Supplementary Online Content

Franz AR, Engel C, Bassler D, et al; ETTNO Investigators. Effects of liberal vs restrictive transfusion thresholds on survival and neurocognitive outcomes in extremely low-birth-weight infants: the ETTNO randomized clinical trial. *JAMA*. Published August 11, 2020. doi:10.1001/jama.2020.10690

This supplementary material has been provided by the authors to give readers additional information about their work.

eTable 1. Red Blood Cell Transfusion Hemoglobin Trigger Thresholds (4 Centers, 36 Patients)

	Liberal Red Blood Cell Transfusion Thresholds		Restrictive Red Blood Cell Transfusion Thresholds	
Postnatal Age	State of Health		State of Health	
	'Critical'	'Non-Critical'	'Critical'	'Non-Critical'
Randomization - 7 days	<8.5mmol/l	<7.2mmol/l	<7.0mmol/l	<5.8mmol/l
8 - 21days	<7.6mmol/l	<6.4mmol/l	<6.2mmol/l	<5.0mmol/l
>21 days	<7.0mmol/l	<5.8mmol/l	<5.6mmol/l	<4.3mmol/l

Four centers (36 patients) used the above hemoglobin trigger thresholds, with hemoglobin concentrations also determined by co-oximetry/photometry from capillary blood samples. Centers had to choose either hematocrit or hemoglobin triggers before starting enrolment to preclude ambiguity.

eTable 2. Number of Red Blood Cell Transfusions and Volumes Transfused After Randomization Until Postmenstrual Age 36 Weeks [a]

Endpoint		Liberal (n=492)	Restrictive (n=521)
Number of RBCT per Infant [a]	Median (1^{st}Q-3^{rd}Q) (n)	2 (1 - 4) (492)	1 (0 - 3) (521)
Number of Infants with **0 RBCT** **1 RBCT** **2 RBCT** **3 RBCT** **4 RBCT** **5 RBCT** **more than 5 RBCT** [a, b, c]	 n (%) n (%) n (%) n (%) n (%) n (%) n (%)	 101 (21%) 95 (19%) 84 (17%) 70 (14%) 44 (9%) 50 (10%) 48 (10%)	 210 (40%) 101 (19%) 74 (14%) 59 (11%) 33 (6%) 10 (2%) 34 (7%)
Cumulative Volume Transfused per Infant [ml] [a]	Median (1^{st}Q-3^{rd}Q) (n)	40 (16 - 73) (486)	19 (0 - 46) (519)
Number of Infants who Received at least 1 RBCT Not Given According to Protocol	n/N (%)	47 / 492 (10%)	97 / 521 (19%)
Number of Infants who Received all RBCT According to Protocol but at least 1 RBCT Not According to the Hct Trigger, but According to 'Exceptional Indications'	n/N (%)	34 / 492 (7%)	66 / 521 (13%)
Number of Infants who Received all RBCT According to Hct Trigger (Including Those Without any Documented RBCT)	n/N (%)	411 / 492 (84%)	358 / 521 (69%)
Number of Infants who did Not Receive an Indicated RBCT Despite a Hct Value Below the Trigger Threshold	n/N (%)	65 / 492 (13%)	5 / 521 (1%)
Number of RBCT Not Given According to Protocol[a]	n/N (%)	60 / 1258 (5%)	137 / 904 (15%)
Number of RBCT Not According to Hct Trigger but According to 'Exceptional Indications' [a]	n/N (%)	72 / 1258 (6%)	193 / 904 (21%)
No. of RBCT Not According to Hct-Trigger with Indication "Major Surgery" [d]	n/N (%)	37 /1258 (3%)	89 / 904 (10%)
No. of RBCT Not According to Hct-Trigger with Indication "Other Emergency" [d]	n/N (%)	32 /1258 (3%)	88 / 904 (10%)

No. of RBCT Not According to Hct-Trigger with Indication "Lactic Acidosis" [d]	n/N (%)	15 / 1258 (1%)	40 / 904 (4%)
No. of RBCT Not According to Hct-Trigger with Indication "Major Bleeding" [d]	n/N (%)	6 / 1258 (0.5%)	18 / 904 (2%)
No. of RBCT Administered in Critical Condition [a, e]	n/N (%)	1028 / 1258 (82%)	823 / 904 (91%)
No. of RBCT Administered in Critical Condition According to Postnatal Age [e]	≤d7 / n	206	129
	d8-d21 / n	386	284
	>d21 / n	436	410
No. of RBCT Administered in Non-Critical Condition [a, e]	n/N (%)	230 / 1258 (18%)	80 / 904 (9%)
No. of RBCT Administered in Non-Critical Condition According to Postnatal Age [e]	≤d7 / n	13	2
	d8-d21 / n	45	17
	>d21 / n	172	61
Storage Age of RBCs at Time of RBCT [days]	Median (Q1-Q3) (n)	8 (4 – 13) (1092)	7 (4 – 13) (790)
No. of RBCT with irradiated RBCs	n/N (%)	786 / 1242 (63%)	608 / 898 (68%)
Time between Irradiation of RBCs and Administration [days]	Median (Q1-Q3) (n)	0 (0 – 1) (711)	0 (0 – 1) (506)

RBCT = red blood cell transfusion; RBCs = red blood cells

[a] This table summarizes the 2162 RBCT administered from randomization until 36 weeks postmenstrual age (PMA).

[b] There were additional 13 infants who had their first RBCT(s) after randomization beyond 36 weeks PMA (9 in the liberal and 4 in the restrictive group), together 16 RBCTs. Furthermore, another 201 RBCTs were administered after PMA 36 weeks in 132 infants (87 in the liberal and 45 infants in the restrictive triggers group) who had already been transfused previously. The total number of RBCT documented in ETTNO sums up to 2379.

[c] 24 infants (6 in liberal and 18 in the restrictive group) of the 311 infants listed herein without RBCT from randomization to 36 weeks PMA had received at least 1 RBCT before randomization (of these 3 (2 liberal, 1 restrictive) had additional RBCT after 36weeks PMA). In total, 88 infants in the liberal group and 189 infants in the restrictive group had no RBCT at all.

[d] More than 1 exceptional indication possible. RBCTs according to exceptional indication did not result in exclusion from the per-protocol population.

[e] In 1 / 2062 RBCT from randomization to 36 weeks postmenstrual age, the clinical condition was not documented.

eTable 3. Study Population Versus Patients Enrolled in German Neonatal Network 2011-2014

		ETTNO Study Population	German Neonatal Network
Inclusion period		07-14-2011 - 11-14-2014	01-01-2011 - 12-31-2014
No. of inclusions	n	1013	4412
Gestational Age at Birth [weeks]	Mean (SD)	26.3 (1.7)	26.2 (1.8)

The German Neonatal Network database was consulted searching infants born between 01-01-2011 and 12-31-2014 (the recruitment period of this trial), and filtered for infants meeting the inclusion criteria of this trial (birth weight <1000g, gestational age at birth <30 weeks and full medical support). Data kindly provided by Wolfgang Göpel, University of Lübeck, Germany, personal communication 07-10-2019.

