Variations in patterns of care across neonatal units and their associations with outcomes in very preterm infants: the French EPIPAGE-2 cohort study

Veronique Pierrat,1,2 Antoine Burguet,3 Laetitia Marchand-Martin,1 Gilles Cambonie,4 Anaëlle Coquelin,1 JC Roze,5,6 Melanie Durox,7,1 Bernard Guillotis,7 Andrei S Morgan,8,1 Monique Kaminski1,1 on behalf of the Neurodevelopmental care study group of EPIPAGE-2

ABSTRACT

Objectives To describe patterns of care for very preterm (VP) babies across neonatal intensive care units (NICUs) and associations with outcomes.

Design Prospective cohort study, EPIPAGE-2.

Setting France, 2011.

Participants 53 (NICUs); 2135 VP neonates born at 27 to 31 weeks.

Outcome measures Clusters of units, defined by the association of practices in five neonatal care domains – respiratory, cardiovascular, nutrition, pain management and neurodevelopmental care. Mortality at 2 years corrected age (CA) or severe/moderate neuro-motor or sensory disabilities and proportion of children with scores below threshold on the neurodevelopmental Ages and Stages Questionnaire (ASQ).

Methods Hierarchical cluster analysis to identify clusters of units. Comparison of outcomes between clusters, after adjustment for potential confounders.

Results Three clusters were identified: Cluster 1 with higher proportions of neonates free of mechanical ventilation at 24 hours of life, receiving early enteral feeding, and neurodevelopmental care practices (26 units; n=1118 babies); Cluster 2 with higher levels of patent ductus arteriosus and pain screening (11 units; n=398 babies); Cluster 3 with higher use of respiratory, cardiovascular and pain treatments (16 units; n=619 babies). No difference was observed between clusters for the baseline maternal and babies’ characteristics. No differences in outcomes were observed between Clusters 1 and 3. Compared with Cluster 1, mortality at 2 years CA or severe/mild neuro-motor or sensory disabilities was lower in Cluster 2 (adjusted OR 0.46, 95% CI 0.25 to 0.84) but with higher proportion of children with an ASQ below threshold (adjusted OR 1.49, 95% CI 1.07 to 2.08).

Conclusion In French NICUs, care practices for VP babies were non-randomly associated. Differences between clusters were poorly explained by unit or population differences, but were associated with mortality and development at 2 years. Better understanding these variations may help to improve outcomes for VPT babies, as it is likely that some of these discrepancies are unwarranted.

INTRODUCTION

It is well described that some variations in clinical care are unwarranted because they cannot be explained by type or severity of illness or by patient preferences.1 Local medical opinion appears more important than science in determining how medical care is delivered. In the field of neonatal care, most of the decisions neonatologists have to take are for care where the evidence of benefit is not well established, or where possible benefit is accompanied by significant risk of adverse effects.2 3 Neonatal intensive care is an extremely complex care system requiring expertise in conventional fields of medicine as well as in ethics, in babies’ and parents’ physiological and emotional needs, and also in the development of a preterm neonate.1 With increasing knowledge on
the role of environmental exposures on newborn neurodevelopment and maternal-newborn bonding, identifying overuse of treatment with potential adverse effects has become more critical for neonatologists. Thanks to numerous collaborative quality improvement initiatives, uncontrolled variation has been described in the use of health services or conventional care practices, for example, variations in the use of antibiotics or ventilator treatment in Norway. To our knowledge, no study has tried to investigate how care practices in different areas of neonatal care are intertwined within units. In addition, practices are usually reported for babies born extremely preterm although babies born between 27 and 31 weeks of gestation (WG) represent a higher proportion of preterm babies at high risk of mortality and disabilities.

The EPiPAGE-2 cohort study was designed to measure survival and morbidity after very preterm birth in France. This study is a secondary analysis of EPiPAGE-2 data. We wanted to know if there were inter-relationships between the use of neurodevelopmental care and more conventional care practices—for example, if higher use of neurodevelopmental care was associated with less frequent use of invasive practices—and, if so, if this was associated with subsequent outcomes. Our first objective was thus to explore if patterns of units could be identified for babies born between 27 and 31 WG. The second objective was to report outcomes at discharge and at 2 years corrected age (CA) in relation to any clusters identified. We hypothesised that patterns of care within units are not distributed at random and that observing differences in outcomes could reveal opportunities to decrease adverse effects of unnecessary care.

**POPULATION AND METHODS**

EPiPAGE-2 study is a national population-based cohort study launched in France in 2011 and scheduled to follow children up to the age of 12 years. Eligible participants included all babies live or stillborn, and all terminations of pregnancy between 22 and 34 completed WG. Infants discharged alive were included in follow-up and evaluated at 2 years CA.

