Sepsis is a frequent cause of death in very-low-birthweight infants and often results in neurological impairment. Its attributable risk of sequelae has not been systematically assessed. To establish an outcome tree for mapping the burden of neonatal sepsis, we performed systematic literature searches to identify systematic reviews addressing sequelae of neonatal sepsis. We included cohort studies and performed meta-analyses of attributable risks. Evidence quality was assessed using GRADE. Two systematic reviews met inclusion criteria. The first included nine cohort studies with 5,620 participants and five outcomes (neurodevelopmental impairment, cerebral palsy, vision impairment, hearing impairment, death). Pooled risk differences varied between 4% (95% confidence interval (CI): 2–10) and 13% (95% CI: 5–20). From the second review we analysed four studies with 472 infants. Positive predictive value of neurodevelopmental impairment for later cognitive impairment ranged between 67% (95% CI: 22–96) and 83% (95% CI: 36–100). Neonatal sepsis increases risk of permanent neurological impairment. Effect size varies by outcome, with evidence quality being low to very low. Data were used to construct an outcome tree for neonatal sepsis. Attributable risk estimates for sequelae following neonatal sepsis are suitable for burden estimation and may serve as outcome parameters in interventional studies.

Introduction

Sepsis is a major cause of death in neonates [1]. The majority of sepsis episodes (80%) occurs in preterm neonates [2]. Among very low birth weight infants (VLBW; <1,500g), rates of sepsis range between 11% and 46% [3]. Sepsis in this high-risk population is mostly acquired during hospital stay with a late onset beyond 48–72 hours of life. Early onset sepsis, which becomes apparent within the first 48–72 hours of life is ‘nosocomial’ in the sense that it occurs in the hospital but should not be considered healthcare-associated because its origin is linked to childbirth and/or maternal-fetal transmission of pathogens [4,5].

Neonatal sepsis and systemic inflammatory response syndrome (SIRS) are associated with brain damage that results in disability, particularly among preterm and VLBW infants [6-9]. However, adverse neurological outcomes frequently occur in VLBW infants for reasons other than sepsis [10]. Therefore, the impact of healthcare-associated neonatal sepsis on adverse outcome is difficult to establish.

A European consortium, as part of a project initiated and funded by the European Centre for Disease Prevention and Control (ECDC), recently developed an incidence- and pathogen-based approach for estimating the burden of communicable diseases expressed in Disability Adjusted Life Years (DALYs) [11]. Relevant health outcomes of communicable diseases are represented by outcome trees which map the weighted progressions of diseases over time by ordering the conditional probabilities of associated health outcomes [11]. Outcome trees take into account probabilities and duration of health outcomes. The burden of healthcare-associated infections (HAIs) was not yet addressed by ECDC for two reasons: (i) Patients with HAIs differ from the general population in terms of comorbidities [7] and may be different regarding other factors as for example certain risk behaviour and social determinants; (ii) HAIs cannot be allocated to a specific pathogen, and thus the pathogen-based approach is not applicable.
Statements about the burden of communicable diseases must be based on evidence-based medicine principles as in other fields of medicine. Systematic reviews have become the gold standard of assessing the evidence in medicine but they are time-consuming and expensive. Systematic reviews of systematic reviews (so-called ‘umbrella reviews’) offer a time-saving alternative to identify and exploit the current state of evidence in a field [12,13]. The aim of the study was to identify the relevant sources for the construction of an evidence-based outcome tree for neurological sequelae due to healthcare-associated neonatal sepsis, umbrella review on neurological sequelae of healthcare-associated sepsis in very-low-birthweight infants, date of search 25 September 2013

**Methods**

We followed the approach described by Whitlock et al. [14] and Robinson et al. [15] to identify and exploit existing systematic reviews by re-analysing their data. Our study was performed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement (PRISMA) [16].

**Identification of studies**

In a first step, we performed a systematic review of systematic reviews (i.e. an umbrella review) on the association between neonatal sepsis and neurodevelopment in later life. To identify relevant systematic reviews we performed a systematic literature search using MEDLINE, EMBASE and Cochrane Central Register of Controlled Trials (CENTRAL), without language restrictions (Box 1). All systematic reviews published from 1 January 2000 until 25 September 2013 were eligible if meeting predefined inclusion criteria (see below).