eTable 4. Patient Recruitment by Study Site

Recruiting Hospital	Country Code	Number of Patients Recruited
University Hospital Ulm	DE	70
Charité Universitätsmedizin Berlin	DE	61
University Hospital Münster	DE	58
Klinikum Neukölln, Berlin	DE	49
Helios Klinikum Erfurt	DE	48
Vestische Kinder-und Jugendklinik Datteln Universität Witten/Herdecke	DE	47
University Hospital Düsseldorf	DE	43
Department of General Pediatrics and Neonatology	DE	42
University Children's Hospital Tübingen	DE	42
University Hospital Magdeburg	DE	41
University Hospital Schleswig-Holstein, Campus Lübeck	DE	38
Klinikum Links der Weser, Bremen	DE	33
Department of Pediatrics I, University Duisburg-Essen	DE	32
University Hospital Frankfurt	DE	32
University Hospital Marburg	DE	30
Klinikum Stuttgart, Olgahospital	DE	29
University Hospital Leipzig	DE	29
Universitätsklinik für Kinder- und Jugendmedizin, Klinikum Oldenburg AöR	DE	24
University Hospital Munich	DE	24
Rigshospitalet Copenhagen	DK	23
DRK-Kinderklinik Siegen	DE	23
Medizinische Hochschule Hannover	DE	23
University Hospital Carl Gustav Carus, Dresden	DE	23

Department of Neonatology, University Children's Hospital Regensburg (KUNO), Regensburg	DE	21
Universitätskinderklinik Bochum, St.-Elisabeth-Hospital	DE	22
University Medical Center Hamburg-Eppendorf	DE	18
University Hospital Aachen	DE	16
Klinikum Augsburg	DE	16
Universitätsmedizin Greifswald	DE	14
Klinikum Nürnberg Süd	DE	12
University Hospital Erlangen	DE	9
Aarhus Universitetshospital	DK	7
The Medical University Ostrava	CZ	5
University Hospital Cologne	DE	5
Altonaer Kinderkrankenhaus Hamburg	DE	3
Tartu University Hospital	EE	1

Recruiting hospitals with number of patients contributed in the order of number of patients enrolled (Country codes are: EE = Estonia, CZ = Czech Republic, DK = Denmark, DE = Germany)

eTable 5. Additional Patient Characteristics Before Randomization

		Liberal (n=492)	Restrictive (n=521)
Courses of Antenatal Corticosteroids **0 (None)** **1 (one incomplete)** **2 (one complete)** **3 (two complete)** **4 (>2 complete)** **unknown**	 n/N (%) n/N (%) n/N (%) n/N (%) n/N (%) n	 52/484 (11%) 129/484 (27%) 248/484 (51%) 53/484 (11%) 2/484 (0.4%) 8	 64/515 (12%) 121/515 (23%) 270/515 (52%) 58/515 (11%) 2/515 (0.4%) 6
Any Surfactant before Randomization	n/N (%)	402/492 (82%)	413/520 (79%)
Any Caffeine before Randomization	n/N (%)	443/492 (90%)	460/521 (88%)
Any Catecholamines before Randomization	n/N (%)	72/492 (15%)	88/521 (17%)
Any Systemic Corticosteroids [a] **before Randomization**	n/N (%)	20/492 (4%)	20/521 (4%)
Any Inhaled Corticosteroids before Randomization	n/N (%)	5/492 (1%)	6/521 (1%)
Any Inhaled Nitric Oxide before Randomization	n/N (%)	18/492 (4%)	17/521 (3%)
Any Prophylactic Indomethacine / Ibuprofen before Randomization	n/N (%)	62/492 (13%)	58/521 (11%)
Any Therapeutic Indomethacine / Ibuprofen before Randomization	n/N (%)	41/492 (8%)	40/521 (8%)
Any Erythropoietin before Randomization	n/N (%)	0/492 (0%)	0/521 (0%)
Focal Intestinal Perforation before Randomization	n/N (%)	4/492 (1%)	3/521 (1%)
NEC > IIa before Randomization	n/N (%)	0/492 (0%)	0/521 (0%)
Pulmonary Hemorrhage Requiring Transfusion before Randomization	n/N (%)	6/492 (1%)	8/521 (2%)
Any I/PVH [b] **before Randomization**	n/N (%)	57/483 (12%)	50/509 (10%)

I/PVH 3°/4° [b] **before Randomization**	n/N (%)	13/483 (3%)	13/509 (3%)
Cystic Periventricular Leukomalacia before Randomization [b]	n/N (%)	1/483 (0.2%)	5/509 (1%)
FiO_2 at Randomization	Median (1^{st}Q-3^{rd}Q) (n)	0.22 (0.21-0.27) (n=491)	0.21 (0.21-0.28) (n=519)
Respiratory Support at Randomization			
None	n/N (%)	4/492 (1%)	6/521 (1%)
CPAP	n/N (%)	236/492 (48%)	235/521 (45%)
Noninvasive Ventilation	n/N (%)	57/492 (12%)	62/521 (12%)
Invasive Ventilation	n/N (%)	195/492 (40%)	218/521 (42%)
Any Tocolytic Therapy	n/N (%)	220/490 (45%)	246/516 (48%)
Any Maternal Antibiotics	n/N (%)	256/489 (52%)	262/518 (51%)
Preterm Premature Rupture of Membranes	n/N (%)	125/485 (26%)	110/511 (22%)
Clinical Chorioamnionitis [c]	n/N (%)	153/491 (31%)	144/518 (28%)
Histological Chorioamnionitis			
Chorioamnionitis with Funisitis	n/N (%)	35/483 (7%)	29/515 (6%)
Chorioamnionitis without Funisitis	n/N (%)	34/483 (7%)	34/515 (7%)
No Chorioamnionitis on histology	n/N (%)	104/483 (22%)	116/515 (23%)
no histology	n/N (%)	310/483 (64%)	336/515 (65%)
unknown	n	9	6
Placental Abruption	n/N (%)	52/458 (11%)	48/492 (10%)
Cause of Preterm Delivery (multiple answers permitted):			
Pre-eclampsia / HELLP / PIH [d]	n/N (%)	107/491 (22%)	123/520 (24%)
Abnormal CTG/pathological Doppler	n/N (%)	184/490 (38%)	210/519 (40%)
Unsupressible Labor	n/N (%)	166/491 (34%)	193/520 (37%)
Chorioamnionitis	n/N (%)	123/491 (25%)	114/519 (22%)

[a] 18 infants in the liberal and 19 infants in the restrictive transfusion trigger group received systemic corticosteroids for treatment of arterial hypotension, all remaining for prevention of bronchopulmonary dysplasia.

[b] I/PVH = Intra-/ Periventricular Hemorrhage / Hemorrhagic Infarction. Data on I/PVH include results of 250 head ultrasounds performed at the day of randomization (i.e., which may have been done shortly after randomization).

