BMJ Open

Using cluster analysis to describe phenotypical heterogeneity in extremely preterm infants: a retrospective whole-population study

Theodore Dassios ,[1,2] Emma E Williams,[1] Christopher Harris,[1,2] Anne Greenough [1,3]

¹Department of Women and Children's Health, School of Life Sciences, Faculty of Life Science and Medicine, King's College London, London, UK
²Neonatal Intensive Care Centre, King's College Hospital NHS Foundation Trust, London, UK
³Asthma UK Centre for Allergic Mechanisms in Asthma, King's College London, London, UK

Correspondence to
Dr Theodore Dassios;
theodore.dassios@kcl.ac.uk

ABSTRACT

Objective To use cluster analysis to identify discrete phenotypic groups of extremely preterm infants.

Design Secondary analysis of a retrospective whole population study.

Setting All neonatal units in England between 2014 and 2019.

Participants Infants live-born at less than 28 weeks of gestation and admitted to a neonatal unit.

Interventions K-means cluster analysis was performed with the gestational age, Apgar score at 5 min and duration of mechanical ventilation as input variables.

Primary and secondary outcome measures Bronchopulmonary dysplasia, discharge on home oxygen, intraventricular haemorrhage, death before discharge from neonatal care.

Results Ten thousand one hundred and ninety-seven infants (53% male) were classified into four clusters: Cluster 1 contained infants with intermediate gestation and duration of ventilation and had an intermediate mortality and incidence of bronchopulmonary dysplasia. Cluster 2 contained infants with the highest gestation, a shorter duration of ventilation and the lowest mortality. Cluster 3 contained infants with the lowest Apgar score and highest mortality and incidence of intraventricular haemorrhage. Cluster 4 contained infants with the lowest gestation, longest duration of ventilation and highest incidence of bronchopulmonary dysplasia.

Conclusion Clinical parameters can classify extremely preterm infants into discrete phenotypic groups with differing subsequent neonatal outcomes.

Strengths and limitations of this study

► First study to use cluster analysis to classify extremely preterm infants in different phenotypical groups.
► We used readily available clinical parameters to classify extremely preterm infants into distinct phenotypic groups.
► The ensuing groups had differing neonatal outcomes such as survival to discharge from neonatal care and bronchopulmonary dysplasia.
► We used the whole population rather than a representative sample, making our results more generalisable.
► Our cluster analysis is not an early prediction study as we have included as input variable the duration of ventilation, which might be prolonged in the life of an extremely preterm infant.

and the associated implications for patients and families are only expected to increase in the future.

Although many extremely preterm infants suffer significant perinatal and long-term morbidity,[7] there is substantial variation in the severity of the neonatal outcomes with a sizeable proportion escaping major morbidity. For example, we have recently reported the overall mortality and incidence of bronchopulmonary dysplasia (BPD) of extremely preterm infants in England to be 18.9% and 57.5%, respectively, which means that 81.1% of those infants survived to discharge and 23.6% were free of BPD.[8] Although the multifactorial origin of adverse neonatal outcomes has been well described,[9] the heterogeneity in neonatal outcomes is poorly understood.

The use of unsupervised learning based on big-data analytics has been used in the clinical domain for the prediction of individual risk factors and for clinical decision support.[10] Unsupervised learning could also be helpful

INTRODUCTION

Preterm birth is a major cause of childhood morbidity and mortality with a global incidence of approximately 10.6%.[1] Infants born at less than 28 completed weeks of gestation (extremely preterm) suffer significant multi-system morbidity lasting into adolescence and adulthood,[2][3] with important financial implications for health systems[4] and an increased physical, emotional and financial burden for the families.[5] The survival of these infants has increased over the past decades[6]

to identify different phenotypic groups (clusters) corresponding to distinct clinical phenotypes of extreme preterm infants. Cluster analysis as a method of detecting subgroups within multidimensional data sets could also be helpful in identifying subgroups of patients that might benefit from targeted interventions.[11 12] To our knowledge, such unsupervised learning approaches have not been previously employed to describe different phenotypical groups in extreme prematurity.

We aimed to derive and describe clusters of extremely preterm infants based on readily available clinical information and to explore whether these distinct phenotypes were associated with different neonatal outcomes.

MATERIALS AND METHODS
Study design and participants
All infants live-born before 28 completed weeks of gestation and admitted to a neonatal unit in England between 1 January 2014 and 1 January 2019 (study period) were included. This was an analysis of data previously acquired to investigate the relationship of growth impairment with the development of BPD in a retrospective, whole-population study.[8] A predefined data set was acquired from the National Neonatal Research Database (NNRD), Imperial College London, UK. As the study used data held in an existing database, participation did not require approval from individual Trusts, but only from the NHS Trust holding the database (Chelsea and Westminster NHS Foundation Trust) which was obtained.

