

BMJ Open 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 ¹, Bernard Guillois,⁷ Andrei S Morgan ^{1,8}, Monique Kaminski ¹, on behalf of the Neurodevelopmental care study group of EPIPAGE-2

To cite: Pierrat V, Burguet A, Marchand-Martin L, *et al*. Variations in patterns of care across neonatal units and their associations with outcomes in very preterm infants: the French EPIPAGE-2 cohort study. *BMJ Open* 2020;**10**:e035075. doi:10.1136/bmjopen-2019-035075

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2019-035075>).

VP and AB contributed equally.

Received 17 October 2019

Revised 13 April 2020

Accepted 18 May 2020



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Dr Veronique Pierrat;
veronique.pierrat@inserm.fr

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 cofounders.

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/moderate 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.

Strengths and limitations of this study

- EPIPAGE-2 is the first national cohort study to report variability of neonatal care practices for very preterm babies born at 27 to 31 weeks' gestation and examine how they are related in neonatal intensive care units.
- Practices were analysed using five domains of care: respiratory, cardiovascular, nutrition, pain and neurodevelopmental care.
- Hierarchical cluster analysis was used to examine the association between domains within units in France and clusters' relationships with outcomes at hospital discharge and at 2 years corrected age are reported.
- The description of care domains was limited to variables collected for the study and specific pathways of care implementation were not explored.
- A lack of detailed information on organisational and unit cultural factors limit the understanding of pathways leading to different care patterns.

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.⁴ With increasing knowledge on

the role of environmental exposures on newborn neurodevelopment and maternal-newborn bonding,⁵ identifying overuse of treatment with potential adverse effects has become more critical for neonatologists. Thanks to numerous collaborative quality improvement initiatives,^{6,7} unwarranted variation has been described in the use of health services or conventional care practices,^{2,3,8,9} for example, variations in the use of antibiotics or ventilator treatment in Norway.³ 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^{10–12} 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.¹³ 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.¹³ 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,¹⁴ 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 'necessary care'² for some infants (treatment rate reflects the prevalence of a clinical condition in the population), but as 'preference-sensitive care'² 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.^{15,16} Greater variations between units were thus expected in this domain. For ethical reasons, assessment of pain, in this highly vulnerable population, should be