[c] Clinical chorioamnionitis was defined as at least one of the following criteria: maternal rectal temperature > 38.0° C, maternal axillary temperature > 38.5° C, maternal leukocytosis > 15000/µl, maternal CRP > 3.0 mg/dl, fetal tachycardia > 170/min

[d] HELLP = High Liver Enzymes, Low Platelets; PIH = Pregnancy-induced Hypertension

eTable 6. Weekly Mean Hematocrit Values [in %] by Treatment Group

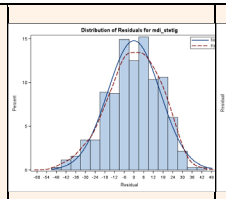
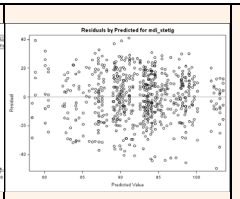
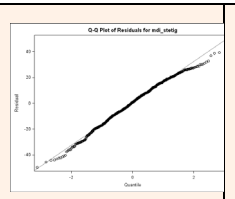
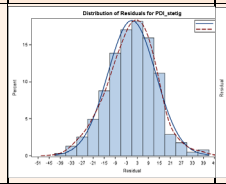
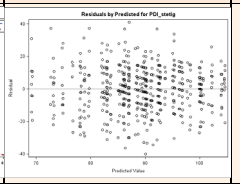
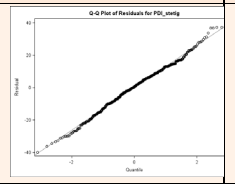
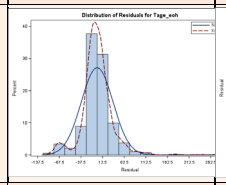
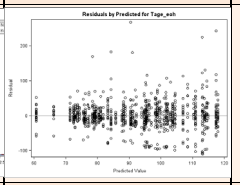
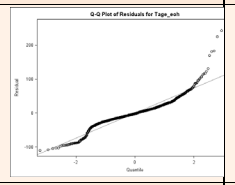
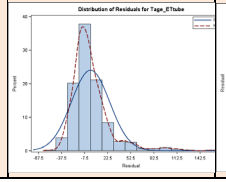
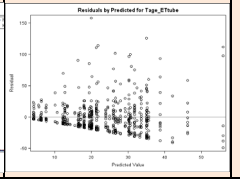
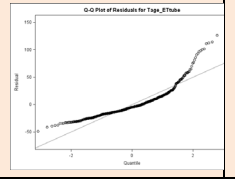
	Liberal	Restrictive		
	Mean (SD) (n)	Mean (SD) (n)	Mean-Difference (95%CI) [a]	p-Value [b]
Week 1 Before Randomization	46.8 (7.2) (n=472)	47.3 (7.1) (n=494)	-0.6 (-1.5 – +0.4)	0.230
Week 1 After Randomization	45.1 (5.7) (n=463)	43.6 (6.5) (n=505)	+1.5 (+0.7 – +2.3)	<0.0001
Week 2	41.9 (5.2) (n=461)	39.5 (5.5) (n=489)	+2.4 (+1.7 – +3.1)	<0.0001
Week 3	39.5 (4.8) (n=443)	36.2 (5.0) (n=468)	+3.2 (+2.6 – +3.8)	<0.0001
Week 4	37.6 (4.6) (n=411)	34.3 (4.8) (n=445)	+3.3 (+2.7 – +4.0)	<0.0001
Week 5	36.1 (4.7) (n=407)	33.4 (4.8) (n=436)	+2.7 (+2.0 – +3.3)	<0.0001
Week 6	35.8 (4.9) (n=385)	32.7 (4.9) (n=399)	+3.1 (+2.4 – +3.8)	<0.0001
Week 7	35.3 (4.6) (n=350)	32.0 (5.0) (n=365)	+3.2 (+2.5 – +4.0)	<0.0001
Week 8	35.0 (4.8) (n=298)	32.0 (5.0) (n=325)	+3.0 (+2.3 – +3.8)	<0.0001
Week 9	35.1 (5.2) (n=240)	31.7 (4.6) (n=258)	+3.4 (+2.5 – +4.3)	<0.0001
Week 10	34.9 (4.5) (n=193)	31.8 (4.4) (n=189)	+3.1 (+2.2 – +4.0)	<0.0001
Week 11	34.3 (5.0) (n=124)	32.2 (4.7) (n=124)	+2.0 (+0.8 – +3.2)	0.0013
Week 12	34.8 (6.0) (n=66)	33.9 (6.0) (n=73)	+0.9 (-1.1 – +2.9)	0.36
Week 13	36.6 (6.8) (n=15)	33.0 (3.5) (n=18)	+3.6 (-0.2 – +7.3)	0.06

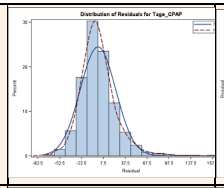
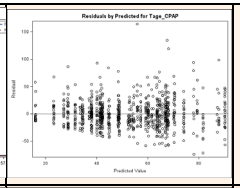
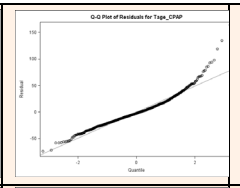
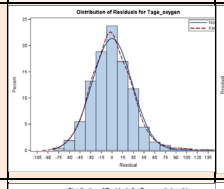
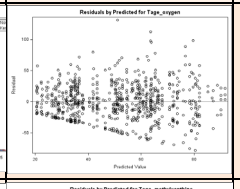
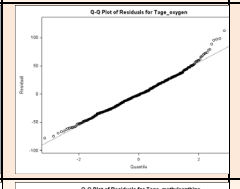
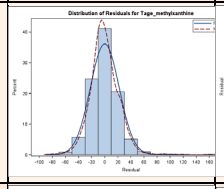
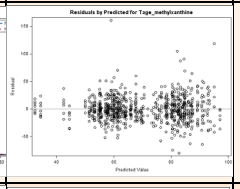
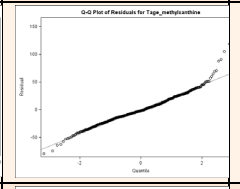
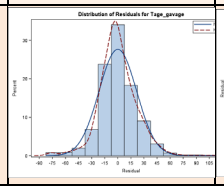
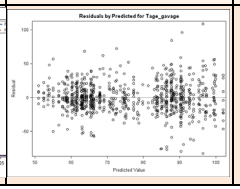
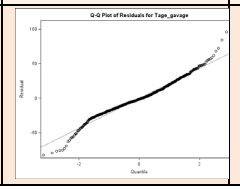
[a] Slight deviations to numerical difference between mean values due to rounding

[b] By two-sided t-test

eTable 7. Model Diagnostics for Primary and Secondary Outcome Analyses in the Main Publication

	Type of Analysis	Factors	Hosmer-Lemeshow 'Goodness of Fit' [a]	Contingency Coefficient C	Residual Histogram	Residual by Predicted Plot	QQ-Plot	R-Square
Primary Endpoint								
Death or Neurodevelopmental Impairment by 24 Months [b]	Logistic Regression	treatment, center and birth weight stratum	p=0.89	0.691		-	-	-
Secondary Endpoints								
Death by 24 Months	Logistic Regression	treatment	-	0.511	-	-	-	-
Cognitive Deficit	Logistic Regression	treatment, center and birth weight stratum	p=0.96	0.688	-	-	-	-
Cognitive Deficit defined as Mental Developmental Index <85	Logistic Regression	treatment	-	0.510	-	-	-	-
Cognitive Deficit defined as Mental Developmental Index <70	Logistic Regression	treatment	-	0.524	-	-	-	-

Mental Developmental Index –Score	ANOVA	treatment, center and birth weight stratum)	-					0.09
Cerebral Palsy	Logistic Regression	treatment	-	0.536	-	-	-	-
Psychomotor Developmental Index –Score	ANOVA	treatment, center and birth weight stratum	-					0.26
Gross Motor Function Classification System-Score I-V	Logistic Regression	treatment	-	0.521	-	-	-	-
Severe Visual Impairment	Logistic Regression	treatment	-	0.516	-	-	-	-
Severe Hearing Impairment	Logistic Regression	treatment	-	0.543	-	-	-	-
Length of Hospital Stay [days]	ANOVA	treatment, center and birth weight stratum	-					0.14
Postnatal Age at End of Invasive Ventilatory Support [days]	ANOVA	treatment, center and birth weight stratum	-	-				0.14

Postnatal Age at Last Positive Pressure Respiratory Support [days]	ANOVA	treatment, center and birth weight stratum	-	-				0.32
Postnatal Age at Last Supplemental Oxygen [days]	ANOVA	treatment, center and birth weight stratum	-	-				0.28
Postnatal Age at Last Caffeine Administration [days]	ANOVA	treatment, center and birth weight stratum	-	-				0.28
Postnatal Age at End of Gavage Feeding [days]	ANOVA	treatment, center and birth weight stratum	-	-				0.29

[a] higher p-values in the Hosmer-Lemeshow test indicate better ‘fit’ of the underlying model. The test was only applied to logistic regressions with more than one factor.

