Type of study	Study	Study participation	Study attrition	Prognostic factor measurement	Outcome measurement	Study confounding
In-utero vs. ex-utero transfer to level 3 or regional perinatal centre	Lamont et al.	• Single network study • Defined exclusion criteria (lethal congenital anomalies, infants transferred for surgical correction of congenital anomaly) • Comparison of baseline characteristics (GA, BW)	• Unclear whether retrospective or prospective • Completeness of data on demographic/confounding factors 100% • Outcome analysis for all babies 28-31 weeks meeting inclusion criteria	• Undefined birth location for transferred babies (from all referring hospitals to the single University Hospital)	• Survival to discharge	• Unadjusted for confounding factors
	Truffert et al.	• Population based study • Defined exclusion criteria (GA<25 and >33 weeks) • Comparison of baseline characteristics (private hospital, >1200 births/year, staff present at delivery, spontaneous labour, mode of delivery, presentation, multiple pregnancy, gender, Apgar score, temperature on transfer from delivery room)	• Retrospective • Completeness of data on demographic/confounding factors 99.5% (comparison only conducted for babies born between 31-32 weeks) • Outcome analysis for all babies 27-30 weeks meeting inclusion criteria	• Undefined birth location for transferred babies (included random selection of all hospitals with maternity units)	• Mortality, disability (CP, deafness, Brunet Lezine developmental score <80), survival without disability at 2 years	• Unadjusted for confounding factors
	Hauspy et al.	• Single network study • Defined exclusion criteria (GA<24 and >35 weeks) • Comparison of baseline characteristics (preterm labour, preterm PROM, IUGR/SGA, preeclampsia, placental abruption, placenta praevia, vaginal bleeding, gender, GA, BW)	• Retrospective • Completeness of data on demographic/confounding factors 100% • Outcome analysis for all babies 28-31 weeks meeting inclusion criteria	• Undefined birth location for transferred babies (from all referring hospitals to the single University Hospital)	• Neonatal mortality, RDS	• Unadjusted for confounding factors
	Lee et al.	• Population based study • Defined exclusion criteria (GA≥32 weeks, moribund, admission to NICU >4 days) • Comparison of baseline characteristics (GA, BW, Apgar score, SNAP-II score, SGA, multiple gestation, maternal hypertension, antenatal care, mode of delivery, antenatal corticosteroids, presentation)	• Prospective • Completeness of data on demographic/confounding factors 100% • Outcome analysis for all babies 27-31 weeks meeting inclusion criteria • Cranial US (and therefore analysis for IVH) available for 82% of infants	• Defined birth location for transferred babies (including level 1 hospitals where care can be provided by 'family physicians')	• Pre-discharge mortality, IVH (≥grade 3), ROP (≥stage 3), RDS, CLD, NEC, survival without major morbidity (IVH, CLD, NEC, ROP)	• Adjusted for confounding factors (GA, Apgar score, SGA, mode of delivery, multiple gestation, maternal hypertension, presentation, antenatal corticosteroids, antenatal care, SNAP-II score)
	Boland et al.	• Population based study • Defined exclusion criteria (GA<22 and >32 weeks, lethal congenital anomalies) • Comparison of baseline characteristics (maternal age, multigravida, multiple pregnancy, APH, hypertensive disorders of pregnancy, prelabour ROM, spontaneous preterm labour, mode of delivery, gender, GA, BW)	• Retrospective • Completeness of data on demographic/confounding factors 99.9-100% • Outcome analysis for all babies 28-31 weeks meeting inclusion criteria	• Defined birth location for transferred babies (including hospitals without special care or obstetric units and births before arrival at hospital)	• Infant mortality	• Adjusted for confounding factors (GA, BW, gender)

Level of unit of birth (level 3 or perinatal regional centre vs. lower level or local unit)	Holmgren et al.	• Population based study • Defined exclusion criteria (GA>37 weeks) • No comparison of baseline characteristics of population by level of unit	• Retrospective • Completeness of data on demographic/confounding factors 100% • Outcome analysis not carried out on entire population of babies meeting inclusion criteria (at 28-31 weeks possible lack of data on 13.4% of babies for neonatal and infant mortality)	• Comparing outcomes of tertiary and secondary level units • Explanation given of facilities available (e.g. level 2 units provide IC or have neonatal 'care units')	• Perinatal (up to 1 week), neonatal, and infant mortality, severe asphyxia (Apgar score <5 at 10 minutes of age)	• Unadjusted for confounding factors
	Johansson et al.	• Population based study • Defined exclusion criteria (GA<24 and >32 weeks, births at units without paediatric services) • Comparison of baseline characteristics (maternal BMI, smoking status, cohabitation with father, age, country of birth, placenta praevia, abruptio-placenta, APH, preeclampsia, hypertension, GDM, chronic diseases, GA, BW for GA, mode of delivery, gender, presentation, SGA, LGA, major congenital anomalies)	• Retrospective • Completeness of data on demographic/confounding factors 97.1-100% • Outcome analysis for 99.6% of babies 28-31 weeks meeting inclusion criteria (hospital of birth unknown for 9 babies)	• Comparing outcomes of university and general hospitals • Explanation of facilities available (e.g. general hospitals have 1000-5000 deliveries/year, similar obstetric and anaesthetic facilities to university hospitals, provide neonatal IC before transfer)	• Infant mortality	• Adjusted for confounding factors (mode of delivery, GA, BW for GA, gender, presentation, placental complications, maternal hypertension)
Level of care (level 3 or perinatal regional centre vs. lower level or local unit)	Field et al.	• Population based study • No exclusion criteria • Comparison of baseline characteristics (GA, BW, RDS, presentation, Apgar score, multiple pregnancy)	• Prospective • Unable to determine completeness of data on demographic/confounding factors • Unable to determine proportion of babies for which outcome analysed	• Comparing outcomes of 'large'/IC units and 'small'/SC units • Explanation of facilities available (e.g. some SC units provided IC to their own babies, others transfer out, none had out of hours middle-grade paediatricians on site, 5-420 ventilation days annually)	• Mortality period unspecified	• Unadjusted for confounding factors
	Jonas et al.	• Population based study • Defined exclusion criteria (GA <20 and >32 weeks, known BW) • No comparison of baseline characteristics of population by level of unit	• Retrospective • Completeness of data on demographic/confounding factors 93.3-100% • Outcome analysis for all babies 28-31 weeks meeting inclusion criteria	• Comparing outcomes of level 3 and non-level 3 units • No explanation of facilities available in different levels of units	• Neonatal mortality	• Adjusted for confounding factors (gender, mode of delivery, intubation, year of birth, multiple gestation, GA, BW, presentation, maternal age, parity, marital status)

Table S1 Quality assessment of studies characterising neonates by gestational age using modified QUIPS tool

GA (gestational age), BW (birthweight), NICU (neonatal intensive care unit), IC (intensive care), SC (special care), ROM (rupture of membranes), SGA (small for gestational age), IUGR (intrauterine growth retardation), LGA (large for gestational age), SNAP-II (Score for Neonatal Acute Physiology [REF]), US (ultrasound), CP (cerebral palsy), RDS (respiratory distress syndrome), IVH (intraventricular haemorrhage), CLD (chronic lung disease), NEC (necrotising enterocolitis), ROP (retinopathy of prematurity), APH (antepartum haemorrhage), BMI (body mass index), GDM (gestational diabetes mellitus)