For this study, inclusion criteria were live birth between 27 and 31 WG, hospitalised in the same level III neonatal intensive care units (NICUs) until day 7 of age. Level III NICUs are located in centres that provide obstetric and ongoing neonatal intensive care. In addition, some of these NICUs also provide surgical care. Neonates were included between March and October 2011. Exclusion criteria were death in the delivery room, presence of severe congenital malformations that might affect survival, transfer to another unit before day 7, and admission to a NICU with fewer than 20 neonates included in the study.

**Patient and public involvement**

Patients were not involved in setting the research question or the outcome measures, nor were they involved in developing plans for design of the study. Parents demonstrated overwhelming support for the study through high follow-up rates. EPiPAGE-2 maintains contact with parents in the cohort through letters, newsletters and its website (https://epipage2.insERM.fr/index.php/fr/cote-parents/temoignages). National parents’ associations assisted with the dissemination of the results.

**Data collection**

Data were obtained through questionnaires completed in maternity units and throughout the neonatal hospitalisation by perinatal teams, and through medical and parental questionnaires at 2 years of age.

**Practices**

Evaluated care practices, collected during the first week of life, were categorised into five domains: three related to conventional care (respiratory, cardiovascular, nutrition), and two to neonates’ and parents’ developmental and emotional needs (pain and neurodevelopmental care). Practices, considered as markers of interest for these different domains, were: administration of surfactant and mechanical ventilation at 24 hours of life for the respiratory domain; systematic echocardiographic screening of patent ductus arteriosus (PDA) before day 3, treatment with vasoactive amines, and PDA treatment with ibuprofen for the cardiovascular domain; early enteral feeding (before day 2) for the nutrition domain; treatment with opioids, sedatives-hypnotics, or general anaesthetics (O-SH-GA), and at least one assessment of procedural or prolonged pain, for the pain domain; permanent incubator cover, kangaroo care during the first 3 days of life, parental involvement in feeding support (feed with support or swaddling by parents, or during skin-to-skin contact, or opportunity for the baby to suck a dummy offered by parents during tube feeding) and breast contact (with or without nutritive or non-nutritive sucking) for neurodevelopmental care. Most care practices that were studied were considered markers of evidence-based quality during the time period of the study but appropriate utilisation rates are unknown. In the group of conventional care, all can be considered as ‘necessity care’ for some infants (treatment rate reflects the prevalence of a clinical condition in the population), but as ‘preference-sensitive care’ for others (that is, indications and health benefits are unclear or controversial within the medical community). For example, mechanical ventilation at 24 hours of life is dependent not only on the child’s respiratory condition but also on medical opinion towards early weaning from mechanical ventilation. Variations between units in conventional care may be observed but should be limited. Cares studied in the neurodevelopmental care domain respond mainly to the definition of ‘preference-sensitive care’. For example, kangaroo care before day 3 depends on the clinical condition of the child and is highly dependent on the team opinion. Greater variations between units were thus expected in this domain. For ethical reasons, assessment of pain, in this highly vulnerable population, should be
close to 100% but treatment with O-SH-GA is dependent on unit culture and case-mix.

Maternal, obstetric and neonatal characteristics

Maternal characteristics were: age (years), birth in France, parents’ socio-economic status (professional, intermediate, administrative or public service, self-employed or student, shop assistants or service workers, manual workers and unknown occupation); obstetric characteristics: singleton pregnancy, antenatal steroids and vaginal delivery; neonatal characteristics: gestational age (GA, weeks), sex (male/female) and small-for-gestational age (SGA) defined as birth weight less than the 10th percentile for GA and sex based on French intrauterine growth curves17 and severe neonatal morbidity,18 including any of the following complications: severe bronchopulmonary dysplasia (administration of oxygen for at least 28 days plus need for 30% or more oxygen and/or mechanical ventilation or continuous positive airway pressure at 36 weeks’ postmenstrual age), necrotising enterocolitis stage 2 to 3, severe retinopathy of prematurity stage >3 or any of the following severe cerebral abnormalities on cranial ultrasonography: intraventricular haemorrhage grade III or IV or cystic periventricular leukomalacia.

At 2 years of age

A medical questionnaire collected information on cerebral palsy (CP) and sensory deficits (bilateral or unilateral blindness or deafness).19 CP was defined according to the Surveillance of Cerebral Palsy in Europe network,20 and severity classified with Gross Motor Function Classification System (GMFCS).21 Severe neuro-motor or sensory disabilities were defined as non-ambulatory CP (GMFCS level 3 to 5) or severe visual or auditory impairment; moderate disability included GMFCS level 2 CP and/or moderate visual or auditory impairment.19 The parental questionnaire included the second version of the 24-month Ages and Stages Questionnaire (ASQ),22 covering five developmental domains. ASQs were analysed if completed between 22 and 26 months CA in children without CP, deafness or blindness. Results are reported as ASQ score below threshold, defined as a score lower than two SD from the mean for any of the five ASQ domains.22
Outcomes
Outcomes are reported for babies admitted to NICUs and for survivors at 2 years CA. We first consider mortality and mortality or severe neonatal morbidities at hospital discharge, and mortality and mortality or severe/moderate neuro-motor or sensory disabilities at 2 years CA. We also report proportions of children with CP, and proportions of children with an ASQ below threshold at 2 years CA.