The following variables were collected: maternal age (years), administration of antenatal steroids (yes/no), gestational age at birth calculated from the last menstrual period or ultrasonographically (weeks), birth weight (kg), Apgar score at 5 min of age, sex (male/female), duration of invasive ventilation (days), BPD development defined as any need for respiratory support at 36 weeks postmenstrual age (PMA),[13 14] administration of postnatal corticosteroids (defined as parenteral administration of dexamethasone or hydrocortisone for more than 5 consecutive days—yes/no), surgical intervention for necrotising enterocolitis (yes/no),[15] surgical ligation of patent ductus arteriosus (yes/no),[16] intraventricular haemorrhage (IVH) grade 3–4 (yes/no),[17] periventricular leucomalacia (yes/no),[18] death before discharge from neonatal care (yes/no), age at death (days), weight at 36 weeks PMA (kg), PMA at discharge (weeks), weight at discharge (kg), discharged on home oxygen (yes/no). We calculated the birth weight z-score (ΔWz) using the UK-WHO preterm reference chart[19] and the Microsoft Excel add-in LMS Growth (V.2.77; www.healthforchildren.co.uk). The birth weight z-score was not calculated for infants born <23 completed weeks of gestation, as there were no reference data.

PPI statement
This is a retrospective study, so patients and public could not be involved.

Analysis
Cluster analysis was performed with the smallest possible number of continuous variables necessary to adequately characterise neonatal outcomes, since inclusion of a large number of variables might degrade the final classification.[11 20] K-means cluster analysis was performed with gestational age, Apgar score at 5 min and duration of mechanical ventilation as input variables. These parameters were selected, as they were independent, contributory to neonatal outcomes and pathophysiologically distinct between them. All results were standardised as z-scores prior to clustering. To select the number of clusters, clustering was performed for solutions comprising 2 to 10 clusters and the optimal solution was selected based on the minimum number of iterations that changed the cluster centres and the optimal F factor and significance of the corresponding Analysis of Variance (ANOVA) analysis. The F factor for each ANOVA cluster solution was calculated as the ratio of two variances, and a higher F factor corresponded to a stronger separation of the clusters.[21 22] Subjects with missing values in any of the three input variables were excluded from the analysis. Continuous variables were compared across clusters using Kruskal-Wallis with Mann-Whitney U test as a post hoc test for pairwise comparisons. Categorical variables were compared across clusters using χ^2 test with Bonferroni adjustment as a post hoc test for pairwise comparisons. Cluster profiles were presented graphically using radial plots, where the length of each 'spoke' was proportional to the magnitude of the standardised variable. The clustering was also visualised using a discriminant-coordinates biplot, generated by canonical variate analysis, which projects multidimensional data into a lower dimensional space while preserving as much information as possible, to provide an easily interpretable two-dimensional representation of cluster separation and density.

Cluster and statistical analysis were performed using SPSS software, V.26.0 (IBM, Armonk, New York).

RESULTS
During the period of the study, 11 806 infants were born alive below 28 completed weeks of gestation and admitted to a neonatal unit in England. One thousand five hundred and eighty-five infants were excluded for missing data on Apgar score in 5 min. Twenty-one infants were removed for missing data on the duration of mechanical ventilation and a further three infants because of indeterminate sex (figure 1). A total of 10 197 infants (53% male) were included in the final analysis with a median (IQR) gestational age of 26.0 (24.9–27.1) weeks and birth weight of 0.81 (0.68–0.96) kg (table 1). They were ventilated for a median of 10 (3–26) days and the ones that survived to discharge from neonatal care (81%) were discharged at a median (IQR) PMA of 39.6 (37.6–42.1) weeks (table 1).

A four-cluster solution was found to be optimal (table 1). The clinical profiles for the four clusters are presented graphically in figures 2 and 3. The boxplots of
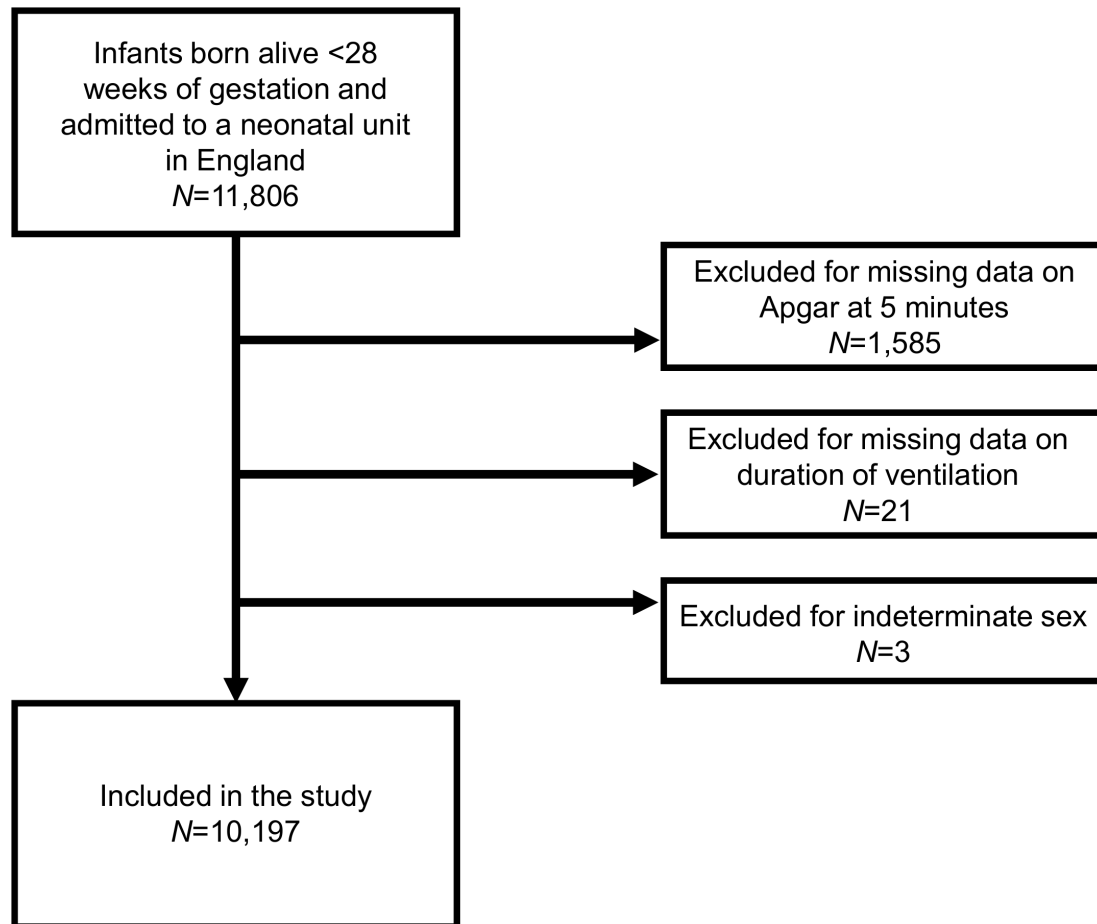
Figure 1 Flow diagram of the included infants.