A further search for planned, ongoing and published systematic reviews was performed in the Prospective International Register of Systematic Reviews (PROSPERO).

In a second step, we performed an umbrella review on the positive predictive value of neurodevelopmental impairment for later cognitive function. To identify appropriate systematic reviews, we performed a systematic literature search (date of last search: 2 July 2014) using MEDLINE, EMBASE and the Cochrane Central Register of Controlled Trials (CENTRAL) (Box 2).

Electronic search was complemented by manually checking the reference lists of identified reviews and studies.

**Study eligibility**

In a first step, we searched for systematic reviews on the association between neonatal sepsis and neurodevelopment in later life. These systematic reviews had to capture primary studies which fulfilled the following inclusion criteria: (i) study population had to be neonates; (ii) the exposure had to be sepsis acquired in a healthcare setting; (iii) the comparator (or control) had to be participants without sepsis; (iv) the studies included had to be cohort studies or clinical trials; (v) the studies had to investigate at least one neurodevelopmental outcome during follow-up, and (vi) the studies had to be conducted in a healthcare setting within an upper-middle- or high-income country [17]. An expert panel discussed and agreed on the addressed outcomes to be relevant (for names and affiliations of the members of the expert panel, see Acknowledgements).

In a second step, the systematic reviews were searched for positive predictive values of neurodevelopmental impairment for later cognitive function. To be eligible,
Two independent reviewers (SH and TH) screened the systematic reviews, located the primary studies analysed in the review and extracted the following data: citation, study period, study design, demographics, sex, ethnicity, definition of sepsis, definition of outcome, length of follow-up, number of exposed and non-exposed with outcome, test used for first and second examination, positive predictive value and prevalence of condition. Discrepancies between the reviewers were solved by discussion until a consensus was reached.

Risk of bias assessment

The assessment of multiple systematic reviews (AMSTAR) tool was used to determine the methodological quality of the systematic reviews included [20]. Risk of bias in the included cohort studies was assessed using the Newcastle Ottawa Scale [21]. Following the suggestions of the Cochrane Collaboration, we assessed risk of bias separately for each outcome in each study [22]. For the studies reporting positive predictive values, the Scottish Intercollegiate Guidelines Network (SIGN) checklist for diagnostic accuracy studies was used to assess risk of bias [23]. The results of the risk of bias assessments were expressed in terms of ‘high risk of bias’, ‘low risk of bias’ and ‘unclear risk of bias’. All risk of bias assessments were conducted by two independent investigators (SH and TH).

Assessment of the quality of the body of evidence

We adapted the methodology of the Grading of Recommendations Assessment, Development and Evaluation (GRADE) working group to assess the quality of the body of evidence [24,25]. The GRADE methodology was initially developed to assess intervention studies. According to GRADE, the quality of evidence indicates the extent to which one can be confident that the estimate of effect is correct. Taking into account the entire body of evidence on one outcome, four levels of evidence quality are applied: +++, very low; ++, low; +++, moderate; and +++++, high. Adapting the original GRADE approach and considering the proposal by Huguet et al. [26], all bodies of evidence were initially graded as high quality of evidence. Considering the following criteria led to decreasing evidence quality: (i) risk of bias, (ii) inconsistency, (iii) indirectness, (iv) imprecision and (v) publication bias (for details on the criteria, see [24]).

Data synthesis

Extracted study characteristics and data were summarised in tables, together with risk of bias assessments. For data synthesis, the following two effect measures were used:

- Risk differences were applied to calculate attributable risk to sepsis exposure as follows: Using data of the individual studies we subtracted the absolute risk of developing sequelae in controls from the risk of developing sequelae in cases (infected minus uninfected and corresponding 95% confidence intervals).