[b] Please refer to eFigure 1 for standardized residual plots for the analysis of the primary outcome.

eTable 8. Timing and Causes of Death

Timing			***Liberal (n=492)***	***Restrictive (n=521)***
Before Discharge [a]	n / N(%)		36 / 488 (7.4%)	38 / 515 (7.4%)
After Discharge to 24 months [b]	n / N (%)		2 / 460 (0.4%)	6 / 491 (1.2%)
Postnatal Age at Death [days]	Median (1^{st}Q-3^{rd}Q) (Min. – Max.)		19 (9 – 72) (3 – 360)	19 (8 – 125) (3 – 346)
Cause				
Pulmonary	n (%)		10 (2.0%)	16 (3.1%)
Cardiac	n (%)		2 (0.4%)	3 (0.6%)
Gastrointestinal	n (%)		14 (2.8%)	11 (2.1%)
Infectious	n (%)		6 (1.2%)	9 (1.7%)
IVH/PVL	n (%)		1 (0.2%)	2 (0.4%)
Multiorgan Failure	n (%)		5 (1.0%)	3 (0.6%)
Total	n (%)		38 (7.7%)	44 (8.4%)

[a] Denominator: all infants followed through discharge (i.e., all infants enrolled except those in whom consent was withdrawn before discharge)

[b] Denominator: all infants followed through 24 months corrected age (i.e., all infants enrolled except those lost to follow-up and in whom consent was withdrawn before 24 months)

eTable 9. Growth Data [a]

		Liberal (n=492)	*Restrictive (n=521)*	*Difference in LS Means (95% CI) [b]*	*p-Value [b]*
Birth Weight [g]	Mean (SD) (n)	753 ± 164 (n=492)	750 ± 163 (n=521)	-[c]	
SDS$_{(Birth\ Weight)}$	Mean (SD) (n)	-0.52 ± 0.93 (n=492)	-0.57 ± 0.84 (n=521)	-[c]	
Weight$_{(PMA36)}$ [g]	Mean (SD) (n)	2113 ± 356 (n=459)	2068 ± 361 (n=483)	+44 (+3 - +85)	0.04
SDS$_{(Weight\ PMA36)}$	Mean (SD) (n)	-1.33 ± 0.88 (n=459)	-1.44 ± 0.92 (n=483)	+0.11 (+0.01 - +0.21)	0.04
Weight$_{(24M\ FU)}$ [kg]	Mean (SD) (n)	10.93 ± 1.50 (n=417)	10.93 ± 1.59 (n=442)	+0.0 (-0.2 - +0.2)	0.98
SDS$_{(Weight\ 24M\ FU)}$	Mean (SD) (n)	-0.71 ± 1.13 (n=417)	-0.78 ± 1.22 (n=442)	+0.06 (-0.09 - +0.21)	0.45
Birth HCU [cm]	Mean (SD) (n)	23.2 ± 1.8 (n=486)	23.2 ± 1.8 (n=514)	-[c]	
SDS$_{(Birth\ HCU)}$	Mean (SD) (n)	-0.48 ± 1.29 (n=480)	-0.51 ± 1.30 (n=510)	-[c]	
HCU$_{(PMA36)}$ [cm]	Mean (SD) (n)	30.6 ± 1.8 (n=452)	30.5 ± 1.8 (n=476)	+0.1 (-0.1 - +0.3)	0.59
SDS$_{(HCU\ PMA36)}$	Mean (SD) (n)	-1.27 ± 1.17 (n=452)	-1.29 ± 1.21 (n=476)	+0.04 (-0.10 - +0.17)	0.62
HCU$_{(24M\ FU)}$ [cm]	Mean (SD) (n)	47.1 ± 1.9 (n=415)	47.2 ± 1.9 (n=441)	-0.0 (-0.3 - +0.2)	0.74
SDS$_{(HCU\ 24M\ FU)}$	Mean (SD) (n)	-1.50 ± 1.54 (n=415)	-1.49 ± 1.64 (n=441)	+0.01 (-0.19 - +0.22)	0.91
Birth Length [cm]	Mean (SD) (n)	32.8 ± 2.7 (n=489)	32.8 ± 2.7 (n=511)	-[c]	
SDS$_{(Birth\ length)}$	Mean (SD) (n)	-0.48 ± 1.34 (n=483)	-0.50 ± 1.39 (n=507)	-[c]	
Length$_{(PMA36)}$ [cm]	Mean (SD) (n)	42.8 ± 2.7 (n=450)	42.5 ± 2.7 (n=475)	+0.2 (-0.1 - +0.6)	0.13
SDS$_{(Length\ PMA36)}$	Mean (SD) (n)	-1.62 ± 1.07 (n=450)	-1.70 ± 1.10 (n=475)	+0.10 (-0.03 - +0.22)	0.13
Length$_{(24M\ FU)}$ [cm]	Mean (SD) (n)	84.7 ± 4.2 (n=417)	84.7 ± 4.2 (n=442)	+0.1 (-0.4 - +0.7)	0.67
SDS$_{(Length\ 24M\ FU)}$	Mean (SD) (n)	-0.58 ± 1.25 (n=417)	-0.66 ± 1.27 (n=442)	+0.09 (-0.07 - +0.26)	0.27

[a] Anthropometric measures were recorded at birth, at 36 weeks postmenstrual age (PMA 36) and at the 24 months follow-up (24M FU). Standard deviation scores were computed based on the growth charts adapted by Fenton and Kim[2].