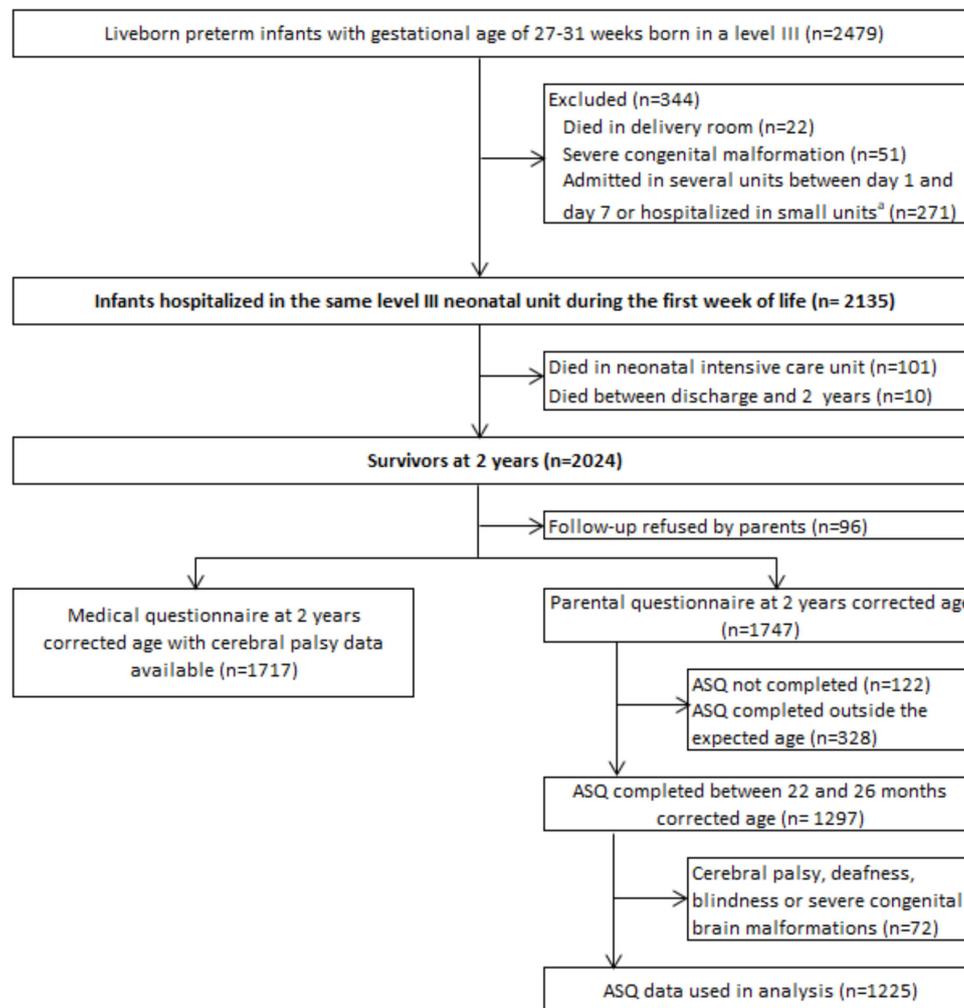


Figure 1 Flowchart of the study population. ASQ, Ages and Stages Questionnaire. ^aSmall units are units with less than 20 inclusions in EPIPAGE-2.