the input parameters in the four clusters along with the birth weight a-score, the duration of PN and the PMA age at discharge are presented in figure 4.

Cluster 1 (intermediate) contained 27% of the population. The infants of cluster 1 had a lower median (IQR) gestational age (24.8 (24.1–25.2) weeks) compared with clusters 2 and 3 (27.0 (26.6–27.6) weeks and 25.4 (24.4–26.4) weeks respectively, p<0.001 for both). The median Apgar score for infants of cluster 1 was higher compared with the infants of clusters 3 and 4 (p<0.001 for both), but not of cluster 2 infants. Infants in cluster 1 were ventilated for a median (IQR) period of 21 (10–33) days, significantly longer that infants of clusters 2 and 3 (4 (2–10) and 9 (3–21) days, respectively, p<0.001 for both) but significantly shorter than the infants of cluster 4 (60 (52–73) days, p<0.001). Infants of cluster 1 developed more often BPD (59%) compared with infants of clusters 2 and 3 (48% and 42%, respectively, p<0.001) but less often than infants of cluster 4 (89%, p<0.001). Infants of cluster 1 had lower mortality (27%) compared with cluster 3 (40%, p<0.001) but higher mortality compared with cluster 2 (8%, p<0.001).

Cluster 2 (favourable) contained 47% of the population and had the highest median (IQR) gestational age (27.0 (26.6–27.6) weeks) compared with groups 1, 3 (25.4 (24.4–26.4) weeks, p<0.001) and 4 (24.4 (23.9–25.3)

weeks, p<0.001). The median Apgar score for infants of cluster 2 was higher compared with the infants of clusters 3 and 4 (p<0.001 for both). Infants in cluster 2 had the shorter duration of ventilation with a median (IQR) period of 4 (2–10) days, compared with infants of cluster 1 (21 (10–33) days)], cluster 3 (9 (3–21) days) and cluster 4 (60 (52–73) days, p<0.001 for both). Infants of cluster 2 developed less often BPD (48%) compared with infants of clusters 1 and 4 (59% and 89%, respectively, p<0.001 for both), but more often than infants of cluster 3 (42%, p<0.001). Infants of cluster 2 had the lowest mortality (8%) compared with all other clusters (p<0.001).

Cluster 3 (low Apgar) contained 17% of the population and the infants with the lowest median (IQR) Apgar score (4 (2–5)) compared with all other groups (p<0.001 for all). Their median (IQR) gestation (25.4 (24.4–26.4) weeks) was higher than clusters 1 and 4, but lower than cluster 2 (p<0.001 for all). Infants of cluster 3 had a shorter median (IQR) duration of ventilation (9 (3–21)) days, compared with infants of clusters 1 and 4 (p<0.001 for both), but longer than infants of cluster 2 (p<0.001). Infants of cluster 3 had the lower incidence of BPD (42%) compared with infants of clusters 1, 2 and 4 (59%, 48% and 89%, respectively, p<0.001 for all). Infants of cluster 3 had the highest mortality (40%) compared with all other clusters (p<0.001).

