sobre Nascidos Vivos/ Live Birth Information System)

This system is updated using the registration of a live birth. This is a legal document, created in 1990 and used throughout the country. The forms are pre numbered and in three copies identified by colours (white- the form kept by the local health council that digitizes the information and sends it to the Brazilian Information System headquarters; yellow- kept by the local registry office that generates a birth certificate; pink- kept in the health records of the pregnant women or neonate in the facility). This document has to be completed by a health professional who was present at the delivery.

This form is divided in eight blocks (appendix III). I -characteristics of the newborn: sex; Apgar score in the 1 and 5 minute, birth weight (the child should be weighed in the first five hours of life), presence of abnormality. II- identification of the place of birth: home, centre or hospital birth and includes the address of the place. III- characteristics of the mother: name, age, marital status, education, race, place of residence. IV- identification of the father: name and age of the father (this is a new block introduced after 2011). V- characteristics of pregnancy and delivery: previous pregnancy- there is information on the number of previous pregnancies of live births, stillbirth or abortion; about the current pregnancy- there is information on the length of gestation (estimated from the day of the last period), type of delivery, number of fetus, number of visits to prenatal care facilities, which health professionals were present at the delivery.
VI- characteristics of congenital anomalies: this block should be filled in when congenital anomalies are identified at birth using the ICD-10 code. VII- identification of the professional completing the notification. VIII- registry office identification (Table 1) ⁸.

Table 1: Information available on SINASC

<table>
<thead>
<tr>
<th>Blocks</th>
<th>Contents</th>
</tr>
</thead>
<tbody>
<tr>
<td>I- characteristics of the newborn</td>
<td>Sex; Apgar score in the 1 and 5 minute, birth weight, presence of abnormality</td>
</tr>
<tr>
<td>II- identification of the place of birth</td>
<td>Home, centre or hospital birth and includes the address of the place</td>
</tr>
<tr>
<td>III- characteristics of the mother</td>
<td>Name, age, marital status, education, race, place of residence</td>
</tr>
<tr>
<td>IV- identification of the father</td>
<td>Name and age of the father</td>
</tr>
<tr>
<td>V- characteristics of pregnancy and delivery</td>
<td>Previous pregnancy- number of previous pregnancies of live births, stillbirth or abortion; Current pregnancy- length of gestation, type of delivery, number of fetus, number of visits to prenatal care facilities, which health professionals were present at the delivery.</td>
</tr>
<tr>
<td>VI- characteristics of congenital anomalies</td>
<td>Congenital anomalies identified at birth using the ICD-10 code</td>
</tr>
<tr>
<td>VII- identification of the professional</td>
<td></td>
</tr>
<tr>
<td>VIII- registry office identification</td>
<td></td>
</tr>
</tbody>
</table>

The Ministry of Health of Brazil uses the WHO definition of a live birth (Table 2) ⁹:

The complete expulsion or extraction from the body of the pregnant woman of a product of conception, independent of the duration of pregnancy, who, after the separation, breathes or shows any other signs of life, such as heartbeat, umbilical cord pulsation, or definite movement of voluntary muscles, whether or not the cord is cut and whether or not the placenta is attached.
Table 2: Brazilian Ministry of Health definition of the terms

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Live birth</td>
<td>The complete expulsion or extraction from the body of the pregnant woman of a product of conception, independent of the duration of pregnancy, who, after the separation, breathes or shows any other signs of life, such as heartbeat, umbilical cord pulsation, or definite movement of voluntary muscles, whether or not the cord is cut and whether or not the placenta is attached.</td>
</tr>
<tr>
<td>Stillbirth</td>
<td>Stillbirth is the death of a product of conception before the expulsion or complete extraction from the body of the pregnant woman, occurring after 20 week of pregnancy or 154 days or fetus weighing more than 500 g. The fetus, must not after the separation, breathe or show any other sign of life, such as heartbeat, umbilical cord pulsation, or definite movement of voluntary muscles.</td>
</tr>
<tr>
<td>Maternal deaths</td>
<td>Maternal death is the death of a woman while pregnant or within 42 days of termination of pregnancy, irrespective of the duration and site of the pregnancy, from any cause related to or aggravated by the pregnancy or its management but not from accidental or incidental causes.</td>
</tr>
</tbody>
</table>

Most of the studies using SINASC show that the coverage rates of the system in the country is higher than 90% \(^ {10}\) varying from 85.7% to 100% between 1999 and 2011 \(^ {5,11-14}\). Regarding the accuracy of the information recorded in SINASC, no routine evaluation is done \(^ {15}\), however, some studies have compared information from SINASC with information obtained from medical records or interviews with mothers. The results of these studies have shown that the birth weight registered in SINASC are similar, almost 100%, to medical records data \(^ {16}\). However, the percentage of preterm births recorded in SINAN was found to be underestimated by 15% \(^ {16,17}\). The completeness of the data is very high, missing data do not exceed 10% in most of the variables, except the number of previous stillbirths or abortions experienced (19% in 2002 and 11.1% in 2010).

b) SIM (Sistema de Informação de Mortalidade/ Mortality Information System).

From this dataset, I retained information on stillbirths and maternal deaths. This system uses the death certificate, a legal document that can only be completed by a
The forms are in three copies identified by colours (white- this form is kept by the local health council that digitizes the information and sends it to the Brazilian Information System headquarters; yellow- kept by the local registry office that generates a death certificate; pink- kept in the health records of the women)\textsuperscript{18}.

This form is divided into nine blocks (appendix IV). I – identification: in this block first you notify if it is a fetal death and then provide information on date of death and birth, name of the dead individual, name of the mother and father, sex, race, marital status, occupation and education of the dead person; in the case of fetal death some of the variables will not be applicable. II- identification of the place of residence: complete address. III- identification of the place of death: hospital, home, public place and includes the address of the place. IV- characteristics of the mother: name, age, marital status, education, occupation, race, number of births, place of residence, length of gestation, number of previous stillbirths or abortions, type of delivery, number of fetus in the current pregnancy, birth weight (this block should be filled only in the case of fetal deaths or infant mortality). V- cause of death: identify the cause of death using ICD-10 code. VI- identification of the professional completing the notification. VII- identification of violent deaths. VIII- registry office identification. IX - This block must be filled in when the death occurred in a place without a medical doctor (Table 3)\textsuperscript{18}.\textsuperscript{9}
### Table 3: Information available on SIM

<table>
<thead>
<tr>
<th>Blocks</th>
<th>Contents</th>
</tr>
</thead>
<tbody>
<tr>
<td>I – identification</td>
<td>Notify if it is a fetal death, date of death and birth, name of the dead individual, name of the mother and father, sex, race, marital status, occupation and education of the dead person; in the case of fetal death some of the variables will not be applicable.</td>
</tr>
<tr>
<td>II- identification of the place of residence</td>
<td>Complete address</td>
</tr>
<tr>
<td>III- identification of the place of death</td>
<td>Hospital, home, public place and includes the address of the place</td>
</tr>
<tr>
<td>IV- characteristics of the mother</td>
<td>Name, age, marital status, education, occupation, race, number of births, place of residence, length of gestation, number of previous stillbirths or abortions, type of delivery, number of fetus in the current pregnancy, birth weight (this block should be filled only in the case of fetal deaths or infant mortality)</td>
</tr>
<tr>
<td>V- cause of death</td>
<td>Cause of death using ICD-10 code</td>
</tr>
<tr>
<td>VI- identification of the professional completing the notification</td>
<td></td>
</tr>
<tr>
<td>VII- identification of violent deaths</td>
<td></td>
</tr>
<tr>
<td>VIII- registry office identification</td>
<td></td>
</tr>
<tr>
<td>IX - This block must be filled in when the death occurred in a place without a medical doctor</td>
<td></td>
</tr>
</tbody>
</table>

The Brazilian Ministry of Health uses the WHO definition of stillbirth (Table 2)\(^{19}\):

Stillbirth is the death of a product of conception before the expulsion or complete extraction from the body of the pregnant woman, occurring after 22 week of pregnancy or 154 days or fetus weighing more than 500 g. The fetus, must not after the separation, breathe or show any other sign of life, such as heartbeat, umbilical cord pulsation, or definite movement of voluntary muscles.

Although Brazil uses the WHO definition, in which the gestational age limit is 22 weeks, the doctor can fill a death certificate when the stillbirth occurs from 20 weeks\(^{18}\).

The Ministry of Health of Brazil uses the WHO definition of maternal deaths (Table 2):
Maternal death is the death of a woman while pregnant or within 42 days of termination of pregnancy, irrespective of the duration and site of the pregnancy, from any cause related to or aggravated by the pregnancy or its management but not from accidental or incidental causes.

In Brazil, for investigation purposes, maternal death is considered the death of a woman while pregnant or up to a year of termination of pregnancy, irrespective of the duration and site of the pregnancy, from any cause related to or aggravated by the pregnancy or its management but not from accidental or incidental causes. However, the Ministry of Health calculates the official estimates using the 42 days definition.

In this thesis, I only used maternal deaths coded under the chapter XV (pregnancy, childbirth and puerperium) of the ICD-10 using the one year rule, because there was not enough information to classify whether deaths due to diseases coded under different chapters could be a maternal death (Table 4).

Table 4: International classification of disease-10 chapter XV (pregnancy, childbirth and puerperium)

<table>
<thead>
<tr>
<th>Code</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>O00-O08</td>
<td>Pregnancy with abortive outcome</td>
</tr>
<tr>
<td>O10-O16</td>
<td>Edema, proteinuria and hypertensive disorders in pregnancy, childbirth and puerperium</td>
</tr>
<tr>
<td>O20-O29</td>
<td>Other maternal disorders predominantly related to pregnancy</td>
</tr>
<tr>
<td>O30-O48</td>
<td>Maternal care related to the fetus and amniotic cavity and possible problems</td>
</tr>
<tr>
<td>O60-O75</td>
<td>Complications of labour and delivery</td>
</tr>
<tr>
<td>O80-O84</td>
<td>Delivery</td>
</tr>
<tr>
<td>O85-O92</td>
<td>Complications predominantly related to puerperium</td>
</tr>
<tr>
<td>O95-099</td>
<td>Other obstetric conditions</td>
</tr>
</tbody>
</table>

The completeness of the variables from SIM database differs among stillbirths and non-stillbirths. For stillbirths there is a high frequency of missing variables, the lack of data varies according to variables and exceeded 15% for most of the variables, except maternal education, which was more than 25% in 2010. The completeness of maternal
deaths information is very high. In our data, completeness of the variables of interest was higher than 95%.

c) SINAN (Sistema de Informação sobre Agravos de Notificação/ Information System for Notifiable Diseases)

From this database, information on women who were notified and confirmed as a dengue cases was retained (appendix V). In Brazil compulsory notification of cases of dengue is required since 1998, in accordance with Ordinance No 1,271/2014. Suspected and/or confirmed cases must be reported to the Epidemiological Surveillance service on a specific numerated notification form (appendix V), which is available in any local health facility. This form can be filled in by any health professional who suspects dengue. It collects information on date of notification, date of onset of symptoms, date of birth, name of the patient, age, sex, and address.

The Epidemiological Surveillance Centre \textsuperscript{20} then investigates to confirm or discard the suspicion based on the Brazilian definition of clinical epidemiological cases and/or laboratory results. The Brazilian definition of clinical epidemiological cases of dengue is the presence of fever and two or more of the following, retro-orbital or ocular pain, headache, rash, myalgia, arthralgia, leukopenia, or haemorrhagic manifestations. In addition to these symptoms, the patient must have been in areas where dengue is being transmitted or where Aedes is present \textsuperscript{21} in the past fifteen days. Laboratory confirmed cases are those with the presence of clinical symptoms and a positive test that could be IgM antibodies detection by enzyme-linked immunosorbent assay (ELISA), viral Ribonucleic acid (RNA) detection via Polymerase chain reaction (PCR), Nonstructural protein 1 (NS1) viral antigen detection, or positive viral culture. Suspected cases are discarded based on 1) negative laboratory test, especially IgM (once confirmed that the
sample was collected in the appropriate time, in dengue cases, IgM may remain elevated for two to three months); 2) does not have clinical-epidemiological link criteria, that is there is no way to establish the chain of transmission; 3) laboratory confirmation for another disease; 4) case without laboratory confirmation and with signs and symptoms compatible with another pathology. The case must be closed (confirmed or discarded) within 60 days.

After investigation, information on laboratory results and clinical progress of the case is collected and then, if confirmed, it can be classified into three clinical categories: “dengue fever” as a self-limiting disease (fever, with a severe headache, pain behind the eyes, muscle and joint pain, and rash), “complicated dengue”, and “dengue haemorrhagic fever/ dengue shock syndrome”. Complicated dengue is a Brazilian definition of severe cases of dengue that do not meet the WHO criteria for dengue haemorrhagic fever (fever, haemorrhagic evidence, thrombocytopenia and evidence of plasma leakage) and could not be classified as a mild self-limited disease due to its severity. Complicated dengue is used when a probable case of dengue presents one of the following: severe changes in the nervous system, cardiorespiratory dysfunction; insufficient hepatic function; gastrointestinal bleeding, cavity spills, or thrombocytopenia equal or less than 50,000/mm³ leucometry less than 1000/mm. Dengue haemorrhagic fever follows the WHO criteria and is characterized by fever, haemorrhagic evidence, thrombocytopenia and evidence of plasma leakage.

After 2009, the WHO changed the dengue classification into dengue fever (DF), dengue haemorrhagic fever (DHF) and dengue shock syndrome (DSS) to non-severe (with and without warning signs) and severe. This new classification started to be implemented in Brazil in 2014, after the study period of this thesis.
Studies have shown that dengue surveillance substantially underestimates the disease burden, mainly during periods of low-transmission. It is estimated that there were 12 cases of dengue per reported case in the community, however, during low dengue transmission the ratio can reach 17:1\(^{22}\). The completeness of the data in the used variables were reasonable. Less than 0.05% of the records were excluded because the name of the patient was missing and 5% of the dengue cases did not have a final classification of severity of the case.

**Statistical analysis**

Specific information is included in the methods section of the papers. After concluding the linkage process, I used logistic regression to model the variables, as the study outcome is a dichotomous variable, including birth weight that was categorized as low birth weight and normal birth weight. In general, I estimated the crude and adjusted odds/risk ratios using logistic regression (conditional or not, depends on how the comparison group was selected). However, empty and small cells can make the model unstable; in this case, I applied the Firth method (to reduce the small sample bias in maximum likelihood estimation). The Firth method can be used to overcome the condition in which maximum likelihood estimation tends to infinity called separation. In this method, the coefficient of interest is constrained to zero and left in the model to allow its contribution to the penalization. The penalization allows convergence to finite estimates with very sparse data and over-corrects for bias \(^{23}\).

The variables used as confounders followed the criteria: available in good quality in the datasets; associated with the exposure; risk factor for the outcome and not in the causal pathway. The used confounders vary according to the outcome, but the full list are maternal age (categorized as \(<=20\); between 20-35; \(>=35\) years old), education (less than
3 years; between 4-7 years; more than 8 years) and marital status (single; widow; divorced or married or stable union) and mode of delivery (vaginal; caesarean section).

A sensitivity analysis of the validity of clinical/epidemiologic diagnosis was carried out; I repeated the analyses using only laboratory confirmed dengue. This was because misclassification of dengue based on the lack of laboratory confirmation is possible. I investigated the effect of time between disease onset and maternal/birth outcomes, since dengue is an acute disease, with few reports of chronicity. I would expect that the magnitude of the association would be higher in the first days of disease onset, since dengue is known as an acute disease. The time between disease onset and the outcome was calculated using the date of the disease onset (information available from SINAN) and the date when the outcome occurred (date of birth/death). Finally, I performed analyses for dengue severity (mild dengue, dengue with complications and dengue haemorrhagic fever). When necessary, the potential effects of missing data were also investigated, assuming that all missing data for confounding variables in the cases were in the low-risk groups and all missing data for the comparison group were in the high-risk groups.

In one study, the population attributable fraction was calculated, (PAF; PAF=[p1(OR−1)] /OR) using the punafcc package in Stata, which uses a logistic regression method and provides PAF (and 95% CIs) with adjustment for confounding variables by combining adjusted ORs and the observed incidence of dengue among cases.

For all the statistical analyses Stata version 14.1 software was used.

The linkage procedures will be discussed in the following two chapters.
# Chapter 4. Linkage procedures and evaluation - Description of the stillbirth and live birth linkage process

## Cover sheet

**RESEARCH PAPER COVER SHEET**

**PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED FOR EACH RESEARCH PAPER INCLUDED IN A THESIS.**

### SECTION A – Student Details

<table>
<thead>
<tr>
<th>Student</th>
<th>Enny da Paixao Cruz (Enny S Paixao)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Principal Supervisor</td>
<td>Elizabeth Brickley</td>
</tr>
<tr>
<td>Thesis Title</td>
<td>Symptomatic dengue and adverse pregnancy outcomes: a population-based record linkage study</td>
</tr>
</tbody>
</table>

*If the Research Paper has previously been published please complete Section B. If not please move to Section C*

### SECTION B – Paper already published

<table>
<thead>
<tr>
<th>Where was the work published?</th>
<th>BMC Medical Informatics and Decision Making</th>
</tr>
</thead>
<tbody>
<tr>
<td>When was the work published?</td>
<td>2017</td>
</tr>
<tr>
<td>If the work was published prior to registration for your research degree, give a brief rationale for its inclusion</td>
<td></td>
</tr>
<tr>
<td>Have you retained the copyright for the work?</td>
<td>No</td>
</tr>
<tr>
<td>Was the work subject to academic peer review?</td>
<td>Yes</td>
</tr>
</tbody>
</table>

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.*

### SECTION C – Prepared for publication, but not yet published

<table>
<thead>
<tr>
<th>Where is the work intended to be published?</th>
</tr>
</thead>
</table>

Please list the paper’s authors in the intended authorship order:

Stage of publication: Choose an item.

### SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)

Student Signature: [Signature]

Date: 8/03/19

Supervisor Signature: [Signature]

Date: 8/03/19

Improve health worldwide

www.lshtm.ac.uk
Paper 2 title: Evaluation of record linkage of two large administrative databases in a middle income country: stillbirth and notifications of dengue during pregnancy in Brazil

Authors: Enny S. Paixao, Katie Harron, Kleydson Andrade, Maria Gloria Teixeira, Rosemeire L. Fiaccone, Maria da Conceicao N Costa, Laura C. Rodrigues

Author contribution:

ESP carried out the analysis and wrote the first draft of the article. LCR, MGT conceived the study. KH, KA, MCNC and RLF contributed to the study design and interpretation. All authors revised the manuscript and approved the final version.

Permission from copyright holder to include this work:
Evaluation of record linkage of two large administrative databases in a middle income country: stillbirths and notifications of dengue during pregnancy in Brazil

Enny S Paixo, Katie Harron, Kleidson Andrade, Maria Gloria Teixeira, Rosemerie L Fiaccone, Maria da Conceição N. Costa and Laura C. Rodrigues

Abstract

Background: Due to the increasing availability of individual-level information across different electronic datasets, record linkage has become an efficient and important research tool. High quality linkage is essential for producing robust results. The objective of this study was to describe the process of preparing and linking national Brazilian datasets, and to compare the accuracy of different linkage methods for assessing the risk of stillbirth due to dengue in pregnancy.

Methods: We linked mothers and stillbirths in two routinely collected datasets from Brazil for 2009–2010: for dengue in pregnancy, notifications of infectious diseases (SINAN); for stillbirths, mortality (SIM). Since there was no unique identifier, we used probabilistic linkage based on maternal name, age and municipality. We compared two probabilistic approaches, each with two thresholds: 1) a bespoke linkage algorithm; 2) a standard linkage software widely used in Brazil (RedimKII), and used manual review to identify further links. Sensitivity and positive predictive value (PPV) were estimated using a subset of gold-standard data created through manual review. We examined the characteristics of false matches and missed matches to identify any sources of bias.

Results: From records of 678,999 dengue cases and 62,373 stillbirths, the gold-standard linkage identified 191 cases. The bespoke linkage algorithm with a conservative threshold produced 131 links, with sensitivity = 64.6% (68 missed-matches) and PPV = 92.5% (82 false-matches). Manual review of uncertain links identified an additional 37 links, increasing sensitivity to 83.7%. The bespoke algorithm with a relaxed threshold identified 132 true matches (sensitivity = 69.1%), but introduced 61 false-matches (PPV = 68.4%). RedimKII produced lower sensitivity and PPV than the bespoke linkage algorithm. Linkage error was not associated with any recorded study variables.

Conclusion: Despite a lack of unique identifiers for linking mothers and stillbirths, we demonstrate a high standard of linkage of large routine databases from a middle income country. Probabilistic linkage and manual review were essential for accurately identifying cases for a case-control study, but this approach may not be feasible for larger databases or for linkage of more common outcomes.

Keywords: Data linkage, Routine data, Electronic health records, Linkage quality, Linkage accuracy, Stillbirth, Dengue

*Correspondence: enny.spaixo@ucl.ac.uk
1. London School of Hygiene and Tropical Medicine, Keppel St, Bloomsbury, London WC1E 7HT, UK
Full list of author information is available at the end of the article
Background

Record linkage is the process used to bring together information recorded in different sources about the same individual or group of individuals [1]. Due to the growing availability of administrative population-based health databases, linkage has become an efficient and important research tool [2]. One research area for which linkage is particularly important is maternal and infant health, where linkage can facilitate increased understanding of how maternal health trajectories prior to and during pregnancy are associated with birth and later childhood outcomes [3]. Many countries use linkage of records from mothers and their babies to underpin both research and service evaluation [4–6].

High quality linkage is required so that robust results can be obtained from linked data. However, there are specific challenges in mother-baby linkage, particularly concerning data from low-middle income countries. Firstly, quality of linkage is limited by the availability of unique or well-completed identifiers, and by data quality (including truncated records and absent or ambiguous information) [7, 8]. A further complication is that linkage of records belonging to different individuals is required (e.g. mothers and their babies), which limits the availability of common identifiers for linkage. Finally, when datasets do not overlap exactly, the expected number of matches is unknown, making it difficult to establish the expected number of records that should be linked [3].

There are two main approaches to linkage: deterministic and probabilistic. Deterministic linkage is a rule-based approach, typically using a unique identifier or set of common identifiers present in both files. Probabilistic linkage is useful when the quality of identifiers is not sufficient for a strict deterministic linkage due to missing values or typographical errors [9]. Probabilistic linkage combines evidence across a number of identifiers such as name, age and place of residence to calculate a match weight, representing the likelihood that two records belong to the same person, i.e., that they are a "true match" [10]. Match weights are used to classify records as links, non-links and uncertain links, generally by defining two thresholds [11]. The choice of threshold affects the number of "false-matches" (records from different individuals that are linked) and "missed-matches" (records from the same individual that fail to link) [12]. Probabilistic linkage has facilitated many studies in countries without a unique identifier [4–6].

The impact of linkage errors, in terms of bias in results of analysis, depends on the study in question [13]. For example, in many studies, it may be important to achieve a high match rate, so that the resulting linked data is representative of the source population and to avoid selection bias (especially if missed-matches are non-random, i.e. are more likely to occur in specific subgroups of records). In others, it may be more important to avoid false-matches. For example, in case-control studies, we may be more concerned with accurately establishing exposure status through linkage with a disease registry. This would require certainty that linked records really should have been linked, and unlinked records did not have the exposure.