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

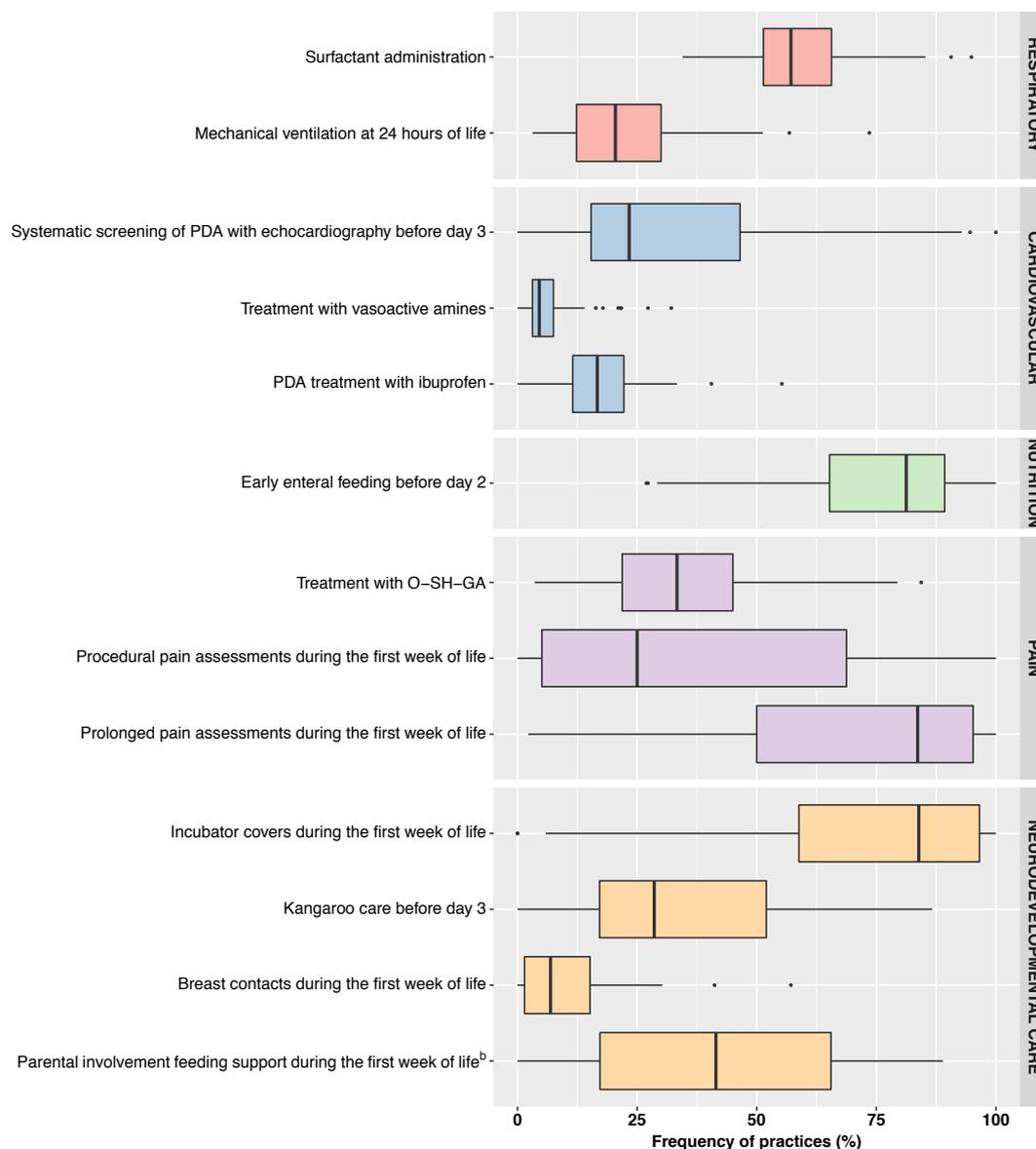


Figure 2 Distribution of the frequency of care practices in units for the study population a. The statistical unit is the neonatal unit (n=53). The vertical bar inside each box is the median, the right and left of the box indicate the IQR, the - bars indicate the 95th percentiles, and the circles indicate outliers. ^aNeonates born between 27 and 31 weeks of gestation, admitted between day 0 and 7 in a single level III neonatal intensive care unit and after exclusion of neonates with severe congenital malformations, as well as neonates born in units with less than 20 inclusions in EPIPAGE-2. ^bParental involvement in feeding 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 proposed by parents during tube feeding. O-SH-GA, opioids, sedatives-hypnotics, or general anaesthetics; PDA, patent ductus arteriosus.

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,²³ 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

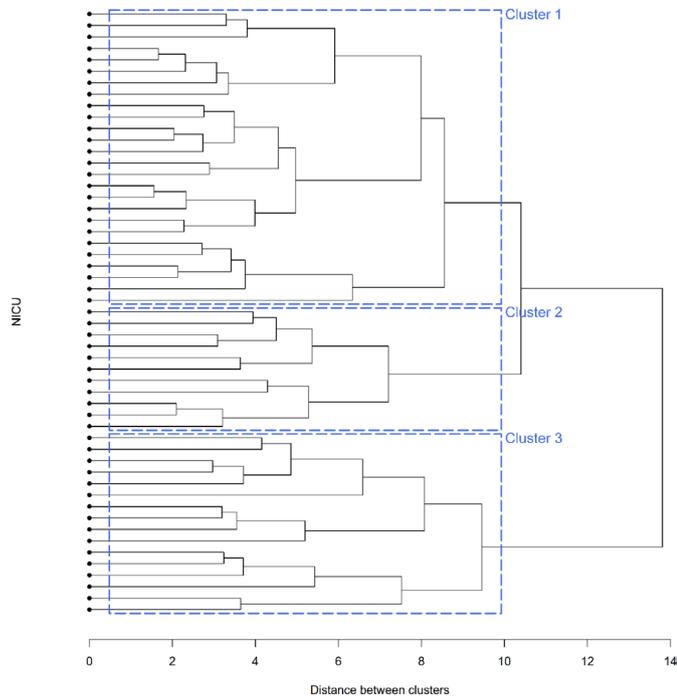


Figure 3 Dendrogram showing the distribution of NICUs among three clusters. Hierarchical cluster analysis was used to classify NICUs on the 13 ratios ‘observed / expected’ percentages of practices. The classification was performed using Ward’s method with Euclidean distance. The dendrogram illustrates the results of the cluster analysis. Three main clusters were identified. NICU, neonatal intensive care unit.

maternal, obstetric and infant characteristics for children with and without missing data for CP and ASQ as well as the proportions of missing data for CP and ASQ in each cluster. We then compared outcome measures between clusters 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 equations using the Statistical Analysis System (SAS) Multiple Imputation (MI) procedure.²⁴ Imputation model variables 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 echocardiographic 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 practices 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 availability of neonatal surgery and training in neurodevelopmental 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 2 years 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 supplementary table 2); proportions of children with missing data were also similar among clusters (online supplementary table 3). At discharge, no difference in outcomes was