[b] by ANOVA with factors treatment, birth weight stratum and center

[c] Differences in Least Square Means were not calculated for baseline variables.

eTable 10. Post Hoc Analysis of Cognitive Deficit by Mode of Classification in Survivors

Mode of Classification		***Liberal*** *(n=454 Presumed Survivors)* [a]	***Restrictive*** *(n=477 Presumed Survivors)* [a]
No Follow-up (Lost / Withdrawn)	n/N (%)	32/454 (7%)	30/477 (6%)
Incomplete Follow-up, Unable to Assess Cognitive Deficit	n/N (%)	12/454 (3%)	17/477 (4%)
Cognitive Deficit Classified [b]	n/N (%)	410/454 (90%)	430/477 (90%)
Cognitive Deficit (Bayley II MDI ≥70 and <85)	n/N (%)	72/410 (18%)	77/430 (18%)
Cognitive Deficit (Bayley II MDI ≥50 and <70)	n/N (%)	29/410 (7%)	35/430 (8%)
Cognitive Deficit (Bayley II raw score resulted in MDI<50)	n/N (%)	34/410 (8%)	16/430 (4%)
Cognitive Deficit (Bayley II Attempted but Severe Impairment of the Child Precluded Completion)	n/N (%)	8/410 (2%)	11/430 (3%)
Cognitive Deficit (Bayley III [c] ***Cognitive or Language Composite <85)***	n/N (%)	2/410 (0.5%)	4/430 (1%)
Cognitive Deficit (other Cognitive Test or Rating of Family Pediatrician = Abnormal)	n/N (%)	9/410 (2%)	5/430 (0.9%)
Normal Cognitive Development (Bayley II MDI ≥85)	n/N (%)	235/410 (57%)	248/430 (56%)
Normal Cognitive Development (Bayley III [c]	n/N (%)	0/410 (0%)	4/430 (1%)

Cognitive and Language composite ≥85)			
Normal Cognitive Development **(other Cognitive Test or Rating of Family Pediatrician = Normal)**	n/N (%)	21/410 (5%)	30/430 (7%)

[a] "Presumed Survivors" refers to the number of infants in whom death was not reported – including those lost to follow-up and withdrawn.

[b] In infants not tested by Bayley II, in whom a different neurodevelopmental test result (e.g., Bayley III) or an assessment by the family pediatrician during scheduled well-baby-assessments according to the German Preventive Medical Check-up Program was available[1], cognitive deficit was deemed to be present if cognitive *or* language composite scores were <(Mean-1SD) or the pediatrician's rating was "definitely abnormal" (n=20)), absent, if cognitive *and* language scores were ≥(Mean-1SD) or (in the absence of an abnormal cognitive or language test) the pediatrician's rating was "normal" (n=55)), and missing if results could not be allocated (n=11) based on a review of these data by a team consisting of an independent psychologist, the coordinating investigator and a monitor, all blinded to treatment group assignment.

[c] The German reference data was applied for the conversion of raw scores to scaled scores in Bayley III [3].

eTable 11. Sensitivity Analysis of the Primary Outcome [a]

Primary Endpoint		**Liberal**	**Restrictive**	**p-Value**
Death or Neurodevelopmental Impairment	n / N (%)	242 / 492 (49.2%)	248 / 521 (47.6%)	
	Adj. OR (95%CI)[b,c]	1.04 (0.80 - 1.35)		0.78
	Absolute Difference (95% CI) [d]	+1.6% (-4.6% - +7.7%)		

[a] As described in the study protocol, a pre-defined analysis of a worst case scenario was performed, for which all infants with unknown 2 year outcome were counted as if they had an adverse outcome (i.e. as if the primary outcome of death or neurodevelopmental impairment had been present):

In the restrictive threshold group 43 additional infants were counted and in the liberal threshold group 42 additional infants were counted as having reached the primary outcome.

[b] Logistic Regression with factors treatment, center and birth weight stratum (400-749g / 750-999g).

[c] Model quality: p=0.99 according to Hosmer-Lemeshow ‘goodness of fit’, where higher p-values indicate better fit of the model.

[d] Absolute differences were calculated post hoc for binary outcomes, without adjustment for center and birth weight stratum

eTable 12. Primary and Secondary Endpoints in the Per Protocol Population

Primary Endpoint		**Liberal**	**Restrictive**	**p-Value**
Death or Neurodevelopmental Impairment by 24 months	n / N (%)	156 / 343 (45.5%)	144 / 374 (38.5%)	
	OR (95%CI)[a]	1.25 (0.90 - 1.72)		0.18
	Risk Difference (95%CI)	+7.0% (-0.2% - +14.2%)		
Secondary Endpoints		**Liberal**	**Restrictive**	**p-Value**
Death by24 Months	n / N (%)	33 / 350 (9.4%)	30 / 385 (7.8%)	
	OR (95%CI) [b]	1.25 (0.74 – 2.08)		0.43
	Risk Difference (95%CI)	+1.6% (-2.4% - +5.7%)		
Cognitive Deficit[d]	n / N (%)	118 / 308 (38.3%)	106 / 341 (31.1%)	
	OR (95%CI) [a]	1.27 (0.89 - 1.79)		0.19
	Risk Difference (95%CI)	+7.2% (-0.1% - +14.5%)		
Cognitive Deficit Defined as MDI<85[d]	n / N (%)	110 / 285 (38.6%)	100 / 312 (32.1%)	
	OR (95%CI) [b]	1.33 (0.95 – 1.85)		0.10

	Risk Difference (95%CI)	+6.5% (-1.1% - +14.2%)		
Cognitive Deficit Defined as MDI<70[d]	n / N (%)	52 / 285 (18.2%)	40 / 312 (12.8%)	
	OR (95%CI)[b]	1.52 (0.97 – 2.38)		0.07
	Risk Difference (95%CI)	+5.4% (-0.4% - +11.2%)		
MDI-Score[d]	Mean (SD) (n)	92.4 (16.8) (256)	93.4 (17.1) (297)	
	Difference in LS Means (95%CI)[c]	+0.3 (-3.1 - +2.5)		0.83
Cerebral Palsy	n / N (%)	14 / 314 (4.5%)	14 / 352 (3.9%)	
	OR (95%CI)[b]	1.14 (0.53 – 2.38)		0.76
	Risk Difference (95%CI)	+0.5% (-2.6% - +3.5%)		
PDI-Score[d]	Mean (SD) (n)	90.3 (14.5) (236)	91.0 (14.6) (254)	
	Difference in LS Means (95%CI)[c]	+0.4 (-2.9 - +2.0)		0.71
GMFCS-Score I-V[e]	n / N (%)	24 / 313 (7.7%)	23 / 349 (6.6%)	
	OR (95%CI)[b]	1.18 (0.65 – 2.13)		0.59

	Risk Difference (95%CI)	+1.1% (-2.9% - +5.0%)		
Severe Visual Impairment	n / N (%)	6 / 313 (1.9%)	8 / 352 (2.3%)	
	OR (95%CI) [b]	0.84 (0.28 – 2.44)		0.75
	Risk Difference (95%CI)	-0.4% (-2.5% - +1.8%)		
Severe Hearing Impairment	n / N (%)	1 / 313 (0.3%)	3 / 352 (0.9%)	
	OR (95%CI) [b]	- not analyzed -		-
	Risk Difference (95%CI)	-0.5% (-1.7% - +0.6%)		
Length of Hospital Stay [days]	Mean (SD) (n)	91 (40) (368)	87 (34) (400)	
	Difference in LS Means (95%CI) [c]	+1.6 (-3.4 – +6.5)		0.54
Postnatal Age at End of Invasive Ventilatory Support[f] [days]	Mean (SD) (n)	21 (27) (224)	20 (24) (236)	
	Difference in LS Means (95%CI) [c]	-0.4 (-4.9 – +4.0)		0.86
Postnatal Age at Last Positive Pressure Respiratory Support[f] [days]	Mean (SD) (n)	53 (31) (337)	50 (28) (374)	

	Difference in LS Means (95%CI) [c]	+1.1 (-2.6 – +4.8)		0.56
Postnatal Age at Last Supplemental Oxygen[f] [days]	Mean (SD) (n)	52 (33) (273)	44 (31) (298)	
	Difference in LS Means (95%CI) [c]	+5.0 (-0.3 – -9.7)		0.04
Postnatal Age at Last Caffeine Administration[f] [days]	Mean (SD) (n)	69 (27) (313)	68 (25) (340)	
	Difference in LS Means (95%CI) [c]	-1.3 (-4.7 – +2.2)		0.48
Postnatal Age at End of Gavage Feeding[f] [days]	Mean (SD) (n)	75 (25) (314)	72 (24) (351)	
	Difference in LS Means (95%CI) [c]	-0.2 (-3.4 – +3.1)		0.91

OR = odds ratio, LS = least square, CI = confidence interval

[a] Logistic Regression with factors treatment, center and birth weight stratum (400-749g / 750-999g).