In this study, we aimed to use linkage of national data sources to facilitate a case-control study of stillbirths in women who had notified symptomatic dengue during pregnancy. Linkage was required to establish exposure status in cases. In this scenario, it was important to: 1) prioritize true matches (i.e. high specificity of the exposure), whilst high sensitivity was less important; 2) retain sufficient cases for a reasonable sample size; and 3) verify absence of bias (i.e. understand whether any groups were more or less likely to be linked). This study presents an approach to preparing and linking mortality and morbidity data from routine data sources in Brazil, comparing the performance of different linkage methods for identifying exposure of pregnant women to dengue.

Methods

Datasets

We linked two routinely collected data sets: for dengue, notifications of infectious diseases (SINAN); for stillbirths, mortality (SIM), for 2009 and 2010.

1. SINAN: Notifiable Diseases Information System (Sistema de Informação de Agravos de Notificações/SINAN), containing individual-level data on all notified diseases.

Source

The Brazilian Ministry of Health has required notification of all cases of dengue seen in health facilities in Brazil. SINAN captures clinical cases of disease, through forms completed by any health professional who suspects dengue; this notification is compulsory in the country. After a dengue suspected case is identified, the Epidemiological Surveillance Service investigates cases in order to confirm or discard the suspicion based on laboratory results and the Brazilian definition of clinical epidemiological criteria of dengue: presence of fever and two or more of the following symptoms (retro-orbital or ocular pain, headache, rash, myalgia, arthralgia, leukopenia, or hemorrhagic manifestations) [14]. Forms include personal information on the patient (name, place of residence, age, marital status, and education) and on their disease (symptoms, laboratory tests and severity).
Completeness of linkage variables
Maternal name was 99.5% complete; municipality was 100% complete. Where age in years was missing (1.3% of records), we derived age from date of birth.

Data extraction
Data were extracted from SINAN for all suspected dengue cases (n = 1,981,912 individuals). We excluded records for 900,054 men, and 398,710 cases discarded by the Epidemiological Surveillance services (Fig. 1).

2. SIM: Mortality Information System (Sistema de Informação sobre Mortalidade), containing individual-level data on all deaths including stillbirths.

Source
The Brazilian Ministry of Health requires notification of all patients who die, irrespective of place of death. Deaths are recorded using a death certificate, which is a legal document completed by physicians [15]. The definition of stillbirth as recorded in SIM is the death of a product of conception before the expulsion or complete extraction from the body of the pregnant woman, occurring from 22 weeks or weighing more than 500 g [15]. The form includes information on the mother (name, place of residence, age, marital status, education, whether she had a previous stillbirth or a child who died); and the pregnancy (length of gestation, type of delivery).

Completeness of linkage variables
Maternal name was complete in 98.8% of records; municipality was 100% complete. Age was complete in 84.6% of records.

Data extraction
Data on all deaths were extracted from SIM (n = 2,303,893 records). We kept only the stillbirths (n = 63,108).

The process of linking
Data pre-processing
We excluded records without names, those with generic names such as "unknown" or "stillbirth", records with only one name (e.g. "Maria"), and those name recorded as numbers; our final study population comprised 678,999 dengue notifications from SINAN and 62,373 stillbirths from SIM (Fig. 1). We searched automatically for improbable ages (e.g. 99 or age of the mother 1 year

Fig. 1 Number of records from Brazilian Information System of Notifiable Disease and Brazilian Information System of Mortality

1,981,912 records of dengue suspected cases - Brazilian Information System of Notifiable Disease - Dengue (January 1, 2009- December 31, 2010)
2,303,893 records of death from the Brazilian Information System of Mortality (January 1, 2009- December 31, 2010)
1,262,513 were excluded which: 946,064 were males
386,710 were discarded
4,149 unnameable name
2,241,520 were excluded which: 2,240,785 no stillbirth death were excluded
735 unnameable name
678,999 eligible dengue cases
62,373 eligible stillbirth
Conservative threshold bespoke algorithm
123 true links
8 false matches
Conservative threshold Recklink algorithm
102 true links
23 false matches
old) and set to null. We removed punctuation, deleted consecutive spaces, transformed known prefixes in the real names (e.g. Mr. → Maria), deleted unknown prefixes, replaced names in upper case and dropped middle initials.

In Brazil, full names usually have several components, and we aimed to retain the discriminatory power of this variable. In our data, the mean number of names per mother was three (maximum eight). We derived distinct variables for first name, second name, and last name. Similarity between names recorded in SIM and SINAN was compared using the Jaro-Winkler string comparator [16]. The Jaro-Winkler string comparator counts the number of common characters between two strings and the number of transpositions of these common characters, producing similarity values varying between 1 (perfectly similar) and 0 [16]. We categorized the string comparator score as (0.85), (0.85/0.9), (0.9/0.95), (0.95/0.99) and (0.99/1).

**Blocking**

The total number of pairwise comparisons between SINAN and SIM would be prohibitively high: $184,922 \times 31,867 = 5,892,908,574$ in 2009 and $494,077 \times 30,506 = 15,072,312,962$ in 2010 (a dengue epidemic year). To decrease the number of pairwise comparisons, we only attempted to link women that resided in the same municipality ($n = 5,565$ municipalities in Brazil in 2010). This process is called blocking. Blocking by municipality assumes that records from different municipalities do not belong to the same women. We used this blocking scheme because we considered municipality to be very reliable, and unlikely to have changed between dengue in pregnancy and end of pregnancy.

**Gold-standard data**

We created a “gold-standard” dataset using nine steps (Table 1). Firstly, we linked SIM and SINAN using deterministic linkage with exact agreement on full name and age. Subsequent steps relaxed the rules. Since mothers could be diagnosed during pregnancy in 1 year and give birth in the following year, and to allow for errors in recording, we allowed matches where there were differences of up to 2 years in age. To account for differences in the ways names were recorded, we separately looked for agreement on first and last name, first and second name, and second and last name. Each step was followed by manual review to exclude false-matches. Since some names were very common in the database (e.g. Maria), records with these names were only considered as a match where there was exact agreement on name and municipality, and age differed by no more than 1 year.

<table>
<thead>
<tr>
<th>Table 1 Deterministic rules used to create the gold-standard database</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linkage rule</td>
</tr>
<tr>
<td>---------------</td>
</tr>
<tr>
<td>Full name and age*</td>
</tr>
<tr>
<td>Age¹ and combination of first and last name</td>
</tr>
<tr>
<td>Age² and combination of first and second name</td>
</tr>
<tr>
<td>Age³ and combination of second and last name</td>
</tr>
<tr>
<td>Full name</td>
</tr>
<tr>
<td>First and last name</td>
</tr>
<tr>
<td>First and second name</td>
</tr>
<tr>
<td>Second and last name</td>
</tr>
<tr>
<td>Age¹ and Jaro-Winkler using comparator ≤0.95</td>
</tr>
<tr>
<td><strong>Total</strong></td>
</tr>
</tbody>
</table>

*Age in years as recorded in the data or as derived from data of birth

**Bespoke algorithm**

For each record pair we calculated a probabilistic match weight based on two conditional probabilities: the probability of agreement given records belong to the same mother-baby pair (m-probability: P (agreement | match)) and the probability of agreement given records belong to different mother-baby pairs (u-probability: P (agreement | non-match)).

M-probabilities for each identifier were estimated from the true matches in the gold-standard dataset. U-probabilities were calculated based on a list of non-matches, created from all pairwise comparisons of records within SINAN, excluding those belonging to the same individual. We used the SINAN database to calculate the u-probabilities, as it was the larger of the two databases.

Frequency-based weights were calculated for each category of Jaro-Winkler score comparator [16] (for name of the mother) and year of age. Weights were also separately calculated for the five most frequently occurring names in the data (Maria, Ana, Santos, Souza and Oliveira).

Since we were linking different subjects (the mother and her stillbirth), and because the exposure (dengue during pregnancy) could have happened up to 9 months before the outcome (stillbirth) there were timing issues to consider. Two records could differ in time by 9 months and could bridge over the calendar year; some mothers would have the birthday between the data of dengue and the date of the stillbirth. To allow for this, we estimated different weights according to the similarity of age across datasets: equal ages, age differing by 1 year, age differing by 2 years, and ages differing by more than 2 years (Table 2).

Match weights were calculated by summing the ratio of m-probabilities and u-probabilities across different identifiers [1]. The algorithm was implemented in Stata and R.
Table 2: Comparison of linkage strategies for the bespoke algorithm and Reclink

<table>
<thead>
<tr>
<th></th>
<th>Bespoke algorithm</th>
<th>Reclink</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manipulation of names</td>
<td>Multiple variables created for first name, second name, and last name</td>
<td>Variables created for first and last name</td>
</tr>
<tr>
<td>Blocking</td>
<td>Municipality</td>
<td>Soundex for name + municipality</td>
</tr>
<tr>
<td>Calculation of m and u probabilities</td>
<td>m probability: calculated using true matches in gold-standard&lt;br&gt;u probability: calculated using non-matches in SINAN</td>
<td>m probabilities = 0.9&lt;br&gt;u probabilities = 0.1</td>
</tr>
<tr>
<td>Match weight calculation</td>
<td>Separate weights calculated for the five most common names&lt;br&gt;Agreement on name classified using Jaro-Winkler string comparator&lt;br&gt;Different weights calculated according to distance of age</td>
<td>Did not account for common names&lt;br&gt;Levenshtein string comparator&lt;br&gt;Did not account for timing issues</td>
</tr>
</tbody>
</table>

**Reclink** algorithm

As we were using Brazilian data, we compared the accuracy of our bespoke linkage algorithm with a widely used software for linking data in Brazil, called Reclink version III.

Reclink III calculates match weights in 3 steps: 1) Manipulation of names; 2) Blocking (Reclink matches pairs within blocks of similar names defined by soundex; in addition we programmed Reclink III to block by municipality); 3) Match weight calculation (the final score is the sum of the weighted scores of each field, e.g. name and age).

Reclink III applies the Levenshtein string comparator to compare names [17]. The Levenshtein string comparator is defined as the minimum number of insertions, deletions, or substitutions necessary to change one string into the other, the values varying between 1 (perfect similarity) and 0 (total disagreement). The m-probabilities and u-probabilities were based on default values as suggested by the software m-probabilities = 0.9 and u-probabilities = 0.1 for all identifiers (Table 2) [18].

**Classification of links**

Records pairs were ordered by match weight and manually inspected to identify threshold values to classify comparison pairs as non-links, links and uncertain links. The number of expected matches was unknown, because we did not know a priori how many of the mothers in SIM should link to a stillbirth in SINAN. Therefore, we explored two different threshold choices for each algorithm. We first chose a conservative threshold, aiming to exclude as many as false-matches as possible (high positive predictive value). We then chose a relaxed threshold, aiming to capture as many of the true matches as possible (high sensitivity). Any records above the cut-off threshold were classified as links. For the best performing approach, we manually inspected uncertain links to determine whether or not they belonged to the same mother-baby pair.

**Statistical analysis**

For both the bespoke and the Reclink III algorithm, and for each threshold (conservative and relaxed), we estimated the sensitivity and positive predictive value (PPV) by comparing linkage results with the gold-standard dataset. Since we expected the number of links to be very small in comparison to the size of the datasets, we did not calculate specificity or negative predictive value (as these measures would be consistently high). To account for in-sample optimism, we also present average estimates based on leave one out cross classification.

For the best performing algorithm, we examined which characteristics were associated with false-matches and missed matches. Categorical variables were compared between groups with Chi^2 test or Fisher’s exact test. A two-sided P value of less than 0.05 was considered to indicate statistical significance. We examined maternal age (<20, 20–35, >35 years), maternal education (illiterate, 1–3 years, 4–7 years, 8–11 years and more than 11 years), previous stillbirths or abortions (yes/no), gestational age (less than 22 weeks, 22–27 weeks, 28–31 weeks, 32–36 weeks, 37–41 weeks, more than 42 weeks) and weight when the stillbirth occurred (> = 2500, 1500–2500, <1500 g). Stata version 14.1 was used for the statistical analyses.

**Results**

**Gold-standard**

Of the 678,999 eligible dengue cases in SINAN for 2009–2010, 191 were linked to a stillbirth record using the nine-step gold-standard algorithm (Table 1).

**Bespoke algorithm**

The conservative threshold was set to a match weight of 21, and resulted in 131 links (Fig. 1). Comparison with the gold-standard identified 8 false matches and 68 missed-matches, giving a sensitivity of 122/191 = 64.4% and a PPV of 122/131 = 93.9% (Table 3). Adjusting for in-sample optimism gave values of 64.4% and 94.1% respectively.

The relaxed threshold was set at 20 and resulted in 193 links. Comparison with the gold-standard identified 132 true links, giving a sensitivity of 132/191 = 69.1%, and 61 false-matches, giving a PPV of 132/193 = 68.4%
Table 3 Performance of linkage algorithms and thresholds

<table>
<thead>
<tr>
<th></th>
<th>Bespoke algorithm</th>
<th></th>
<th>ReCLink algorithm</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Conservative threshold = 21</td>
<td>Relaxed threshold = 20</td>
<td>Conservative threshold = 12</td>
<td>Relaxed threshold = 10</td>
</tr>
<tr>
<td>N linked</td>
<td>131</td>
<td>193</td>
<td>125</td>
<td>788</td>
</tr>
<tr>
<td>N true links</td>
<td>132</td>
<td>132</td>
<td>102</td>
<td>114</td>
</tr>
<tr>
<td>N false matches</td>
<td>8</td>
<td>6</td>
<td>23</td>
<td>6/4</td>
</tr>
<tr>
<td>N missed matches</td>
<td>60</td>
<td>59</td>
<td>8</td>
<td>77</td>
</tr>
<tr>
<td>Sensitivity % (95% CI)</td>
<td>64.4 (61.2-71.2)</td>
<td>69.1 (62.0-76.5)</td>
<td>53.4 (46.1-60.6)</td>
<td>59.7 (57.4-66.7)</td>
</tr>
<tr>
<td>Positive predictive value % (95% CI)</td>
<td>93.9 (66.9-96.3)</td>
<td>68.4 (61.3-74.8)</td>
<td>81.6 (73.2-87.9)</td>
<td>14.5 (12.1-17.1)</td>
</tr>
</tbody>
</table>

(95). Adjusting for in-sample optimism gave values of 69.1% and 68.4% respectively.

ReCLinkIII

The conservative threshold was set at 12 and resulted in 125 links (Fig. 1). Comparison with the gold-standard identified 102 true links, giving a sensitivity of 102/191 = 53.4%, and 23 false matches, giving a PPV of 102/125 = 81.6% (Table 3). Adjusting for in-sample optimism gave values of 53.4% and 81.8% respectively.

The relaxed threshold was set at 10 and resulted in 788 links. Comparison with the gold-standard identified 114 true links, giving a sensitivity of 114/191 = 59.7%, and 674 false matches, giving a PPV of 114/788 = 14.5% (Table 3). Adjusting for in-sample optimism gave values of 59.7% and 14.4% respectively.

Linkage errors

Missed matches and false matches had a higher proportion of missing data (Table 4). Linkage errors were not associated with any recorded study variables, although there was a suggestion that mothers aged <20 were slightly less likely to link (p = 0.047) (Table 4).

Manual review

For the bespoke linkage algorithm with the conservative threshold, uncertain links were defined as those with weights between 16 and 21. Records with these weights were classified through manually inspecting each record to determine whether or not they belonged to the same person. This added 37 links, resulting in 160 true links, increasing the sensitivity to 83.7% and bringing the total number of linked records to 168.

Discussion

Our study demonstrates that high-quality linkage of records belonging to different individuals in large routine databases from a middle-income country can be achieved, without unique identifiers. Importantly, for the purpose of establishing exposure to dengue for a case-control study, we were able to accurately identify links with a low false-match rate by using a bespoke linkage algorithm designed to overcome the challenges of linking different individuals in national data from Brazil. Although the restricted number of common variables for mothers and stillbirths limited the number of links that could be automatically detected, manual inspection allowed us to greatly improve sensitivity; however, this approach is resource-intensive and may not be feasible for larger databases, such as live births in Brazil with 3 million records a year. Our comparison of missed matches and true matches indicate that linkage errors occur randomly, and are unlikely to introduce bias into our analyses.

We show that linkage between large administrative datasets is complex and requires a number of steps. Our description of these steps (and commands listed in the annex) are available as guidance to others aiming to link similar data sources. Although we did not have a readily available training dataset (where the true match status of each record pair was known), we were able to create a gold-standard dataset from which to derive an appropriate match weight algorithm, and to evaluate the accuracy of both linkage methodologies. This was possible in our study because we had access to identifiable data, and could examine records manually, but is not always the case, when data from clinical and identifiers information are separated to protect the patient privacy [19].

The bespoke algorithm created specifically to link this dataset achieved higher linkage quality that the off-the-shelf program ReCLinkIII, which has been widely used to perform linkage in Brazil and which has previously been shown to have high sensitivity and specificity [7, 20-22]. ReCLinkIII appears to work less well when there are a limited number of variables available to perform the linkage, as demonstrated in our study and a similar study by Coutinho et al. [23], which obtained a sensitivity of 60.9% and 72.8% (without and including the uncertain area respectively). The ReCLinkIII program uses the Levenshtein string comparator, which according to Freire [24] is not the most effective option to compare names in Brazil. The Jaro-Winkler string comparator, used within our bespoke algorithm, has been shown to give the best results when compared with other string comparators to link names in Brazil [24]. Other string
Table 4 Associations between linkage accuracy (using the bespoke algorithm) and characteristics of the cohort

<table>
<thead>
<tr>
<th>Age of the mother in years</th>
<th>True matches N = 123, n (%)</th>
<th>Mixed-matches N = 68, n (%)</th>
<th>OR (95% CI)</th>
<th>p-value</th>
<th>False-matches N = 8, n (%)</th>
<th>OR (95% CI)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 20</td>
<td>25 (20.3)</td>
<td>19 (27.9)</td>
<td>1</td>
<td>p = 0.047</td>
<td></td>
<td></td>
<td>p = 0.095</td>
</tr>
<tr>
<td>20-35</td>
<td>67 (24.3)</td>
<td>30 (44.1)</td>
<td>1.7 (0.8-3.5)</td>
<td>8 (100)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt; 35</td>
<td>27 (17.9)</td>
<td>7 (13.3)</td>
<td>2.3 (0.8-6.7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>9 (7.3)</td>
<td>12 (17.3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maternal literacy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Illiterate</td>
<td>7 (5.7)</td>
<td>5 (7.1)</td>
<td>1</td>
<td>p = 0.113</td>
<td>1</td>
<td>p = 0.17</td>
<td></td>
</tr>
<tr>
<td>1-3 years</td>
<td>8 (6.5)</td>
<td>4 (5.9)</td>
<td>1.4 (0.3-7.5)</td>
<td>1 (12.5)</td>
<td>0.9 (0.4-1.8)²</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-7 years</td>
<td>32 (25.0)</td>
<td>14 (20.6)</td>
<td>1.6 (0.4-6.0)</td>
<td>2 (25.0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt; 8 years</td>
<td>34 (27.6)</td>
<td>20 (29.4)</td>
<td>1.2 (0.3-4.3)</td>
<td>4 (50.0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt; 11 years</td>
<td>19 (15.4)</td>
<td>3 (4.4)</td>
<td>4.5 (0.8-34.1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>23 (18.7)</td>
<td>22 (32.3)</td>
<td>1</td>
<td>p = 0.248</td>
<td>2 (25.0)</td>
<td>1</td>
<td>p = 0.248</td>
</tr>
<tr>
<td>Previous fetal death or abortion</td>
<td>36 (29.3)</td>
<td>17 (25.0)</td>
<td>1</td>
<td>p = 0.374</td>
<td>2 (25.0)</td>
<td>1</td>
<td>p = 0.374</td>
</tr>
<tr>
<td>Yes</td>
<td>59 (48.0)</td>
<td>31 (46.3)</td>
<td>0.9 (0.4-1.8)</td>
<td>4 (50.0)</td>
<td>1.2 (0.2-7)²</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>28 (22.7)</td>
<td>20 (29.4)</td>
<td>2</td>
<td>p = 0.248</td>
<td>2 (25.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gestational age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 22 weeks</td>
<td>9 (7.3)</td>
<td>2 (27.9)</td>
<td>1</td>
<td>p = 0.248</td>
<td></td>
<td></td>
<td>p = 0.402</td>
</tr>
<tr>
<td>22-27 weeks</td>
<td>29 (23.6)</td>
<td>18 (26.5)</td>
<td>0.3 (0.1-1.8)</td>
<td>3 (37.5)</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>28-31 weeks</td>
<td>21 (17.3)</td>
<td>11 (16.2)</td>
<td>0.4 (0.1-2.3)</td>
<td>3 (37.5)</td>
<td>0.7 (0.4-1.3)²</td>
<td></td>
<td></td>
</tr>
<tr>
<td>32-35 weeks</td>
<td>20 (16.1)</td>
<td>19 (27.9)</td>
<td>0.3 (0.1-1.5)</td>
<td>2 (25.0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37-41 weeks</td>
<td>28 (22.8)</td>
<td>9 (13.2)</td>
<td>0.7 (0.1-3.8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥ 42 weeks</td>
<td>2 (1.6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>8 (6.5)</td>
<td>9 (13.2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birth or death weight</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥ 2500</td>
<td>23 (18.7)</td>
<td>16 (23.3)</td>
<td>1</td>
<td>p = 0.798</td>
<td></td>
<td></td>
<td>p = 0.165</td>
</tr>
<tr>
<td>1500-2500</td>
<td>26 (21.9)</td>
<td>16 (23.5)</td>
<td>1.1 (0.4-2.7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 1500</td>
<td>64 (52.0)</td>
<td>31 (45.0)</td>
<td>1.4 (0.6-3.0)</td>
<td>7 (87.5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>10 (8.1)</td>
<td>7 (7.3)</td>
<td>1</td>
<td>p = 0.248</td>
<td></td>
<td></td>
<td>p = 0.248</td>
</tr>
</tbody>
</table>

*Due to the small number of observations, we used only two categories, the first category without missing value as the reference one.

Comparators may be more appropriate in other situations [25]. The improved performance of the bespoke linkage algorithm was also likely due to our derivation of frequency-based match weights and the fact that we allowed weights to differ for common names and differences in age. Additional strengths of our approach were that the data were rigorously prepared to ensure that information belonging to the same mothers and babies could be linked; we performed the validation study in a large sample size that allowed evaluation of two thresholds and comparison with a commercially available software, and because we had additional information about the characteristics of the mothers included in the study we were able to check for bias.

An important aspect of linkage quality is the choice of threshold. The investigator, keeping in mind that different thresholds are more appropriate for different study questions, must make this choice. In our study, where linkage aimed to provide information on exposure (dengue during pregnancy) [36] we chose a conservative threshold to minimise false-matches. We showed that more relaxed threshold added a number of false-matches, drastically decreasing the positive predictive value, without substantially increasing sensitivity. This is because stillbirth is a rare outcome, and there was a large number of real non-matches. This study has a number of limitations. Although a gold-standard dataset was used to measure the linkage
accuracy, there remains scope for linkage errors to occur: we may not have identified all missed matches due to missing data on some records. Our estimates of sensitivity should therefore be interpreted with caution, and should not be assumed to apply to other datasets. Due to the low sensitivity of our linkage approach, our study should not be used to estimate stillbirth rates for women exposed to dengue during pregnancy. Our analysis of the association between study characteristics and linkage error may have been limited due to low power. A further consideration for case-control studies is whether linkage of controls and cases (to the exposure) has similar accuracy. We did not address the accuracy of linkage to live births in this study, but will evaluate this in future research.