Table 1 Subject characteristics by cluster

	Whole population N=10.197	Cluster 1 N=2782 27%	Cluster 2 N=4773 47%	Cluster 3 N=1777 17%	Cluster 4 N=865 8%	P value*
Maternal age (years)	31 (26–35)	31 (26–35)	31 (26–35)	31 (26–35)	31 (26–35)	0.352
Tertiary neonatal unit at birth	7223 (71)	2093 (75)	3255 (68)	1180 (66)	695 (80)	<0.001
Antenatal steroids	9154 (90)	2501 (90)	4.371 (92)	1515 (85)	767 (89)	<0.001
Gestational age (weeks)	26.0 (24.9–27.1)	24.8 (24.1–25.4)	27.0 (26.6–27.6)	25.4 (24.4–26.4)	24.4 (23.9–25.3)	<0.001
Male gender	5441 (53)	1464 (53)	2508 (53)	1022 (58)	491 (57)	0.001
Apgar score at 5min	7 (6–9)	8 (7–9)	8 (7–9)	4 (2–5)	6 (5–8)	<0.001
Birth weight (kg)	0.81 (0.68–0.96)	0.70 (0.62–0.79)	0.94 (0.83–1.05)	0.77 (0.65–0.90)	0.64 (0.57–0.73)	<0.001
Birth weight z-score	−0.38 (−0.89–0.07)	−0.44 (−0.87 to −0.02)	−0.30 (−0.86 to 1.05)	−0.39 (−0.87 to 0.05)	−0.61 (−1.10 to −0.14)	<0.001
Duration of ventilation (days)	10 (3–26)	21 (10–33)	4 (2–10)	9 (3–21)	60 (52–73)	<0.001
Postnatal corticosteroids	1995 (20)	771 (28)	338 (7)	287 (16)	599 (69)	<0.001
Duration of PN (days)	15 (9–25)	17 (10–28)	13 (9–20)	13 (5–22)	36 (20–61)	<0.001
BPD at 36 weeks	5431 (53)	1629 (59)	2280 (48)	750 (42)	772 (89)	<0.001
Home oxygen	3010 (30)	937 (34)	1245 (26)	441 (25)	387 (45)	<0.001
Died before discharge	1985 (19)	753 (27)	366 (8)	705 (40)	161 (19)	<0.001
Age at death (days)	7 (2–25)	10 (3–24)	7 (2–23)	2 (1–10)	91 (60–144)	<0.001
ROP treatment	1197 (12)	502 (18)	219 (5)	151 (8)	325 (38)	<0.001
PDA Ligation	389 (4)	134 (5)	62 (1)	40 (2)	153 (18)	<0.001
IVH grade III–IV	1654 (16)	614 (22)	421 (9)	420 (24)	199 (23)	<0.001
PVL	399 (4)	110 (4)	143 (3)	60 (3)	86 (10)	<0.001
Operated for NEC	582 (6)	208 (7)	152 (3)	72 (4)	150 (17)	<0.001
Weight at 36 weeks (kg)	2.15 (1.89–2.41)	2.12 (1.88–2.38)	2.20 (1.95–2.45)	2.20 (1.93–2.47)	1.90 (1.63–2.18)	<0.001
PMA at discharge†	39.6 (37.6–42.1)	40.7 (38.7–43.1)	38.5 (36.9–40.4)	40.0 (38.1–42.3)	44.9 (41.9–49.0)	<0.001
Discharge weight (kg)†	2.68 (2.26–3.24)	2.85 (2.40–3.38)	2.50 (2.16–2.98)	2.80 (2.35–3.35)	3.41 (2.82–4.17)	<0.001

Data are presented as median (IQR) or N (%).

*Comparison between clusters 1 and 4.

†Data for the alive at discharge only.

BPD, bronchopulmonary dysplasia; IVH, intraventricular haemorrhage; NEC, necrotising enterocolitis; PDA, patent ductus arteriosus; PMA, postmenstrual age; PN, parenteral nutrition; PVL, periventricular leucomalacia; ROP, retinopathy of prematurity.
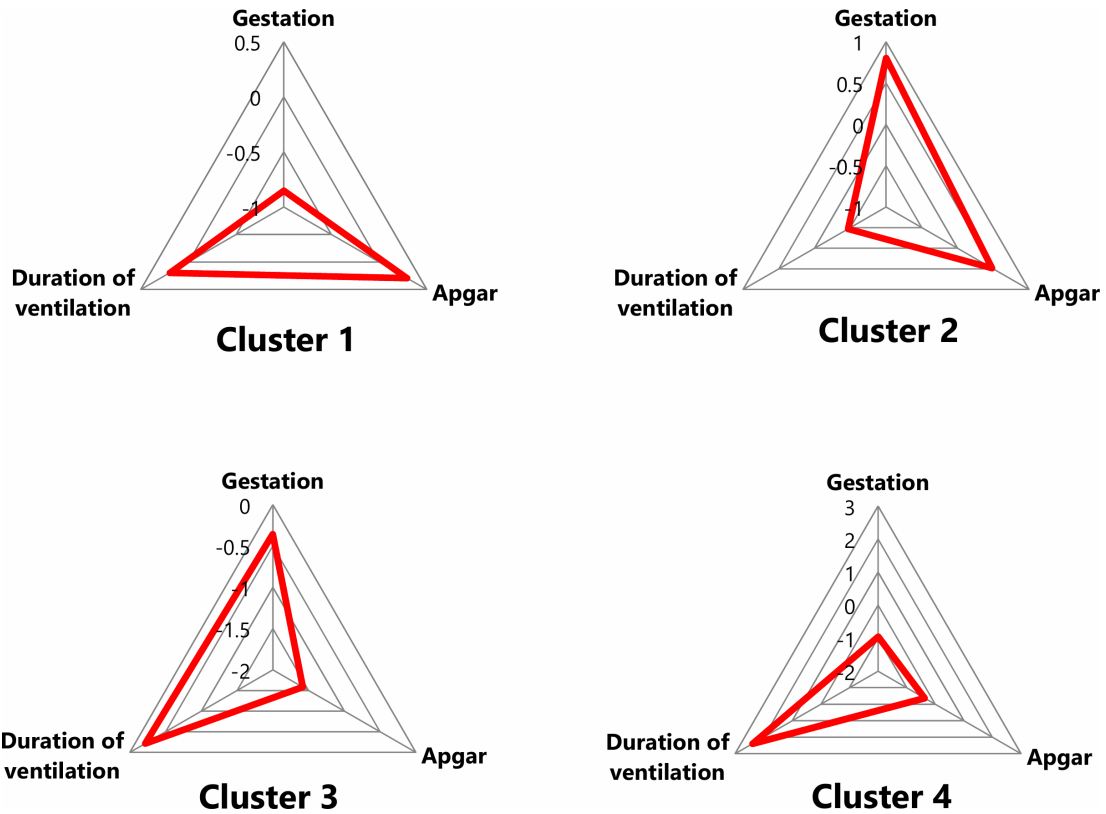
Figure 2 Radial plots showing the clinical profiles for the four clusters. Data are standardised (expressed as z-scores referenced to the whole cohort) and the spike points represent medians.