Statistical analysis
To identify clusters of units, we first calculated observed proportions of each practice in each unit, using estimated expected proportions to take into account differences in the populations cared for in each unit. Expected proportions were obtained using logistic regression models including a priori identified confounders (online supplementary table 1). Units were then classified into clusters using ascending hierarchical analysis, carried out on observed/expected rather than observed proportions. Second, we compared practices for clusters of units, after adjustment for potential confounders. To help understand differences between the clusters, we present comparisons of unit and individual (maternal, obstetrical and neonatal) characteristics. Third, we describe
rubin's rule.25 All tests were two-30 iterations each. Estimates were pooled according to

we generated 50 independent imputed data sets with

SAS software (V.9.4).

considered significant. All analyses were performed with


Imputation (MI) procedure.24 Imputation model vari-
tions using the Statistical Analysis System (SAS) Multiple

imputation. Missing data were imputed by chained equa-

tions after adjustment for a priori identified potential

confounders (maternal age, maternal country of birth,

parents’ socio-economic status, singleton pregnancy,

antenatal corticosteroids, mode of delivery, GA, sex and

SGA). To account for the non-independence of babies

within units, generalised estimating equations were used.
Results are given for complete cases and after multiple

imputation. Missing data were imputed by chained equa-
tions using the Statistical Analysis System (SAS) Multiple

Imputation (MI) procedure.24 Imputation model vari-
ables included both those potentially predicting non-

response and/or outcomes (maternal age and country of

birth, parity, parental socio-economic status, antenatal

steroids, caesarean section, multiple pregnancy, GA, sex,

SGA, inborn status, surfactant, postnatal steroids, severe

neonatal morbidities and use of breast milk at discharge),
and outcomes (CP, neuro-motor or sensory disabilities

and ASQ score below threshold), as previously reported.19
We generated 50 independent imputed data sets with
30 iterations each. Estimates were pooled according to
Rubin’s rule.25 All tests were two-sided with P values<0.05
considered significant. All analyses were performed with
SAS software (V.9.4).

results
population
Among the overall cohort, 2479 neonates were born alive
in a level III NICU between 27 and 31 WG. After applying
exclusion criteria, 2135 were included in the study (figure
1). At 2 years CA, 2024 children were eligible for follow-up; medical and parental questionnaires were
available for 1717 and 1747, respectively, with ASQ data
suitable for analysis for 1225 children.

distribution of care between units
Of the 66 level III NICUs existing in France in 2011, 13
were excluded because <20 babies were eligible for
inclusion in this study. Large variabilities were observed
between units in the administration of care in the five
evaluated domains (figure 2). For example, median
and (IQR) were 23% (15 to 46) for systematic echocar-
diographic screening of PDA before day 3, 33% (22 to
45) for treatment with O-SH-GA, and 29% (17 to 52) for
kangaroo care during the first 3 days of life. Systematic
PDA screening was never reported for babies born after
29 weeks. In the hierarchical analysis, three clusters of
units were identified (figure 3). Half of the units (26/53)
were in Cluster 1. The distribution of the studied practices
in each investigated domain and by cluster is reported
in table 1. Higher proportions of infants weaned from
mechanical ventilation before 24 hours of life, receiving
early enteral feeding and neurodevelopmental care prac-
tices were observed in Cluster 1, higher screening of PDA
and of pain in Cluster 2, and higher use of respiratory,
cardiovascular and pain treatments in Cluster 3. The
mean length of stay in the first unit was 49 days (SD 31),
44 days (SD 27) and 45 days (SD 29) in Clusters 1, 2 and
3, respectively (p=0.001).

units' characteristics by cluster
Differences between clusters were observed for the avail-
ability of neonatal surgery and training in neurodevelop-
mental care (table 2).
In Clusters 1 and 2, similar proportions of units
provided neurodevelopmental care training to staff, but
the types of training were different. Units in Cluster 3 had
a lower availability of neonatal surgery, and nearly 60% of
did not provide any training in neurodevelopmental care.

maternal and infant characteristics by clusters
Differences between clusters were observed for maternal
place of birth, mode of delivery and babies’ sex; the
GA distribution between clusters was not significantly
different (table 3).

outcomes
At 2years CA, children without missing data for CP or
ASQ were born more frequently to mothers with higher
socio-economic status than children with missing data,
but neonatal characteristics were similar (online supple-
mentary table 2); proportions of children with missing
data were also similar among clusters (online supplemen-
tary table 3). At discharge, no difference in outcomes was