Conclusion
To make best use of linked data, it is important to evaluate the quality of linkage processes and to understand the limitations and bias that errors in linkage could introduce in the research results [27, 28]. Validation studies are therefore useful for assessing whether probabilistic matching of such records is effective and whether results are reliable. We present this validation study to add to the limited existing evaluations of data linkage from middle-income countries using a limited number of identifiers. We show that it is possible to achieve a high standard of linkage between different individuals within administrative data from Brazil, specifically for the purposes of accurately stillbirths exposed to dengue [20]. Our results highlight that bespoke linkage algorithms perform better than off-the-shelf software, and that manual review can be a valuable tool for improving sensitivity.

Abbreviations
PPV: Positive predictive value; REDDIT: Software to link data used in Brazil; SIM: Mortality Information system; SINAN: Notifiable disease information system

Acknowledgements
The authors would like to thank Daniela Martins Neto, Tiago Mendes Cruz and Yara Rodrigues for the valuable help in the data cleaning process.

Funding
RP is funded by National Council for Scientific and Technological Development (CNPq-Brazil). LLCI is partly funded by the European Union’s Horizon 2020 research and innovation program under 7th Framework Program Grant Agreement No. 723324. RP is funded by the Welcome Trust (grant number 093757/Z/14/Z). However the funder of this study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Availability of data and materials
The data that support the findings of this study are available from Brazilian Ministry of Health but restrictions apply to the availability of these data, which were used under license for the current study and so are not publicly available. Data are however available from the Brazilian Ministry of Health upon reasonable request.

Authors’ contributions
RP carried out the analysis and wrote the first draft of the article. LLCI conceived the study. KH, BE, MC and FF contributed to the study design and interpretation. All authors revised the manuscript and approved the final version.

Ethics approval and consent to participate
Ethical approval was obtained from the Research Ethics Committee, Public Health Institute, Federal University of Bahia, Salvador, Brazil (CAAE: 30707.814.7.000.0030. CEP: 03) and from London School of Hygiene and Tropical Medicine (ethics ref.10269).

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
1. London School of Hygiene and Tropical Medicine, Kappel St. Rebornst, London, WC1E 7HT, UK; Instituto de Saúde Coletiva, Rua Basílio de Carne, no. 36, Salvador, Bahia CEP: 40160-000, Brazil; 2. Universidade de São Paulo, Av. Ademar de Barros, s/n, Onix, Salvador, Bahia CEP: 40170-110, Brazil.

Received: 10 April 2017 Accepted: 10 July 2017 Published online: 17 July 2017

References


Submit your next manuscript to BioMed Central and we will help you at every step:
- We accept pre-submission inquiries
- Our selector tool helps you find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at www.biomedcentral.com/submt
Evaluation of records linkage of two large administrative databases: live births and notifications of dengue

Background

In the paper above, I presented and discussed the methods used in the linkage process and evaluated the quality of the linkage between dengue cases and stillbirth (SIM-SINAN). The next linkage carried out was the linkage between dengue cases and live births (SINASC-SINAN). This linkage faced the same challenges as dengue and stillbirth linkage, such as the availability of well-completed identifiers, linkage of records belonging to different individuals and the unknown number of expected matches. Because of this, I implemented the same technique described in the paper above.

Although I applied the same technique used for stillbirths and used the same variable in the linkage, there are two main reasons to explain the importance of conducting an evaluation of the quality of the linkage between live births and dengue. First, for any linkage it is important to incorporate information on linkage error into results. Second, to estimate the association between dengue and stillbirth and dengue and maternal deaths, live births is the control group. Therefore, I have to ensure the validity of the linkage to measure the exposure in a similar way for cases and controls. In other words, this is to guarantee that the linkage error is a non-differential error so that misclassification of dengue does not vary according to outcome (stillbirth or live birth). For example, if the sensitivity (true links among the matches) of the linkage were higher among stillbirth, the measure of association would be overestimated, and a false assumption about a positive relationship would be presented.

To link and evaluate the linkage between dengue notification and live births, the same technique described in the paper above was used. First data pre-processing, blocking
by municipality, the creation of a Gold-standard data and classification of the links. I selected only two Brazilian states to work with in the linkage evaluation, because there are approximately 3 million live births in Brazil and limited time and resources for this project. In the data pre-processing step, my final study population comprised 23,190 dengue notifications from SINAN and 180,438 live births from SINASC in two Brazilian states (Ceara and Espirito Santo) in 2010. The subsequent steps (blocking, gold standard, match calculation, classification of links and statistical analysis) were performed as described previously.

**Results**

Of the 23,190 eligible dengue cases in SINAN for 2010, 369 were linked to a live birth record using the nine-step gold-standard algorithm. The threshold was set to a match weight of 21.9, and resulted in 241 links. Comparison with the gold-standard identified 11 false matches and 139 missed-matches, giving a sensitivity of 230/369=62% and a PPV of 230/241=95%. Linkage errors were not associated with any recorded study variable (Table 1).

The threshold for uncertain links was defined as those with weights between 21.9-17. Records with a match between these thresholds were classified through manually inspection to determine whether they belonged to the same person. This added 95 true links increasing the sensitivity to 88% and bringing the total number of linked records to 325.

**Discussion**

I applied the same method to both datasets (live births and stillbirth). As shown in the stillbirth linkage, the sensitivity was not very high, probably due to the lower
number of identifiers. However, when I compared the results obtained in both linkages the accuracy was similar. The stillbirth-dengue linkage presented a sensitivity of 64% and PPV of 93.9, the live birth-dengue linkage showed a sensitivity of 62% and PPV of 95%. Therefore, the linkage used as an instrument to assess the exposure in this study introduced error, (the sensitivity was around 60%). However, this error was independent of the outcome.

The analyses of linkage error were not associated with any maternal or birth characteristic (Table 1). These analyses showed that the missed matched was not associated with a specific group, so it is unlikely that the linkage process introduced bias in the analyses that will be carried out in the subsequent studies.
Table 1. Associations between linkage accuracy and characteristics of the cohort

<table>
<thead>
<tr>
<th></th>
<th>True matches N=230</th>
<th>Missed-matches N=139</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age of the mother in years</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;20</td>
<td>50 (21.7)</td>
<td>35 (25.2)</td>
<td></td>
</tr>
<tr>
<td>20-35</td>
<td>168 (73.0)</td>
<td>96 (69.1)</td>
<td>p=0.668</td>
</tr>
<tr>
<td>&gt;35</td>
<td>12 (5.2)</td>
<td>8 (5.8)</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Maternal literacy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Illiterate</td>
<td>3 (1.3)</td>
<td>1 (0.7)</td>
<td></td>
</tr>
<tr>
<td>1-3 years</td>
<td>13 (5.6)</td>
<td>6 (4.3)</td>
<td></td>
</tr>
<tr>
<td>4-7 years</td>
<td>81 (35.2)</td>
<td>42 (30.2)</td>
<td>p=0.793</td>
</tr>
<tr>
<td>&gt;8 years s</td>
<td>96 (41.7)</td>
<td>69 (49.6)</td>
<td></td>
</tr>
<tr>
<td>&gt;11 years</td>
<td>31 (13.5)</td>
<td>17 (12.2)</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>6 (2.6)</td>
<td>4 (2.9)</td>
<td></td>
</tr>
<tr>
<td>Maternal status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single/Widow/Divorced</td>
<td>155 (67.4)</td>
<td>91 (65.5)</td>
<td>p=0.912</td>
</tr>
<tr>
<td>Married/Union</td>
<td>71 (30.9)</td>
<td>46 (33.1)</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>4 (1.7)</td>
<td>2 (1.4)</td>
<td></td>
</tr>
<tr>
<td>Gestational age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22-27 weeks</td>
<td>1 (0.4)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>28-31 weeks</td>
<td>2 (0.9)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>32-36 weeks</td>
<td>17 (7.4)</td>
<td>6 (4.3)</td>
<td></td>
</tr>
<tr>
<td>37-41 weeks</td>
<td>208 (90.4)</td>
<td>129 (92.8)</td>
<td>p=0.324</td>
</tr>
<tr>
<td>≥42 weeks</td>
<td>2 (0.9)</td>
<td>3 (2.2)</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>0</td>
<td>1 (0.7)</td>
<td></td>
</tr>
<tr>
<td>Birth or death weight</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥2,500</td>
<td>210 (91.3)</td>
<td>134 (96.4)</td>
<td>p= 0.064</td>
</tr>
<tr>
<td>1,500-2,500</td>
<td>19 (8.3)</td>
<td>4 (2.9)</td>
<td></td>
</tr>
<tr>
<td>&lt;1,500</td>
<td>1 (0.4)</td>
<td>1 (0.7)</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Chapter 5. Linkage procedures and validation – description of the maternal death linkage process

The linkage between maternal deaths and women notified with dengue could have been done by linking information belonging to the same person woman-woman (woman notified with dengue and woman who died). It would have increased the number of common identifiers in the datasets improving the quality of the linkage. However, as discussed before, to estimate the association between dengue during pregnancy and maternal deaths, the comparison group is live births. In this case, if the linkage were performed directly (cases of dengue and maternal deaths), there would be a variation in the quality of the linkage to assess the exposure (dengue status) between maternal deaths and live births. In this case, the linkage would perform better among the cases (maternal deaths) compared with the comparison group (live births), introducing a differential error. In these circumstances, the linkage of maternal deaths would have a better sensitivity than the linkage among live births; therefore, the measure of association would be overestimated.

To avoid this issue, I proposed an indirect linkage. In brief, I used the composite file (dengue notifications and live births/stillbirths) of the linkage described above to link with maternal death records. This method will be explained in detail in the following paper. Another advantage of this is the overlapping between the datasets (maternal deaths and fetal outcomes –livebirth/stillbirth); consequently, there was an expected number of known matched records.
RESEARCH PAPER COVER SHEET

PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED FOR EACH RESEARCH PAPER INCLUDED IN A THESIS.

SECTION A – Student Details

<table>
<thead>
<tr>
<th>Student</th>
<th>Enny da Paixao Cruz (Enny S Paixao)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Principal Supervisor</td>
<td>Elizabeth Brickley</td>
</tr>
<tr>
<td>Thesis Title</td>
<td>Symptomatic dengue and adverse pregnancy outcomes: a population-based record linkage study</td>
</tr>
</tbody>
</table>

If the research paper has previously been published, please complete Section B. If not, please move to Section C.

SECTION B – Paper already published

<table>
<thead>
<tr>
<th>Where was the work published?</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>When was the work published?</td>
<td></td>
</tr>
</tbody>
</table>

If the work was published prior to registration for your research degree, give a brief rationale for its inclusion.

Have you retained the copyright for the work?*  
Choose an item.  
Was the work subject to academic peer review?  
Choose an item.

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

<table>
<thead>
<tr>
<th>Where is the work intended to be published?</th>
<th>Plos One</th>
</tr>
</thead>
<tbody>
<tr>
<td>Please list the paper's authors in the intended authorship order:</td>
<td>Enny S. Paixao, Oona Campbell, Laura C. Rodrigues, Maria Gloria Teixeira, Maria da Conceicao N Costa, Elizabeth B. Brickley, Katie Harron</td>
</tr>
<tr>
<td>Stage of publication</td>
<td>Submitted</td>
</tr>
</tbody>
</table>

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)

Student Signature: [Signature]  
Date: 3/03/18

www.lshtm.ac.uk
Paper 3 title: Validating linkage of multiple population-based administrative databases in Brazil

Authors: Enny S. Paixao, Oona Campbell, Laura C. Rodrigues, Maria Gloria Teixeira, Maria da Conceicao N Costa, Elizabeth B. Brickley, Katie Harron

Author contribution:

ESP carried out the analysis and wrote the first draft of the article. LCR, KH conceived the study. OC, MCNC and MGT contributed to the study design and interpretation. EBB contributed with the revision of the first draft. All authors revised the manuscript and approved the final version.
Validating linkage of multiple population-based administrative databases in Brazil

Enny S Paixão*, MSc¹

Oona Campbell, PhD¹ email address: oona.campbell@lshtm.ac.uk

Laura C. Rodrigues, PhD¹ email address: laura.rodrigues@lshtm.ac.uk

Maria Glória Teixeira, PhD² email address: magloria@ufba.br

Maria da Conceição N. Costa, PhD² email address: mncosta@ufba.br

Elizabeth B. Brickley, PhD¹ email address: elizabeth.brickley@lshtm.ac.uk

Katie Harron, PhD¹ email address: katie.harron@lshtm.ac.uk

* Corresponding author email: enny.cruz@lshtm.ac.uk

1 London School of Hygiene and Tropical Medicine. Keppel St, Bloomsbury, London WC1E 7HT, United Kingdom

2 Instituto de Saúde Coletiva. Rua Basílio da Gama, s/n.Canela. CEP 40110040. Salvador, Bahia, Brazil.
Abstract

Background: Linking routinely-collected data provides an opportunity to measure the effects of exposures that occur before birth on maternal, fetal and infant outcomes. High quality linkage is a prerequisite for producing reliable results, and there are specific challenges in mother-baby linkage. Using population-based administrative databases from Brazil, this study aims to estimate the accuracy of linkage between maternal deaths and birth outcomes linked with dengue notifications, and to identify potential sources of bias for assessing the risk of maternal death due to dengue in pregnancy.

Methods: Women with dengue during pregnancy were identified from a previously linked dataset of dengue notifications in women who had live births or stillbirths during 2007-2012. We linked this dataset with maternal death records using probabilistic linkage with maternal name, age and municipality. We estimated the accuracy of the linkage and examined the characteristics of false-matches and missed-matches to identify any sources of bias.

Results: Of 10,259 maternal deaths recorded in 2007-2012, 6717 were linked: 5444 to a live birth record, 1306 to a stillbirth record, and 33 to both a live and stillbirth record. After identifying 2620 missed-matches and 124 false-matches, estimated sensitivity was 72%, specificity was 88%, and positive predictive value was 98%. Linkage errors were associated with maternal education and skin colour; women who had more than 7 years of education or self-declared as Caucasian were more likely to link. Dengue status was not associated with error.

Conclusion: Despite a lack of unique identifiers for linking mothers and birth outcomes, we demonstrate a high standard of linkage, with sensitivity and specificity values comparable with previous literature. Although there were no differences in the
characteristics of dengue cases missed or included in our linked dataset, linkage error occurred disproportionally according to some social-demographic characteristics, which should be taken into account in future analyses.
Background

Research involving record linkage has increased in recent years with the rising availability of administrative population-based health databases and the relatively low costs of bringing together information from different sources. Linkage has been particularly important for research on maternal, fetal and infant health. Information on maternal health and women’s social contexts prior to and during pregnancy can be linked with fetal and childhood outcomes beyond the limits of traditional cohort studies. Due to the large scale and availability of population-based datasets and because these data are collected routinely (i.e., do not require a large cohort with many years of follow up to be established), rare or long-term outcomes can also be investigated. In the past decade, many developed countries have used linkage to investigate the health of mothers and their babies 1-5. However, for developing countries, record linkage methods have only been developed more recently 6.

The quality of the linkage is important for producing reliable results, and there are specific challenges in mother-baby linkage, mainly due to the limited availability of common identifiers for linkage 1. The impact of linkage error depends on the question of interest 7. For example, linkages with low sensitivity cannot be used to measure the incidence of a given disease, such as cancer, in a population, otherwise the measure calculated will be underestimated 8. However if the study presents a low sensitivity randomly distributed between groups, the measure of association will not be biased, so it can be useful under this circumstance, as shown in the study of dengue during pregnancy and stillbirth 8. In other studies, it may be more important to prioritise specificity over sensitivity. For example, in case-control studies, we may be more concerned with precisely establishing exposure status through linkage, rather than capturing all links 6.
this case, it is also important to avoid selection bias (if missed-matches are non-random, i.e. are more likely to occur in specific subgroups of records).

This linkage study is part of a series of studies measuring the association between dengue and pregnancy outcomes \textsuperscript{6,9}. In this particular article, we aimed to validate our linkage strategy discussed in Paixao et al 2017\textsuperscript{6}, to facilitate a cohort study on the effects of dengue on maternal and birth outcomes, in which linkage was required to establish exposure status. Evaluating a linkage strategy with a dataset for which we have, \textit{a priori}, information on the expected number of matches (a maternal mortality dataset) gives us information about the likely quality of the same linkage strategy when applied to linkage between datasets where the expected number of matches is not known (linkage between live births, stillbirths and dengue notifications). The purpose of the analysis presented here (linkage of maternal death records with live births and stillbirths in the context of maternal dengue infections) was to investigate whether any linkage errors were disproportionately distributed amongst the group with the outcome (maternal death) and the comparison group (mothers who did not die). This was important to establish any potential bias in the linked dataset, which might influence subsequent analyses. We aimed to estimate the accuracy of the linkage and identify potential sources of bias, where any subgroups of records were more or less likely to link.

\textbf{Methods}

\textbf{Datasets}

We linked three routinely collected Brazilian datasets: a) notifications of infectious disease (SINAN) for dengue; b) Live Births Information System (SINASC) for live births; and c) Mortality Information System (SIM) for stillbirths and maternal deaths. The datasets were linked in two stages (Figure 1). First, we probabilistically linked
dengue notifications in women (SINAN) with records of women who had live births (SINASC) or stillbirths (SIM), to identify those women who had dengue during pregnancy during 2007-2012. This first linkage (between dengue notifications and live births/stillbirths) has been described in detail elsewhere. Second, we linked this composite file (women with a live birth/stillbirth linked with a dengue notification) with maternal death records for 2007-2012. In this paper, we focus on the results of this second stage of the linkage.

In both linkage processes we used the common variables in all three datasets: maternal name, maternal age, and place of residence. For linkage with death records, we also used information on time between the maternal death and birth outcome in the algorithm, because maternal deaths are likely to occur in the same period (often on the same day) as the birth outcome. Therefore, the timing of these dates contributed information on the likelihood of two records being a match.
Data Source 1: Brazilian Notifiable Diseases Information System - SINAN

The Brazilian Notifiable Diseases Information System (Sistema de Informação de Agravos de Notificação; SINAN) contains records on notifiable diseases, characteristics of the individual with the disease (name, place of residence, age, sex, and years of education), symptoms of the disease, laboratory tests, and disease severity.

After excluding men, and non-dengue records, we retained 1,725,943 cases of dengue from 2007-2012. This is likely an underestimate of the burden of disease, as most dengue cases are not reported to the National Surveillance system in Brazil. For every notified case, there are an estimated 12 dengue cases in the community 10.

Data Source 2: Brazilian Live Birth Information System - SINASC

The Brazilian Live Birth Information System (Sistema de Informação sobre Nascimentos; SINASC) contains records of all live births in Brazil; these data come from birth registration, a legal document completed by the health worker who attended the birth. It includes information on the mother (name, place of residence, age, marital status, education); the pregnancy (length of gestation, type of delivery); and the neonate (birth weight, the presence of congenital anomalies). According to an evaluation of the birth registration system in Brazil, 97% of Brazilian live births are registered.

Data Source 3: Brazilian Mortality Information System - SIM

The Brazilian Mortality Information System (Sistema de Informação sobre Mortalidade; SIM) contains records of all deaths in Brazil, including stillbirths (stillbirth in Brazil is defined as the death of a product of conception before the expulsion or
complete extraction from the body of the pregnant woman, occurring from 22 weeks or weighing more than 500g). These data come from the Death Certificate, a required legal document. We retained all maternal deaths, defined as those coded under obstetric causes of death by ICD-10, the "O" group), and all stillbirths.

It is possible that stillbirths are under-reported in the national system\textsuperscript{11}. However, it is difficult to estimate the rate of under-ascertainment. After exclusion of records without the name of the mothers, we retained 187,487 stillbirth records and 10,259 maternal deaths (with an obstetric code in ICD-10 recorded as the cause of death) extracted from SIM (2007-2012).

*Completeness of linkage variables*

Table 1 contains information on the number of missing values in the variables used in the linkage process in the three datasets (SIM, SINAN and SINASC).

<table>
<thead>
<tr>
<th>Databases</th>
<th>Records</th>
<th>Percentage missing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Maternal name</td>
<td>Maternal age</td>
</tr>
<tr>
<td>Notifiable Diseases Information System (Sistema de Informação de Agravos de Notificação; SINAN)</td>
<td>1,725,943</td>
<td>0.4%</td>
</tr>
<tr>
<td>Live Birth Information System (Sistema de Informação sobre Nascimentos; SINASC)</td>
<td>17,387,267</td>
<td>0.03%</td>
</tr>
<tr>
<td>Mortality Information System (Sistema de Informação sobre Mortalidade; SIM). Stillbirths</td>
<td>187,487</td>
<td>2%</td>
</tr>
<tr>
<td>Mortality Information System (Sistema de Informação sobre Mortalidade; SIM). Maternal Deaths</td>
<td>10,259</td>
<td>0%</td>
</tr>
</tbody>
</table>
Ethics approval and consent to participate

Ethical approval was obtained from the Research Ethics Committee, Public Health Institute, Federal University of Bahia, Salvador, Brazil (CAAE: 26797814.7.0000.5030 CEP-ISC) and from London School of Hygiene and Tropical Medicine (Ethics Ref: 10269).

Linkage strategy

We derived our pre-processing and blocking schemes, gold-standard data and match weights based on previous linkage of Brazilian birth data, as described by Paixao et al (2017)\(^6\). Firstly we cleaned and prepared the data: we excluded records without names, those with generic names such as “ignorado, hospital name”, and deleted punctuation and consecutive spaces, and unknown prefixes. We transformed known abbreviations for names (e.g. Ap → Aparecida) into full names, and all names into upper case, dropping middle initials. We then blocked the records by municipality and used the Jaro-Winkler string comparator to calculate the similarity between names recorded in each dataset\(^12\).

To estimate parameters for linkage weights and to validate the quality of the linkage, we created a gold-standard dataset. Firstly, we linked the dataset using deterministic linkage with exact agreement on full name and age. Subsequent steps relaxed the rules by allowing matches with differences in the age and for differences in the ways names were recorded. Each step was followed by manual review to exclude false-matches\(^6\).

Match weight calculations were based on the Fellegi-Sunter method\(^13\). For each record pair we calculated a probabilistic match weight based on two conditional
probabilities: m-probability, P (agreement|match) and u-probability, P (agreement|non-match). We estimated m-probabilities for each identifier from the true matches in the gold-standard dataset. We calculated u-probabilities based on a list of non-matches, created from all pairwise comparisons of records within SINAN, excluding those belonging to the same individual.

Frequency-based weights were calculated for each category of Jaro-Winkler score comparator (for name of the woman) and year of age, so that rarer values were given higher weights. Separate weights were also calculated for the five most frequently occurring names in the data (Maria, Ana, Santos, Souza and Oliveira).