Cluster 4 (prolonged ventilation) had the smallest number of subjects (8%), the lowest gestational age (24.4 (23.9–25.3) weeks) and the longest median (IQR) duration of ventilation (60 (52–73) days) compared with all other clusters (p<0.001 for all). The median Apgar score for infants of cluster 4 was 6 (5–8), which was higher compared with the infants of cluster 3 (p<0.001) but lower than clusters 1 and 2 infants (p<0.001 for both).
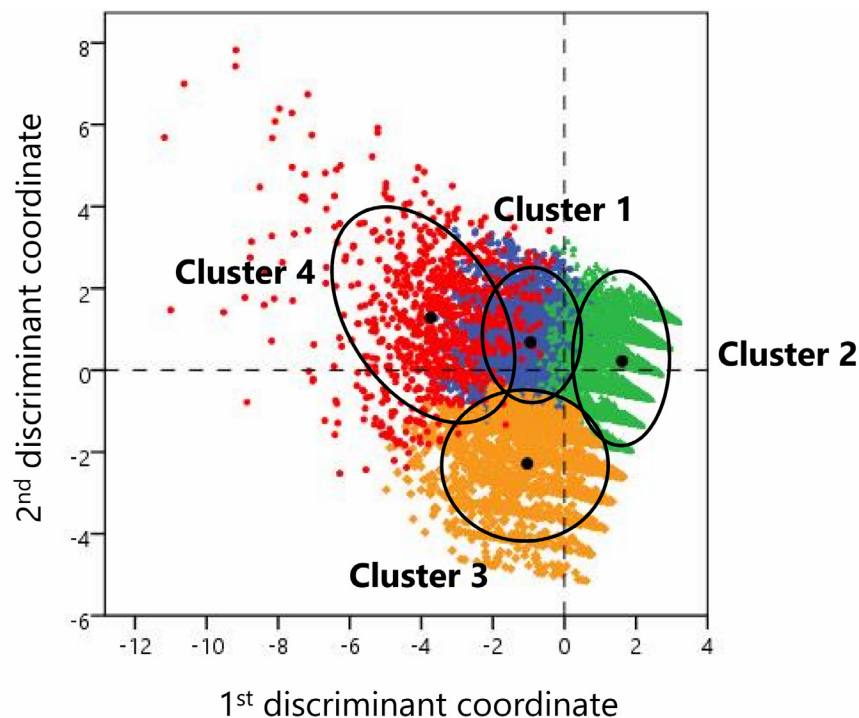

Figure 3 Discriminant two-dimensional plot of the clustering solution. The cluster centroids are presented as black circles.
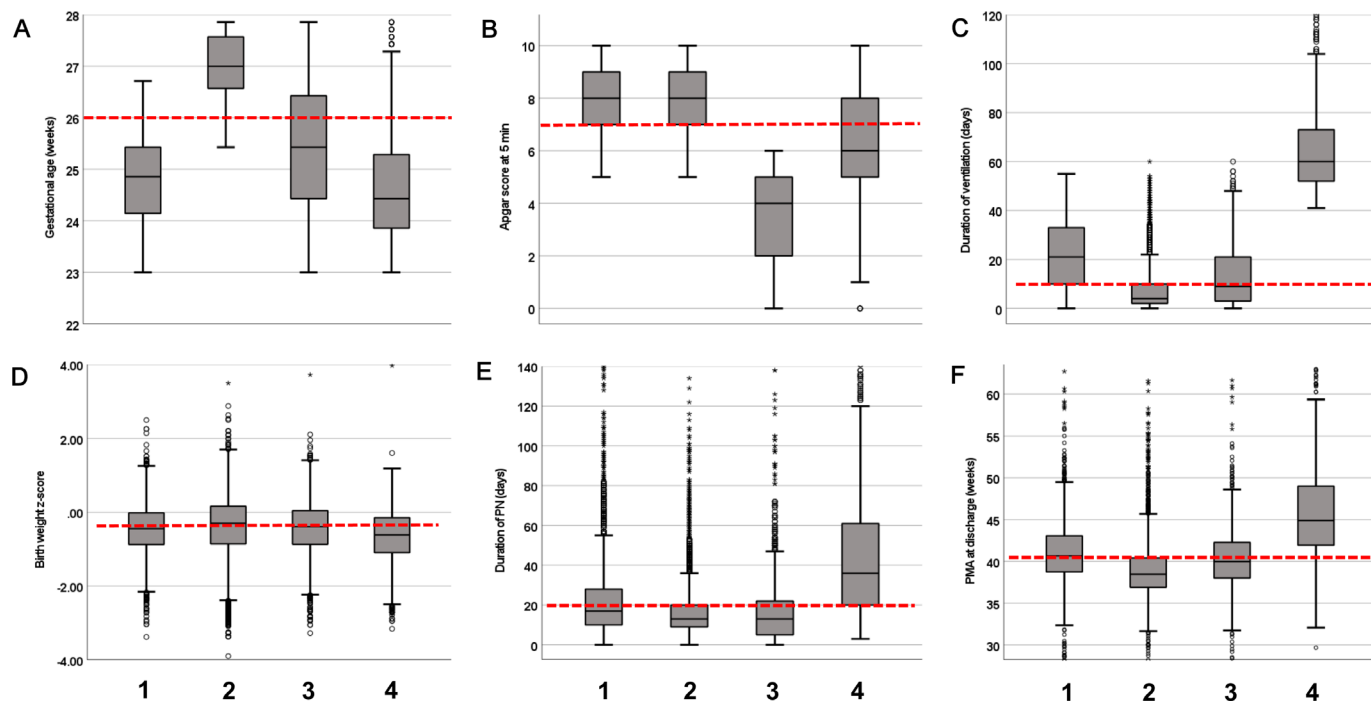
Figure 4 Boxplots of the gestational age (A), Apgar score at 5 min (B), duration of ventilation (C), birth weight z-score (D), duration of parenteral nutrition (PN) (E) and postmenstrual age (PMA) at discharge (F) in the four clusters. The median value for each parameter for the whole population is depicted as a dashed line. The whiskers represent the maximum and minimum values. The asterisks are extremes (at least three box lengths from the median) and the circles are outliers (1.5 box lengths from the median).