**Table 1** Proportions of practices in each investigated domain for the study population and by cluster of units

	Cluster * 1 n=1118 infants		Cluster * 2 n=398 infants		Cluster * 3 n=619 infants		Adjusted p value
Respiratory							
Surfactant administration	612/1108	55.2	213/395	53.9	423/615	68.8	<0.001†
Mechanical ventilation at 24 hours of life	177/1096	16.2	86/389	22.1	219/601	36.4	<0.001†
Cardiovascular							
Systematic screening of PDA with echocardiography before day 3	240/1088	22.1	182/388	46.9	246/598	41.1	<0.001‡
Treatment with vasoactive amines	44/1103	4.0	22/391	5.6	78/611	12.8	<0.001§
PDA treatment with ibuprofen	186/1087	17.1	70/390	17.9	123/610	20.2	0.15‡
Nutrition							
Early enteral feeding (before day 2)	877/1069	82.0	263/372	70.7	429/598	71.7	<0.001¶
Pain							
Treatment with O-SH-GA	342/1113	30.7	109/396	27.5	307/614	50.0	<0.001‡
Procedural pain assessment (at least one assessment during the first week of life)	389/1118	34.8	170/398	42.7	181/619	29.2	<0.001‡
Prolonged pain assessment (at least one assessment during the first week of life)	785/1118	70.2	366/398	92.0	327/619	52.8	<0.001‡
Neurodevelopmental care							
Permanent incubator cover during the first week of life	889/1011	87.9	122/355	34.4	426/536	79.5	<0.001‡
Kangaroo care during the first 3 days of life	443/1038	42.7	95/388	24.5	106/563	18.7	<0.001**
Parental involvement in feeding support †† during the first week of life	495/987	50.2	139/379	36.7	166/563	29.5	<0.001**
Breast contact during the first week of life	138/1031	13.4	27/381	7.1	29/584	5.0	<0.001**

*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.

§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

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

	Cluster * 1		Cluster * 2		Cluster * 3		P value†
	n=26 units		n=11 units		n=16 units		
University hospital	21/26	80.8	7/11	63.6	8/16	50.0	0.11
NICU only ‡	14/26	53.8	8/11	72.7	12/16	75.0	0.31
Number of beds, median (IQR)	10 (8 to 14)		11 (8 to 16)		12 (8 to 16)		0.61
Annual number of admission before 31 weeks in 2011, median (IQR)	91 (75 to 118)		90 (64 to 112)		86 (64 to 119)		0.74
Neonatal surgery available for							
Necrotising enterocolitis	22/26	84.6	9/11	81.8	7/16	43.8	0.012
Patent ductus arteriosus	20/26	76.9	8/11	72.7	4/16	25.0	0.002
Ventriculoperitoneal shunt	13/26	50.0	4/11	36.4	2/16	12.5	0.048
Neurodevelopmental care training							
NIDCAP	7/26	26.9	2/11	18.2	1/16	6.3	0.023
Sensory motor programme	2/26	7.7	5/11	45.5	1/16	6.3	
Introductory course	9/26	34.6	1/11	9.1	5/16	31.3	
No training	8/26	30.8	3/11	27.3	9/16	56.3	

NICU only without paediatric intensive care unit

*Clusters are defined according to the distribution of care within units using a hierarchical analysis.

† χ^2 test for categorical variables or Kruskal-Wallis tests for quantitative variable.

‡Units admitting neonates only, without a paediatric intensive care unit.

NICU, neonatal intensive care unit; NIDCAP, Neonatal Individualised Developmental Care and Assessment Program.

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.

EPIPAGE-2 is a large, national cohort study with prospective enrolment of preterm babies that enabled us to focus on babies born between 27 and 31 WG. This is important as this population includes a larger number of babies when compared with the extremely preterm population but has been less well studied. Updated data on care practices and outcomes for these neonates may have an impact on public health by enabling neonatal teams to reconsider strategies for care provision. We included inborn babies only, as birth outside a level III unit is associated with an increased likelihood of death before discharge. Thus, outcomes are more likely to be related to units' practices than to characteristics of the populations admitted in each cluster. In addition, this strategy may help to reduce unmeasured variability, for example, due to differing clinical experiences of staff members. We also report issues at both discharge and 2 years CA; this enabled us to observe that the lower mortality at discharge in Cluster 2 was associated with a higher proportion of children with an ASQ score below threshold at 2 years CA, and thus at risk of having developmental or cognitive delay.^{26 27} However, the use of

parental questionnaires rather than objective assessment 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.²⁸ 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 individualise 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',²⁹ 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 3 Maternal and infant characteristics for the study population and by cluster units