[b] Logistic regression with factor treatment only (a reduced model was fitted because iterations did not converge).

[c] ANOVA with factors treatment, center and birth weight stratum (400-749g / 750-999g).

[d] Whereas for the cognitive deficit as component of the primary outcome other cognitive assessments were taken into account in case no Bayley test was available as described in 'methods', the rates of MDI<85 / <70 and PDI<85 are based on infants with attempted Bayley II only (still including infants whose raw scores were so low that MDI/PDI was <50). Analyses of the MDI/PDI as continuous variable are based on complete Bayley II data, only.

[e] GMFCS had been defined as further endpoint, not as secondary endpoint in the study protocol. The presentation of the GMFCS data herein deviates from the statistical analysis plan, which foresaw presentation as median (Q1-Q3), due to the low rate of GMFCS>0.

[f] Analyses of duration of various forms of support are limited to those infants who a) received this therapeutic intervention (e.g., 158 infants in the liberal group and 155 infants in the restrictive group never received invasive respiratory support through an endotracheal tube) *and* b) discontinued the intervention before the day of discharge home.

This table shows numerical differences in favor of the restrictive group, which were all not statistically significant but more pronounced than in the analysis of the population of all randomized patients. It must be noted that ‘non-protocol-justified’ transfusion predominantly occurred in the restrictive group and there preferentially in the low birth weight stratum, and, by contrast, missed transfusions predominantly occurred in the liberal group and there in the high birth weight stratum, the proportions of low birth weight versus high birth weight were skewed in the per-protocol population, i.e., the proportions for low and high birth weight stratum were 190 / 370 (51%) and 180 / 370 (49%) in the liberal group versus 177 / 411 (43%) and 234 / 411 (57%) in the restrictive group in the per-protocol population.

eTable 13. Pre-Defined Subgroup Analysis 1: Primary Outcome and Key Secondary Outcomes by Birth Weight Stratum and Transfusion Thresholds

		Birth Weight <750g		Birth Weight 750-999g	
		Liberal	Restrictive	Liberal	Restrictive
Primary Outcome	n/N (%)	125 / 230 (54.3%)	131 / 243 (53.9%)	75 / 220 (34.0%)	74 / 235 (31.4%)
	OR (95%CI) p-Value[a]	1.02 (0.71 – 1.54) p=0.82		1.06 (0.70 – 1.61) p=0.78	
Cerebral Palsy	n/N (%)	7 / 203 (3.4%)	15 / 209 (7.1%)	11 / 216 (5.1%)	10 / 234 (4.3%)
	OR (95%CI) p-Value [b]	0.46 (0.18 – 1.16) p=0.10		1.20 (0.50 – 2.85) p=0.68	
MDI Score [d]	Mean (SD) (n)	91 (18) (n=158)	90 (18) (n=163)	94 (15) (n=178)	95 (17) (n=197)
	Difference in LS Means (95%CI) p-Value [c]	+0.46 (-3.44 – +4.37) p=0.82		-0.06 (-3.30 – +3.18) p=0.97	

[a] Logistic Regression with factors treatment and center; [b] Logistic Regression with factor treatment; [c] ANOVA with factors treatment and center; [d] MDI = mental developmental index score of the Bayley Scales of Infant and Toddler Development (2nd edition)

There was no significant interaction, i.e., difference in treatment effect between subgroups, as assessed by Breslow-Day test (p=0.71).

eTable 14. Pre-Defined Subgroup Analysis 2: Primary Outcome and Key Secondary Outcomes by Gender and Transfusion Thresholds

		Male		Female	
		Liberal	Restrictive	Liberal	Restrictive
Primary Outcome	n/N (%)	116 / 224 (51.8%)	119 / 238 (50.0%)	84 / 226 (37.2%)	86 / 240 (35.8%)
	OR (95%CI) p-Value [a]	1.19 (0.80 – 1.78) p=0.38		1.00 (0.66 – 1.52) p=0.99	
Cerebral Palsy	n/N (%)	11 / 199 (5.5%)	15 / 215 (6.9%)	7 / 220 (3.2%)	10 / 228 (4.4%)
	OR (95%CI) p-Value [b]	0.78 (0.42 – 1.75) p=0.54		0.71 (0.27 – 1.92) p=0.51	
MDI Score [d]	Mean (SD) (n)	90 (16) (n=155)	89 (18) (n=170)	95 (17) (n=181)	95 (17) (n=190)
	Difference in LS Means (95%CI) p-Value [c]	+0.20 (-3.58 – +3.98) p=0.92		-0.11 (-3.52 – +3.30) p=0.95	

[a] Logistic Regression with factors treatment, center and birth weight stratum (400-749g / 750-999g); [b] Logistic Regression with factor treatment; [c] ANOVA with factors treatment, center and birth weight stratum (400-749g / 750-999g); [d] MDI = mental developmental index score of the Bayley Scales of Infant and Toddler Development (2nd edition)

There was no significant interaction, i.e., difference in treatment effect between subgroups, as assessed by Breslow-Day test (p=0.96).

eTable 15. Pre-Defined Subgroup Analysis 3: Primary Outcome and Key Secondary Outcomes by Institutional SpO_2-Target Range and Transfusion Thresholds