---

**Forest plot of risk differences of (panel A) impaired neurodevelopment (mental development index < 70) and (panel B) cerebral palsy in neonates with sepsis compared with those without sepsis, umbrella review on neurological sequelae of healthcare-associated sepsis in very-low-birthweight infants, date of search 25 September 2013**

<table>
<thead>
<tr>
<th>Study</th>
<th>Decreased risk after sepsis</th>
<th>Increased risk after sepsis</th>
<th>Risk difference (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Addison et al. (2009)</td>
<td>-0.09 (-0.09 to 0.27)</td>
<td>0.01 (-0.11 to 0.24)</td>
<td>0.07 (-0.11 to 0.24)</td>
</tr>
<tr>
<td>Chen et al. (2008)</td>
<td>0.21 (-0.01 to 0.43)</td>
<td>0.29 (0.10 to 0.48)</td>
<td>0.03 (0.10 to 0.37)</td>
</tr>
<tr>
<td>Friedman et al. (2000)</td>
<td>0.21 (-0.11 to 0.53)</td>
<td>0.29 (0.10 to 0.48)</td>
<td>0.06 (0.03 to 0.19)</td>
</tr>
<tr>
<td>Göcer et al. (2011)</td>
<td>0.28 (0.07 to 0.50)</td>
<td>0.29 (0.10 to 0.48)</td>
<td>0.06 (0.03 to 0.19)</td>
</tr>
<tr>
<td>Hack et al. (2000)</td>
<td>0.10 (-0.23 to 0.44)</td>
<td>0.06 (0.03 to 0.19)</td>
<td>0.01 (-0.09 to 0.11)</td>
</tr>
<tr>
<td>Mualli et al. (1996)</td>
<td>0.20 (0.03 to 0.37)</td>
<td>0.15 (0.12 to 0.18)</td>
<td>0.05 (0.03 to 0.17)</td>
</tr>
<tr>
<td>Schlapbach et al. (2011)</td>
<td>0.06 (0.03 to 0.19)</td>
<td>0.15 (0.12 to 0.18)</td>
<td>0.05 (0.03 to 0.17)</td>
</tr>
<tr>
<td>Stoll et al. (2004)</td>
<td>0.13 (0.05 to 0.20)</td>
<td>0.13 (0.05 to 0.20)</td>
<td>0.05 (0.03 to 0.17)</td>
</tr>
</tbody>
</table>

Overall (P=0.05%, p=0.003)

Risk difference (95% CI)

<table>
<thead>
<tr>
<th>Study</th>
<th>Decreased risk after sepsis</th>
<th>Increased risk after sepsis</th>
<th>Risk difference (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Addison et al. (2009)</td>
<td>-0.10 (-0.23 to 0.04)</td>
<td>0.29 (0.10 to 0.48)</td>
<td>0.19 (0.10 to 0.30)</td>
</tr>
<tr>
<td>Schlapbach et al. (2011)</td>
<td>0.06 (0.03 to 0.19)</td>
<td>0.15 (0.12 to 0.18)</td>
<td>0.05 (0.03 to 0.17)</td>
</tr>
<tr>
<td>Shah et al. (2008)</td>
<td>0.04 (0.02 to 0.16)</td>
<td>0.09 (0.07 to 0.16)</td>
<td>0.02 (0.01 to 0.04)</td>
</tr>
<tr>
<td>Stoll et al. (2004)</td>
<td>0.09 (0.07 to 0.16)</td>
<td>0.08 (0.06 to 0.10)</td>
<td>0.05 (0.03 to 0.17)</td>
</tr>
</tbody>
</table>

Overall (P=0.0%, p=0.55)

Cl: confidence interval; MDI: Mental Developmental Index.

Studies are ordered alphabetically by first author. The pooled risk differences (overall; diamonds) were calculated by means of a random-effects model. Ninety-five percent CIs are shown in parentheses and as horizontal bars.
Because the numbers of studies were too small (< 10), publication bias was not investigated. Positive predictive values were not pooled but rather presented as a range of values to account for heterogeneity [27].

Results

Neurodevelopmental sequelae of neonatal sepsis

Our search identified a total of 207 titles (Figure 1A). After eliminating duplicates and screening of titles and abstracts four publications were left for full text evaluation. Of these, only one systematic review fulfilled the inclusion criteria and thus, was eligible for further analysis [3], whereas the remaining three publications were not eligible [28-30]. This systematic review was of acceptable methodological quality (AMSTAR summary score: 7/11). The review included 17 original studies reporting data on neurodevelopmental sequelae of neonatal sepsis in VLBW infants. We screened the abstracts and full texts of these studies and identified nine of them to be eligible (see Figure 1B), whereas seven did not fulfil the inclusion criteria [6,31-36] and one citation could not be located in data banks or libraries.