	Cluster * 1		Cluster * 2		Cluster * 3		P value
	n=1118 infants		n=398 infants		n=619 infants		
Maternal characteristics							
Maternal age							
<25 years	221/1118	19.8	67/398	16.8	111/619	17.9	0.29
25 to 34 years	640/1118	57.2	248/398	62.3	353/619	57.0	
≥35 years	257/1118	23.0	83/398	20.9	155/619	25.0	
Birth in France	879/1100	79.9	286/394	72.6	454/615	73.8	0.002
Parents' socio-economic status †							
Professional	238/1055	22.6	97/382	25.4	122/596	20.5	0.053
Intermediate	213/1055	20.2	82/382	21.5	129/596	21.6	
Administrative, public service, self-employed and students	288/1055	27.3	122/382	31.9	160/596	26.8	
Shop assistants and service workers	132/1055	12.5	35/382	9.2	92/596	15.4	
Manual workers	143/1055	13.6	35/382	9.2	71/596	11.9	
Unknown occupation	41/1055	3.9	11/382	2.9	22/596	3.7	
Obstetric factors							
Singleton pregnancy	753/1118	67.4	269/398	67.6	406/619	65.6	0.72
Antenatal steroids	995/1104	90.1	359/394	91.1	531/607	87.5	0.12
Vaginal delivery	288/1113	25.9	136/393	34.6	169/615	27.5	0.004
Neonatal characteristics							
Gestational age, weeks							
27	168/1118	15.0	47/398	11.8	87/619	14.1	0.26
28	185/1118	16.5	65/398	16.3	105/619	17.0	
29	209/1118	18.7	74/398	18.6	113/619	18.3	
30	275/1118	24.6	113/398	28.4	132/619	21.3	
31	281/1118	25.1	99/398	24.9	182/619	29.4	
Male	557/1118	49.8	228/398	57.3	347/619	56.1	0.008
Small-for-gestational age ‡	445/1118	39.8	170/398	42.7	248/619	40.1	0.58

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).

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)² such as healthcare professionals' availability³⁰ was an obvious limitation.

The magnitude of absolute difference in care practices between clusters is difficult to interpret.³¹ 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.³² Grouping the units provides an opportunity to observe differences and to reflect on

Table 4 Outcome at discharge from NICUs and at 2 years CA in the study population by cluster of units

	Cluster * 1		Cluster * 2		Cluster * 3		P value
Infants admitted to NICU							
At discharge							
Mortality	63/1118	5.6 %	9/398	2.3 %	29/619	4.7 %	
aOR (95% CI)†		1		0.43 (0.20 to 0.93)		0.88 (0.51 to 1.54)	0.010
Mortality or severe neonatal morbidity‡	197/1118	17.7 %	58/398	14.7 %	100/619	16.2 %	
aOR (95% CI)†		1		0.84 (0.49 to 1.42)		0.93 (0.62 to 1.39)	0.72
At 2 years CA							
Mortality	68/1118	6.1 %	9/398	2.3 %	34/619	5.5 %	
aOR (95% CI)†		1		0.39 (0.18 to 0.85)		0.96 (0.56 to 1.67)	0.060
Mortality or severe/moderate neuro-motor or sensory disabilities§	103/1118	9.3 %	17/398	4.3 %	61/619	9.9 %	
aOR (95% CI)†		1		0.46 (0.25 to 0.84)		1.13 (0.78 to 1.63)	0.010
Survivors at 2 years CA							
Cerebral palsy	54/1050	5.1%	14/389	3.6 %	25/585	4.1 %	
aOR (95% CI)†		1		0.57 (0.21 to 1.55)		0.81 (0.44 to 1.50)	0.41
ASQ below threshold ¶	420/989	42.5 %	195/375	52.7 %	260/553	47.6 %	
aOR (95% CI)†		1		1.49 (1.07 to 2.08)		1.17 (0.89 to 1.55)	0.042
Complete cases analysis							
Cerebral palsy	43/884	4.9%	11/327	3.4%	20/506	4.0%	
aOR (95% CI)†		1		0.63 (0.23 to 1.74)		0.86 (0.47 to 1.56)	0.64
ASQ below threshold ¶	245/650	37.7%	109/232	47.0%	141/343	41.1%	
aOR (95% CI)†		1		1.51 (1.07 to 2.14)		1.19 (0.86 to 1.64)	0.063

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.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 WG³⁴ 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.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.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 EPOPé, 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 Womens' Health, University College London, London, UK

Acknowledgements We are grateful for the participation of all families of preterm infants in the EPIPAGE-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 EPIPAGE-2 Study Group for its substantial contribution to the conception, design and acquisition of data.