Figure 3  Dendrogram showing the distribution of NICUs among three clusters. Hierarchical cluster analysis was used to classify NICUs on the 13 ratios ‘observed / expected’ percentages of practices. The classification was performed using Ward’s method with Euclidean distance. The dendrogram illustrates the results of the cluster analysis. Three main clusters were identified. NICU, neonatal intensive care unit.
<table>
<thead>
<tr>
<th>Table 1</th>
<th>Proportions of practices in each investigated domain for the study population and by cluster of units</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cluster * 1 n=1118 infants</td>
</tr>
<tr>
<td>Respiratory</td>
<td></td>
</tr>
<tr>
<td>Surfactant administration</td>
<td>612/1108 55.2</td>
</tr>
<tr>
<td>Mechanical ventilation at 24 hours of life</td>
<td>177/1096 16.2</td>
</tr>
<tr>
<td>Cardiovascular</td>
<td></td>
</tr>
<tr>
<td>Systematic screening of PDA with echocardiography before day 3</td>
<td>240/1088 22.1</td>
</tr>
<tr>
<td>Treatment with vasoactive amines</td>
<td>44/1103 4.0</td>
</tr>
<tr>
<td>PDA treatment with ibuprofen</td>
<td>186/1087 17.1</td>
</tr>
<tr>
<td>Nutrition</td>
<td></td>
</tr>
<tr>
<td>Early enteral feeding (before day 2)</td>
<td>877/1069 82.0</td>
</tr>
<tr>
<td>Pain</td>
<td></td>
</tr>
<tr>
<td>Treatment with O-SH-GA</td>
<td>342/1113 30.7</td>
</tr>
<tr>
<td>Procedural pain assessment (at least one assessment during the first week of life)</td>
<td>389/1118 34.8</td>
</tr>
<tr>
<td>Prolonged pain assessment (at least one assessment during the first week of life)</td>
<td>785/1118 70.2</td>
</tr>
<tr>
<td>Neurodevelopmental care</td>
<td></td>
</tr>
<tr>
<td>Permanent incubator cover during the first week of life</td>
<td>889/1011 87.9</td>
</tr>
<tr>
<td>Kangaroo care during the first 3 days of life</td>
<td>443/1038 42.7</td>
</tr>
<tr>
<td>Parental involvement in feeding support †† during the first week of life</td>
<td>495/987 50.2</td>
</tr>
<tr>
<td>Breast contact during the first week of life</td>
<td>138/1031 13.4</td>
</tr>
</tbody>
</table>

*Clusters are defined according to the distribution of care within units using a hierarchical analysis.
†Adjusted for gestational age, antenatal corticosteroids and small-for-gestational age.
‡Adjusted for gestational age and hypotension.
§Adjusted for gestational age, small-for-gestational age and normal bowel frequency defined as at least one stool per day.
¶Adjusted for gestational age, type of pregnancy and mode of delivery.
††Parental involvement in feeding support was defined as a feed with support or swaddling by parents, or during skin-to-skin contact, or opportunity for the baby to suck a dummy offered by parents during tube feeding. Data are number of events/number in group and percentages. P values were estimated from a logistic model adjusted for specified variables.
O-SH-GA, opioids, sedatives-hypnotics, or general anaesthetics; PDA, patent ductus arteriosus.

observed between Clusters 1 and 3 (table 4). Mortality was lowest in Cluster 2, with no difference between clusters in proportions of children who died or had severe neonatal morbidity. At 2 years CA, proportions of CP were no different between clusters but a higher proportion of children with an ASQ below threshold was observed in Cluster 2. After multiple imputation rates of CP were only slightly modified, a consistent increase was observed in each cluster in rates of ASQ scores below threshold.

**DISCUSSION**

In this population-based cohort of babies born between 27 and 31 WG, we found variability in care practices between
units. This occurred not only in the use of individual practices but also in which combinations of practices were used within units. Three clusters were identified with few differences between them in terms of baseline population characteristics. Despite different strategies of care, similar outcomes were observed between Clusters 1 and 3.

Cluster 2 had the lowest mortality at discharge but also the highest proportion of children with an ASQ below threshold at 2 years CA. This may be viewed as a limitation. Therefore results of the ASQ were not included in a composite outcome at 2 years CA to describe children with intact survival. Of note, unlike mortality, having an ASQ below threshold is not a rare event and the OR slightly overestimates the relative risk. Another limitation is that the investigation was restricted to care delivered during the first week of life as we were limited to practices collected in EPIPAGE-2. On the other hand, this also targets the most vulnerable time period for VPT babies. Particularly, the respiratory and cardiovascular practices studied are most reflective of intensive care provided during the first week of life. We were unable to quantify early non-invasive respiratory support. Recommendations, published after data collection commenced, are that protocols should be directed at avoiding mechanical ventilation where possible. Hence low rates of mechanical ventilation at 24 hours of life may suggest the use of less invasive strategies in line with the implementation of these recommendations. Defining neurodevelopmental care with practices only is not ideal and does not consider whether units individually care or have a family-centred care philosophy—both core concepts of neurodevelopmental care. Conversely, a high level of implementation of neurodevelopmental care practices has been considered as a marker for a unit’s ‘state of mind’, and our strategy to describe implementation of neurodevelopmental care may be helpful at a population level. We also did not consider whether babies were transferred to another hospital. However, the mean length of stay for babies included in our study was