All nine included studies [7,8,37-43] were cohort studies. Details are shown in Table 1. The studies included a total of 5,620 neonates born between 1983 and 2007 and were conducted in six upper-middle and high-income countries. Three studies provided data on infants with extremely low birth weight (ELBW; < 1,000g) [7,40,44]. A further three studies reported on neonates with VLBW [37,38,41], whereas the remaining three studies based their inclusion criteria on gestational age [8,42,43]. Eight studies provided a definition of neonatal sepsis that was based on clinical and/or laboratory parameters. One study did not provide a definition [41]. One study reported on invasive Candida spp. infections only [40]. Duration of follow-up varied between 12 and 52 months.

From the reported outcomes we considered the following five outcomes as clinically relevant: neurodevelopmental impairment, cerebral palsy, vision impairment, hearing impairment and death. Neurodevelopmental impairment was defined as having a Mental Developmental Index (MDI) < 70 [18]. For vision and hearing impairment, varying definitions were used in the studies. According to the Newcastle-Ottawa Scale [21], three studies had high risk of bias, while the remaining six showed a low risk of bias (Table 1).
Candida spp. infection (95% CI: -13 to 17), while Msall et al. [42] provided data to calculate a risk difference of 14% for bacterial sepsis (95% CI: 5–20). Meta-analysis was not conducted because exposure and results were too heterogeneous.

Neither grouping of primary studies by birth weight, nor by publication date had an influence on risk differences of neurodevelopmental impairment or cerebral palsy. The number of studies was too small to allow stratified meta-analysis.

To systematically assess the quality of evidence for each outcome we applied the GRADE methodology. For neurodevelopmental impairment, the quality of the evidence had to be graded down by three levels: (i) for serious risk of bias, (ii) for serious inconsistency due to widely differing point estimates of the single studies, and (iii) for serious imprecision due to a wide CI around the pooled estimate. Therefore, evidence quality was only ‘very low’ for this outcome. Regarding cerebral palsy, very serious risk of bias (grading down two levels) led to an evidence quality of ‘low’ for this outcome. For the outcome vision impairment, evidence quality was graded down to ‘moderate’ due to serious risk of bias. Accounting for serious risk of bias and serious imprecision, evidence quality was graded down to ‘low’ for the outcome hearing impairment.

Predictive value of early neurodevelopmental impairment for later cognitive function
Our search identified three potentially eligible reviews. After title and abstract screening, only one publication remained for full text screening [45]. This systematic review fulfilled our inclusion criteria and was therefore used as a database for further analysis.

The review was of acceptable methodological quality (AMSTAR summary score: 7/11). It contained a total of 18 publications that reported data on the relation between MDI scores during the first three years of life and cognitive function measured later in life in VLBW infants. After abstract and full text screening, four studies were eligible for further analysis [39,46-48], whereas the remaining 14 did not meet the inclusion criteria [49-62].

All included studies were cohort studies. Details are shown in Table 2. Studies accumulated a total of 472 infants of either VLBW (n=2 studies) or ELBW (n=2 studies) who were born between 1977 and 2004 in three different high-income countries. All four studies used the Bayley Scale of infant development to assess the proportion of infants with neurodevelopmental impairment (i.e. MDI<70) at 12 to 24 months, and re-evaluated the study sample at 3.4 to 8.6 years of age, using three different test batteries. According to the SIGN50 checklist [23], all four studies had a low risk of bias.
### Table 1
Characteristics of included studies on neurodevelopment after neonatal sepsis, umbrella review on neurological sequelae of healthcare-associated sepsis in very-low-birthweight infants, date of search 25 September 2013