Since we were linking different subjects (the woman and her live or stillbirth), and because the exposure (dengue during pregnancy) could have happened up to nine months before the outcome (birth/ maternal death), there were timing issues to consider. Two records could differ in time by nine months and could bridge over a calendar year; some mothers would have a birthday between the data of dengue notification and the date of the live birth/death. To allow for this, we estimated different weights according to the similarity of age across datasets: equal ages, age differing by one year, age differing by two years, and ages differing by more than two years. In the maternal mortality linkage, we included the time between the birth outcome and death of the mother as linkage variable (occurred at the same day as the birth outcome; between 2-7 days; between 8-15 days; between 16-30 days; after 30 days).

Match weights were calculated by summing the log of the ratio of m-probabilities and u-probabilities across different identifiers. The algorithm was implemented in Stata 14.1 and R 3.4.1.
Records pairs were ordered by match weight and manually inspected to identify a conservative threshold values aiming to exclude as many as false-matches as possible (prioritising a high positive predictive value). Any records above the cut-off threshold were classified as links.

**Statistical analysis and evaluation of linkage quality**

For the first linkage stage (of women with dengue notification with those with live births/stillbirths), the expected number of matches was unknown *a priori*, as we did not know how many pregnant women were expected to link with a dengue notification registered in SINAN. These findings are described by Paixao 2017.

For the second linkage stage (of maternal mortality), we expected all maternal deaths would link with a live birth or a stillbirth after excluding women with maternal deaths coded under pregnancy with abortive outcome (ICD-10 codes O00-O08; 10% of maternal mortality notifications). We therefore expected around 90% of maternal morbidity records to have a pregnancy outcome (stillbirth or live birth) and therefore be linked with one of our datasets. We used this value to identify the number of missed matches (records from the same mother-baby pair that failed to link) and to estimate the sensitivity of the linkage.

To estimate the proportion of false matches (records from different mother-baby pairs that are linked), we looked at women who were coded as having died in a pregnancy with an abortive outcome (i.e. those who should not have linked with a live birth/stillbirth record). Those who did link were assumed to be false-matches, as were those who linked to a live birth and a stillbirth simultaneously, unless they were multiple births.
We then examined which characteristics were associated with missed matches in the maternal mortality dataset (since we expected more missed matches than false matches). We compared the characteristics of maternal deaths classified as having dengue through linkage, with maternal deaths where dengue was coded as an underlying cause of death (using ICD-10 codes) but that were missed from our linkage. We examined maternal age in years (continuous variable), maternal education (illiterate to 3 years or education, 4-7 years of education, more than 7 years or education), maternal marital status (married, divorced/widow or single), self-declared skin colour (Caucasian, mixed/black, and other) and classification of deaths occurring during pregnancy or puerperium. Categorical variables were compared between groups with Chi² test or Fisher’s exact test, and means were compared with a t-test. A two-sided P value of less than 0.05 was considered to indicate statistical significance. Stata version 14.1 was used for the statistical analyses.

Results

Maternal Mortality linkage

Of the 10,259 maternal deaths in SIM from 2007-2012, 6717 were linked: 5,444 (53%) were linked to a live birth record, and 1,306 (13%) were linked to a stillbirth record. Of these linked records, 33 linked to both a live and stillbirth simultaneously: 20 were multiple pregnancies (with a live birth and stillbirth outcome); 3 were duplicate records (stillbirths misreported as live births and notified in both datasets); and 10 were classified as false matches. After excluding these, 3542 maternal deaths remained unlinked (Figure 2).
Of the 10,259 maternal deaths, 1,046 (10%) were coded in SIM as being pregnancies with an abortive outcome, which we did not expect to link to a live or stillbirth. Of these pregnancies with an abortive outcome, 124 (12%) linked (i.e. were false matches); 55 linked with a live birth and 71 linked with a stillbirth (two linked with both). This gave an estimated specificity of 922/1046=88% (Table 2). However, since
misclassification of abortion and stillbirth is very common (with either a stillbirth being
coded as an abortion or an abortion as a stillbirth), we also estimated the specificity for
linkage with live births only, which gave a specificity of 991/1046=95%.

Table 2: Linkage accuracy

<table>
<thead>
<tr>
<th>Link</th>
<th>Match</th>
<th>Non-match</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6593</td>
<td>124</td>
</tr>
<tr>
<td>Non-link</td>
<td>2620</td>
<td>922</td>
</tr>
</tbody>
</table>

|           | 9213  | 1046      | 10259 |

After excluding the 1,046 maternal deaths with abortive outcomes from the
10,259 coded pregnancies, we expected that the remaining 9,213 maternal death records
would all have linked with a live birth record or a stillbirth record. Instead, we observed
only 6717 linked records, of which 124 were the aforementioned false matches and 6,593
were true links (Table 2). This is an estimated sensitivity of 6,593/9213=72% and a
positive predictive value (PPV) of 6593/6717=98%.

Maternal mortality and live births/stillbirths linkage errors

Missed matches were not associated with age or marital status (Table 3). Records
were more likely to link if women had more than 7 years of education or self-declared as
Caucasian. Records were also more likely to link if maternal death occurred in the
puerperium than if records classified the death as occurring during pregnancy.
Table 3. Associations between linkage accuracy and maternal characteristics

<table>
<thead>
<tr>
<th></th>
<th>True matches N=6593 n (%)</th>
<th>Missed-matches N=2620 n (%)</th>
<th>OR(95% CI)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age of the mother in years</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean age</td>
<td>28.4</td>
<td>28.7</td>
<td>-</td>
<td>0.127</td>
</tr>
<tr>
<td><strong>Maternal literacy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>More than 8 years</td>
<td>2,236 (33.9)</td>
<td>750 (28.6)</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>4-7 years</td>
<td>1,773 (26.9)</td>
<td>699 (26.7)</td>
<td>0.85 (0.75-0.96)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Less than 3 years</td>
<td>1,009 (15.3)</td>
<td>510 (19.5)</td>
<td>0.66 (0.58-0.76)</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>1,575 (23.9)</td>
<td>661 (25.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Marital status</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>3,615 (54.8)</td>
<td>1,427 (54.5)</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>2,264 (34.3)</td>
<td>914 (34.9)</td>
<td>0.98 (0.89-1.1)</td>
<td>0.611</td>
</tr>
<tr>
<td>Divorced/widow</td>
<td>148 (2.3)</td>
<td>68 (2.6)</td>
<td>0.85 (0.64-1.1)</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>566 (8.6)</td>
<td>211 (8.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Self-declared skin colour</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>2,322 (35.2)</td>
<td>846 (32.3)</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Mixed and black</td>
<td>3,787 (57.4)</td>
<td>1,583 (60.4)</td>
<td>0.87 (0.79-0.96)</td>
<td>0.005</td>
</tr>
<tr>
<td>Others</td>
<td>91 (1.4)</td>
<td>51 (1.9)</td>
<td>0.65 (0.45-0.92)</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>393 (6.0)</td>
<td>140 (5.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Pregnancy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pregnant</td>
<td>1,173 (17.8)</td>
<td>1,331 (50.8)</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Puerperium</td>
<td>4,870 (73.9)</td>
<td>993 (37.9)</td>
<td>5.5 (5.0-6.1)</td>
<td>0.001</td>
</tr>
<tr>
<td>Missing</td>
<td>550 (8.3)</td>
<td>296 (11.3)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Maternal Mortality linkage and dengue**

We identified 23 maternal deaths that were coded in SIM with ICD-10 codes for dengue as the underlying cause of death, but that had not linked to the dengue notification cases in our linkage. The reasons were: a) the maternal death was in a pregnancy with an abortive outcome (the record should have been excluded from the previous linkage, n=3 cases); b) the maternal death record did not link with live birth or stillbirth (missed match, n=8 cases); c) the maternal death linked with a live birth or stillbirth, but not to the dengue notifications (misclassification of dengue status, n=12 cases). We investigated these latter
12 cases further, finding that in 5/12 cases, the symptoms of dengue began in the puerperium, more than 42 days after the fetus was born, and therefore were classified as no dengue during pregnancy in the previous linkage; 1/12 cases were classified as an uncertain link; the remaining 6/12 were false-matches in the original linkage.

Excluding the three abortive outcomes, there were no differences in socio-demographic characteristics between the 20 records that our linkage failed to identify, and the 14 cases of dengue identified in our linkage of maternal mortality and dengue datasets.

Discussion

We implemented the linkage methods described by Paixao et al (2017) in a dataset with a known number of expected matches, and consequently were able to quantify measures of linkage accuracy and validate our previous linkage strategy. We therefore demonstrate that the quality of a particular linkage strategy can be validated using a dataset where the expected number of matches is known. Due to the lack of common variables for linkage, it was not possible to increase the sensitivity without increasing the number of false matches. However, our linkage achieved a PPV of 98%, indicating that the majority of links that were made were correct. In our evaluation of linkage error, we observed that mothers with fewer years of education and mixed ethnic background were more likely to be missed from the linkage. However, there were no differences in the socio-demographic characteristics between the cases of dengue identified in our linked dataset, and those that were missed, to dengue status groups.

Linkage to bring together information from two different sources about the same individual can achieve high rates of sensitivity, even when using records containing
truncated or ambiguous matching variables. However, bringing together information from two different people (in this case, mother and baby) has been considered a more difficult task, due to a lack of common variables. Sensitivity in these cases tends to be lower, e.g. 38% missed links in Georgia and 20% in New Jersey; rates that are comparable with our results (28% missed records). Specificity in our study was estimated at between 88% (pregnancies with an abortive outcome that linked with either live or stillbirths) and 95% (pregnancies with an abortive outcome that linked with live births only), but it was difficult to distinguish between false matches and misclassification of stillbirths/abortive outcomes.

A very important aspect of linkage is how any linkage error might impact on inferences from the linked dataset. The analyst should know and report any evaluation of linkage accuracy and be aware of groups disproportionately affected by linkage error. In this study, we found differences between the matches and the residual (non-linked) records: the matches were more likely to have more than 7 years of education and self-declare as Caucasian compared with the non-links. Adams et al in USA found that incomplete records were related to social disadvantage and consequently these records were less likely to match.

Beyond socioeconomic status, there are other possible explanations for the distinctive characteristics observed in the non-linked group. First, completion of the forms by health care professionals and the process of digitalization in private and public health facilities may result in differential quality of records. Second, in Brazil, induced unsafe abortion has been associated with low education and income, and because it is illegal, it may be that some missed maternal death matches had an abortive outcome resulting from an illegal and unsafe abortion, which were coded under a different cause.
of death in the legal document. Therefore the missed-matches may not all be linkage errors but coding errors, which would explain the differences observed between these two groups. The fact that most of the missed-matches occurred during pregnancy rather than postpartum supports this hypothesis.

To further evaluate the linkage accuracy on the dengue status of the pregnant women (exposure variable), we compared the dengue cases that the algorithm was capable of capturing, with those that it could not (i.e., those identified in post-hoc analyses as individuals coded with dengue as cause of death), and found no difference in socio-demographic characteristics between the two groups. It is therefore unlikely that the linkage process introduced bias in our previous linkage of live births/stillbirths and dengue, since missed matches occurred randomly in relation to the exposure variable. However, this analysis had limited power because of the sample size (34 cases of dengue). Therefore, results showing no difference between the two groups should be interpreted with caution.

This study has a number of limitations. We did not have information about the gestational age at which the maternal death occurred, and so we cannot conclude if the missed-matches occurred due to linkage errors, or if they should not have been linked (i.e., were abortive outcomes). It is possible that we have under estimated the number of false-matches, but we expected all the maternal deaths to link, it is unlikely that records have linked to the wrong record, because that record would also have been linked. Due to the restricted number of variables, we could not improve the sensitivity of the algorithm, and so our linked dataset should not be used to estimate rates of live births or maternal mortality rates for women exposed to dengue during pregnancy. Our analysis of the association between study characteristics and linkage error may have been limited due to low power. Finally, we assumed that linkage quality metrics in our maternal mortality
linkage would be representative of those in our initial linkage between birth outcomes and dengue notifications, since the data are derived from similar sources and contain the same matching variables.

It is important to understand the quality of the linkage and the potential bias that can be introduced in results of analyses of linked data. In this study, sensitivity and specificity of our linkage strategy were comparable with previous literature. Although there were no differences in the characteristics of dengue cases missed or included in our linked dataset, linkage error occurred disproportionately according to some social-demographic characteristics, which should be taken into account in future analyses. These results reinforce the need to evaluate linkage quality and to take linkage error into account within analyses of the following studies when necessary.


11. Vieira, M. S. M., Vieira, F. M., Fröde, T. S. & d’Orsi, E. Fetal Deaths in Brazil:


Section 2 reference


Section 3-RESULTS
Results Section

Overview

This section presents the estimated association between maternal dengue infection and pregnancy outcomes. I analysed the data obtained after the linkage process, and each chapter is one paper and discuss one specific outcome. The chapter 6 (paper 4) describe the association between symptomatic dengue and stillbirth. Chapter 7 (paper 5) estimate the association between maternal dengue and live birth outcome (premature birth, low birth weight and small for gestational age), and finally chapter 8 (paper 6) presents the results for maternal deaths. As a paper style thesis there is some repetition of information.
Chapter 6. Description of the association between symptomatic dengue during pregnancy and stillbirth

Cover sheet

<table>
<thead>
<tr>
<th>Student</th>
<th>Enny da Paixao Cruz (Enny S Paixao)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Principal Supervisor</td>
<td>Elizabeth Brickley</td>
</tr>
<tr>
<td>Thesis Title</td>
<td>Symptomatic dengue and adverse pregnancy outcomes: a population-based record linkage study</td>
</tr>
</tbody>
</table>

If the Research Paper has previously been published please complete Section B, if not please move to Section C

<table>
<thead>
<tr>
<th>Where was the work published?</th>
<th>Lancet Infectious Disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>When was the work published?</td>
<td>September 2017</td>
</tr>
<tr>
<td>If the work was published prior to registration for your research degree, give a brief rationale for its inclusion</td>
<td></td>
</tr>
<tr>
<td>Have you retained the copyright for the work?*</td>
<td>No</td>
</tr>
</tbody>
</table>

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

<table>
<thead>
<tr>
<th>Where is the work intended to be published?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Please list the paper’s authors in the intended authorship order:</td>
</tr>
<tr>
<td>Stage of publication</td>
</tr>
</tbody>
</table>

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)

Student Signature: [Signature] Date: 8/03/18
Supervisor Signature: [Signature] Date: 8/17/18
Paper 4 title: Symptomatic dengue infection during pregnancy and the risk of stillbirth in Brazil, 2006–12: a matched case-control study

Authors: Enny S. Paixao, Maria da Conceicao N Costa, Maria Gloria Teixeira, Katie Harron, Marcia Furquim de Almeida, Mauricio L. Barreto, Laura C. Rodrigues

Author contribution: ESP, MGT, and LCR designed the study. ESP, MdCNC, KH, MFdA, and MLB contributed to the analysis and interpretation. All authors revised the manuscript and approved the final version.
Symptomatic dengue infection during pregnancy and the risk of stillbirth in Brazil, 2006–12: a matched case-control study

Euny S. Patanó, Maria do Conceição N. Costa, Maria Cavaleiro Teixeira, Keila V. Hohner, Marcio F. Giramini de Almeida, Mauricio L. Romero, Laura C. Rodrigues

Summary
Background Maternal infections during pregnancy can increase the risk of fetal death. Dengue infection is common, but little is known about its role in fetal mortality. We aimed to investigate the association between symptomatic dengue infection during pregnancy and fetal death.

Methods We did a nested case-control study using obstetrician-collected data from the Brazilian livebirth information system (SINASC), the mortality information system (SIM), and the national reportable disease information system (SINAN). We identified all pregnancies ending in stillbirth and a random sample of livebirths between Jan 1, 2006, and Dec 31, 2012. We did linkage to determine which mothers were diagnosed with dengue infection during pregnancy. By use of stillbirths as cases and a sample of matched livebirths as a control, we calculated matched odds ratios (mORs) using conditional logistic regression adjusted for maternal age and education.

Findings 275 (0.2%) of 162 138 women who had stillbirths and 1507 (0.1%) of 1 536 105 women who had livebirths were diagnosed with dengue infection during pregnancy. Symptomatic dengue infection during pregnancy almost doubled the odds of fetal death (mOR 1.9, 95% CI 1.5–2.3). The increase in risk was similar when analyses were restricted to laboratory-confirmed cases of dengue infection (1.8, 1.4–2.2). Severe dengue infection increased the risk of fetal death by about five times (4.9, 2.3–10.2).

Interpretation Symptomatic dengue infection during pregnancy is associated with an increased risk of fetal death. We recommend further epidemiological and biological studies of the association between dengue and poor birth outcomes to measure the burden of subclinical infections and elucidate pathological mechanisms.

Introduction Fetal death is a common adverse outcome of pregnancy, even in high-income countries. Although the global burden of stillbirths declined by 25–5% between 2000 and 2015, there were still 2.6 million stillbirths worldwide in 2015. The incidence of stillbirth varies from an estimated four in 1000 births in high-income countries to 26 in 1000 births in low-income and middle-income countries.1

Infection during pregnancy is believed to cause 10–25% of all fetal deaths in high-income countries,2 with infection in early pregnancy carrying an increased risk.3 Infection during the first 28 weeks of gestation causes 15% of all fetal deaths, while infection after 28 weeks of gestation causes 6% of all fetal deaths.4 Viral infections are estimated to contribute to 14–5% of all fetal deaths, although some unexplained stillbirths could be due to undiagnosed viral infections.3

The confirmation of Zika virus infection as a cause of congenital anomalies and stillbirth highlighted the importance of original research on the effect of flaviviruses on stillbirths.5 A review by McClure and colleagues6 in 2014 mentioned dengue infection in pregnant women as a cause of fetal death, yet no population-based studies showing this association have been done. In 2010, a systematic review7 concluded that it was unclear whether dengue infection during pregnancy had any negative effect on fetal outcomes. Another systematic review8 found that only one study9 of dengue infection during pregnancy and fetal death had a comparison group. That study9 included 13 stillbirths (nine of which had been exposed to dengue virus), and the crude relative risk of stillbirth was 6.7 (95% CI 2.1–21.3) in women with symptomatic dengue infection compared with women without. All other evidence to date comes from case reports or case series, none of which have included appropriate controlling for confounding variables or population control groups.

The mechanisms whereby dengue infection could lead to stillbirth are unknown, but three main hypotheses exist: the symptoms of dengue infection in the mother have a direct effect on the fetus (severe dengue); dengue infection causes changes that affect the placenta; and the dengue virus has a direct effect on the fetus.10 The incidence of dengue infection is increasing, and about half of the global population is at risk, including women of reproductive age.11 However, the effect of dengue infection during pregnancy on fetal mortality remains unclear. Therefore, we investigated the association between stillbirth and symptomatic dengue infection during pregnancy.
Methods

Study design
We did a population-based nested case-control study with linkage of routine data for all pregnant women whose outcomes were livebirth or stillbirth in Brazil from Jan 1, 2006, to Dec 31, 2012. Data were extracted from three Brazilian databases: the Brazilian livebirth information system (SINASC), the mortality information system (SIM), and the national reportable disease information system (SINAN). SINASC contains records of all livebirths in Brazil. SIM contains records of all stillbirths in Brazil, and SINAN contains records of all notifiable diseases in Brazil.

SINASC consists of data from livebirth notifications, which is a legal document completed by the health professional who assisted the delivery. This document includes information about the mother (name, place of residence, age, marital status, education, whether she had a stillbirth or a child who died), the pregnancy (length of gestation, type of delivery), and the newborn baby (birthweight, presence of congenital anomalies). The completeness of the data from SINASC is very high, with most variables completed in more than 90% of cases; the exception being information about whether the mother had a previous stillbirth or abortion.

Data in SIM come from the death certificate, which is also a legal document. The proportion of records with missing data varies by variable and is around 20% for maternal education and number of previous fetal deaths or abortions. Data about notifiable diseases, including personal information about the patient (name, place of residence, age, education), symptoms of disease, laboratory tests, and disease severity, are captured in SINAN, which contained reasonably complete data for the variables we used for linkage and in the final analyses (<0.05% of records were excluded because of missing name). SINAN contains two sources of information about maternal age (age and date of birth). About 5% of dengue cases did not have a final classification of severity. Laboratory confirmation was not required for notification of dengue cases because dengue was the main (and for some years the only) vector-borne disease circulating in Brazil (yellow fever and malaria occur in restricted areas; notification of Zika and chikungunya viruses was not introduced until 2014).

This study was approved by the Federal University of Bahia (protocol number 2697/84700000500) and the London School of Hygiene & Tropical Medicine (ethics reference number 10216).

Procedures
In SIM, stillbirth was defined as the death of a product of conception before the expulsion or complete extraction from the body of the pregnant woman, occurring within 22 weeks or weighing more than 500 g, according to the Brazilian definition. In SINASC, livebirth was defined as a product of conception that, after separation from the mother’s body (independent of the duration of pregnancy), breathes or shows any other signs of life, such as a heartbeat, umbilical cord pulsation, or definite movement of voluntary muscles, according to the Brazilian definition.

Records with missing or implausible names for the mother, multiple pregnancies, congenital anomalies, birth or stillbirth from municipalities without dengue infection notification, and an uncertain link status were excluded from case and control groups. Multiple pregnancies and congenital anomalies were excluded because they might be causes or related to causes of stillbirth.

Cases were stillbirths recorded in SIM. As controls, we used a random sample of ten livebirths (without replacement), matched to cases by month and calendar year of birth, and municipality of residence. We randomly ordered cases and controls within month and year and
place of birth groups, and selected the first case in the group to be matched to the first ten controls in the group. The second case was matched to controls 11-20 and so on, forming pairwise combinations. We selected ten controls per case to enable us to control for potential confounders, such as dengue seasonality, intensity of virus circulation, and social environment.

The exposure in this study was a notified confirmed case of dengue infection during pregnancy that resulted in the birth of a case or control. Dengue in Brazil can be confirmed on the basis of clinical epidemiological criteria, including the presence of clinical symptoms of dengue in the same area and time as other confirmed cases of dengue, or on the basis of clinical laboratory criteria, including the presence of clinical symptoms and laboratory confirmation (a positive result from one of IgM detection by ELISA, viral RNA detection by PCR, NS1 viral antigen detection, or positive viral culture).

We considered a fetus to be exposed if their mother was notified to SINAN as confirmed as a dengue case (and therefore their records were linked to a notification of dengue during the pregnancy). Dengue infection during pregnancy refers to all confirmed cases of dengue (based on clinical epidemiological and clinical laboratory data); dengue infection during pregnancy, laboratory confirmed refers to only the laboratory-confirmed cases. We classified dengue into two clinical categories: dengue fever as a self-limiting disease (fever with severe headache, pain behind the eye, muscle and joint pain, and rash) or severe dengue. We classified severe dengue as all cases of dengue that the Brazilian Ministry of Health classified as dengue haemorrhagic fever, dengue shock syndrome, or complicated dengue. Complicated dengue is a Brazilian definition for all severe cases of dengue that do not meet the WHO criteria for dengue haemorrhagic fever (fever, hemorrhagic evidence, thrombocytopenia, and evidence of plasma leakage) and cannot be classified as mild self-limited disease because of their severity. The term complicated dengue was used when a probable case of dengue presented with severe changes in the nervous system, cardiorespiratory dysfunction, insufficient hepatic function, gastrointestinal bleeding, cavity spills, thrombocytopenia of 50000 platelets per ml or fewer, or leukocytosis 3000 cells per ml or fewer. We used only two clinical categories because of the small number of observations and because dengue haemorrhagic fever and dengue shock syndromes are rare.