### Table 2 Units' characteristics according to the three clusters of units

<table>
<thead>
<tr>
<th></th>
<th>Cluster * 1: n=26 units</th>
<th>Cluster * 2: n=11 units</th>
<th>Cluster * 3: n=16 units</th>
<th>P value†</th>
</tr>
</thead>
<tbody>
<tr>
<td>NICU only ‡</td>
<td>14/26 53.8</td>
<td>8/11 72.7</td>
<td>12/16 75.0</td>
<td>0.31</td>
</tr>
<tr>
<td>Annual number of admission before 31 weeks in 2011, median (IQR)</td>
<td>10 (75 to 118)</td>
<td>11 (64 to 112)</td>
<td>12 (64 to 119)</td>
<td>0.74</td>
</tr>
<tr>
<td>NICU only without paediatric intensive care unit</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Clusters are defined according to the distribution of care within units using a hierarchical analysis.
†χ² test for categorical variables or Kruskal-Wallis tests for quantitative variable.
‡Units admitting neonates only, without a paediatric intensive care unit.
NIDCAP, Neonatal Individualised Developmental Care and Assessment Program.
The magnitude of absolute difference in care practices between clusters is difficult to interpret. For example, a 15% difference for surfactant between Cluster 2 and 3 may be considered a small difference, but from an economic perspective, with regard to the cost of the surfactant, it could be considered big; more than 20% difference for kangaroo care during the first 3 days of life may be viewed as important for the infant neurodevelopment but also for parental bonding; and variations in the use of vasopressors was interesting as this situation is rare. Treatment of shock and hypotension is an area of neonatology where there is great uncertainty in identifying which patients would benefit from treatment. Grouping the units provides an opportunity to observe differences and to reflect on relatively high, and fewer than 50% were transferred after the first week of life (data not shown). The rate of loss to follow-up was another limitation, although the follow-up rate was high if one considers the size and the geographical dispersion of the cohort. We used multiple imputation to account for missing data; ORs were in the same direction in the complete cases analysis and after multiple imputation. We thus find it plausible that the results we observed are valid and that health outcome reflect units’ policies. Finally, the paucity of information we had on ‘supply-sensitive care’ (referring to medical services for which usage rates are sensitive to the local availability of healthcare resources) such as healthcare professionals’ availability was an obvious limitation.

Table 3  Maternal and infant characteristics for the study population and by cluster units

<table>
<thead>
<tr>
<th>Maternal characteristics</th>
<th>Cluster * 1 n=1118 infants</th>
<th>Cluster * 2 n=398 infants</th>
<th>Cluster * 3 n=619 infants</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal age</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;25 years</td>
<td>221/1118 19.8</td>
<td>67/398 16.8</td>
<td>111/619 17.9</td>
<td>0.29</td>
</tr>
<tr>
<td>25 to 34 years</td>
<td>640/1118 57.2</td>
<td>248/398 62.3</td>
<td>353/619 57.0</td>
<td></td>
</tr>
<tr>
<td>&gt;35 years</td>
<td>257/1118 23.0</td>
<td>83/398 20.9</td>
<td>155/619 25.0</td>
<td></td>
</tr>
<tr>
<td>Birth in France</td>
<td>879/1100 79.9</td>
<td>286/394 72.6</td>
<td>454/615 73.8</td>
<td>0.002</td>
</tr>
<tr>
<td>Parents’ socio-economic status †</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Professional</td>
<td>238/1055 22.6</td>
<td>97/382 25.4</td>
<td>122/596 20.5</td>
<td>0.053</td>
</tr>
<tr>
<td>Intermediate</td>
<td>213/1055 20.2</td>
<td>82/382 21.5</td>
<td>129/596 21.6</td>
<td></td>
</tr>
<tr>
<td>Administrative, public service, self-employed and students</td>
<td>288/1055 27.3</td>
<td>122/382 31.9</td>
<td>160/596 26.8</td>
<td></td>
</tr>
<tr>
<td>Shop assistants and service workers</td>
<td>132/1055 12.5</td>
<td>35/382 9.2</td>
<td>92/596 15.4</td>
<td>0.72</td>
</tr>
<tr>
<td>Manual workers</td>
<td>143/1055 13.6</td>
<td>35/382 9.2</td>
<td>71/596 11.9</td>
<td></td>
</tr>
<tr>
<td>Unknown occupation</td>
<td>41/1055 3.9</td>
<td>11/382 2.9</td>
<td>22/596 3.7</td>
<td></td>
</tr>
<tr>
<td>Obstetric factors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Singleton pregnancy</td>
<td>753/1118 67.4</td>
<td>269/398 67.6</td>
<td>406/619 65.6</td>
<td></td>
</tr>
<tr>
<td>Antenatal steroids</td>
<td>995/1104 90.1</td>
<td>359/394 91.1</td>
<td>531/607 87.5</td>
<td></td>
</tr>
<tr>
<td>Vaginal delivery</td>
<td>288/1113 25.9</td>
<td>136/393 34.6</td>
<td>169/615 27.5</td>
<td></td>
</tr>
<tr>
<td>Neonatal characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gestational age, weeks</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>168/1118 15.0</td>
<td>47/398 11.8</td>
<td>87/619 14.1</td>
<td>0.26</td>
</tr>
<tr>
<td>28</td>
<td>185/1118 16.5</td>
<td>65/398 16.3</td>
<td>105/619 17.0</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>209/1118 18.7</td>
<td>74/398 18.6</td>
<td>113/619 18.3</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>275/1118 24.6</td>
<td>113/398 28.4</td>
<td>132/619 21.3</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>281/1118 25.1</td>
<td>99/398 24.9</td>
<td>182/619 29.4</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>557/1118 49.8</td>
<td>228/398 57.3</td>
<td>347/615 56.1</td>
<td>0.008</td>
</tr>
<tr>
<td>Small-for-gestational age ‡</td>
<td>445/1118 39.8</td>
<td>170/398 42.7</td>
<td>248/619 40.1</td>
<td>0.58</td>
</tr>
</tbody>
</table>