<table>
<thead>
<tr>
<th>Reference</th>
<th>Location</th>
<th>Birth year(s)</th>
<th>Population</th>
<th>Definition of sepsis</th>
<th>Duration of follow-up</th>
<th>n⁰</th>
<th>Risk of bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Addison et al. (2009) [37]</td>
<td>US</td>
<td>1999–2001</td>
<td>VLBW</td>
<td>Positive blood culture; at least 7 days of antibiotics</td>
<td>12–18 months</td>
<td>65</td>
<td>High</td>
</tr>
<tr>
<td>Chen et al. (2008) [38]</td>
<td>Taiwan</td>
<td>1998–2005</td>
<td>VLBW</td>
<td>Culture-proven sepsis and unstable vital signs</td>
<td>18–39 months</td>
<td>122</td>
<td>Low</td>
</tr>
<tr>
<td>Friedman et al. (2000) [40]</td>
<td>Canada</td>
<td>1988–1996</td>
<td>ELBW</td>
<td>Positive-culture <em>Candida</em> spp. or supportive brain autopsy</td>
<td>24 months</td>
<td>299</td>
<td>Low</td>
</tr>
<tr>
<td>Göcer et al. (2011) [41]</td>
<td>Turkey</td>
<td>2002</td>
<td>VLBW</td>
<td>Not defined</td>
<td>33–45 months</td>
<td>117</td>
<td>Low</td>
</tr>
<tr>
<td>Msall et al. (1994) [42]</td>
<td>US</td>
<td>1983–1986</td>
<td>ELBW</td>
<td>Positive blood culture and 14–21 days antibiotics</td>
<td>52 months</td>
<td>149</td>
<td>Low</td>
</tr>
<tr>
<td>Schlappbach et al. (2011) [65]</td>
<td>Switzerland</td>
<td>2000–2007</td>
<td>ELBW</td>
<td>Positive blood culture and clinical signs and antibiotics for ≥ 5 days</td>
<td>18–24 months</td>
<td>372</td>
<td>Low</td>
</tr>
</tbody>
</table>

ELBW: extremely low birth weight; GA: gestational age; US: United States; VLBW: very low birth weight.

⁰ Study size.

### Table 2
Characteristics of included studies on predictive value of early neurodevelopmental impairment for later cognitive function, umbrella review on neurological sequelae of healthcare-associated sepsis in very-low-birthweight infants, date of search 2 July 2014

<table>
<thead>
<tr>
<th>Reference</th>
<th>Location</th>
<th>Birth year(s)</th>
<th>Population</th>
<th>Test at first examination</th>
<th>Test at second examination</th>
<th>Age at first test</th>
<th>Age at second test</th>
<th>n⁰</th>
<th>PPV (95% CI)</th>
<th>Prevalence of MDI &lt; 70</th>
<th>Risk of bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Munck et al. (2012) [47]</td>
<td>Finland</td>
<td>2001–2004</td>
<td>VLBW</td>
<td>Bayley Scale vers. Two (MDI &lt; 70)</td>
<td>Wechsler Scale (FSIQ &lt; 70)</td>
<td>2 years (corrected)</td>
<td>5 years</td>
<td>124</td>
<td>83% (36–100)</td>
<td>4% Low</td>
<td></td>
</tr>
<tr>
<td>Hack et al. (2005) [39]</td>
<td>US</td>
<td>1992–1995</td>
<td>ELBW</td>
<td>Bayley Scale vers. Two (MDI &lt; 70)</td>
<td>Kaufman Assessment Battery for Children, Mental Processing Composite (MPC &lt; 70)</td>
<td>20 months</td>
<td>8.6 years</td>
<td>200</td>
<td>37% (27–49)</td>
<td>15% Low</td>
<td></td>
</tr>
<tr>
<td>Kitchen et al. (1987) [46]</td>
<td>Australia</td>
<td>1977–1980</td>
<td>ELBW</td>
<td>Bayley Scale (MDI &lt; 70)</td>
<td>Wechsler Preschool and Primary Scale (FSIQ &lt; 70)</td>
<td>2 years (corrected)</td>
<td>5.5 years (corrected)</td>
<td>54</td>
<td>67% (22–96)</td>
<td>7% Low</td>
<td></td>
</tr>
<tr>
<td>Ross et al. (1985) [48]</td>
<td>US</td>
<td>1978–1979</td>
<td>VLBW</td>
<td>Bayley Scale (MDI &lt; 70)</td>
<td>Stanford-Binet Intelligence Scale (IQ &lt; 70)</td>
<td>12 months</td>
<td>3.4 years</td>
<td>94</td>
<td>75% (35–97)</td>
<td>6% Low</td>
<td></td>
</tr>
</tbody>
</table>

CI: confidence interval; ELBW: extremely low birth weight; FSIQ: full-scale intelligence quotient; IQ: intelligence quotient; MDI: mental development index; MPC: mental processing composite; PPV: positive predictive value; US: United States; VLBW: very low birth weight.