Linkage procedure
SINAN dengue records were linked probabilistically with the cohort of pregnant women from SINASC and SIM to identify those who had dengue infection during pregnancy. Name, age, and place of residence of the mother at the time of delivery or diagnosis were used in the matching process. We excluded records with missing or implausible names and duplicates. Match weight calculations were based on the Fellegi-Sunter method.

We created a gold-standard dataset using all stillbirth records for 2 years of data (2009 and 2010) and all livebirth records within two states (Ceará and Espírito

---

**Figure 1:** Study profile
One case was matched to ten controls by month and calendar year and municipality of the mother.

---

www.thesource.com/infection Vol 17 September 2017
with a high degree of certainty were retained for the case-control study to avoid misclassification; uncertain links (when we could not determine whether or not a woman had dengue infection) were excluded from the analyses.

To evaluate the quality of the linkage algorithm, we compared the links obtained using the probabilistic algorithm with the gold standard created by manual review. This validation showed that sensitivity was 64% for stillbirths and 62% for livebirths, positive predictive value was 94% for stillbirths and 95% for livebirths, and specificity was more than 99% for both. The missed matches occurred randomly in cases and controls (appendix). The procedures and evaluation of the matching process are the subject of a separate paper (unpublished).

Statistical analysis

To calculate the sample size, we did two preliminary deterministic linkages of records with 100% agreement in the three identifiers (name, age, and municipality of the mother): the first between SINAN and SISM, to calculate the minimum proportion of mothers of stillbirths who had dengue infection during pregnancy (0.05%), and the second between SINAN and SINASC, to calculate the minimum proportion of mothers of livebirths who had dengue infection during pregnancy (0.01%). We estimated that, with 9475 cases, we would be able to detect a matched odds ratio (mOR) of 1.6 using the parameters proportion exposed (0.05% for stillbirths and 0.03% for livebirths), 80% power, 95% confidence level, and a ratio of one case to ten controls (appendix).

For the overall association, we estimated crude mORs with univariate conditional logistic regression and adjusted mORs with conditional logistic regression, controlling for maternal age and maternal education (as a proxy for socioeconomic status). For a sensitivity analysis of the validity of dengue clinical diagnosis, we repeated the analysis with laboratory-confirmed dengue cases only. We investigated the effect of dengue by disease severity and time between the disease and outcome. We did an analysis stratified by report of a previous fetal death or abortion to explore whether dengue infection had a different effect in pregnancies with a high risk of fetal death. To investigate the potential effect of missing data, we did another sensitivity analysis in which we assumed that all missing data for confounding variables in the cases were in the low-risk groups and all missing data for the controls were in the high-risk groups.

The population attributable fraction (PAF; PAF = [p(1-OR)]/OR) was calculated with the pafcalc package in Stata, which uses a logistic regression method and provides PAF (and 95% CIs) with adjustment for confounding variables by combining adjusted ORs and the observed incidence of dengue among cases (stillbirths). We used Stata version 14.1 software for the statistical analyses.

Santo) for 1 year (2010) of data. We did a manual review of record pairs to identify true matches (records belonging to the same mother) and estimate m probabilities (Pr[Agreement(match)], which is the probability of a match given agreement on each identifier. To calculate u-probabilities (Pr[Agreement(non-match)], we used a subset of SINAN to create a set of non-matches by cross-joining all record pairs and removing record pairs belonging to the same individual. Record pairs were then ordered by match weight and manually inspected to identify obvious non-links that had high weights and probable links with low weights. Only links and non-links
Role of the funding source

The funder of this study had no role in study design, data collection, data analysis, interpretation, or writing of the report. All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

SIM recorded 224582 stillbirths and SINASC recorded 20333482 livebirths during the study period. After exclusions, 162781 stillbirths were eligible for the study. No matched control was found for 593 (0.4%) stillbirths and so these were excluded. The final study population included 162188 stillbirths and 1586158 livebirths, with 275 stillbirths and 1507 livebirths exposed to dengue (figure 1).

The characteristics of cases and controls are shown in table 1. The proportion of missing values for all variables was higher among cases than among controls, including values for classic risk factors for stillbirth, such as previous fetal death or abortion (28.3% vs 16.1%) and maternal education (27.5% vs 2.3%). This finding was also true for age, birth or death weight, and gestational age.

The risk of stillbirth among all births recorded in the Information System was 11 in 1000 livebirths. The risk of stillbirth in women with dengue infection during pregnancy was 15 (95% CI 13–17) in 1000 pregnancies. Overall, dengue infection was laboratory-confirmed in more than 30% of the dengue cases recorded. Dengue infection during pregnancy was more common in mothers of cases (0.2%) than in mothers of controls (0.1%); table 1). Potential confounding variables (i.e., variables associated with dengue infection and stillbirth) were maternal education (p=0.001) and previous fetal death or abortion (p=0.02; table 2).

The crude association between symptomatic dengue infection during pregnancy and stillbirth was mOR 1.8 (95% CI 1.6–2.0), and was similar when adjusted for maternal education and maternal age (mOR 1.9, 95% CI 1.6–2.2). The analysis restricted to laboratory-confirmed dengue gave a similar adjusted mOR (1.8, 1.4–2.3), We investigated whether the association between dengue infection and stillbirth depended on previous fetal death or abortion, but the mOR did not differ according to this variable (table 3). The risk of fetal death was dependent on the time between the first symptoms of dengue infection and the date of livebirth or stillbirth; the risk appeared to be bimodal. The risk of fetal death peaked in the first 20 days after the onset of symptoms (4.1, 9.2–7.5); to a greater extent when dengue was severe (table 3 and figure 2). After 20 days, the risk of mortality remained increased (1.7, 1.4–2.4), but was roughly constant until the end of pregnancy (figure 2). Severe dengue during pregnancy increased the risk of stillbirth about five times (4.9, 2.3–10.2), almost three times that of mild dengue (1.7, 1.5–2.0; table 3).

A sensitivity analysis was done because the proportion of missing data was higher among cases than among controls. We assumed that all missing data for confounding variables in the cases were in the low-risk groups and that all missing data in the controls were in

---

**Table 2:** Maternal characteristics for 1,585/105 live singleton births in Brazil by dengue status

<table>
<thead>
<tr>
<th>With notified dengue (n=1585)</th>
<th>Without notified dengue (n=1058)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age of the mother</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;20 years</td>
<td>355 (25.0%)</td>
<td>428 (38.2%)</td>
</tr>
<tr>
<td>20-35 years</td>
<td>1055 (66.8%)</td>
<td>1072 (95.3%)</td>
</tr>
<tr>
<td>&gt;35 years</td>
<td>130 (8.7%)</td>
<td>113 (10.2%)</td>
</tr>
<tr>
<td>Missing</td>
<td>0 (0.0%)</td>
<td>6 (0.5%)</td>
</tr>
</tbody>
</table>

**Table 3:** Association between symptomatic dengue infection during pregnancy and stillbirth

<table>
<thead>
<tr>
<th></th>
<th>Crude odds ratio (95% CI)</th>
<th>Adjusted odds ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All cases of dengue infection during pregnancy</td>
<td>1.8 (1.6–2.0)</td>
<td>1.8 (1.6–2.2)</td>
</tr>
<tr>
<td>Laboratory-confirmed dengue infection only</td>
<td>1.7 (1.3–2.0)</td>
<td>1.8 (1.4–2.3)</td>
</tr>
<tr>
<td>Previous fetal death or abortion</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1.8 (1.4–2.3)</td>
<td>2.8 (1.4–4.2)</td>
</tr>
<tr>
<td>Yes</td>
<td>1.7 (1.3–2.1)</td>
<td>1.6 (1.2–2.1)</td>
</tr>
<tr>
<td>Severity of dengue infection</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mild</td>
<td>5.7 (4.1–9.9)</td>
<td>5.7 (4.5–7.2)</td>
</tr>
<tr>
<td>Severe</td>
<td>4.6 (2.4–9.5)</td>
<td>4.9 (2.3–9.2)</td>
</tr>
<tr>
<td>Length of time between onset of symptoms and outcome</td>
<td></td>
<td></td>
</tr>
<tr>
<td>First 30 days</td>
<td>4.3 (1.9–9.5)</td>
<td>4.9 (2.7–7.5)</td>
</tr>
<tr>
<td>After 20 days</td>
<td>1.6 (1.4–1.9)</td>
<td>1.7 (1.4–2.0)</td>
</tr>
</tbody>
</table>

*Adjusted for maternal education and maternal age.
the high-risk groups. The magnitude of the association was very similar (mOR 1.8, 95% CI 1.5–2.0).

The prevalence of symptomatic dengue infection during pregnancy was low (0–1% of mothers of controls were notified of dengue infection during pregnancy), so the FAF was also low; 0–0.8% of all stillbirths during the period were attributed to dengue infection, with marked yearly variation and a maximum of 0.13% during epidemic years (appendix).

Discussion

We found that symptomatic dengue infection during pregnancy roughly doubled the risk of stillbirth, with the highest risk observed in the first 20 days after disease onset and in women with severe disease. The increase in risk was observed for all cases of dengue (either based on clinical epidemiological data or on clinical laboratory data) and when the analysis was restricted to only laboratory-confirmed dengue. Dengue infection during pregnancy is reasonably rare, and the proportion of all fetal deaths attributable to dengue during pregnancy in Brazil during the period was relatively small, but was larger during epidemic years.

To our knowledge, this is the first study to investigate the risk of stillbirth among women with dengue infection during pregnancy by use of a population-based approach, with a sufficiently large sample size and controlling for confounders. However, we were not able to adjust for all possible confounding factors because we were dependent on routine information obtained from notification system records. The results of our study were consistent with those of a study done in French Guiana, which found that the risk of stillbirth was increased in pregnancies with symptomatic dengue. Stillbirth has been associated with other viral maternal infections and the magnitude of this association varies depending on the virus involved, the severity of maternal illness, gestational age, and other factors. For example, the risk of fetal death in a pregnant woman with influenza virus infection requiring admission to hospital was 4.2 (95% CI 1.4–12.4), whereas the risk of stillbirth in pregnant women with a mild influenza virus infection was 1.9 (1.1–3.4). The risk of stillbirth in women with HIV infection was 1.67 (1.0–2.6).

Although many viral infections increase the risk of fetal death, the mechanism underlying this association is not clear. The analyses of the length of time between symptom onset and stillbirth and of disease severity showed that the risk was highest during acute maternal illness and in women with severe disease, suggesting that the mechanisms involved in the association between dengue infection and stillbirth might be through maternal illness; if the mother becomes severely ill, the fetus could die in the days after symptom onset because of a high fever or other systemic manifestations in the mother. After this first period, in a small proportion of cases, other mechanisms could be relevant. The proposed mechanism for the effect of some viral infections on pregnancy outcomes includes direct fetus infection leading to damage to vital organs, such as the brain and heart. However, vertical transmission of dengue is not thought to be common, although the virus and antibodies against the virus have been found in placenta, in the cord blood of infants, and in the lung and kidney cells of an aborted fetus. In the absence of vertical transmission, the fetus might be harmed by alterations that remain after the infection is over, such as pathological changes in the placenta that lead to hypoxia, as proposed by Ribeiro and colleagues.

However, we could not elucidate this mechanism in this study, which could be clarified by further research with detailed clinical information (including information about medication and comorbidities) and full investigation by use of appropriate instruments and specific laboratory techniques. These studies could be done with a much smaller number of patients than was used here.

According to a systematic review of studies of small sample size, a comment by Carles, and a study of...
similar Brazilian data, dengue infection during pregnancy is associated with adverse fetal outcomes, and severe forms of the disease carry a higher risk than less severe forms. This finding was supported by our study, in which the risk of stillbirth among women who had severe forms of dengue was almost three times that of those with mild disease. Because a history of fetal loss is a risk factor for stillbirth, we investigated whether this variable acted as an effect modifier, but we found that the association was similar for women with and without a previous stillbirth.

Our study had limitations and strengths inherent to the linkage process and to use of secondary data. The main strength of our study was the very large sample size, including all confirmed cases of notified dengue in a country where dengue incidence is high and variable over time. We also had a rigorously selected group of controls and were able to control for confounding. The sensitivity analyses showed the robustness of our findings.

With regard to the use of secondary data, we showed very good agreement on the estimates of risk in cases with and without laboratory confirmation. Our dataset presented some challenges for linkage, and the process and validation of the linkage is discussed in a separate paper. However, it is unlikely that the linkage process introduced bias, because linkage errors occurred randomly across cases and controls. Missing data were more common in cases than in controls, which might have resulted in better linkage among controls than cases, decreasing the magnitude of the observed association. However, the validation study showed a linkage sensitivity of nearly 60% for both cases and controls, so the magnitude of the association should not have been affected by linkage error, although the PAF would have been underestimated.

Despite these limitations, we provide evidence that symptomatic dengue infection during pregnancy is a risk factor for stillbirth, with increased risk in severe forms of the disease. We recommend further research in different settings to confirm our results and to explore other negative outcomes of pregnancy (preterm birth, intrauterine growth restriction, congenital anomalies, and maternal mortality) and studies of other vector-borne diseases contributing to fetal death. Additional research is required to measure the burden of subclinical maternal viral infections in stillbirth and to elucidate the pathological mechanisms involved in this association. If the association between dengue infection and stillbirth is shown in other studies, recommendations should be made to closely monitor pregnant women with dengue symptoms, to incorporate dengue control into programmes to reduce fetal mortality, and to include pregnant women as an at-risk population in dengue control programmes, to strengthen the health education actions for individual protection of pregnant women.

Contributors
ESP, MG2, and LCR designed the study. ESP, Mc2NG, KH, MFLA, and NLMB contributed to the analysis and interpretation. All authors revised the manuscript and approved the final version.

Declaration of interests
We declare no competing interests.

References
24 Verrito AK, Frenkel RF, Rosembuhl D. Placental passage of antibodies to dengue virus in persons living in a region of hyperendemic dengue virus infection. J Infect Dis 1975; 131: 64-68.
Chapter 7. Description of the association between symptomatic dengue during pregnancy and birth outcomes

Cover sheet

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
www.lshtm.ac.uk

Registry
T: +44(0)20 7299 4646
F: +44(0)20 7299 4656
E: registry@lshtm.ac.uk

RESEARCH PAPER COVER SHEET

PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED FOR EACH RESEARCH PAPER INCLUDED IN A THESIS.

SECTION A – Student Details

<table>
<thead>
<tr>
<th>Student</th>
<th>Enny da Paixao Cruz (Enny S Paixao)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Principal Supervisor</td>
<td>Elizabeth Brickley</td>
</tr>
<tr>
<td>Thesis Title</td>
<td>Symptomatic dengue and adverse pregnancy outcomes: a population-based record linkage study</td>
</tr>
</tbody>
</table>

If the Research Paper has previously been published please complete Section B, if not please move to Section C

SECTION B – Paper already published

<table>
<thead>
<tr>
<th>Where was the work published?</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>When was the work published?</td>
<td></td>
</tr>
<tr>
<td>If the work was published prior to registration for your research degree, give a brief rationale for its inclusion</td>
<td></td>
</tr>
<tr>
<td>Have you retained the copyright for the work?*</td>
<td>Choose an item.</td>
</tr>
</tbody>
</table>

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

<table>
<thead>
<tr>
<th>Where is the work intended to be published?</th>
<th>Journal of Epidemiology &amp; Community Health</th>
</tr>
</thead>
<tbody>
<tr>
<td>Please list the paper’s authors in the intended authorship order.</td>
<td>Enny S. Paixao, Oona M R Campbell, Maria Gloria Teixeira, Maria da Conceicao N Costa, Katie Harron, Mauricio L. Barreto, Marcia Furquim de Almeida, Laura C. Rodrigues</td>
</tr>
<tr>
<td>Stage of publication</td>
<td>Submitted</td>
</tr>
</tbody>
</table>

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)

Student Signature: [Signature] Date: 9/03/18
Supervisor Signature: [Signature]  Date: 8/3/15
Paper 5 title: Dengue during pregnancy and adverse birth outcomes: a cohort analysis using routine data

Authors: Enny S. Paixao, Oona M R Campbell, Maria Gloria Teixeira, Maria da Conceicao N Costa, Katie Harron, Mauricio L. Barreto, Marcia Furquim de Almeida, Laura C. Rodrigues

Author contribution: ESP and KH carried out the analysis. ESP wrote the first draft of the article. LCR, MGT conceived the study. MCNC, MLB, OC, MFA contributed to the study design and interpretation. All authors revised the manuscript and approved the final version.
Dengue during pregnancy and adverse birth outcomes: a cohort analysis using routine data

Running title: Dengue and adverse birth outcomes

Enny S Paixão*, MScabc
Oona M.R.Campbell, PhDd
Maria Glória Teixeira, PhDbc
Maria da Conceição N. Costa, PhDb
Katie Harron, PhDd
Mauricio L. BarretoPhDbc
Marcia Furquim de AlmeidaPhDd
Laura C. Rodrigues, PhDac

* Corresponding author email: enny.cruz@lshtm.ac.uk; Phone 44 02079588171

a London School of Hygiene and Tropical Medicine. Keppel St, Bloomsbury, London WC1E 7HT, United Kingdom

b Instituto de Saúde Coletiva. Rua Basílio da Gama, s/n.Canela. CEP 40110040. Salvador, Bahia, Brasil.

c Center of Data and Knowledge Integration for Health (CIDACS), Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, CEP 41745-715 Salvador-Bahia, Brazil

d Universidade de São Paulo.Av. Dr. Arnaldo, 715 Cerqueira Cesar, 01246904 - São Paulo, SP – Brasild
Abstract

Background: Dengue is the most common viral mosquito-borne disease, and women of reproductive age who live in or travel to endemic areas are at risk. Little is known about the effects of dengue during pregnancy on birth outcomes.

Methods: We conducted a population-based cohort study using routinely-collected Brazilian data from 2006-2012. We used birth registrations to identify pregnancies and linked to identify mothers of live births with dengue notified during pregnancy. Using multinomial logistic regression and Firth method, we estimated risk and odds ratios for preterm birth, low birth weight and small for gestational age associated with dengue. We also investigated the effect of time between the onset of the disease and the outcome.

Findings: We included 16,738,000 live births. Dengue haemorrhagic fever doubled the risk of preterm birth OR=2.4;95%CI 1.3-4.4 and low birth weight OR=2.1;95%CI 1.1-4.0 and did not have an effect on small for gestational age OR=2.1;95% CI 0.4-12.2). The magnitude of the effects was higher in the acute disease period.

Conclusion: This study showed an increased risk of adverse fetal outcomes in women with severe dengue during pregnancy. It appears that one contributing factor may be medical intervention to mitigate maternal risk with severe acute dengue.

Key words: Dengue, pregnancy, birth outcomes
Summary

In this paper, we analyzed the effects of symptomatic dengue during pregnancy on birth outcomes (preterm birth, low birth weight, small for gestational age) by dengue severity and time between the onset of disease and birth outcome. Our results have shown that dengue haemorrhagic fever in pregnant women doubled the risk of preterm birth and of low birth weight; however we did not find an affect in small for gestational age for either severe or mild disease. The main dengue effect on birth outcomes occurred during acute disease, 10 days between dengue onset and birth outcome and possibly one contributing factor to these adverse birth outcomes may be medical interventions triggered to mitigate mothers risk associated with dengue haemorrhagic fever. Therefore, in areas where dengue is circulating, the health of birth outcome should be a concern and this disease highlighted as public health priority.
Introduction

Dengue is the most common viral vector-borne disease worldwide, with an estimated 390 million people infected each year, 96 million of whom will develop clinical symptoms. It is endemic in over 100 countries (mostly in South and Central American and Southeast Asia), and is spreading to new areas; according to the World Health Organization (WHO), approximately half of the world’s population is at risk \(^1\). Since women of reproductive age who live in or travel to endemic areas are at risk of dengue, and since the effect of this disease during pregnancy is still unclear, it needs to be investigated. The majority of studies published to date are cases series\(^2\); the seven studies published before 2016 with comparison groups show conflicting results \(^3\) \(^4\) \(^5\) \(^6\) \(^7\) \(^8\) \(^9\) and a meta-analysis published in 2016 found dengue during pregnancy was only associated with low birth weight and preterm birth in symptomatic women\(^2\). In 2017, a large Brazilian study found that notification of mild dengue during pregnancy was not associated with higher rates of preterm birth or lower birth weight in neonates compared to rates in a random sample of all neonates. However, the study found an increase in the risk of preterm birth among women with confirmed dengue compared with reported maternal dengue cases with a dengue negative test\(^10\).

Maternal dengue, especially severe dengue, has been associated with fetal deaths\(^11\) but the effects on those who survive, especially when the mothers progress to severe disease, needs to be better investigated. We analysed a large, population-based, retrospective cohort to investigate the association between symptomatic dengue notified during pregnancy and adverse fetal outcomes, and to examine the effect of maternal dengue severity on outcomes.
Methods

We conducted a population-based retrospective cohort study by linking routine records of live births with records of women notified and confirmed with dengue disease, in Brazil from January 1, 2006 until December 31, 2012.

Data sources

We extracted routinely collected data from two Brazilian databases.

1) The Live Births Information System (Sistema de Informação sobre Nascimentos; SINASC), which records live births in Brazil; these data are derived from the birth registration, a legal document completed by the health worker who attended the birth. It includes information on the mother (name, place of residence, age, marital status, education); the pregnancy (length of gestation, mode of delivery); and the neonate (birth weight, presence of congenital anomalies)\(^1\). Data completeness is very high, with 97% of Brazilian births registered\(^1\), and most variables were >90% complete\(^1\).

2) Notifiable Diseases Information System (Sistema de Informação de Agravos de Notificação; SINAN), which records notifiable diseases. Dengue notifications information on the individual (name, place of residence, age, sex, and years of education), symptoms of the disease, laboratory tests, and disease severity, assessed clinically and laboratory according with the Brazilian dengue manual\(^15\). These data are reasonably complete on the variables used for linkage: <0.05% records were excluded because of a missing name. Around 5% of dengue cases did not have a final classification of severity. Laboratory confirmation was not required for notification or confirmation of dengue. During the study period (2006-2016), dengue was the main (and sometimes the only) vector-borne disease circulating in Brazil, as yellow fever and malaria occurred in restricted areas, and Zika and Chikungunya did not begin to circulate until 2014.
Procedures

SINASC defines a live birth as the product of conception that, independent of the duration of pregnancy, breathes or shows any other signs of life (such as a heartbeat, umbilical cord pulsation, or definite movement of voluntary muscles), after the separation from the mother's body \(^2\). Live birth records with a missing or implausible name of the mother (e.g., the name of a hospital as the name of the mother, twin or other multiple pregnancies, or congenital anomalies, or coming from municipalities without dengue notifications, and those with uncertain link status) were excluded (Figure 1).