Data are number of events/number in each group and percentage.

*Clusters are defined according to the distribution of care within units using a hierarchical analysis.
†Defined as the highest occupational status of the mother and father, or mother only if living alone.
‡Small-for-gestational age was defined as birth weight less than the 10th percentile for gestational age and sex based on French intrauterine growth curves (Ego 2016).
Table 4: Outcome at discharge from NICUs and at 2 years CA in the study population by cluster of units

<table>
<thead>
<tr>
<th>Cluster * 1</th>
<th>Cluster * 2</th>
<th>Cluster * 3</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infants admitted to NICU</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>At discharge</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mortality</td>
<td>63/1118</td>
<td>5.6 %</td>
<td>9/398</td>
</tr>
<tr>
<td>aOR (95% CI)†</td>
<td>1</td>
<td>0.43 (0.20 to 0.93)</td>
<td>0.88 (0.51 to 1.54)</td>
</tr>
<tr>
<td>Mortality or severe neonatal morbidity‡</td>
<td>197/1118</td>
<td>17.7 %</td>
<td>58/398</td>
</tr>
<tr>
<td>aOR (95% CI)†</td>
<td>1</td>
<td>0.84 (0.49 to 1.42)</td>
<td>0.93 (0.62 to 1.39)</td>
</tr>
<tr>
<td><strong>At 2 years CA</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mortality</td>
<td>68/1118</td>
<td>6.1 %</td>
<td>9/398</td>
</tr>
<tr>
<td>aOR (95% CI)†</td>
<td>1</td>
<td>0.39 (0.18 to 0.85)</td>
<td>0.96 (0.56 to 1.67)</td>
</tr>
<tr>
<td>Mortality or severe/moderate neuro-motor or sensory disabilities§</td>
<td>103/1118</td>
<td>9.3 %</td>
<td>17/398</td>
</tr>
<tr>
<td>aOR (95% CI)†</td>
<td>1</td>
<td>0.46 (0.25 to 0.84)</td>
<td>1.13 (0.78 to 1.63)</td>
</tr>
<tr>
<td>Survivors at 2 years CA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cerebral palsy</td>
<td>54/1050</td>
<td>5.1 %</td>
<td>14/389</td>
</tr>
<tr>
<td>aOR (95% CI)†</td>
<td>1</td>
<td>0.57 (0.21 to 1.55)</td>
<td>0.81 (0.44 to 1.50)</td>
</tr>
<tr>
<td>ASQ below threshold ¶</td>
<td>420/989</td>
<td>42.5 %</td>
<td>195/375</td>
</tr>
<tr>
<td>aOR (95% CI)†</td>
<td>1</td>
<td>1.49 (1.07 to 2.08)</td>
<td>1.17 (0.89 to 1.55)</td>
</tr>
<tr>
<td><strong>Complete cases analysis</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cerebral palsy</td>
<td>43/884</td>
<td>4.9 %</td>
<td>11/327</td>
</tr>
<tr>
<td>aOR (95% CI)†</td>
<td>1</td>
<td>0.63 (0.23 to 1.74)</td>
<td>0.86 (0.47 to 1.56)</td>
</tr>
<tr>
<td>ASQ below threshold ¶</td>
<td>245/650</td>
<td>37.7 %</td>
<td>109/232</td>
</tr>
<tr>
<td>aOR (95% CI)†</td>
<td>1</td>
<td>1.51 (1.07 to 2.14)</td>
<td>1.19 (0.86 to 1.64)</td>
</tr>
</tbody>
</table>

Data are number of events/number in each group and percentage, unless otherwise noted. Generalised Estimating Equations are used to take into account NICUs effects. Results are based on multiple imputations unless noted.