Infants of cluster 4 had the highest rate of BPD (89%) compared with infants of all other clusters (p<0.001 for all). The mortality of the infants of cluster 4 (19%) was lower than clusters 1 and 3 (p<0.001 for both) and higher than cluster 2 (p<0.001).

DISCUSSION

Using unsupervised learning to detect and describe distinct phenotypical groups of extremely preterm infants, we demonstrated that survival to discharge from neonatal care and other important neonatal outcomes differed significantly across these groups. We used readily available parameters to classify the extremely preterm infants into the four clusters.

Cluster 1 (intermediate) comprised infants with a combination of a low gestational age and a relatively long duration of invasive ventilation whose resuscitation was overall straightforward as demonstrated by a median Apgar of 8 at 5 min after birth. Nevertheless, they had a relatively high mortality and incidence of BPD. Cluster 2 (favourable) was the largest group and exhibited the best outcomes. It contained infants with the highest gestational age and birth weight z-score and the briefest duration of ventilation. Not surprisingly, they also had the lowest mortality and a relatively modest incidence of complications such as BPD. They were also the earliest to be discharged from neonatal care. The major clustering characteristic of the infants in cluster 3 was the low Apgar score. The poor condition at birth might explain why they

had the highest mortality and rate of IVH compared with all other groups, but their relatively larger gestation and brief duration of ventilation might explain why they had the lowest incidence of BPD. Cluster 4 (prolonged ventilation) consisted of the most premature infants, with the longest duration of mechanical ventilation and highest rates of BPD. These infants were also discharged home at a later PMA than all other cluster infants.

Our results highlight two distinct within-cluster associations. First, poor condition at birth with subsequent high rates of mortality and IVH (cluster 3) and, second, the association of prolonged ventilation with BPD (cluster 4). Both these associations have been previously individually described in population studies: Jensen *et al* studied 3343 extremely low-birth weight infants and reported that among the survivors, exposure to a greater number of mechanical ventilation episodes was associated with a progressive increase in the risk of BPD.[23] Interestingly, the highest mortality in our study was not in the infants of the lower gestation at birth, but rather in the ones with the lowest Apgar score. The association of poor condition at birth with IVH and increased mortality is primarily attributed to the intrinsic fragility of the germinal matrix vasculature and the disturbance in the cerebral blood flow.[24 25] It is also interesting to note that the infants of cluster 3, other than the highest mortality and lowest Apgar score, also had the lowest rate of antenatal steroids and the lowest incidence of birth in a tertiary centre. This is in agreement with previous studies that have

highlighted that delivery outside a tertiary centre is associated with higher mortality and a lower uptake of antenatal steroids.[26] Our study complements the literature by presenting these associations within distinct groups of the same population rather than via a unilateral relationship that does not take into account the other outcomes.

The choice of our input variables (gestational age, Apgar score at 5 min and duration of mechanical ventilation) was made on clinical grounds and has evidently major implications for our conclusions. We selected to use these three parameters because they are continuous and reliably available at a population level, while they also represent relatively distinct pathophysiological processes that are associated with impaired later outcomes. It is well described, for example, that the Apgar score is influenced by gestation and cannot be generally considered as evidence of asphyxia nor can it be used at an individual level to predict mortality or neurological outcomes.[27] Alternative continuous variables to assess the condition at birth such as the cord pH or lactate could not be used, however, as these indices were not consistently inserted in the database and had a very high percentage of missing values. Similarly, the duration of ventilation was selected as a proxy for the severity of lung disease[28] which also can be assessed relatively early chronologically as the median duration of ventilation in our population was 10 days.