Collaborators Study group: Neurodevelopmental care study group of EPIPAGE-2: C Arnaud, Inserm 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 Gire, Department of Neonatal Pediatrics and Intensive Care, Nord Hospital, Marseille, France. B Guillois, 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 EPIPAGE-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 CPP SC-2873). 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 EPIPAGE studies are subject to a data sharing policy that may be downloaded from <http://epipage2.inserm.fr/index.php/fr/cote-recherche/acces-aux-donnees-epipage-2>.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Veronique Pierrat <http://orcid.org/0000-0001-7435-5144>
 Gilles Cambonie <http://orcid.org/0000-0002-1360-7915>
 Melanie Durox <http://orcid.org/0000-0002-4038-9036>
 Andrei S Morgan <http://orcid.org/0000-0003-4143-1977>
 Monique Kaminski <http://orcid.org/0000-0002-2548-2304>

REFERENCES

- Wennberg JE. Unwarranted variations in healthcare delivery: implications for academic medical centres. *BMJ* 2002;325:961–4.
- Goodman DC, Little GA, Harrison WN, et al. Dartmouth atlas of neonatal intensive care. Available: https://www.dartmouthatlas.org/Neonatal_Atlas_090419.pdf [Accessed 19 Feb 2020].
- Norwegian neonatal healthcare atlas, 2009–2014. Available: <https://helseatlas.no/sites/default/files/norwegian-neonatal-healthcare.pdf> [Accessed 19 Feb 2020].
- Westrup B. Family-Centered developmentally supportive care. *Neoreviews* 2014;15:e325–35.
- Jobe AH. A risk of sensory deprivation in the neonatal intensive care unit. *J Pediatr* 2014;164:1265–7.
- Horbar JD, Plsek PE, Leahy K, et al. NIC/Q 2000: establishing habits for improvement in neonatal intensive care units. *Pediatrics* 2003;111:e397–410.
- Ellsbury DL, Clark RH. Does quality improvement work in neonatology improve clinical outcomes? *Curr Opin Pediatr* 2017;29:129–34.
- Ohlinger J, Brown MS, Laudert S, et al. Development of potentially better practices for the neonatal intensive care unit as a culture of collaboration: communication, accountability, respect, and empowerment. *Pediatrics* 2003;111:e471–81.
- Horbar JD, Edwards EM, Greenberg LT, et al. Variation in performance of neonatal intensive care units in the United States. *JAMA Pediatr* 2017;171:e164396.
- Lee SK, McMillan DD, Ohlsson A, et al. Variations in practice and outcomes in the Canadian NICU network: 1996–1997. *Pediatrics* 2000;106:1070–9.
- Alleman BW, Bell EF, Li L, et al. Individual and center-level factors affecting mortality among extremely low birth weight infants. *Pediatrics* 2013;132:e175–84.
- Beltempo M, Isayama T, Vento M, et al. Respiratory management of extremely preterm infants: an international survey. *Neonatology* 2018;114:28–36.
- Ancel P-Y, Goffinet F, EPIPAGE 2 Writing Group. EPIPAGE 2: a preterm birth cohort in France in 2011. *BMC Pediatr* 2014;14:97.
- Delorme P, Goffinet F, Ancel P-Y, et al. Cause of preterm birth as a prognostic factor for mortality. *Obstet Gynecol* 2016;127:40–8.
- Pierrat V, Coquelin A, Cuttini M, et al. Translating neurodevelopmental care policies into practice: the experience of neonatal ICUs in France-The EPIPAGE-2 cohort study. *Pediatr Crit Care Med* 2016;17:957–67.
- Seidman G, Unnikrishnan S, Kenny E, et al. Barriers and enablers of kangaroo mother care practice: a systematic review. *PLoS One* 2015;10:e0125643.
- Ego A, Prunet C, Lebreton E, et al. [Customized and non-customized French intrauterine growth curves. I - Methodology]. *J Gynecol Obstet Biol Reprod* 2016;45:155–64.