*Clusters are defined according to the distribution of care within units using a hierarchical analysis.
†Adjusted for maternal age, maternal country of birth, type of pregnancy, mode of delivery, antenatal corticosteroids, GA, sex, small-for-gestational age (defined as birth weight less than the 10th percentile for GA and sex based on French intrauterine growth curves (Ego 2016)) and parents’ socio-economic status.
‡Severe neonatal morbidity was defined as severe bronchopulmonary dysplasia or necrotising enterocolitis stage 2 to 3 or severe retinopathy of prematurity stage >3 or any of the following severe cerebral abnormalities on cranial ultrasonography: intraventricular haemorrhage grade III or IV or cystic periventricular leukomalacia (Ancel 2015).
§Moderate or severe neuro-motor or sensory disabilities. Severe neuro-motor or sensory disabilities: cerebral palsy GMFCS levels 3 to 5 and/or bilateral deafness and/or bilateral blindness; moderate neuro-motor or sensory disabilities: cerebral palsy GMFCS level 2 and/or unilateral deafness and/or unilateral blindness.
¶For each of the five domains of ASQ, a score of less than 2 SD below threshold of the US ASQ-3 reference was identified. If a score was below threshold in at least one domain, the ASQ was considered below threshold. Infants with cerebral palsy, deafness, blindness or severe congenital anomalies were excluded.

aOR, adjusted OR; ASQ, Ages and Stages Questionnaire; CA, corrected age; GA, gestational age; GMFCS, Gross Motor Function Classification System; NICU, neonatal intensive care unit.

practices. Nevertheless, it is interesting to note that for each practice except PDA treatment, differences between clusters, adjusted for the main confounders, were highly significant. Even if differences between each practice may be viewed as minimal, the association of small differences in different practices, leading to a team culture, appears to have an impact on health outcomes. Results also partly support our hypothesis. The highest implementation of neurodevelopmental care was observed in Cluster 1 which was also the cluster with the lowest proportions of conventional respiratory care, as well as low proportions of treatment with vasoactive amines or O-SH-GA. Cluster 3 was characterised by high conventional treatment rates but had the lowest rates of neurodevelopmental care provision. An interesting finding was the absence of differences in outcomes between Clusters 1 and 3. Patterns of care in Cluster 3 could be defined as more invasive than in Cluster 1. This may suggest an overuse of care in Cluster 3 and thus could offer opportunities for decreasing adverse effects and reducing unnecessary spending in such units. This could also mean that some babies are exposed to needless days of intensive care.
increasing the risk of adverse effects associated with care and of interference with bonding and attachment.\textsuperscript{33} Identification of Cluster 2 was less expected. It was characterised by increased use of screening practices for PDA and pain and this could generate new hypotheses. The lower mortality rate observed in Cluster 2 deserves attention. Our group has previously shown that systematic screening of PDA was associated with a lower mortality in neonates born between 24 and 29 WG\textsuperscript{34} and we add a new perspective to this previous study. The difference in mortality should be explored in more detail but is somewhat counterbalanced by the increased number of children at risk of developmental delay at 2 years CA.

**Implications for clinicians and policymakers**

It has been proposed that greater reductions in morbidity may be achieved by concentrating on the best rather than the worst performing hospitals.\textsuperscript{35} Our results highlight the difficulties in defining the ‘best’ hospitals when considering the complexity of neonatal care and interventional strategies to improve care developed in accordance with recently published guidelines should be explored.\textsuperscript{36} Identifying patterns of care across NICUs appears to have the potential to reduce overuse and costs, and improve outcomes through the application of current medical knowledge early in life. The results also emphasise the complexity of neonatal care, demonstrate the difficulty of achieving high quality of care in every domain, and highlight the importance of well-resourced routine data collection and benchmarking.

**CONCLUSION**

This study, derived from a large national cohort, describes variations in patterns of care between NICUs associated with differences in outcomes for children born between 27 and 31 WG. Most of these variations are likely due to hospital organisations and clinical styles of practices. The interaction between patterns of care and regulatory, organisational and unit cultural factors should be investigated in more detail to better understand pathways of care implementation in everyday practice.