⁰ Study size.

According to Newcastle-Ottawa scale [21].
The positive predictive value, i.e. the probability of having a positive test result at the second examination when the first test result was positive, varied between 37% (95% CI: 27–49) and 83% (95% CI: 36–100). Heterogeneity between estimates was mainly due to the study by Hack et al. [39], which was the only study that did not use an externally validated test battery to assess IQ at follow-up. Excluding their estimate from the study pool resulted in positive predictive values between 67% (95% CI: 22–96) and 83% (95% CI: 36–100).

The results of the systematic review were used to develop the outcome tree (Figure 4). Risk differences obtained from meta-analyses were used to estimate the transitional probabilities for acquiring neurodevelopmental impairment, cerebral palsy, vision impairment, hearing impairment and death after having experienced sepsis during neonatal life. Furthermore, we used the positive predictive values identified in the second systematic review to estimate the probability of having a permanently impaired cognitive function after early neurodevelopmental impairment.

Box 1

Search strategy for umbrella review on the association between neonatal sepsis and neurodevelopment in later life, date of search 25 September 2013

We used the following search strategy:
#1 “outcome”
#2 “follow-up”
#3 “sequelae”
#4 “consequence”
#5 “death”
#6 “cerebral palsy”
#7 “retinopathy”
#8 “necrotizing enterocolitis”
#9 “bronchopulmonary dysplasia”
#10 “neurodevelopmental impairment”
#11 “periventricular leukomalacia”
#12 “intraventricular haemorrhage”
#13 neonat*”
#14 “newborn”
#15 “sepsis”

#16 (#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12) AND (#13 OR #14) AND #15

Filters: Publication date from 2000/01/01; Humans; “Medline” OR “systematic review” OR “meta-analysis” OR “intervention”

Box 2

Search strategy for umbrella review on the positive predictive value of neurodevelopmental impairment for later cognitive function, date of search 2 July 2014

The search strategy was the following:
#1 “predictive value”
#2 “Bayley scale”
#3 “systematic review”
#4 “meta-analysis”
#5 #3 OR #4
#1 AND #2 AND #5 (no filters)

Discussion

We developed an outcome tree for neurological sequelae of neonatal sepsis in VLBW infants using the methods of evidence-based medicine. Our study shows that 4–14% of neurological sequelae in ELBW and VLBW are attributable to neonatal sepsis. Although this may be lower than anticipated in this high-risk group, about three-quarters of infants with early neurodevelopmental impairment suffer from persistent cognitive impairment later in life. Evidence quality was low to very low, mainly due to high risk of bias in the single studies as well as imprecision of estimates.

Due to the inclusion criteria of the systematic review, primary studies which used different definitions of neonatal sepsis were analysed together. While some authors defined sepsis as culture-proven sepsis plus clinical signs [38,44], others applied a definition that included antibiotic treatment in addition to positive blood culture [37,42,43]. Moreover, one study did not provide any definition of neonatal sepsis [41]. Definition issues also applied for outcomes. Vision impairment was defined differently in the primary studies: while both studies defined vision impairment as uni- or bilateral blindness, one study also allowed the need for corrective lenses [44]. To estimate the impact of definition of exposure and/or outcome on the associations of interest, a larger number of carefully conducted prospective studies with subgroup analyses of sufficient power would be needed.