**Figure 1. Flowchart with the linkage information Brazil, 2006-2012.**

- 20,332,313 live births
- Brazilian Information System of Births
- January 1, 2006 - December 31, 2012

- 1,892,835 births from municipalities without dengue notification;
  - 397,194 multiple pregnancies;
  - 118,653 congenital anomalies

- 20,332,313 live births eligible for linkage with dengue records

- 1,185,631 uncertain links

- Probing the linkage with 1,866,929 dengue notifications

- 16,738,000 live births included

- 17,673 live births of mothers notified with dengue (0.1%)

- 16,720,327 live births no linked with dengue notification
Outcomes

We assessed three birth outcomes: preterm birth (<37 weeks); low birth weight (<2500g) and small for gestational age), calculated according to the Intergrowth scale using the weight and gestational age in weeks at birth. This scale is sex specific, and those babies with a birth weight less than the 10th centile were considered small for gestational age. The data to calculate this outcome were available only for 2011 and 2012 (prior to this, gestational age was categorized).

Exposure

The exposure was being a confirmed case of dengue notified to SINAN during a pregnancy that resulted in a live birth. In Brazil, confirmation of dengue is based either on clinical and epidemiological criteria, namely presence of clinical symptoms of dengue in the same area and time as other confirmed cases of dengue, or on clinical and laboratory criteria, namely the presence of clinical symptoms and a positive test for one of: a) IgM detection by ELISA, b) viral RNA detection via PCR, c) NS1 viral antigen detection, or d) positive viral culture. Laboratory confirmed cases used in the analyses were referred to as “dengue during pregnancy, laboratory confirmed”. The notification does not necessarily state the woman notified was pregnant – there is a field for this variable, but it is largely missing. Rather, we considered a woman to be a case of dengue in pregnancy, and her newborn to be “exposed”, if the woman was notified to SINAN and confirmed as a dengue case and she linked to a live birth in the nine months after notification. We used the clinical classification valid in Brazil in the period in which data were collected, which classified dengue into three clinical categories: a) “dengue fever”, b) “complicated dengue”, and c) “dengue haemorrhagic fever/ dengue shock syndrome”. a) Dengue fever is a self-limiting disease, characterized as fever, with a severe headache, pain behind the
eyes, muscle and joint pain, and rash; b) Dengue haemorrhagic fever follows the WHO criteria and is characterized by fever, haemorrhagic evidence, thrombocytopenia and evidence of plasma leakage; c) Complicated dengue is a probable case of dengue characterized by one of the following: severe changes in the nervous system, cardiorespiratory dysfunction; insufficient hepatic function; gastrointestinal bleeding, cavity spills, or thrombocytopenia equal or less than 50,000/mm$^3$, or leucometry less than 1000/mm.$^{17}$ Complicated dengue is a Brazilian definition for severe cases of dengue that do not meet WHO criteria for dengue haemorrhagic fever and cannot be classified as mild self-limited disease due to their severity;

**Linkage process**

The data used in the linkage process were identified because we probabilistically linked records of women with confirmed dengue (SINAN) with records of live births (SINASC) up to nine months after the SINAN notification, to identify those mothers who had dengue during pregnancy. For this, we used name, age or date of birth, and the place of residence of the mother at the time of delivery and of notification. We excluded records with missing or implausible names (Figure 1). Match weight calculations used the Fellegi-Sunter method.$^{18}$

Linkage followed the same procedure as previously described for linking stillbirth records and the dengue database in Brazil, in which we demonstrated a sensitivity of 62% and a positive predictive value of 95%.$^{19}$ To estimate parameters for linkage weights and to validate the quality of the linkage, we created a gold-standard dataset. For the dengue-live birth linkage, we used all live birth records in Ceará and Espirito Santo States for 2010. Record pairs were manually reviewed to establish true matches (records belonging to the same mother) and to derive match weights. Records pairs were then ordered by match weight and manually inspected to identify any errors in linkage (false matches or
missed matches). Live birth records linked to confirmed dengue notifications with high certainty were classified as dengue in pregnancy cases, and birth records we were confident had no link were classified as non-dengue cases. Uncertain links (where we could not establish whether a woman had dengue) were excluded to avoid misclassification.

**Statistical analysis**

The data used in the analysis were de-identified. We estimated the crude and adjusted risk ratios from multinomial logistic regression, controlling for maternal age (categorized as <=20; between 20-35 and >=35 years old), education (less than 3 years; between 4-7 years and more than 8 years) and marital status (single/widow/divorced or married/stable union). For a sensitivity analysis of the validity of clinical/epidemiologic diagnosis, we repeated the analyses using laboratory confirmed dengue only. We investigated the effect of time between disease onset and live birth outcomes of preterm birth, low birth weight, and small for gestational age. In general, dengue is an acute disease with rapid recovery. The time between disease onset and the outcome was calculated using the date of the disease onset (information available in SINAN) and the date when the outcome occurred (date of live birth); we categorized this difference as being less than or equal to ten days or greater than 10 days. For the analyses by dengue severity (mild dengue, dengue with complications and dengue haemorrhagic fever), we calculated the odds ratio because we used the Firth method (to reduce the small sample bias in maximum likelihood estimation) as we were analysing rare events, and controlled for maternal age, education and marital status.20
Ethics Statements

Ethical approval was obtained from The Federal University of Bahia, Salvador, Brazil (CAAE: 26797814.7.0000.5030) and from The London School of Hygiene and Tropical Medicine (Ethics Ref:10269).

Results

The SINASC recorded 20,333,482 live births from 2006-2012. After exclusions, 16,738,000 live births were included in the study, and 17,673 (0.1%) of their mothers were linked with a dengue notification record (Figure 1).

The cohort characteristics by dengue status are described in Table 1. Compared to pregnant women without dengue, pregnant women with dengue were more likely to have more years of formal education and to have delivered by caesarean-section.

Table 1. Maternal characteristics, delivery details, and birth outcomes in relation to dengue status, Brazil, 2006-2012

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Notified with confirmed dengue in pregnancy</th>
<th>Without dengue notification</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n (%)</td>
<td>n (%)</td>
<td></td>
</tr>
<tr>
<td><strong>Age of the mother</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 20</td>
<td>4,499 (25.5)</td>
<td>4,384,159 (26.2)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>20-35</td>
<td>11,981 (67.8)</td>
<td>11,042,342 (66.0)</td>
<td></td>
</tr>
<tr>
<td>&gt;35</td>
<td>1,193 (6.7)</td>
<td>1,293,171 (7.8)</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>-</td>
<td>655 (0.0)</td>
<td></td>
</tr>
<tr>
<td><strong>Maternal education</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than 3 years</td>
<td>1,065 (6.1)</td>
<td>1,358,815 (8.3)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>4-7 years</td>
<td>4,661 (27.0)</td>
<td>4,592,979 (28.1)</td>
<td></td>
</tr>
<tr>
<td>More than 8 years</td>
<td>11,578 (66.9)</td>
<td>10,399,703 (63.6)</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>369 (2.1)</td>
<td>368,830 (2.2)</td>
<td></td>
</tr>
<tr>
<td><strong>Marital status</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single/Widow/Divorced</td>
<td>11,518 (66.1)</td>
<td>9,660,306 (58.7)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Married/Union</td>
<td>5,899 (33.9)</td>
<td>6,797,658 (41.3)</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>256 (1.4)</td>
<td>262,363 (1.6)</td>
<td></td>
</tr>
</tbody>
</table>
Maternal dengue slightly increased the risk of preterm birth from 7.3% to 7.9%, RR 1.1 (95% CI 1.0-1.2) and low birth weight from 7.2% to 8.4% RR1.2 (95% CI 1.1-1.2), although the confidence interval was borderline. There was no association between maternal dengue and being small for gestational age (Table 2). Restricting analyses to 33%, (5,755/17,388) of laboratory confirmed dengue cases did not change the magnitude of the associations.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Notified with confirmed dengue in pregnancy</th>
<th>Without dengue notification</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n (%)</td>
<td>n (%)</td>
<td></td>
</tr>
<tr>
<td>Numbers of pre-natal visits</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inadequate (less than 7 times)</td>
<td>6,991 (39.9)</td>
<td>6,817,679 (41.2)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Adequate (7 or more times)</td>
<td>10,515 (60.1)</td>
<td>9,712,338 (58.8)</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>167 (0.9)</td>
<td>190,310 (1.1)</td>
<td></td>
</tr>
<tr>
<td>Delivery</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vaginal</td>
<td>8,332 (47.2)</td>
<td>8,383,714 (50.2)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>C-section</td>
<td>9,320 (52.8)</td>
<td>8,307,072 (49.8)</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>21 (0.1)</td>
<td>29,541 (0.2)</td>
<td></td>
</tr>
<tr>
<td>Gestational age in the delivery</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than 22 weeks</td>
<td>17 (0.1)</td>
<td>8,844 (0.0)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>22-27 weeks</td>
<td>68 (0.4)</td>
<td>58,760 (0.4)</td>
<td></td>
</tr>
<tr>
<td>28-31 weeks</td>
<td>147 (0.8)</td>
<td>112,407 (0.7)</td>
<td></td>
</tr>
<tr>
<td>32-36 weeks</td>
<td>1,157 (6.7)</td>
<td>1,011,633 (6.1)</td>
<td></td>
</tr>
<tr>
<td>More than 37 weeks</td>
<td>15,994 (92.0)</td>
<td>15,215,440 (92.8)</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>290 (1.6)</td>
<td>313,243 (1.9)</td>
<td></td>
</tr>
<tr>
<td>Birth weight</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;=3,000</td>
<td>12,171 (69.1)</td>
<td>11,659,487 (69.9)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>3,000-2,500</td>
<td>3,973(22.5)</td>
<td>3,816,201 (22.9)</td>
<td></td>
</tr>
<tr>
<td>1,500-2,500</td>
<td>1,258 (7.1)</td>
<td>1,026,248 (6.1)</td>
<td></td>
</tr>
<tr>
<td>&lt;1,500</td>
<td>226 (1.3)</td>
<td>178,643 (1.1)</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>38 (0.2)</td>
<td>39,748 (0.2)</td>
<td></td>
</tr>
<tr>
<td>Small for gestational age (10th centile)#</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>3,681 (91.7)</td>
<td>3,289,125 (92.0)</td>
<td>0.608</td>
</tr>
<tr>
<td>Small</td>
<td>331 (8.2)</td>
<td>287,173 (8.0)</td>
<td></td>
</tr>
<tr>
<td>Data not collected</td>
<td>13,661 (77.3)</td>
<td>13,144,029 (78.6)</td>
<td></td>
</tr>
</tbody>
</table>

*% of each category without the missing value
\# data available only in 2011 and 2012
Table 2. Number of cases notified with dengue during pregnancy (N), risk ratio crude and adjusted for the association among dengue during pregnancy and preterm birth, low birth weight, small for gestational age. Brazil, 2006-2012.

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>All dengue cases</th>
<th>Lab confirmed dengue cases</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Crude Risk Ratio</td>
<td>Adjusted Risk Ratio</td>
</tr>
<tr>
<td></td>
<td>(95% CI)</td>
<td>(95% CI)</td>
</tr>
<tr>
<td>Preterm birth</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N=1,193,033</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N=1,389 (0.12%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall (&lt;37 weeks)</td>
<td>1.1 (1.0-1.2)</td>
<td>1.1 (1.0-1.2)</td>
</tr>
<tr>
<td>Less than 28 weeks</td>
<td>1.1 (1.0-1.1)</td>
<td>1.1 (1.0-1.2)</td>
</tr>
<tr>
<td>28-32 weeks</td>
<td>1.2 (1.1-1.5)</td>
<td>1.2 (1.1-1.4)</td>
</tr>
<tr>
<td>32-36 weeks</td>
<td>1.2 (1.0-1.5)</td>
<td>1.1 (0.9-1.4)</td>
</tr>
<tr>
<td>Low birth weight</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N=1,206,375</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N=1484 (0.12%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall (&lt;2500g)</td>
<td>1.2 (1.1-1.2)</td>
<td>1.2 (1.1-1.2)</td>
</tr>
<tr>
<td>1500g-2499g</td>
<td>1.2 (1.1-1.2)</td>
<td>1.2 (1.1-1.2)</td>
</tr>
<tr>
<td>Less than 1500g</td>
<td>1.2 (1.1-1.4)</td>
<td>1.2 (1.0-1.3)</td>
</tr>
<tr>
<td>Small for gestational age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N=287,504</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N=331 (0.12%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall (&lt;10th centile)</td>
<td>1.0 (0.9-1.1)</td>
<td>1.0 (0.9-1.1)</td>
</tr>
</tbody>
</table>

Table 3 shows the dose-response effect of the severity of dengue on preterm birth, low birth weight and small for gestational age. Dengue haemorrhagic fever in pregnancy doubled the odds of preterm birth from 1,191,644/16,407,084 (7.3%) to 12/79 (15.2%), OR 2.4 95% CI 1.3-4.4) and of low birth weight from 1,204,891/16,680,579 (7.2%) to 11/79 (13.9%), OR 2.1 95% IC 1.1-4.0), but we did not see a significant effect of dengue on small for gestational age even among pregnant women who developed haemorrhagic disease (from 287,173/3,576,298 (8.0%) to 1/9 (11.1%), OR 2.5 95% CI 0.4-12.2).
Table 3. Number of cases with dengue during pregnancy (N) and Odds Ratio for the association Dengue during pregnancy by severity of disease and adverse birth outcomes (preterm birth, low birth weight, small for gestational). Brazil, 2006-2012.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Mild dengue Odds Ratio (95% Confidence Interval)</th>
<th>Complicated dengue Odds Ratio (95% Confidence Interval)</th>
<th>Haemorrhagic fever Odds Ratio (95% Confidence Interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preterm birth (&lt;37 weeks)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N=1,192,920</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N (%)</td>
<td>1.234 (0.103%)</td>
<td>30 (0.002%)</td>
<td>12 (0.001%)</td>
</tr>
<tr>
<td>Crude</td>
<td>1.1 (1.0-1.1)</td>
<td>1.4 (0.9-1.9)</td>
<td>2.4 (1.3-4.3)</td>
</tr>
<tr>
<td>Adjusted</td>
<td>1.1 (1.0-1.1)</td>
<td>1.4 (0.9-2.0)</td>
<td>2.4 (1.3-4.4)</td>
</tr>
<tr>
<td>Low birth weight (&lt;2500g)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N=1,206,265</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>1,327 (0.110%)</td>
<td>36 (0.003%)</td>
<td>11 (0.001%)</td>
</tr>
<tr>
<td>Crude</td>
<td>1.1 (1.1-1.2)</td>
<td>1.6 (1.2-2.3)</td>
<td>2.1 (1.1-4.0)</td>
</tr>
<tr>
<td>Adjusted</td>
<td>1.1 (1.1-1.2)</td>
<td>1.6 (1.1-2.3)</td>
<td>2.1 (1.1-4.0)</td>
</tr>
<tr>
<td>Small for gestational age (10th centile)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N=287,474</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>294 (0.102%)</td>
<td>6 (0.002%)</td>
<td>1 (0.000%)</td>
</tr>
<tr>
<td>Crude</td>
<td>1.0 (0.9-1.1)</td>
<td>2.2 (0.9-5.1)</td>
<td>2.0 (0.3-11.4)</td>
</tr>
<tr>
<td>Adjusted</td>
<td>1.0 (0.9-1.1)</td>
<td>2.3 (1.0-5.3)</td>
<td>2.1 (0.4-12.2)</td>
</tr>
</tbody>
</table>

Adjusted for maternal age, education, marital status
*Because of the small numbers, we included extra decimal places.

The risk of dengue in pregnancy on fetal outcomes depended on the time between dengue onset and the date of live birth: the magnitude of the effect of dengue on adverse birth outcomes was higher during the acute disease period, with some residual effects remaining after the first 10 days for preterm birth and low birth weight (Table 4).
### Discussion

Cases of dengue haemorrhagic fever in pregnancy doubled the risk of preterm birth (OR: 2.4) and of low birth weight (OR: 2.1); by contrast the effect of mild dengue fever only raised the risk by 10 to 20 percent for preterm birth and low birth weight respectively. There was no increase in the risk of small for gestational age for either severe or mild disease. The main dengue effect on birth outcomes occurred during acute disease, within the first ten days of disease onset.

Our results are consistent those with Nascimento, et al.\textsuperscript{10} who showed that symptomatic dengue infection did not greatly affect the risk of low birthweight (1.00 95% CI 0.85-1.17) or preterm birth (OR: 0.98 95% CI 0.83-1.16) when compared to a random sample of live births. However, dengue has been associated with an increased risk of

---

**Table 4. Number of cases with dengue during pregnancy (N) and risk Ratio for the association Dengue during pregnancy by timing of disease and adverse birth outcomes (preterm birth, low birth weight, small for gestational). Brazil, 2006-2012.**

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Risk Ratio and frequency of outcomes within 10 days of disease onset (95% Confidence Interval)</th>
<th>Risk Ratio and frequency of outcomes after 10 days from disease onset (95% Confidence Interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Preterm birth</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N=1,191,719</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>75 (0.006%)</td>
<td>1,314 (0.110%)</td>
</tr>
<tr>
<td>Crude</td>
<td>2.1 (1.7-2.7)</td>
<td>1.1 (1.0-1.1)</td>
</tr>
<tr>
<td>Adjusted</td>
<td>2.0 (1.6-2.6)</td>
<td>1.1 (1.0-1.1)</td>
</tr>
<tr>
<td><strong>Low birth weight</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N= 1,206,375</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>74 (0.006%)</td>
<td>1410 (0.117%)</td>
</tr>
<tr>
<td>Crude</td>
<td>2.1 (1.6-2.6)</td>
<td>1.1 (1.1-1.2)</td>
</tr>
<tr>
<td>Adjusted</td>
<td>2.0 (1.6-2.6)</td>
<td>1.1 (1.1-1.2)</td>
</tr>
<tr>
<td><strong>Small for gestational age (10th centile)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N=287,173</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>4 (0.001%)</td>
<td>83 (0.029%)</td>
</tr>
<tr>
<td>Crude</td>
<td>0.5 (0.2-1.5)</td>
<td>1.0 (0.9-1.1)</td>
</tr>
<tr>
<td>Adjusted</td>
<td>0.5 (0.2-1.5)</td>
<td>1.0 (0.9-1.1)</td>
</tr>
</tbody>
</table>

Adjusted for maternal age, education, marital status.  
*Because of the small numbers, we included extra decimal places.  

Discussion

Cases of dengue haemorrhagic fever in pregnancy doubled the risk of preterm birth (OR: 2.4) and of low birth weight (OR: 2.1); by contrast the effect of mild dengue fever only raised the risk by 10 to 20 percent for preterm birth and low birth weight respectively. There was no increase in the risk of small for gestational age for either severe or mild disease. The main dengue effect on birth outcomes occurred during acute disease, within the first ten days of disease onset.

Our results are consistent those with Nascimento, et al.\textsuperscript{10} who showed that symptomatic dengue infection did not greatly affect the risk of low birthweight (1.00 95% CI 0.85-1.17) or preterm birth (OR: 0.98 95% CI 0.83-1.16) when compared to a random sample of live births. However, dengue has been associated with an increased risk of
stillbirth, and by excluding fetal deaths in our analyses, we may be underestimating the risk by only considering birth outcomes in those who survived. Our findings are also consistent with a meta-analyses published in 2016, which found an association between symptomatic dengue and preterm birth and low birth weight within in a hospital based population that sought care, and was likely to include only severe dengue cases. In our study, the association between maternal dengue and fetal outcomes was strongest for women with dengue haemorrhagic fever in pregnancy.

There is evidence that different infectious diseases in pregnancy can lead to adverse outcomes. The evidence is still fragmented, and there is no consistent effect, which appears to vary according to the pathogen involved, and the timing and severity of maternal disease. For instance, studies comparing pregnant women with infection to pregnant women without, found a) severe influenza increased the risk of preterm birth from 2.4 to 4 times, whereas studies based on mild range of illness did not find an effect; b) measles and HIV infection increased the risk of low birth weight by 3.5 and 1.6 times, respectively.

Our data suggest that the main way gestational dengue affects preterm birth and low birth weight is through maternal illness, rather than through a direct effect on the fetus. Since we did not observe an effect on small for gestational age, the effect of dengue in pregnancy on birth weight probably occurred via prematurity of the newborn. Therefore, much of the effect may be explained by early delivery due to medical interventions, such as caesarean section, required because of concern about risk to the mother. The risk of prematurity was higher during the first 10 days after disease onset and among those with haemorrhagic disease. However, data on small for gestational age was only available for 2011-2012, which may have affected our ability to observe an effect.
The potential study limitations are associated with the linkage process and use of secondary data. In relation to the use of secondary data, outcome and exposure identification is susceptible to misclassification, a limitation inherent to this study design. Estimated national prevalence of preterm birth and small for gestational age in low and middle income countries using the available routine collected data are underestimated, however this error probably affected either, those exposed and no exposed to dengue therefore unlikely to bias the results of this study.\textsuperscript{24, 25} Regarding the dengue diagnosis, it is usual practice in outbreaks/ epidemics to test until the origin of the outbreak is clearly established, and after that only test when there is a clinical indication. The proportion of laboratory confirmed cases among general population in Brazil is around 30%,\textsuperscript{26} and therefore misclassification based on the lack of laboratory confirmation is possible but we showed very good agreement with the estimates of risk in subsets of the data with and without laboratory confirmation of dengue. The linkage was rigorously validated.\textsuperscript{19} The results of this showed that despite the low sensitivity (that kept us from making statements about the prevalence of dengue during pregnancy) it is unlikely that the linkage process introduced bias, since missed and false matches occurred randomly. Although we adjusted for confounders, other possible unknown confounders, such as maternal co-morbidities or quality or type of obstetric care, may have contributed to the association between severe dengue and adverse fetal outcomes.

In summary, this study shows a more than doubling in the risk of preterm birth and low birth weight in women with dengue haemorrhagic fever in pregnancy, and increases of 10-20% in women with mild dengue. One contributing factor to adverse fetal outcomes may be medical interventions triggered to mitigate mothers risk associated with dengue haemorrhagic fever. We recommend further research in different settings to
confirm our results, and additional studies of adverse birth outcomes for other vector-borne diseases.

