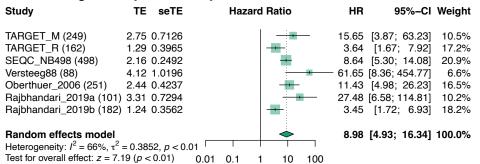
Age >1.5 yr vs. <1.5 yr overall survival association



B MYCN amplified vs. not amplified overall survival association

Study	TE	seTE	Hazard Ratio	HR	95%-CI Weight
TARGET_M (249)	0.44	0.1858		1.55	[1.07; 2.23] 15.3%
TARGET_R (162)	0.32	0.2685	+	1.38	[0.82; 2.34] 14.6%
SEQC_NB498 (498)	2.05	0.2004	-	7.80	[5.26; 11.55] 15.2%
Versteeg88 (88)	1.89	0.3728	- 1	6.64	[3.20; 13.79] 13.5%
Oberthuer_2006 (251)	2.39	0.3408	 	- 10.90	[5.59; 21.26] 13.8%
Rajbhandari_2019a (101)	1.44	0.3321		4.24	[2.21; 8.13] 13.9%
Rajbhandari_2019b (182)	2.19	0.3523	-	8.94	[4.48; 17.83] 13.7%
Random effects model				4.54	[2.29; 9.00] 100.0%
Heterogeneity: $I^2 = 91\%$, τ^2	= 0.76	37, $p < 0.01$			
Test for overall effect: $z = 4.3$	34 (p <	(0.01) 0.	1 0.5 1 2 10		

C INSS stage 4 vs. other stages overall survival association

Study	IE	selE	Hazard Ratio	нн	95%-CI	Weight	
TARGET_M (249)	18.36	1696.0079 ←		→ 93854171.92	[0.00; Inf]	0.0%	
TARGET_R (162)	1.80	0.5135		6.02	[2.20; 16.48]	9.6%	
SEQC_NB498 (498)	2.16	0.2371		8.66	[5.44; 13.78]	45.1%	
Versteeg88 (88)	2.86	0.6117		17.53	[5.29; 58.12]	6.8%	
Oberthuer_2006 (251)	2.35	0.3860		10.44	[4.90; 22.25]	17.0%	
Rajbhandari_2019a (101)	2.73	0.5318		15.32	[5.40; 43.45]	9.0%	
Rajbhandari_2019b (182)	2.23	0.4490		9.30	[3.86; 22.41]	12.6%	
Random effects model			*	9.62	[7.04; 13.14]	100.0%	
Heterogeneity: $I^2 = 0\%$, $\tau^2 =$	0, p = 0).83	1 1 1 1 1	I			
Test for overall effect: $z = 14$.22 (p <	(0.01) 0.10	0.2 0.5 1 2 5	100			

>1.5 years vs. <1.5 years

Study/Source	0	bserved SMD [95% CI]
TARGET_microarray	-	0.29 [-0.08, 0.66]
TARGET_RNA-seq		0.21 [-0.16, 0.59]
SEQC_NB498	-	-0.11 [-0.29, 0.07]
Versteeg88	-	0.02 [-0.40, 0.44]
Oberthuer_2006		0.28 [0.01, 0.54]
Rajbhandari_2019a		0.25 [-0.14, 0.65]
Rajbhandari_2019b		0.15 [-0.15, 0.44]
RE Model Heterogeneity: I ² = 39%, Test for overall effect: z =	t ² = 0.02, p = 1,69, p + 0.0	0.13 [-0.02, 0.28] = 0.16
-1	0 1	

Standardized Mean Difference

E MYCN amplified vs. not amplified

Study/Source	Ob	served SMD [95% CI]		
TARGET_microarray	⊢≡ -	0.13 [-0.15, 0.41]		
TARGET_RNA-seq	-	-0.02 [-0.40, 0.37]		
SEQC_NB498	-	0.40 [0.17, 0.63]		
Versteeg88		-0.64 [-1.19, -0.09]		
Oberthuer_2006		0.07 [-0.31, 0.44]		
Rajbhandari_2019a		0.21 [-0.28, 0.70]		
Rajbhandari_2019b	-	0.09 [-0.28, 0.46]		
RE Model Heterogeneity: $I^2 = 53\%$, $\tau^2 = 0.04$, p = 0.035 Test for overall effect: $r = 0.07$, $r = 0.03$				
-2	-0.5	1		

Standardized Mean Difference

F INSS Stage 4 vs. other stages

Study/Source		Observed SMD [95% CI]		
TARGET_microarray		-0.24 [-0.64, 0.15]		
TARGET_RNA-seq		-0.17 [-0.58, 0.24]		
SEQC_NB498		0.24 [0.04, 0.44]		
Versteeg88		-0.36 [-0.79, 0.08]		
Oberthuer_2006		-0.36 [-0.70, -0.02]		
Rajbhandari_2019a		-0.31 [-0.73, 0.11]		
Rajbhandari_2019b		0.18 [-0.21, 0.56]		
Ackermann_2018	-	-0.47 [-0.74, -0.19]		
RE Model Heterogeneity: $I^2 = 65\%$, $\tau^2 = 0.05$, $p = 0.00085$ Test for overall effect: $z = -1.04$, $p = 0.1$				
	_1 0	1		

Standardized Mean Difference

G relapse/progression vs. no event

Study/Source	C	Observed SMD [95% CI]		
TARGET_microarray TARGET_RNA-seq SEQC_NB498 Versteeg88 Oberthuer_2006 Rajbhandari_2019a Rajbhandari_2019b Asgharzadeh_2006 Ackermann_2018		-0.08 [-0.35, 0.19] -0.04 [-0.39, 0.30] 0.02 [-0.18, 0.22] -0.27 [-0.71, 0.17] -0.21 [-0.53, 0.11] -0.23 [-0.66, 0.19] -0.23 [-0.62, 0.16] -0.77 [-1.18, -0.37] -0.25 [-0.55, 0.05]		
BE Model Beterogeneity: $l_{z}^{2} = 41\%$, $\tau^{2} = 0.\overline{0}^{2}$, $p = 0.709$ [-0.34, -0.05] Test for overall effect: $t = \frac{1}{12} $				
_2	_0.5	1		

Standardized Mean Difference

Figure S2. Known NBL prognostic factors consistently associate with outcome across different NBL studies but not NE scores

A-C. Age at diagnosis (**A**), MYCN amplification (**B**), and INSS stage 4 (**C**) are significantly associated with worse outcome. **D-G.** Meta-analyses of NE scores standardized mean difference between groups stratified by different clinical features. Age at diagnosis (**D**), MYCN amplification (**E**) and INSS stage 4 (**F**) are not significantly associated with NE scores. Lower NE scores are associated with relapse (**G**). Note that although one out of the nine studies showed a significant result, the overall result from the meta-analysis is statistically significant.