To assess attributable risk, we used the original data of the studies included in the systematic reviews to calculate risk differences. This approach may not adjust for potential confounders, which might bias the relation between exposure and outcome. In nearly all analysed studies the infants with neonatal sepsis differed in a number of important prognostic variables from controls such as gestational age, birth weight and co-morbidities. While the original studies did adjust for such variables by applying multivariate analysis, we could not do the same because (i) potential confounders were
not uniform, (ii) some studies adjusted for variables which are not confounders for our purpose, and (iii) we have not had access to the original database with the individual data to do so. We therefore did not follow this approach any further, but considered the problem of confounding in the risk of bias assessments. The study by Hack et al. [44] on ELBW infants born between 1992 and 1996 illustrates that confounding may be outcome specific and leads to surprising results. In this particular study, neonatal exposure to sepsis was associated with hearing impairment, but also with a lower likelihood of neurodevelopmental impairment. This may best be explained by confounding due to postnatal use of corticosteroids administered in that time period in neonates without symptoms of infection to prevent chronic lung disease. Among others, Yeh et al. were able to show the strong side effects of this therapeutic strategy on neurodevelopment when they evaluated long-term neurodevelopmental outcomes of children who had participated in a randomised controlled trial on the effects of dexamethasone therapy [63]. As for the other studies, it was surprising to find that risk differences for neurodevelopmental impairment and cerebral palsy following neonatal sepsis were similar across different settings. Grouping of studies by birth weight or year of publication did not reveal trends for risk differences. More studies would be needed to analyse whether they are independent of birth weight, gestational age, setting and time. It may be concluded, however, that the attributable risks can be used as endpoints for studies evaluating the effectiveness of specific sepsis therapy. Further, it may be hypothesised that sepsis therapy has not, over the years, improved to a similar extent as overall neonatal intensive care.

For meta-analysis, we pooled the risk differences from the individual studies to arrive at a single measure of attributable risk for each outcome. Statistical pooling of risk differences has been reported to cause problems with consistency, with relative risk estimates (including odds ratios) being more consistent than risk differences [64]. For comparison, we pooled the estimates of the calculated relative risks (data available upon request from the authors). Since this analysis did not detect less inconsistency we concluded pooling of risk differences to be an adequate approach.

Our study has several strengths. We based our analyses on a comprehensive systematic review of systematic reviews. By using an outcome-focused approach, we were able to perform a detailed assessment of risk of bias and evidence quality, thereby emphasising the limitations of the current evidence base.

Limitations of our study mainly arise from the limitations in the systematic reviews and primary studies included. In particular, risk of bias and imprecision of the reported estimates might limit the scientific and clinical value of the data summarised here. The search by Alshaikh et al. [3] was last performed June 2012.

We did not conduct a more recent search for primary studies. Thus we may not exclude the possibility that more recent studies could have influenced our findings. Our results may be further improved by using the primary datasets of the included cohort studies and then adjusting for confounders compiled throughout all studies, such as sex, birth weight and gestational age.

In conclusion, this systematic review of systematic reviews shows that VLBW infants with sepsis during neonatal life have an increased risk of developing permanent neurological impairment during later life. The magnitude of this effect varies by outcome, while evidence quality was low to very low. To improve the evidence base, carefully planned and conducted prospective studies are needed.

Acknowledgements

Funding source: This study was conducted within the project ‘Evidence-based research to develop outcome trees for healthcare-associated infections’ initiated and funded by the European Centre for Disease Prevention and Control (ECDC) through a specific service contract (ECDC.4226) to the Robert Koch Institute, Public Federal Institute under jurisdiction of the German Health Ministry, Berlin, Germany. The authors would like to thank Benedetta Allegranzi (World Health Organization), Maria Deja (Charité University Medicine Berlin), and Andreas Jansen, Diamantis Plachouras and Edoardo Colzani (all ECDC), for discussion and comments.

Conflict of Interest

None declared

Authors’ contributions

Sebastian Haller, Thomas Harder conceptualised the study, performed systematic reviews, extracted the data, performed the statistical analysis and drafted the manuscript. Philipp Deindl gave scientific advice, reviewed the data and revised the manuscript. Alessandro Cassini conceptualised the study, reviewed the data and revised the manuscript. Carl Suetens gave scientific advice, reviewed the data and revised the manuscript. Walter Zingg gave scientific advice, reviewed the data and revised the manuscript. Muna Abu Sin reviewed the data, contributed to analysis and interpretation of the data and revised the manuscript. Bettina Weiss reviewed the data, contributed to analysis and interpretation of the data and revised the manuscript. Tanja Ducomble reviewed the data and contributed to analysis and interpretation of the data. Edward Velasco reviewed the data and contributed to analysis and interpretation of the data. Madlen Sixtensson reviewed the data and contributed to analysis and interpretation of the data. Tim Eckmanns conceptualised the study, coordinated the study, reviewed the data and revised the manuscript. All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

References