References


Chapter 8. Description of the association between symptomatic dengue during pregnancy and maternal deaths

Cover sheet
Paper 6 title: Dengue during pregnancy and maternal mortality: a cohort analysis using routine data

Authors: Enny S. Paixao, Oona M R Campbell, Maria Gloria Teixeira, Maria da Conceicao N Costa, Katie Harron, Mauricio L. Barreto, Laura C. Rodrigues

Author Contributions Statement

ESP and KH carried out the analysis. ESP wrote the first draft of the article. LCR, MGT conceived the study. MCNC, MLB, OC contributed to the study design and interpretation. All authors revised the manuscript and approved the final version.
Dengue in pregnancy and maternal mortality: a cohort analysis using routine data

Enny S Paixão*, MSc\textsuperscript{1,2,3}
Katie Harron, PhD\textsuperscript{1}
Oona Campbell, PhD\textsuperscript{1}
Maria Glória Teixeira, PhD\textsuperscript{2,3}
Maria da Conceição N. Costa, PhD\textsuperscript{2}
Mauricio L. BarretoPhD\textsuperscript{2,3}
Laura C. Rodrigues, PhD\textsuperscript{1,3}

* Corresponding author email: enny.cruz@lshtm.ac.uk; Phone 44 7756636748

1 London School of Hygiene and Tropical Medicine. Keppel St, Bloomsbury, London WC1E 7HT, United Kingdom
2 Instituto de Saúde Coletiva. Rua Basílio da Gama, s/n.Canela. CEP 40110040. Salvador, Bahia, Brasil.
3 Center of Data and Knowledge Integration for Health (CIDACS), Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, CEP 41745-715 Salvador-Bahia, Brazil
Abstract

Dengue is a mosquito-borne disease with major public health importance due to its growing incidence and geographical spread. There is a lack of knowledge on its contribution to maternal death. We conducted a population-based cohort to investigate the association between symptomatic dengue during pregnancy and deaths in Brazil from 2007 to 2012. We did this by linking routine records of confirmed dengue cases to records of deaths of women who had a live birth. Using the Firth method, we estimated odds ratios for maternal deaths associated with dengue during pregnancy. Dengue increased the risk of maternal death by 3 times (95% CI, 1.5-5.8) and dengue haemorrhagic fever increased the risk of maternal death by 450 times (95% CI, 186.9-1088.4) when compared to mortality of pregnant women without dengue. The increase in risk occurred mostly during acute dengue 71.5 (95% CI, 32.8-155.8), compared with no dengue cases. This study showed an increased risk of adverse outcomes in pregnant women with dengue. Therefore in areas where dengue is circulating, the health of pregnant women should be not only a public health priority, but health professionals attending pregnant women with dengue should more closely observe these patients to be able to intervene in a timely way and avoid deaths.
**Introduction**

Dengue is a mosquito borne disease with a major importance in the public health arena due to a growing incidence (30-fold rise in the past 50 years) and an expanding geographical range (endemic in more than 100 countries mostly in South American and Southeast Asia and still spreading to new areas, including Europe. According to the World Health Organization (WHO), approximately half of the world’s population is at risk. However the burden of dengue during pregnancy on maternal ill-health is not well understood. Physiological changes that occur during pregnancy (such as hemodilution) can mask the thrombocytopenia, leucopenia, or hemoconcentration associated with dengue, and common obstetric problems can cause haematological and hepatic issues masking the disease. These may make it difficult to differentiate dengue haemorrhagic fever from common obstetric conditions, leading to misdiagnosis.

Dengue during pregnancy has been associated with poor fetal and maternal outcomes. There is some evidence that the risk of severe dengue and of hospitalization due to dengue is higher among pregnant compared with non-pregnant women and a number of cases series have reported maternal deaths associated with dengue, and other complications such as bleeding and increased caesarean section rates; two small cohorts comparing pregnant women exposed and unexposed to dengue in Brazil and Colombia found more maternal deaths among the dengue exposed group.

In this study we analysed a large population-based retrospective cohort to investigate these issues in greater detail and with greater power to explore the the association between symptomatic dengue during pregnancy and maternal mortality.
Methods

We conducted a population-based retrospective cohort study by linking routine records of all women notified and confirmed as having dengue, with records of maternal deaths following delivery of a live birth, in Brazil from January 1, 2007 to December 31, 2012.

Data sources

We extracted routinely collected data from three Brazilian databases.

1) The Live Births Information System (Sistema de Informação sobre Nascimentos; SINASC), which contains records of all live births in Brazil; these data come from birth registration, a legal document completed by the health provider who assisted the delivery. It includes information on the women who gave birth such as name, place of residence, age, marital status, education; the pregnancy (length of gestation, type of delivery); and the newborn (weight at birth, the presence of birth anomalies and gestational age at birth). The data completeness and coverage are very high, with more than 90% completeness for most variables and capturing 97% of Brazilian registered births.

2) Mortality Information System (Sistema de Informação sobre Mortalidade; SIM), which contains records of all deaths in Brazil, including fetal deaths; these data come from the Death Certificate, a required legal document. We retained all deaths of women coded under obstetric causes of death by ICD-10, the "O" group. The proportion of records with missing data varied by variable (e.g. maternal education was missing in 23% of records). Although stillbirth was not an outcome of this study, we retained all fetal deaths to link with maternal deaths as described below.
3) Notifiable Diseases Information System (Sistema de Informação de Agravos de Notificação; SINAN), which contains records on notifiable diseases. The dengue notification system includes information on the individual such as name, place of residence, age, sex, and years of education, on the disease, such as symptoms, laboratory tests, and disease severity. For linkage, <0.05% records were excluded because of missing name, and around 5% of dengue cases did not have a final classification of severity. Laboratory confirmation is not required to confirm dengue in Brazil. According to the Ministry of Health, around 30% of dengue cases were laboratory confirmed at the time. During this period, dengue was the main (and sometimes the only) vector-borne disease circulating in Brazil, as yellow fever and malaria occurred in restricted areas, and Zika and Chikungunya were not circulating until 2014.

**Procedures**

Maternal death records that linked with a stillbirth, a pregnancy with an abortive outcome, or that failed to link were excluded from analysis. Although we linked maternal deaths to livebirths and stillbirths, we only used the maternal deaths linked with live births in this study, because the comparison group was live births that occurred in Brazil during the study period.

Ethical approval was obtained from The Federal University of Bahia, Salvador, Brazil (CAAE: 26797814.7.0000.5030) and from The London School of Hygiene and Tropical Medicine (Ethics Ref:10269).

**Outcomes**

We included female deaths with an obstetric code in ICD-10 recorded as the cause of death. The comparison group in our study was women who had a live birth in the same
period in Brazil who were not excluded during the linkage process described below (Fig 1).

Figure 1. Flowchart with the linkage information Brazil, 2007-2012.

Exposure

The exposure under study was being a confirmed cases of dengue notified during a pregnancy that resulted in a live birth. In Brazil, dengue confirmation can be based on clinical/epidemiological criteria, namely presence of clinical symptoms of dengue in the same area and time as other confirmed cases of dengue, or on clinical/laboratory criteria, namely the presence of clinical symptoms and a positive test from one of IgM detection by ELISA, viral RNA detection via PCR, NS1 viral antigen detection, or positive viral culture. We considered maternal deaths to be “exposed” if the woman’s records linked...
with a live birth that was previously linked to a confirmed dengue case. We defined “dengue during pregnancy” as all confirmed cases of dengue (clinical/epidemiological or clinical/laboratorial). Laboratory confirmed cases were referred to as “dengue during pregnancy, laboratory confirmed”. We used the same three clinical categories of classification as used by the Brazilian Ministry of Health at the time of the study: “dengue fever” (a self-limiting fever, with a severe headache, pain behind the eyes, muscle and joint pain, and rash), “complicated dengue”, and “dengue haemorrhagic fever/ dengue shock syndrome”. Complicated dengue is a Brazilian definition of severe cases of dengue that do not meet the WHO criteria for dengue haemorrhagic fever (i.e. fever, haemorrhagic evidence, thrombocytopenia and evidence of plasma leakage) but that cannot be classified as mild self-limited disease due to their severity. Complicated dengue is used when a probable case of dengue presents with one of the following: severe changes in the nervous system, cardiorespiratory dysfunction, insufficient hepatic function, gastrointestinal bleeding, cavity spills, or thrombocytopenia equal or less than 50,000/mm$^3$ leucometry less than 1000/mm$^3$.

**Linkage process**

We conducted the linkage in two steps. First, we probabilistically linked records of dengue notifications (SINAN) with records of live births (SINASC) and stillbirths (SIM), to identify those women who had dengue during pregnancy and who gave birth. For simplicity, we will refer to these as mothers even if the woman had a stillbirth. To link, we used the name of the mother, two sources of age, and place of residence of the mother at the time of delivery or notification. We excluded records with missing names (Fig 1). Match weight calculations were based on the Fellegi-Sunter method$^{16}$. 
We then linked this composite births file (live births/stillbirth linked with maternal dengue status) with maternal death records using name, age, place of residence of the mother at time of delivery or death, and the time between the birth of the new-born and death of the mother. The procedures and evaluation of the matching process are the subject of a separate paper. However, in brief, we expected that after excluding pregnancies with abortive outcomes (10% of maternal mortality notifications), the large majority of the maternal deaths would have linked with a live birth or stillbirth. Although we did not use the maternal deaths linked to stillbirth for the analyses of this paper, it is important to include this stage in the linkage process section to present our measure of error in a comprehensive way. Women coded as having an abortive outcome that linked to a live birth record were assumed to be a false match. Maternal deaths linking to a live birth and a stillbirth simultaneously were also classified as false matches, unless they were multiple births.

We linked 6593 maternal deaths to the composite births file, of which 65 were identified as false matches. This gave a positive predictive value (PPV) of 6528/6593=99%. Of 9213 maternal deaths without an abortive outcome, we were unable to link 2,675, giving a sensitivity of 6,528/9213=71%. Mothers with more than 7 years of education and self-declared as Caucasian were more likely to link with the composite live birth stillbirth file.

We further evaluated potential linkage error using dengue information obtained from the first linkage (between dengue and live births). We compared the maternal deaths classified as having dengue as a cause of death that linked with a live birth or stillbirth, with maternal deaths where dengue was coded as an underlying case of death-ICD-10 but that were not linked. There was no difference in socio-demographic characteristics between these two groups, suggesting that although we did not capture all of the matches...
in our linkage, there was no selection bias associated with the exposure in our linked cohort.

**Statistical analysis**

Using Chi square test, we compared the characteristics of women by dengue status. We estimated the crude and adjusted odds ratio using the Firth method \(^{18}\) (to reduce the small sample bias in maximum likelihood estimation) as we were analysing rare events, controlling for age, education and mode of delivery. For a sensitivity analysis of the validity of clinical/epidemiologic diagnosis, we repeated the analyses using laboratory confirmed dengue only. We investigated the effect of dengue severity (mild dengue, dengue with complications and dengue haemorrhagic fever) and time between disease onset and maternal deaths, since in general, dengue is an acute disease with rapid recovery. The time between disease onset and the outcome was calculated using the date of the disease onset (information available in SINAN) and the date when the outcome occurred (date of maternal death); we categorized this difference as being less than or equal to ten days or greater than 10 days.

**Data availability**

The data that support the findings of this study are available from Brazilian Ministry of Health but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the Brazilian Ministry of Health upon reasonable request.

**Results**

The Brazilian Information System recorded 10,259 maternal deaths from 2007-2012. After exclusions, 4,053 maternal deaths in women with live births were included in our study population, and 12 (0.3%) had a positive dengue status (Figure 1). The live
birth system recorded 17,391,826 live births from 2007-2012. After exclusions, 14,440,229 women with live births were included in the study, and 16,224 (0.1%) had a dengue notification record.

The cohort characteristics by dengue status are described in Table 1. Compared with women without dengue, women with dengue were more likely to have more years of formal education, and to have a caesarean section as a mode of delivery.

Table 1. Maternal characteristics and delivery details in relation to dengue status, Brazil, 2006-2012

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Notified with confirmed dengue in pregnancy</th>
<th>Without dengue notification</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n (%)</td>
<td>n (%)</td>
<td></td>
</tr>
<tr>
<td>Age of the mother</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 20</td>
<td>4,127 (25.4)</td>
<td>3,738,246 (25.9)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>20-35</td>
<td>11,011 (67.8)</td>
<td>9,560,187 (66.3)</td>
<td></td>
</tr>
<tr>
<td>&gt;35</td>
<td>1,097 (6.8)</td>
<td>1,129,318 (7.8)</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>3 (0.0)</td>
<td>1,488 (0.0)</td>
<td></td>
</tr>
<tr>
<td>Maternal education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than 3 years</td>
<td>946 (5.9)</td>
<td>1,099,923 (7.8)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>4-7 years</td>
<td>4,239 (26.7)</td>
<td>3,867,130 (27.4)</td>
<td></td>
</tr>
<tr>
<td>More than 8 years</td>
<td>10,710 (67.4)</td>
<td>9,153,016 (64.8)</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>343 (2.1)</td>
<td>309,170 (2.1)</td>
<td></td>
</tr>
<tr>
<td>Delivery</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vaginal</td>
<td>7,599 (46.9)</td>
<td>7,117,183 (49.4)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>C-section</td>
<td>8,618 (53.1)</td>
<td>7,286,995 (50.6)</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>21 (0.1)</td>
<td>25,061 (0.2)</td>
<td></td>
</tr>
<tr>
<td>Maternal Mortality*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>12 (0.1)</td>
<td>4,041 (0.03)</td>
<td>0.00</td>
</tr>
<tr>
<td>No</td>
<td>16,224 (99.9)</td>
<td>14,424,005 (99.97)</td>
<td></td>
</tr>
</tbody>
</table>

*% of each category without the missing value

Dengue during pregnancy tripled the risk of maternal death from 0.1% (16,224/14,440,229) to 0.3% (12/4,053), OR 3.0 (95% IC 1.3-5.8) (Table 2). The
The proportion of dengue cases that were laboratory confirmed was much higher among maternal deaths (9/12 = 75%) than among mothers all of live births (5309/16224 = 33%). Restricting analyses to laboratory-confirmed cases showed that dengue increased the risk of maternal death almost eight fold (7.8 95% CI 3.1-15.4).

Table 2: Number of cases of dengue during pregnancy and Odds ratio crude and adjusted for the association between dengue during pregnancy and maternal deaths. Brazil, 2007-2012.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>All dengue cases</th>
<th>Laboratory confirmed dengue cases</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Crude Risk Ratio</td>
<td>Adjusted Risk Ratio</td>
</tr>
<tr>
<td></td>
<td>(95% Confidence Interval)</td>
<td>(95% Confidence Interval)</td>
</tr>
<tr>
<td>Maternal Mortality§</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frequency</td>
<td>12</td>
<td>9</td>
</tr>
<tr>
<td>Overall (OR)</td>
<td>2.7 (1.6-4.8)</td>
<td>3.0 (1.5-5.8)</td>
</tr>
</tbody>
</table>

Adjusted for age, education and mode of delivery

Complicated dengue increased the odds of maternal deaths by 27 fold (95% CI 5.5-136.3), compared with no dengue. However, this number was much higher among the women that developed dengue haemorrhagic fever, for whom the increase in risk of maternal death was 451 fold (95% CI 186-1088).

Table 3: Number of cases of dengue during pregnancy (N) and Odds Ratio crude and adjusted for the association Dengue during pregnancy by severity of disease and maternal deaths. Brazil, 2007-2012.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Mild dengue Odds Ratio</th>
<th>Complicated dengue Odds Ratio</th>
<th>Haemorrhagic fever Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(95% Confidence Interval)</td>
<td>(95% Confidence Interval)</td>
<td>(95% Confidence Interval)</td>
</tr>
<tr>
<td>Maternal Mortality§</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frequency</td>
<td>3</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Crude</td>
<td>0.8 (0.3-2.3)</td>
<td>28.7 (8.2-99.7)</td>
<td>274.5 (115.3-653.8)</td>
</tr>
<tr>
<td>Adjusted</td>
<td>0.95 (0.3-3.3)</td>
<td>27.3 (5.5-136.3)</td>
<td>451 (186.9-1088.4)</td>
</tr>
</tbody>
</table>

Adjusted for age, education and mode of delivery
The risk of maternal death among pregnant women with dengue depended on the time between first symptoms of dengue and the date of death, the deaths occurred mainly during acute disease, 10 days between the disease onset and the date of death (Table 4).

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Odds Ratio and frequency of outcomes within 10 days of disease onset (95% Confidence Interval)</th>
<th>Odds Ratio and frequency of outcomes after 10 days from disease onset (95% Confidence Interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal Mortality³</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frequency</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Crude</td>
<td>62 (31.6-122.6)</td>
<td>1.0 (0.4-2.6)</td>
</tr>
<tr>
<td>Adjusted</td>
<td>71.5 (32.8-155.8)</td>
<td>0.9 (0.3-3.1)</td>
</tr>
</tbody>
</table>

Pre-eclampsia or eclampsia was the registered cause of death in 25% of the patients with dengue; in the group without dengue, but who had a live birth, these causes were responsible for 19% of all deaths (appendix 1).

In this retrospective cohort, we found that pregnant women with symptomatic dengue had a significantly higher risk of maternal death compared to pregnant women without dengue, and this risk was considerably higher when dengue was severe, complicated, or dengue haemorrhagic fever. The proportion of laboratory confirmed cases among women who died was much higher than among the comparison group, and the effect of dengue on death occurred primarily during the acute disease, in the first ten days after disease onset.

The results of this study are consistent with the literature that shows a higher maternal mortality ratio associated with dengue, mainly among those who developed severe dengue. In a Brazilian cohort in Rio Branco the maternal mortality ratio in the dengue exposed group was 13 times the mean maternal mortality ratio of the area ⁸ as
opposed to our finding of 3 times higher. The percentage of maternal deaths among pregnant women with dengue in the case series studies varied from 6.6% in Sri Lanka to 21.7% in the South Sudan, whereas we see 0.07%. This may be because we included less severe disease or excluded women with stillbirth or other abortive outcomes and because of our low sensitivity (71%) to detect dengue. Among women with severe disease and a live birth, the rate was 0.12%. The mechanism for the association between dengue and maternal deaths is not clear: one possible explanation is that the clinical aspects of the disease may be different during pregnancy in a way that increase the susceptibility to dengue haemorrhagic fever. In Brazil pregnant women with dengue were 3 times more likely to develop severe dengue than non-pregnant women. Another hypothesis is that physiological changes occurred during pregnancy such as hemoconcentration and the difficulty in distinguishing between severe dengue and common obstetric conditions may lead to misdiagnosis and delay the disease treatment that can progress to hypovolemic shock and death. A meta-analysis of the effect of infection on pregnancy showed that viral infections can increase the risk of pre-eclampsia, although this study does not mention dengue. It is possible that dengue virus leads to the same etiologic pathway of inflammatory modifications of placental tissues that . Pre-eclampsia was the cause of death in 25% of the patients with dengue, compared with 19% in the comparison group.

There is growing evidence that different infectious diseases during pregnancy can be associated with adverse outcomes. The evidence is still fragmented and the effect of different infections appear to vary according with the pathogen involved and the severity of maternal disease. According to a meta-analysis, pregnant women with active tuberculosis tended to be more likely to suffer a maternal death although this was not significant (OR:4, 95% CI 0.65-25.2), and pregnant women with measles had a 9-fold increased risk of maternal death.
The potential study limitations are associated with the linkage\textsuperscript{17}, although we undertook a rigorous validation process that showed that despite the low sensitivity, that kept us from making statements about the prevalence of dengue during pregnancy, it is unlikely that the linkage process introduced bias in the measure of association since missed and false matches occurred randomly in the population. The maternal death linkage analysis comparing dengue cases that the algorithm was capable of capturing, with those that it could not, but recorded maternal death as cause of death, showed no difference between the two groups, however this result should be interpreted with caution because it had limited power, due to the sample size. There are some limitations associated with the use of secondary data: we only had a limited number of possible confounders to analyze. Although we adjusted for these, unknown confounders, such as maternal co-morbidities or quality of care, may have contributed to the association between severe dengue and maternal deaths.

In summary, this study suggests a marked increase in the risk of maternal deaths in women with dengue during pregnancy. The health of pregnant women is a public health priority, but in places where dengue is circulating and the health professionals attending pregnant women with dengue should observe them more closely to be able to intervene in a timely way, and avoid death. We recommend further research in different settings to confirm our results and studies of negative fetal and maternal outcomes in other vector-borne diseases.
References


Funding

EP funded by National Council for Scientific and Technological Development (CNPq-Brazil); LCR is partially funded by the European Union’s Horizon 2020 research and innovation program under Zika- PLAN grant agreement No. 734584; KH is funded by the Wellcome Trust (grant number 103975/Z/14/Z) However the funder of this study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Conflict of interests

The authors declare that they have no competing interests.

Appendix 1: Table characteristics women that died with dengue

<table>
<thead>
<tr>
<th>Age</th>
<th>Education</th>
<th>Mode of delivery</th>
<th>Cause of death</th>
<th>Lab</th>
<th>severity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>29</td>
<td>More than 12 years of formal education</td>
<td>C-section</td>
<td>Other viral diseases complicating pregnancy, childbirth and the puerperium (dengue registered as underline cause)</td>
<td>Yes</td>
</tr>
<tr>
<td>2</td>
<td>34</td>
<td>4-7 years of formal education</td>
<td>C-section</td>
<td>Other viral diseases complicating pregnancy, childbirth and the puerperium (dengue registered as underline cause)</td>
<td>Yes</td>
</tr>
<tr>
<td>3</td>
<td>17</td>
<td>8-11 years of formal education</td>
<td>C-section</td>
<td>Cardiac complications of anaesthesia during the puerperium</td>
<td>No</td>
</tr>
<tr>
<td>4</td>
<td>19</td>
<td>8-11 years of formal education</td>
<td>C-section</td>
<td>Other viral diseases complicating pregnancy, childbirth and the puerperium</td>
<td>Yes</td>
</tr>
<tr>
<td>5</td>
<td>30</td>
<td>8-11 years of formal education</td>
<td>C-section</td>
<td>Other viral diseases complicating pregnancy, childbirth and the puerperium</td>
<td>Yes</td>
</tr>
<tr>
<td>6</td>
<td>18</td>
<td>8-11 years of formal education</td>
<td>Vaginal</td>
<td>Eclampsia complicating the puerperium</td>
<td>Yes</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>C-section</td>
<td>Diseases of the respiratory system complicating pregnancy, childbirth and the puerperium</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>C-section</td>
<td>Complication of the puerperium, unspecified</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>25</td>
<td>8-11 years of formal education</td>
<td>C-section</td>
<td>Other viral diseases complicating pregnancy, childbirth and the puerperium</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>10</td>
<td>16</td>
<td>8-11 years of formal education</td>
<td>C-section</td>
<td>Other specified diseases and conditions complicating pregnancy, childbirth and the puerperium (dengue registered as underline cause)</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Haemorrhagic disease</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>25</td>
<td></td>
<td>C-section</td>
<td>Eclampsia, unspecified as to time period</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Dengue with complications</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>29</td>
<td></td>
<td>C-section</td>
<td>Severe pre-eclampsia</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mild disease</td>
<td></td>
</tr>
</tbody>
</table>
Section 4- DISCUSSION AND CONCLUSION
Discussion section

Overview

The overall aim of this thesis was to explore the relationship between adverse pregnancy outcomes and symptomatic maternal dengue. This was achieved through taking the following steps: first I developed a linkage algorithm to assess exposure (dengue status) (paper 2 and 3); then I evaluated the quality of the linkage to ensure that it did not affect the measure of association obtained in the following step (paper 2 and 3); and finally I used the linked data to estimate the association between dengue during pregnancy and adverse maternal and fetal outcomes (paper 4, 5, 6).

This chapter contains two main sections. In the first, I summarize the key linkage challenges and main findings of each results chapter and discuss the interpretation and biological plausibility of this relationship. Second, I discuss the implications of this work in the public health arena, as well as make recommendations for future research.
Chapter 9. Final discussion, conclusion and recommendations

Summary of linkage challenges and strength

Data linkage of population level data is an important tool that can be used to answer a range of research questions, especially regarding the relationship between prenatal maternal risk and adverse birth outcomes. In this thesis, as mentioned before, previous studies on dengue during pregnancy have been shown to be inefficient as they lacked sufficient power to perform sensitivity analyses due to insufficient sample size. The use of administrative collected data is a valuable alternative.

Although Brazil has a long tradition in collecting administrative data of relatively good quality, especially for a middle-income country, the linkage pre-processing stage can be challenging for a large datasets, because it is important to inspect, identify and recognize the patterns of error to be able to deal with them. For example, in the present linkage, after inspection I identified known prefixes such as Mr and transformed them into the real name Maria. An additional valuable contribution of data inspection is to identify common names, to be able to attribute a different weight score to them because a linkage error is more likely to occur with these names than with less common names.