**Author affiliations**

1Equipe EPOpE, U 1153, Université de Paris, CRESS, INSERM, INRA, Paris, France
2Department of Neonatal Medicine, Jeanne de Flandre Hospital, CHU Lille, Lille, France
3Department of Neonatal Pediatrics, Dijon University Hospital, Dijon, France
4Neonatology and Neonatal Intensive Care Unit, Montpellier University Hospital Centre, Montpellier cedex 5, France
5Paediatric Intensive Care, University Hospital Centre Nantes Clinic of Medical Paediatrics, Nantes, France
6Centre d’Investigation Clinique (CIC004), University Hospital Centre Nantes, Nantes, France
7Department of Neonatal Pediatrics and Intensive Care, University Hospital, Caen, France
8Institute for Women’s Health, University College London, London, UK

**Acknowledgements** We are grateful for the participation of all families of preterm infants in the EPINAGE-2 cohort study and for the cooperation of all maternity and neonatal units in France. We thank parents’ associations (SOS prema, Collectif interassociatif autour de la naissance (CIANE), Jumeaux et plus) for their overwhelming support and their involvement in the dissemination of the results. We thank the EPINAGE-2 Study Group for its substantial contribution to the conception, design and acquisition of data.

**Collaborators** Study group: Neurodevelopmental care study group of EPINAGE-2: C Arnaud, Inserr U 1027, F-31000 France; Paul-Sabatier University, Toulouse, F-31400 France; Purpan, Clinical Epidemiology Unit, Toulouse, F-31300 France; A Burguet, Department of Neonatal Pediatrics, Dijon University Hospital, Dijon, France; L Caeymaex, Department of Neonatal Medicine, CHIC de Créteil, Centre de recherche clinique CHIC, CEDITEC Paris Est Créteil University, France; G Cambonie, Department of Neonatal Medicine, Montpellier University Hospital, Montpellier, France; V Datin-Dorrière, Department of Neonatal Pediatrics and Intensive Care, University Hospital, Caen, France; C Girre, Department of Neonatal Pediatrics and Intensive Care, Nord Hospital, Marseille, France; B Guillou, Department of Neonatal Pediatrics and Intensive Care, University Hospital, Caen, France; P Kuhn, Department of Neonatal Pediatrics and Intensive Care, Strasbourg University Hospital, Strasbourg, France; A Mitha, CHU Lille, Department of Neonatal Medicine, Jeanne de Flandre Hospital, F-59000 Lille, France. V Pierrat, MD PhD, Obstetrical, Perinatal, and Pediatric Epidemiology Team, Epidemiology and Biostatistics Sorbonne Paris Cité Research Center (U1153), INSERM, Paris, France; Paris Descartes University, Paris, France. CHU Lille, Department of Neonatal Medicine, Jeanne de Flandre Hospital, F-59000 Lille, France. JC Roze, Department of Neonatal Medicine, Nantes University Hospital, Nantes, France. EPIDÉMIologie Clinique, Centre d’Investigation Clinique (CIC004), Nantes University Hospital, Nantes, France. JM Roué, MD PhD, Department of Neonatal Pediatrics and Intensive Care, Pôle de la Femme, de la Mère et de l’Enfant, Brest University Hospital, Brest, France. J Sizun, MD PhD, Department of Neonatal Pediatrics and Intensive Care, Pôle de la Femme, de la Mère et de l’Enfant, Brest University Hospital, Brest, France.

**Contributors** AB, VP and MK conceptualised and designed the study, take responsibility for the integrity of the data and the accuracy of the data analysis, drafted the initial manuscript and reviewed and revised the manuscript. LM-M and AC had full access to all the data in the study, performed the statistical analysis, reviewed and revised the manuscript. MD coordinated data collection, had responsibility for technical support, reviewed and revised the manuscript. ASM contributed to the analysis plan and interpretation of the results and critically reviewed the manuscript for important intellectual content. GC, BG and JCR conceptualised and designed the study, contributed to the analysis plan and interpretation of the results and reviewed and revised the manuscript. All members of the Neurodevelopmental Care Study Group of EPINAGE-2 were involved in the regional organisation for data collection, the design of the study, reviewed and revised the manuscript. All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

**Funding** This work was supported by (1) The French institute of Public Health research/institute of Public Health and its partners: the French Health Ministry, the National Institute of Health and Medical Research (INSERM); (2) The French EQUIPEX program of investments in the future coordinated by the National Research Agency; (3) The Fondation de France. N° 00050329; (4) Fondation pour la Recherche Médicale (N° SPF20160936356).

**Competing interests** None declared.

**Patient consent for publication** Not required.

**Ethics approval** This study was approved by the National Data Protection Authority (CNIL no.911009) and by appropriate ethics committees (Consultative Committee on the Treatment of Data on Personal Health for Research Purposes - reference no. 10.626, Committee for the Protection of People Participating in Biomedical Research - reference OPP SC-2673), Recruitment and data collection occurred only after families had received information and agreed to participate. The need for written consent was waived by the authorising authorities, as this was an observational study only with no active interventions. At hospital discharge following initial hospitalisation, parents of surviving children were given written information about the study, including contact details of the coordinating office, and informed they could withdraw from further follow-up at any stage. This was further approved by the CPP at the time of the 2 years follow-up.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available upon reasonable request. The EPINAGE studies are subject to a data sharing policy that may be downloaded from http://epinage2.inserr.fr/index.php/fr/cote-recherche/acces-aux-donnees-epipage-2.
REFERENCES