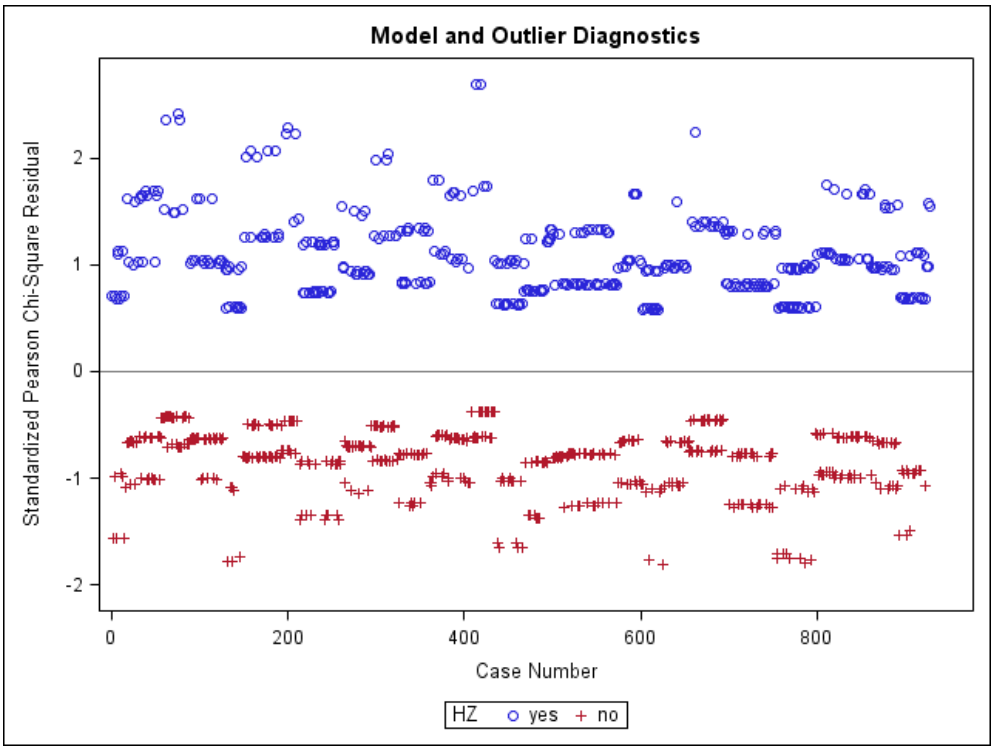
		Lower SpO_2-Target Range		Higher SpO_2-Target Range	
		Liberal	Restrictive	Liberal	Restrictive
Primary Outcome	n/N (%)	80 / 198 (40.4%)	90 / 209 (43.6%)	119 / 251 (47.4%)	111 / 265 (41.9%)
	OR (95%CI) p-Value [a]	0.88 (0.59 – 1.33) p=0.55		1.25 (0.88 – 1.79) p=0.21	
Cerebral Palsy	n/N (%)	9 / 189 (4.8%)	10 / 191 (5.2%)	9 / 229 (3.9%)	13 / 248 (5.2%)
	OR (95%CI) p-Value [b]	0.90 (0.36 - 2.27) p=0.83		0.74 (0.31 – 1.75) p=0.50	
MDI Score [d]	Mean (SD) (n)	93 (16) (n=150)	92 (17) (n=150)	92 (17) (n=186)	93 (18) (n=209)
	Difference in LS Means (95%CI) p-Value [c]	+0.88 (-2.864.62 – +4.62) p=0.64		-0.34 (-3.75 – +3.06) p=0.84	

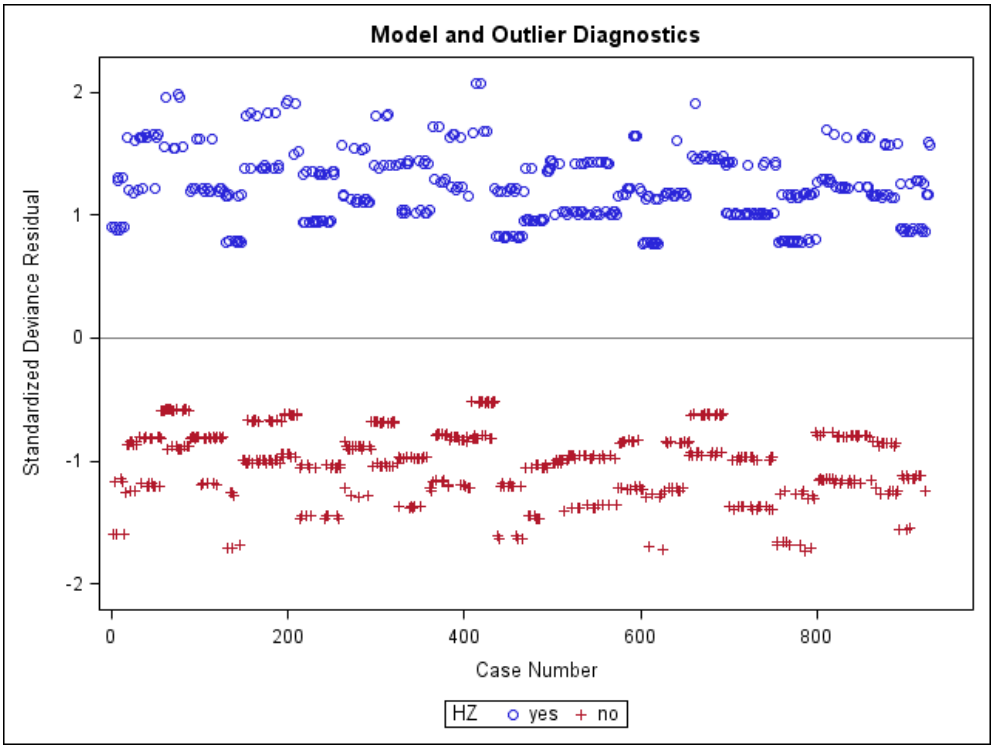
[a] Logistic Regression with factors treatment and birth weight stratum (400-749g / 750-999g); [b] Logistic Regression with factor treatment; [c] ANOVA with factors treatment and birth weight stratum (400-749g / 750-999g); [d] MDI = mental developmental index score of the Bayley Scales of Infant and Toddler Development (2nd edition)

For this table, SpO_2-target ranges reported by each center were dichotomized and classified as 'lower' if the center-value of the institutional SpO_2 target range was <91% and classified as 'higher' if the center-value was ≥91%. In one center, recruiting 5 patients, SpO_2-target ranges could not be allocated, and hence 4 infants from restrictive and 1 from liberal were not taken into account in this analysis.

There was no significant interaction, i.e., difference in treatment effect between subgroups, as assessed by Breslow-Day test (p=0.21).

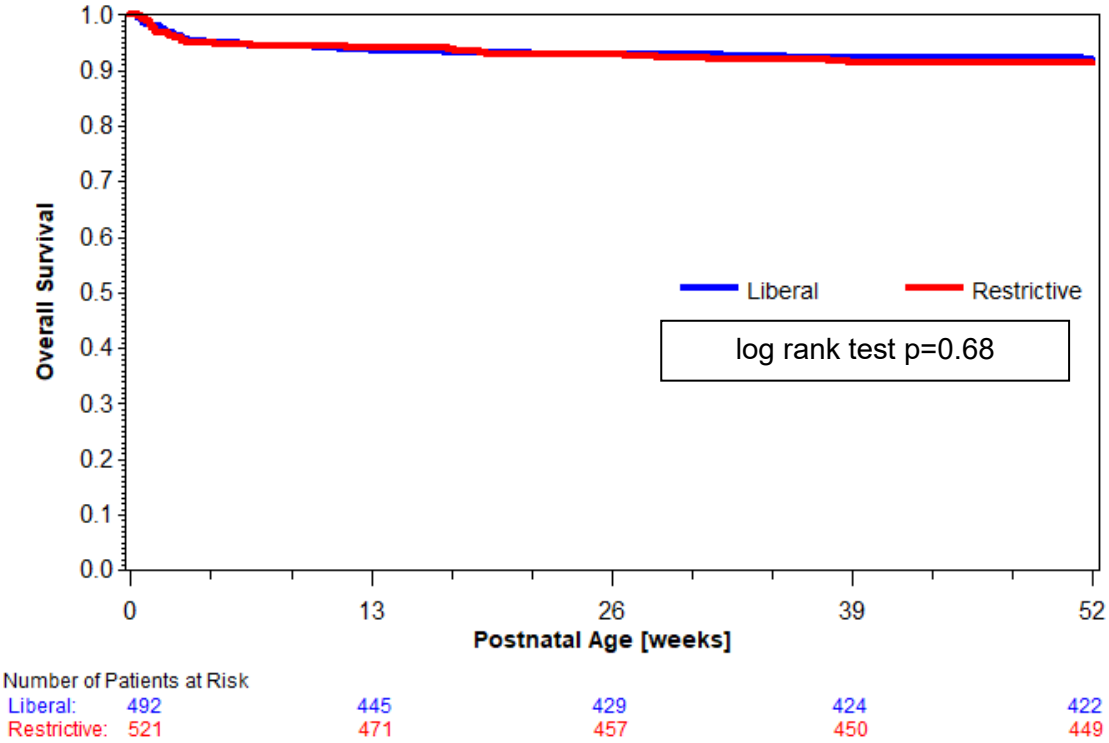
eFigure 1. Standardized Residual Plots (Analysis of the Primary Outcome)





