

Table 1: Coefficients from the regression models for **emergency tracheotomy** performance time and the risk of no injury, mild injury, severe injury and failure with 95% confidence intervals (95% CI) derived from the best piecewise linear models.

a) The effect of attempt on **emergency tracheotomy** performance time expressed as geometric mean ratio (GMR), reflecting the multiplicative change per attempt in **emergency tracheotomy** performance time. **Emergency tracheotomy** performance time decreased in both phases with a steeper decrease in phase I.

b) The effect of attempt on the occurrence of no injury, mild injury, severe injury and failure expressed as relative risk ratio, reflecting the multiplicative changes per attempt in the relative risk vs no injury of minor injury, severe injury and failure, respectively. The relative risks of minor and severe injuries decreased from attempt 1 to 2 and remained constant for all subsequent attempts.

a) ET performance time	Geometric mean ratio (95% CI)	P-value
Phase I (attempts 1-4)	0.89 (0.87 - 0.91)	<0.001
Phase II (attempts 4-10)	0.96 (0.95 - 0.97)	<0.001
b) Injuries	Relative risk ratio (95% CI)	P-value
Attempts 1-2		
Minor vs no injury	0.21 (0.08 - 0.58)	0.002
Severe vs no injury	0.08 (0.03 - 0.18)	<0.001
Failure vs no injury	0.37 (0.08 - 1.82)	0.22
Attempts 2-10		
Minor vs no injury	1.09 (0.98 - 1.21)	0.11
Severe vs no injury	0.97 (0.87 - 1.07)	0.54
Failure vs no injury	1.06 (0.94 - 1.20)	0.34

Table 3: The effects of attempt, age and the interaction of attempt and age on **emergency tracheotomy (ET)** performance time as geometric mean ratio (GMR) with 95% confidence intervals (95% CI) derived from a piecewise linear model with two separate effects for attempts 1-4 (phase I) and 4-10 (phase II). The GMR reflects the multiplicative change in **emergency tracheotomy** performance time per attempt and per decade of age, respectively. The interaction (indicated by #) reflects the effect of age on the learning curve.

	Geometric mean ratio (95% CI)	P-value
phase I	0.88 (0.75 - 1.02)	0.09
phase II	0.87 (0.81 - 0.94)	<0.001
age in decades	0.97 (0.84 - 1.11)	0.64
phase I # age in decades	1.00 (0.97 - 1.04)	0.87
phase II # age in decades	1.02 (1.01 - 1.04)	0.008

Table 4: The effects of attempt, **sex** and the interaction of attempt and **sex** on **emergency tracheotomy (ET)** performance time as geometric mean ratio (GMR) with 95% confidence intervals (95% CI) derived from a piecewise linear model with two separate effects for attempts 1-4 (phase I) and 4-10 (phase II). The GMR reflects the multiplicative change in ET performance time per attempt and of men compared to women, respectively. The interaction (indicated by #) reflects the effect of male **sex** on the learning curve.

	Geometric mean ratio (95% CI)	P-value
phase I	0.85 (0.82 - 0.89)	<0.001
phase II	0.96 (0.94 - 0.98)	<0.001
male <b>sex</b>	0.74 (0.61 - 0.90)	0.003
phase I # male <b>sex</b>	1.07 (1.01 - 1.12)	0.015
phase II # male <b>sex</b>	1.00 (0.97 - 1.02)	0.77

Table 5: The effects of attempt, experience and the interaction of attempt and experience on **emergency tracheotomy (ET)** performance time as geometric mean ratio (GMR) with 95% confidence intervals (95% CI) derived from a piecewise linear model with two separate effects for attempts 1-4 (phase I) and 4-10 (phase II). The GMR reflects the multiplicative change in **emergency tracheotomy** performance time per attempt and per decade of experience, respectively. The interaction (indicated by #) the effect of experience on the learning curve.

	Geometric mean ratio (95% CI)	P-value
phase I	0.89 (0.85 - 0.94)	<0.001
phase II	0.94 (0.92 - 0.96)	<0.001
experience in decades	1.04 (0.91 - 1.19)	0.56
phase I # experience in decades	1.00 (0.96 - 1.03)	0.79
phase II # experience in decades	1.02 (1.00 - 1.04)	0.025

Table 6: Overview of the tested piecewise linear models for **emergency tracheotomy (ET)** performance time. All models with any knots showed a better fit than the linear model with zero knots in likelihood ratio tests ( $p$ -values compared to linear model). The model with one knot at attempt four was selected because it showed the lowest Akaike and Bayesian information criteria. None of the bigger models was significantly better than the selected model in likelihood ratio tests ( $p$ -value compared to selected model).

Number of knots	Location of knot(s)	Degrees of freedom	Log-likelihood	Akaike information criteria	Bayesian information criteria	P-value compared to linear model (*)	P-value compared to selected model (†)
0*		4	-62.78	133.6	150.4		
1†	4.0	5	-52.48	115.0	136.0	<0.001	
1	5.5	5	-55.79	121.6	142.7	<0.001	
2	4.0, 7.0	6	-52.48	117.0	142.2	<0.001	0.95
2	4.0, 8.0	6	-52.43	116.9	142.1	<0.001	0.76
3	3.0, 4.0, 7.0	7	-52.41	118.8	148.3	<0.001	0.94
3	3.0, 5.5, 8.0	7	-53.23	120.5	150.0	<0.001	1.00
4	2.2, 4.4, 6.6, 8.8	8	-52.60	121.2	154.9	<0.001	1.00
8‡	at each attempt	12	-50.68	125.4	175.9	0.002	0.83

\*linear model, †selected model, ‡categorical model

Table 7: Overview of the tested piecewise linear models for the probability of no injury, mild injury, severe injury and failure. All models with any knots fitted better than the linear model with zero knots in likelihood ratio tests ( $p$ -value compared to linear model). The model with one knot at attempt two was selected because it showed the lowest Bayesian and second lowest Akaike information criteria. None of the bigger models was significantly better than the selected model in likelihood ratio tests ( $p$ -value compared to selected model).

Number of knots	Location of knot(s)	Degrees of freedom	Log-likelihood	Akaike information criteria	Bayesian information criteria	P-value compared to linear model (*)	P-value compared to selected model (†)
0*		6	-504.90	1021.8	1047.1		
1†	2.0	9	-487.48	993.0	1030.9	<0.001	
1	5.5	9	-500.40	1018.8	1056.7	0.029	
2	2.0, 3.0	12	-484.22	992.4	1043.0	<0.001	0.09
2	2.0, 5.0	12	-486.73	997.5	1048.0	<0.001	0.69
2	4.0, 7.0	12	-494.20	1012.4	1063.0	0.002	1.00
3	2.0, 3.0, 5.0	15	-483.71	997.4	1060.6	<0.001	0.27
3	3.0, 5.5, 8.0	15	-489.24	1008.5	1071.7	<0.001	1.00
4	2.2, 4.4, 6.6, 8.8	18	-486.62	1009.2	1085.1	<0.001	1.00
8‡	at each attempt	30	-480.61	1021.2	1147.7	0.002	0.88

\*linear model, †selected model, ‡categorical model