Another important challenge of the present linkage was linking records belonging to different individuals, in this case mothers and babies. The most important barrier to linking maternal and baby records is the limited availability of common and complete personal identifiers. This thesis used three identifiers to link the records (name, age and residence of the mother), and this is one of the causes of the relatively low sensitivity, because in same cases it was not possible to tell which record was the correct match, and so no link was made.

A further complication emerged in the linkage performed in this thesis. The linked datasets did not overlap exactly, there was no Gold-Standard data available and the
expected number of matches was unknown because the incidence of dengue during pregnancy can vary from place to place. This makes it difficult to estimate the linkage accuracy and evaluate the error associated with the linkage. To overcome these challenges, I created a “Gold-standard” dataset using nine steps (paper 2); this was possible because I had access to identifiable data. It enabled me to demonstrate the accuracy of the linkage and because I had additional information about the characteristics of the cohort, I also checked for potential bias. This was in fact a real strength of my study, that I had access to the identifiable data and I was able to evaluate the quality of the linkage.

Summary of main findings

Mild dengue during pregnancy roughly doubled the risk of stillbirth, using a set of live births as control group. In contrast, dengue only elevated the risk of preterm birth (PTB) and low birth weight (LBW) respectively by 10 to 20 percent; however, the confidence interval was borderline. There was no increase in the risk of small for gestational age. In the maternal mortality analyses, I showed that pregnant women with symptomatic dengue had a threefold higher risk of maternal death compared to pregnant women without dengue (Table 1).

To address the potential misclassification of dengue due to the absence of laboratory confirmation, I performed an analysis using only laboratory confirmed cases. Laboratory confirmed dengue during pregnancy did not change the magnitude of the association between dengue and the adverse fetal outcomes (stillbirth, preterm birth, low birth weight and small for gestational age). However, the proportion of laboratory confirmed dengue cases among women who died was higher than in the comparison group (women who delivered a live birth and did not die). The risk of maternal death among laboratory confirmed dengue cases was 61% higher than non-laboratory
confirmed cases. We suggest that women with severe dengue were more likely to have had a laboratory test requested by their prenatal carers.

The literature has shown that the effects of infectious disease on maternal and fetal outcomes appear to vary with the severity of the maternal illness \(^1\). The magnitude of the association varies depending on severity of the maternal illness for all pregnancy outcomes. In the stillbirth analyses, the risk of stillbirth among severe cases was 2.9 times higher than mild cases and almost 5 times higher than the controls. Cases of dengue haemorrhagic fever increased the risk of preterm birth and low birth weight twofold compared to non-dengue live births, although even the most severe category of dengue (DHF/DSS) did not have an effect on small for gestational age. The risk of maternal death among women with complicated dengue and DHF/DSS was 451 times higher than in cases of non-dengue pregnant women.

Dengue is an acute infection with low burden of chronicity \(^2\). Because of this the hypothesis was that the magnitude of association would be higher in the first days after the disease onset, so time would act as an effect modifier of the association. I tested this hypothesis and the magnitude of the association was modified by the time between the disease onset and the pregnancy outcome. An excess of negative fetal and maternal outcomes during the acute disease was found. The risk of stillbirth, preterm birth and low birth weight among cases of women with dengue that occurred within 10/20 days of disease onset was more than 2 times higher than cases of women with dengue that occurred after this period. Regarding maternal deaths, the risk of death within 10 days of disease onset was 75 times higher than in the comparison group.
Table 1. Summary of main findings

<table>
<thead>
<tr>
<th>Odds/Risk Ratio of stillbirth (95% Confidence Interval)</th>
<th>Odds/Risk Ratio of PB (95% Confidence Interval)</th>
<th>Odds/Risk Ratio of LBW (95% Confidence Interval)</th>
<th>Odds/Risk Ratio of SGA (95% Confidence Interval)</th>
<th>Odds/Risk Ratio of MM (95% Confidence Interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dengue during pregnancy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.9 (1.6-2.2)</td>
<td>1.1 (1.0-1.2)</td>
<td>1.2 (1.1-1.2)</td>
<td>1.0 (0.9-1.1)</td>
<td>3.0 (1.5-5.8)</td>
</tr>
<tr>
<td>Dengue during pregnancy laboratory confirmed</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.9 (1.6-2.2)</td>
<td>1.2 (1.0-1.3)</td>
<td>1.3 (1.1-1.4)</td>
<td>0.9 (0.7-1.2)</td>
<td>7.8 (3.8-15.9)</td>
</tr>
<tr>
<td>Dengue during pregnancy categorized by severity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mild dengue</td>
<td>1.7 (1.5-2.0)</td>
<td>1.1 (1.0-1.1)</td>
<td>1.2 (1.1-1.3)</td>
<td>1.0 (0.9-1.1)</td>
</tr>
<tr>
<td>Complicated dengue</td>
<td>-</td>
<td>1.4 (0.9-2.1)</td>
<td>1.7 (1.2-2.4)</td>
<td>2.4 (1.0-5.6)</td>
</tr>
<tr>
<td>Dengue haemorrhagic fever</td>
<td>-</td>
<td>2.4 (1.3-4.5)</td>
<td>2.1 (1.1-4.1)</td>
<td>2.5 (0.4-14.4)</td>
</tr>
<tr>
<td>Severe dengue (complicated + haemorrhagic)</td>
<td>4.9 (2.3-10.2)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Dengue during pregnancy categorized by length of time between symptoms and outcome</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than 20 days</td>
<td>4.9 (3.2-7.5)</td>
<td>2.0 (1.6-2.6)</td>
<td>2.0 (1.5-2.6)</td>
<td>0.5 (0.2-1.5)</td>
</tr>
<tr>
<td>More than 20 days</td>
<td>1.7 (1.4-2.0)</td>
<td>1.1 (1.0-1.1)</td>
<td>1.1 (1.1-1.2)</td>
<td>1.0 (0.9-1.1)</td>
</tr>
</tbody>
</table>

Interpretation of main findings

To our knowledge, this is the first study to investigate the risk of stillbirth, small for gestational age and maternal deaths among women with dengue during pregnancy using a population-based approach with a sufficiently large sample size and controlling for confounders and to investigate the effect of disease severity and time between the disease onset on preterm birth and low birth weight.
The results of our study are consistent with the current literature. A hospital-based study conducted in French Guiana found that women with symptomatic dengue had a higher frequency of stillbirth and an increased risk of preterm birth and low birth weight. In this study, the point estimative for PTB and LBW was above two; however, the confidence interval was borderline. In this thesis, although dengue slightly elevated the risk of PTB and LBW by 10 to 20 percent, the confidence interval overlapped with results presented by Nascimento et al. who showed that symptomatic dengue infection did not affect the risk of low birthweight (1.00 95% CI 0.85-1.17) or preterm birth (OR: 0.98 95% CI 0.83-1.16) when compared to a random sample of live births. Our findings are also consistent with a meta-analyses published in 2016, which reported an association between symptomatic dengue and stillbirth, preterm birth and low birth weight in hospital based studies. Hospital based studies selected the population that sought care, therefore they are more likely to include severe dengue cases. As shown in our findings, severe cases had a consistent increase in the risk of fetal outcomes, except SGA. The maternal outcome results were also consistent with the literature. In a Brazilian cohort in Rio Branco, the maternal mortality ratio in the dengue exposed group was 13 times the mean maternal mortality ratio of the area as opposed to our finding of 3 times higher. However, it is important to highlight that this study had only two maternal deaths in the dengue group and their small sample might have overestimated the results.

**Biological Mechanisms**

Investigating the potential mechanism of how dengue may cause adverse pregnancy outcomes is not simple because dengue infection is often not apparent from the case history or physical examination of the mother or fetus. Further information from histological evaluation of the placental and fetal autopsy is difficult to obtain, and there is no data from animal studies to provide a picture of how dengue infection results in adverse pregnancy outcomes.
It is unknown whether these adverse pregnancy outcomes are related to the direct effect of fetal infection or the maternal response to viremia. Stillbirth, for example could be a result of maternal or fetal infection due to a diversity of mechanisms, e.g. direct infection of the fetus, placental damage or severe maternal illness. According to the results obtained in this thesis and publications available on effects of maternal infection on the fetus, there are potential plausible biological pathways from dengue infection, which may result in negative pregnancy outcomes.

The analyses of disease severity and length of time between symptom onset and adverse pregnancy outcome showed that the risk was higher during acute maternal illness suggesting that the main pathological pathway involved in this association is acute maternal illness. In this case, maternal high fever, or systematic symptoms of dengue such as vascular leak syndrome, mucosal bleeding, and lethargy or immunological reactions to fight the virus can be deadly to the fetus without the organisms ever being transmitted to the placenta or fetus. The lack of association between dengue and small for gestational age also corroborates the fact that dengue probably does not cause chronic effects after the mother’s health has been reestablished.

Other mechanisms can also be present such as direct damage to the fetus or placental circulation. The vertical transmission of dengue has not been described frequently, nevertheless, the virus and anti-dengue antibodies have been found in placentas, in cord blood of infants and in the cells of lung and kidney of an aborted fetus. In a few cases, endothelial damage and increased vascular permeability due to DHF may facilitate passage across the placental barrier thus contributing to vertical transmission of dengue infection. Once the virus reaches the placental tissue, pathological changes might be produced such as villous stromal edema, an increase in the formation of syncytial knots and chorangiosis, resulting in hypoxia.
The results of this thesis have not revealed an important association between mild dengue disease and preterm birth, therefore it is unlikely that the immunological reaction of the body against the dengue virus increases the production of pro inflammatory cytokines, including interleukin-6 (IL-6), interleukin-8 (IL-8), tumour necrosis factor α (TNF-α) at levels capable of activating uterine contractions culminating in a preterm delivery. Even among those who had hemorrhagic fever, in part these outcomes may be explained by early delivery due to medical intervention required because of risk to the mother.

I propose three potential explanations for the association between dengue and maternal deaths; however, the analyses performed in this PhD cannot discard nor support any of them. First, the clinical characteristics of the disease may be different during pregnancy such that it increases susceptibility to severe disease. In Brazil pregnant women with dengue were 3 times more likely to develop severe dengue than non-pregnant women. This could be explained by the shifts in the pregnant immunological system, from Th1 to Th2 that can result in infection susceptibility. Another hypothesis is that the physiological changes which occur during pregnancy such as hemoconcentration and the difficulty in distinguishing between severe dengue and common obstetric conditions may lead to misdiagnosis and delay the disease treatment which might lead to hypovolemic shock and death. Finally, studies have shown an increase in the risk of pre-eclampsia among pregnant women with viral infection. It is possible that the dengue virus leads to the same etiologic pathway of inflammatory modifications observed in pregnant women with cytomegalovirus infection, which is upregulated TLR-2/-4mRNA expression and increased levels of IL-6 and TNF- α, and reduced IL-10 compared to matched normal and no pregnancy controls of placental tissues.
Limitations

Specific limitations have been discussed in each paper. Overall, the main limitations to the work presented in this thesis lies in the use of administrative data and the linkage process. Regarding the use of secondary data: the proportion of preterm births recorded in SINAN-Brazil was found to be underestimated by 15%, and misclassification, based on the criteria used to assess the gestational age at birth information (date of the last period), could have occurred. However, these errors probably affected both those exposed and not exposed to dengue therefore; this is unlikely to bias the results of this study. If this assumption is not correct and preterm birth was underestimated differently by exposure status, it either could bias the results in an underestimation (if it was underestimated in the exposed group) or overestimated (if it was underestimated in the non-exposed group).

Regarding dengue, studies have shown that dengue surveillance substantially underestimates the disease burden, mainly during periods of low-transmission. I assume that the underestimation of the disease incidence is a non-differential error, therefore our results could be underestimated, i.e. the magnitude of the association found could be even higher than the magnitude presented in this thesis. Similar to differential error of the outcome measurement, if differential misclassification of dengue status occurred, it could have resulted in either underestimation or overestimation depending on the situation. For example if the comparison group were less likely than the cases to be classified as a dengue case, it would lead to overestimation of the measure of association. Misclassification of exposure based on the lack of laboratory confirmation is also possible but very good agreement with the estimates of risk in subsets of the data with and without laboratory confirmation of dengue were found, except in the maternal mortality analyses.
Another limitation is the restricted availability of potential confounders (e.g. hospital admission, maternal co-morbidities, quality of obstetric care), and some variables available were already categorized, consequently restricting information, such as gestational age, which was recorded as a categorical variable until 2010. Finally, although I used massive numbers of routinely collected data, in some of the analyses, especially for severe disease, the sample sizes were small, although given the large effect size I still have enough power in most of the analyses.

The linkage process posed a number of challenges, such as linking different individuals (mother-baby) due to the limited numbers of identifiers; estimating the linkage accuracy due to the absence of Gold standard dataset or the unknown number of matches that are discussed in detail in chapters 4 and 5. The main limitation inherent to this process is the low sensitivity (that kept us from making statements about the prevalence of dengue during pregnancy). However, it is unlikely that the linkage error introduced bias in the final analyses, since the evaluation of the linkage showed that missed and false matches, used to assess the exposure status, occurred randomly.

**Implications and recommendations**

The linkage process was done very carefully and it proved to be as accurate as possible. The performing linkage of complex data records requires substantial expertise (which I have developed during my PhD) which can be applied to answer different research questions. It can be used to evaluate policy implementation in places where population level identified data are available. An example in Brazil is the recently developed Center of Data and Knowledge for Health (CIDACS), which is developing numerous research projects with routine data and some of the techniques developed in this thesis, can be applied there. Even though the linkage technique used in this thesis can be applied in other studies and different settings, it should be done cautiously, since some
linkage decisions were made based on the characteristics of these particular datasets. For example, the string comparator used was chosen because it performed better in comparing names in Brazil. Names from different places could possibly use another string comparator.

Due to disclosure issues, fragmentation of the data has been a common practice to keep identifiers separate from attributes. This can however, increase the risk of bias in the analyses since linkers and analysts may not be aware of important disproportions that affect the quality of the linkage. Given that I was in a unique position when I performed the linkage and the posterior analyses, I recommend that when possible, the linkage process when not developed by the same person should be closely monitored by the analysts of the linked data. The data linkers should have access to the characteristics of the data, so they can be aware of potential problems and share this information with analysts to be incorporated into results.

This thesis can be reproduced in other setting and it is likely that the results on the association between dengue and pregnancy outcomes will be the same as observed in this cohort or an even higher magnitude of association can be observed if the surveillance system performs better than the Brazilian system. Perhaps in settings where the frequency of severe cases is even higher than in Brazil (e.g. countries in Southeast Asia), the sensitivity analyses by severity can make use of a bigger sample size than that observed in this research and provide greater precision.

Since dengue is a vector borne disease without recommended vaccine for pregnant women, the health of this group should be a public health priority in places where dengue is circulating. Not only this but measures to stimulate self-protection from mosquitos bites should be reinforced. Dengue control programs should strengthen health educational actions on effective methods to protect the population, such as using air conditioning,
screens, or nets when indoors, wearing long sleeves and trousers, using permethrin-treated clothing and gear, and using insect repellents when outdoors. However, there is not enough evidence of the efficacy of these measures.

Mosquito control measures is also important to reduce the incidence of the disease and consequently reduce the risk of adverse pregnancy outcomes. Currently, Aedes mosquito control depends on mechanical breeding-site reduction and chemical pesticides (insecticides and larvicides), measures that have failed to stop dengue and other arbovirus disease from spreading in Brazil. Alternative methods have been proposed and tested in Brazil, such as RIDL (Release of Insects with Dominant Lethality). These have achieved a 95% reduction in local Aedes populations and Wolbachia, that has been tested in open field and mathematical models predict that one strain would reduce the basic reproduction number of DENV transmission by 70%. However, these alternative methods are not part of the arboviruses control program of the Brazilian Ministry of Health.

As dengue prevention is not straightforward and mosquito control has failed in Brazil, the results shown in this thesis point to the crucial role of the clinical management of dengue in pregnant women. Doctors and medical staff in general attending to pregnant women with dengue should closely observe and monitor the patients to be able to intervene timely and avoid death of the mother or fetus, especially during the acute phase. For further investigation, I recommend careful observation and notation of dengue in antenatal records and full dengue investigation of adverse pregnancy outcomes, especially maternal and fetal deaths.

Further research is required in different settings to confirm the results presented in this thesis and examine childhood outcomes. Additional studies are required to measure the burden of subclinical maternal viral infections on pregnancy outcomes, the effect of different dengue serotypes, and the effect of other vector-borne diseases, such as
chikungunya, yellow fever and West Nile disease. Another important research recommendation is to elucidate the pathological mechanisms involved in this association including animal and in vitro models.

**Conclusion**

In summary, this thesis shows there is an increase in the risk of stillbirth and maternal deaths in women with mild dengue during pregnancy. Severe disease increases the magnitude of the association with deadly maternal and fetal outcomes and doubles the risk of preterm birth and low birth weight. Therefore, pregnant women should be included as an at-risk population in dengue control programmes in order to reduce the risk for women and fetal outcomes.

The proposed objective of the thesis was achieved; I was able to develop and evaluate a linkage process addressing the challenges that arose from the complex data and estimate the association between dengue and pregnancy outcomes. This thesis is a document that adds an important piece in the puzzle of the effects of maternal viral infection on pregnancy outcomes, showing that symptomatic dengue during pregnancy is associated with adverse outcomes. During this journey, I also developed expertise in data linkage and built up my skills as an epidemiologist.
Section 4 references


Appendices

Appendix I Ethics approval from Brazil

INSTITUTO DE SAÚDE COLETIVA / UFBA

PARECER CONSUBSTANTIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: Dengue e Desfechos Desfavoráveis da Gravidez
Pesquisador: Maria da Graça Lima Cruz Teixeira
Área Temática:
Versão: 1
CAAE: 26797814.7.0000.5030
Instituição Proponente: Instituto de Saúde Coletiva / UFBA
Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 648.020
Data da Relatoria: 25/03/2014

Apresentação do Projeto:
Trata-se de um projeto envolvendo dois estudos, um de base agregado-populacional e outro individual visando analisar a relação entre infecção pelo vírus do dengue em mulheres grávidas e desfechos desfavoráveis da gestação.

Objetivo da Pesquisa:
Objetivo Primário:
Analisar a relação entre infecção pelo vírus do dengue em mulheres grávidas e desfechos desfavoráveis da gestação.
Objetivo Secundário:
Verificar a existência de associação entre incidência de dengue na população de mulheres em idade reprodutiva e proporção de partos prematuros, abortos espontâneos, natimortos, baixo peso ao nascimento, partos cesáreos, mortalidade materna, hospitalização por aborto espontâneo e por hemorragia na gestação. Examinar se infecção por dengue durante a gestação é fator de risco para prematuridade e baixo-peso ao nascimento. Examinar se infecção por dengue durante a gestação é fator de risco para ocorrência de óbitos maternos e fetais.

Avaliação dos Riscos e Benefícios:
A realização do projeto não implica riscos para a população estudada. Não haverá contato direto com os pacientes.
Benefícios:
O desenvolvimento deste projeto de pesquisa se apresenta como uma possibilidade de contribuição para uma temática que não se encontra totalmente elucidada, apesar de ser de grande interesse para a Saúde Pública. Os resultados deste estudo podem ajudar a ampliar as evidências sobre a associação entre infecções pelo vírus do dengue e resultados da gravidez, saúde materna e intercorrências obstétricas e contribuir para o aprimoramento dos protocolos do Sistema Único de Saúde (SUS) para atenção médica às gestantes e conceptos durante epidemias de dengue e, especificamente, nos casos de manifestações clínicas da doença no curso da gravidez. Os resultados serão apresentados à comunidade científica através da elaboração de artigos científicos e apresentação em Congressos e Seminários.

Comentários e Considerações sobre a Pesquisa:
Trata-se de estudo realizado com dados secundários dos sistemas de informações: a) SINAN b) SIM c) SINASCc) SIH e) IBGE. Para o estudo de agregados será feita uma análise descritiva.
Para verificar a existência de relação entre incidência de dengue em mulheres em idade fértil nos municípios brasileiros e ocorrência de desfechos desfavoráveis na gestação, um banco de dados longitudinal será criado para os anos de 2001 a 2011, de modo que a mesma unidade de análise terá suas observações repetidas ao longo do tempo, para se proceder a Análise em Painel.

Considerações sobre os Termos de apresentação obrigatória:
O estudo não implicará em contato direto com os sujeitos da pesquisa e, portanto, dispensa a utilização do Termo de Consentimento Livre e Esclarecido (TCLE).
Os bancos de dados secundários utilizados serão os dos sistemas de informações em saúde, disponíveis online, dispensando solicitação de autorização de uso.

Recomendações:

Conclusões ou Pendências e Lista de Inadequações:
Não existem pendências ou inadequações

Situção do Parecer:
Aprovado

Necessita Apreciação da CONEP:
Não

Considerações Finais a critério do CEP:
O Comitê de Ética em Pesquisa do Instituto de Saúde Coletiva – UFBA analisou, na sessão do dia 25 de março de 2014, o processo número 069/14, referente ao projeto de pesquisa em tela.
Não tendo apresentado pendências na época da sua primeira avaliação, atendeu de forma adequada e satisfatoriamente às exigências da Resolução nº 466 de 12/12/2012 do Conselho Nacional de Saúde (CNS). Assim, mediante a importância social e científica que o projeto apresenta e a sua aplicabilidade e conformidade com os requisitos éticos, somos de parecer favorável à realização do projeto, classificando-o como APROVADO.
Solicita-se ao pesquisador o envio a este CEP de relatórios parciais sempre quando houver alguma alteração no projeto, bem como o relatório final gravado em CD ROM.

SALVADOR, 14 de Maio de 2014

Assinado por:
Alcione Brilheiro Oliveira Cunha
(Coordenador)
Appendix II Ethics approval from LSHTM

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
United Kingdom
Switchboard: +44 (0)20 7636 8636
www.lshtm.ac.uk

Observational / Interventions Research Ethics Committee

LSHTM
4 September 2015

Dear

Study Title: Dengue infection and adverse pregnancy outcomes

LSHTM Ethics Ref: 10209

Thank you for responding to the Observational Committee’s request for further information on the above research and submitting revised documentation.

The further information has been considered on behalf of the Committee by the Chair.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised, subject to the conditions specified below.

Conditions of the favourable opinion

Approval is dependent on local ethical approval having been received, where relevant.

Approved documents

The final list of documents reviewed and approved by the Committee is as follows:
Appendix III Brazilian forms – live birth form (SINASC)
Appendix IV Brazilian forms – death form (SIM)
Appendix V Brazilian forms – dengue form (SINAN)

The image shows a page with the dengue form (SINAN) from the Brazilian Ministry of Health, detailing the fields for notification of suspected cases of dengue fever. The form contains sections for patient information, notification details, and laboratory results. Each field is labeled and must be filled in to complete the report.
### Informações complementares e observações

**Observações Adicionais**

---

### Investigador

<table>
<thead>
<tr>
<th>Municipal/Unidade de Saúde</th>
<th>Nome</th>
<th>Função</th>
<th>Cód da Unidade de Saúde</th>
<th>Assinatura</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sinan NET / Elean Online</td>
</tr>
</tbody>
</table>