Standardized residual plots are provided for visual analysis of the assumption of ‘no outliers in standardized residuals’ underlying the analysis of the primary outcome.

eFigure 2. Overall Survival



Number of Patients at Risk

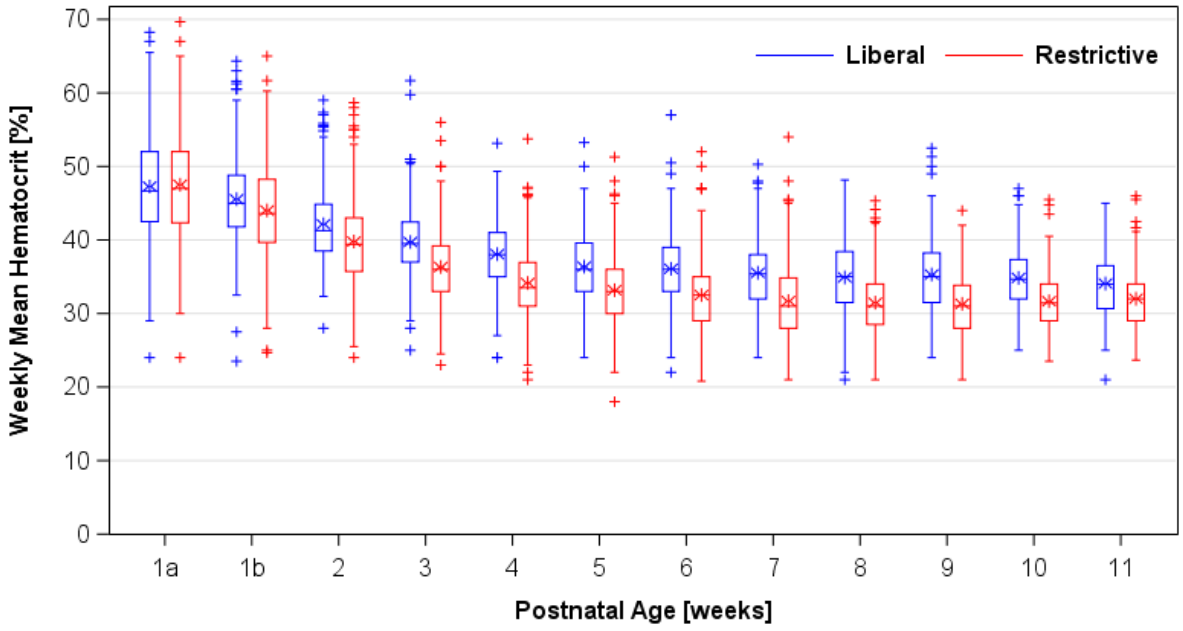
Liberal:	492	445	429	424	422
Restrictive:	521	471	457	450	449

Survival data were analyzed by the method of Kaplan and Meier.

This is the Kaplan-Meier plot indicating the proportion of surviving infants according to treatment group. The table underneath the graph reports the numbers of infants in the study in the respective week of postnatal age by treatment group. No child died beyond 52 weeks postnatal age.

The median (IQR) lengths of observation were 119 (115 – 121) weeks and 119 (115-122) weeks in the liberal and the restrictive threshold group, respectively.

eFigure 3. Treatment Effect on Weekly Mean Hematocrit in the Per-Protocol Population



Number of RBC-Transfusions (Number of Patients transfused)												
Liberal:	121(94)	163(124)	145(120)	166(143)	89(84)	99(94)	69(62)	51(51)	55(50)	30(30)	19(19)	17(16)
Restrictive:	112(86)	89(74)	87(73)	98(87)	61(52)	69(61)	38(33)	32(28)	27(26)	10(10)	13(12)	8(7)
Number of Patients with documented Hematocrit												
Liberal:	353	347	344	330	302	292	278	252	212	169	137	91
Restrictive:	393	397	382	363	344	337	306	280	243	191	131	78

The *weekly mean hematocrit* values of all hematocrit values documented for each infant in the per-protocol population in that week are depicted by treatment group. (Limited to hematocrit values documented until 36 weeks postmenstrual age and truncated when less than 20% of the population remained, i.e., at 11 weeks postnatal age.).

Boxes denote median, first and third quartile, the asterisk within the box indicates the mean, and whiskers indicate highest and lowest values within 1.5 x the interquartile range and dots outside the box outlying data.

The table underneath reports the number of transfusions (number of infants transfused) in the per-protocol population in each week of postnatal age by treatment group in the upper rows (limited to RBC-transfusions administered until 36 weeks postmenstrual age) as well as the numbers of infants in the per-protocol population with at least one documented hematocrit value in that week of postnatal age in the lower rows.

1a refers to the days of the first week of postnatal age before/until randomization and 1b refers to the days of the first week of postnatal age after randomization.

eAppendix. Additional Information on Sample Size Calculations

The sample size calculation were based on the results of the PINT trial's 18 months follow-up report (Whyte RK et al. Pediatrics 2009)[4]. Unfortunately, the rate of the primary outcome selected for the ETTNO trial was not reported by Whyte et al. (they reported death or NDI, where cognitive deficit was defined by MDI<70 (94/208 (45.2%) vs 82/213 (38.5%)), and they reported separately the rates for MDI<70 (38/156 (24.4%) vs. 29/165 (17.6%)) and MDI<85 (70/156 (44.9%) vs. 56/165 (33.9%)). The rates for a composite outcome according to the ETTNO definitions (with cognitive deficit defined as MDI<85) were "estimated" from the available data by adding the difference (number of infants with MDI<85 – the number of infants with MDI<70) to the number of infants with death or NDI in each group: (94+32)/208 (61%) versus (82+27)/213 (51%). These estimates may not exactly reflect the true rates of a composite outcome because some of the infants with mild cognitive deficit could additionally have had cerebral palsy or visual or hearing impairments.

eReferences

1. Maschke C, Ellenrieder B, Hecher K, Bartmann P. National screening program vs. standardized neurodevelopmental follow-up. J Perinat Med 2009;37:87-90.
2. Fenton TR, Kim JH. A systematic review and meta-analysis to revise the Fenton growth chart for preterm infants. BMC Pediatrics 2013;13:59
3. Reuner G, Rosenkranz J. Bayley Scales of Infant Development Third Edition – Deutsche Fassung. Manual zur Durchführung und Auswertung der Bayley III Skalen. Pearsson Assessment & Information GmbH. Frankfurt/Main, Germany, 2014
4. Whyte RK, Kirpalani H, Asztalos EV, et al. Neurodevelopmental outcome of extremely low birth weight infants randomly assigned to restrictive or liberal hemoglobin thresholds for blood transfusion. Pediatrics 2009;123:207-13