- Ancel P-Y, Goffinet F, EPIPAGE-2 Writing Group, et al. Survival and morbidity of preterm children born at 22 through 34 weeks' gestation in France in 2011: results of the EPIPAGE-2 cohort study. *JAMA Pediatr* 2015;169:230–8.
- Pierrat V, Marchand-Martin L, Arnaud C, et al. Neurodevelopmental outcome at 2 years for preterm children born at 22 to 34 weeks' gestation in France in 2011: EPIPAGE-2 cohort study. *BMJ* 2017;358:j3448.
- Surveillance of Cerebral Palsy in Europe. Surveillance of cerebral palsy in Europe: a collaboration of cerebral palsy surveys and registers. surveillance of cerebral palsy in Europe (SCPE). *Dev Med Child Neurol* 2000;42:816–24.
- Ghassabian A, Sundaram R, Bell E, et al. Gross motor milestones and subsequent development. *Pediatrics* 2016;138:e20154372.
- Squire J, Twombly E, Bricker D, et al. ASQ-3: User's guide. Baltimore: Brookes Publishing: Ages & Stages Questionnaires®, Third Edition (ASQ-3TM). Available: <http://products.brookespublishing.com/Ages-Stages-Questionnaires-Third-Edition-ASQ-3-P569.aspx> [Accessed 19 Feb 2020].
- Everitt B, Landau S, Leese M. *Cluster analysis*. 5th edn. New York: Oxford University Press, 2001.
- Multiple imputation of missing data using SAS® - 65370_excerpt.pdf. Available: https://www.sas.com/storefront/aux/en/spmultipleimputation/65370_excerpt.pdf [Accessed 6 Jul 2017].
- Rubin DB, Schenker N. Multiple imputation in health-care databases: an overview and some applications. *Stat Med* 1991;10:585–98.
- Flamant C, Branger B, Nguyen The Tich S, et al. Parent-completed developmental screening in premature children: a valid tool for follow-up programs. *PLoS One* 2011;6:e20004.
- Charkaluk M-L, Rousseau J, Calderon J, et al. Ages and stages questionnaire at 3 years for predicting IQ at 5–6 years. *Pediatrics* 2017;139:e20162798.
- Sweet DG, Carnielli V, Greisen G, et al. European Consensus Guidelines on the Management of Respiratory Distress Syndrome - 2019 Update. *Neonatology* 2019;115:432–50.
- Goldstein RF. Developmental care for premature infants: a state of mind. *Pediatrics* 2012;129:e1322–3.
- Shahroor M, Lehtonen L, Lee SK, et al. Unit-Level Variations in Healthcare Professionals' Availability for Preterm Neonates <29 Weeks' Gestation: An International Survey. *Neonatology* 2019:1–9.
- Zeitlin J, Manktelow BN, Piedvache A, et al. Use of evidence based practices to improve survival without severe morbidity for very preterm infants: results from the EPICE population based cohort. *BMJ* 2016;354:i2976.
- Durmeyer X, Marchand-Martin L, Porcher R, et al. Abstention or intervention for isolated hypotension in the first 3 days of life in extremely preterm infants: association with short-term outcomes in the EPIPAGE 2 cohort study. *Arch Dis Child Fetal Neonatal Ed* 2017;102:490–6.
- Baylis R, Ewald U, Gradin M, et al. First-time events between parents and preterm infants are affected by the designs and routines of neonatal intensive care units. *Acta Paediatr* 2014;103:1045–52.
- Rozé J-C, Cambonie G, Marchand-Martin L, et al. Association between early screening for patent ductus arteriosus and in-hospital mortality among extremely preterm infants. *JAMA* 2015;313:2441–8.
- Simpson JM, Evans N, Gibberd RW, et al. Analysing differences in clinical outcomes between hospitals. *Qual Saf Health Care* 2003;12:257–62.
- O' Cathain A, Croot L, Duncan E, et al. Guidance on how to develop complex interventions to improve health and healthcare. *BMJ Open* 2019;9:e029954.