To our knowledge, non-supervised learning and cluster analysis have not been previously applied exclusively to extremely preterm infants. Souza *et al* used K-means cluster analysis to explore clinical conditions associated with preterm birth and reported that 4150 preterm births were clustered in three groups of women, yet although some maternal characteristics differed among the clusters, maternal and neonatal outcomes did not.[29] This might be explained by population differences, as this was predominantly a maternity study in which the neonatal outcomes were not considered as input variables in the unsupervised models. Greenbury *et al* performed clustering on daily nutritional intakes in premature infants of less than 32 weeks of gestation and reported different nutritional clusters, which were heterogeneous in size, with some showing common interpretable clinical practices. They also identified a relationship between nutritional practice and outcomes such as BPD, with provision of human milk being a protective factor against developing BPD.[30]

Our classification could be clinically useful in further understanding the pathophysiology of extremely preterm morbidity and mortality and possibly in targeting selected subgroups of infants for specific interventions. For example, for cluster 4, such interventions might include administration of postnatal corticosteroids to assist earlier weaning from invasive respiratory support,[31] more gentle resuscitation methods, avoidance of mechanical ventilation by newer non-invasive techniques[32] and use of less invasive surfactant administration, which has been shown to reduce the incidence of BPD.[33] The within-group association of increased mortality and birth outside a tertiary

hospital in cluster 3 could also strengthen the argument for the centralisation of the care of extreme preterms concentrating on high-volume perinatal centres.[34]

Our study has strengths and some limitations. To our knowledge, this is the first study to use unsupervised learning to classify extremely preterm infants in different phenotypical groups. A further strength of our study was that we used the whole population rather than a representative sample and therefore, by avoiding inclusion bias, our results are more generalisable. We should clarify that our analysis is not an early predictive tool that could be applied at birth, as we have included the duration of ventilation as an input parameter, which comes later in the life of an extremely preterm infant compared with the Apgar score or the gestational age. As such, our study is aiming to describe the interaction of different pathophysiological phenomena at a population level rather than provide individualised prediction of outcomes at birth. We should also note that our study only included infants from a recent 5-year period rather than a longer period that would extend more in the past. Our tighter time period, though, would better correspond to current standards of neonatal care. We did not include in our analysis information on whether preterm birth was spontaneous or on the duration of rupture of membranes as the amount of missing data in these categories was very high at a national level, even more so in non-tertiary units.

In conclusion, we have used readily available clinical parameters to classify extremely preterm infants into distinct phenotypic groups. The ensuing groups had differing neonatal outcomes such as survival to discharge from neonatal care and BPD.

and the UK Health Research Authority (HRA) (IRAS project ID: 259225). The research was conducted ethically in accordance with the World Medical Association Declaration of Helsinki.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request.

ORCID iDs
Theodore Dassios http://orcid.org/0000-0001-5258-5301
Anne Greenough http://orcid.org/0000-0002-8672-5349

REFERENCES

1 Chawanpaiboon S, Vogel JP, Moller A-B, *et al*. Global, regional, and national estimates of levels of preterm birth in 2014: a systematic review and modelling analysis. *Lancet Glob Health* 2019;7:e37–46.
2 Stahlmann N, Eisemann N, Thyen U, *et al*. Long-term health outcomes and health-related quality of life in adolescents from a cohort of extremely premature infants born at less than 27 weeks of gestation in northern Germany. *Neuropediatrics* 2016;47:388–98.
3 Glass HC, Costarino AT, Stayer SA. Outcomes for extremely premature infants. *Anes Analg* 2015;120:1337–51.
4 Beam AL, Fried I, Palmer N, *et al*. Estimates of healthcare spending for preterm and low-birthweight infants in a commercially insured population: 2008-2016. *J Perinatol* 2020;40:1091–9.
5 Lakshmanan A, Agni M, Lieu T, *et al*. The impact of preterm birth <37 weeks on parents and families: a cross-sectional study in the 2 years after discharge from the neonatal intensive care unit. *Health Qual Life Outcomes* 2017;15:38.
6 Cheong JL, Spittle AJ, Burnett AC, *et al*. Have outcomes following extremely preterm birth improved over time? *Semin Fetal Neonatal Med* 2020;25:101114.
7 Patel RM. Short- and long-term outcomes for extremely preterm infants. *Am J Perinatol* 2016;33:318–28.
8 Dassios T, Williams EE, Hickey A, *et al*. Bronchopulmonary dysplasia and postnatal growth following extremely preterm birth. *Arch Dis Child Fetal Neonatal Ed* 2021;106:386-391. [Epub ahead of print].
9 Jensen EA, Schmidt B. Epidemiology of bronchopulmonary dysplasia. birth defects research Part a. *Clin Molecular Teratol* 2014;100:145–57.
10 Alonso-Betanzos A, Bolon-Canedo V. Big-data analysis, cluster analysis, and machine-learning approaches. *Adv Exper Med Biol* 2018;1065:607–26.
11 Lunt A, Mortimer L, Rees D, *et al*. Heterogeneity of respiratory disease in children and young adults with sickle cell disease. *Thorax* 2018;73:575–7.
12 MacBean V, Lunt A, Drysdale SB, *et al*. Predicting healthcare outcomes in prematurely born infants using cluster analysis. *Pediatr Pulmonol* 2018;53:1067–72.
13 Jobe AH, Bancalari E. Bronchopulmonary dysplasia. *Am J Respir Crit Care Med* 2001;163:1723–9.
14 Higgins RD, Jobe AH, Koso-Thomas M, *et al*. Bronchopulmonary dysplasia: executive summary of a workshop. *J Pediatr* 2018;197:300–8.
15 Patel RM, Ferguson J, McElroy SJ, *et al*. Defining necrotizing enterocolitis: current difficulties and future opportunities. *Pediatr Res* 2020;88:10–15.

16 Heuchan AM, Clyman RI. Managing the patent ductus arteriosus: current treatment options. *Arch Dis Child Fetal Neonatal Ed* 2014;99:F431–6.
17 Papile LA, Burstein J, Burstein R, *et al*. Incidence and evolution of subependymal and intraventricular hemorrhage: a study of infants with birth weights less than 1,500 GM. *J Pediatr* 1978;92:529–34.
18 Gotardo JW, Volkmer NdeFV, Stangler GP, *et al*. Impact of peri-intraventricular haemorrhage and periventricular leukomalacia in the neurodevelopment of preterms: a systematic review and meta-analysis. *PLoS One* 2019;14:e0223427.
19 Wright CM, Williams AF, Elliman D, *et al*. Using the new UK-WHO growth charts. *BMJ* 2010;340:c1140.
20 Garcia-Aymerich J, Gómez FP, Benet M, *et al*. Identification and prospective validation of clinically relevant chronic obstructive pulmonary disease (COPD) subtypes. *Thorax* 2011;66:430–7.
21 Everitt BS, Landau S, Leese M. *Cluster analysis*. 4th edn. London: Arnold, 2001.
22 Milligan GW, Cooper MC. An examination of procedures for determining the number of clusters in a data set. *Psychometrika* 1985;50:159–79.
23 Jensen EA, DeMauro SB, Kornhauser M, *et al*. Effects of multiple ventilation courses and duration of mechanical ventilation on respiratory outcomes in extremely low-birth-weight infants. *JAMA Pediatr* 2015;169:1011–7.
24 Ballabh P. Intraventricular hemorrhage in premature infants: mechanism of disease. *Pediatr Res* 2010;67:1–8.
25 Szpecht D, Szymankiewicz M, Nowak I, *et al*. Intraventricular hemorrhage in neonates born before 32 weeks of gestation-retrospective analysis of risk factors. *Childs Nerv Syst* 2016;32:1399–404.
26 Helenius K, Longford N, Lehtonen L. Neonatal data analysis unit and the United Kingdom neonatal collaborative. Association of early postnatal transfer and birth outside a tertiary hospital with mortality and severe brain injury in extremely preterm infants: observational cohort study with propensity score matching. *BMJ* 2019;367:l5678.
27 , American Academy of Pediatrics Committee on Fetus and Newborn, American College of Obstetricians and Gynecologists Committee on Obstetric PracticeWatterberg KL, Aucott S, Benitz WE. The Apgar score. *Pediatrics* 2015;136:819–22.
28 Dassios T, Williams EE, Hickey A, *et al*. Duration of mechanical ventilation and prediction of bronchopulmonary dysplasia and home oxygen in extremely preterm infants. *Acta Paediatr* 2021;110:2052–8.
29 Souza RT, Cecatti JG, Passini R, *et al*. Cluster analysis identifying clinical phenotypes of preterm birth and related maternal and neonatal outcomes from the Brazilian multicentre study on preterm birth. *Int J Gynaecol Obstet* 2019;146:110–7.
30 Greenbury SF, Ougham K, Wu J, *et al*. Identification of variation in nutritional practice in neonatal units in England and association with clinical outcomes using agnostic machine learning. *Sci Rep* 2021;11:7178.
31 Doyle LW, Cheong JL, Ehrenkranz RA, *et al*. Late (> 7 days) systemic postnatal corticosteroids for prevention of bronchopulmonary dysplasia in preterm infants. *Cochrane Database Syst Rev* 2017;10:CD001145.
32 Sand L, Szatkowski L, Kwok T'ng Chang, *et al*. Observational cohort study of changing trends in non-invasive ventilation in very preterm infants and associations with clinical outcomes. *Arch Dis Child Fetal Neonatal Ed* 2022;107:150–5.
33 Aldana-Aguirre JC, Pinto M, Featherstone RM, *et al*. Less invasive surfactant administration versus intubation for surfactant delivery in preterm infants with respiratory distress syndrome: a systematic review and meta-analysis. *Arch Dis Child Fetal Neonatal Ed* 2017;102:F17–23.
34 Walther F, Küster DB, Bieber A, *et al*. Impact of regionalisation and case-volume on neonatal and perinatal mortality: an umbrella review. *BMJ Open* 2020;10:e037135.

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4
Objectives	3	State specific objectives, including any prespecified hypotheses	5
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	5
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	5-6
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	*5-6*
Bias	9	Describe any efforts to address potential sources of bias	5-6
Study size	10	Explain how the study size was arrived at	5
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	6-7
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	6-7
		(*b*) Describe any methods used to examine subgroups and interactions	6-7
		(*c*) Explain how missing data were addressed	6-7
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	N/A
		(*e*) Describe any sensitivity analyses	N/A

1

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

BMJ Open

Results

Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	8
		(b) Give reasons for non-participation at each stage	8-12
		(c) Consider use of a flow diagram	8-12
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	8-12
		(b) Indicate number of participants with missing data for each variable of interest	8-12
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	N/A
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	*8-12*
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	*N/A*
		Cross-sectional study—Report numbers of outcome events or summary measures	*N/A*
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	8-12
		(*b*) Report category boundaries when continuous variables were categorized	8-12
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	8-12

Discussion

Key results	18	Summarise key results with reference to study objectives	13-15
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	15
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	13-15
Generalisability	21	Discuss the generalisability (external validity) of the study results	13-15

Other information

Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	16-17

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.