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)
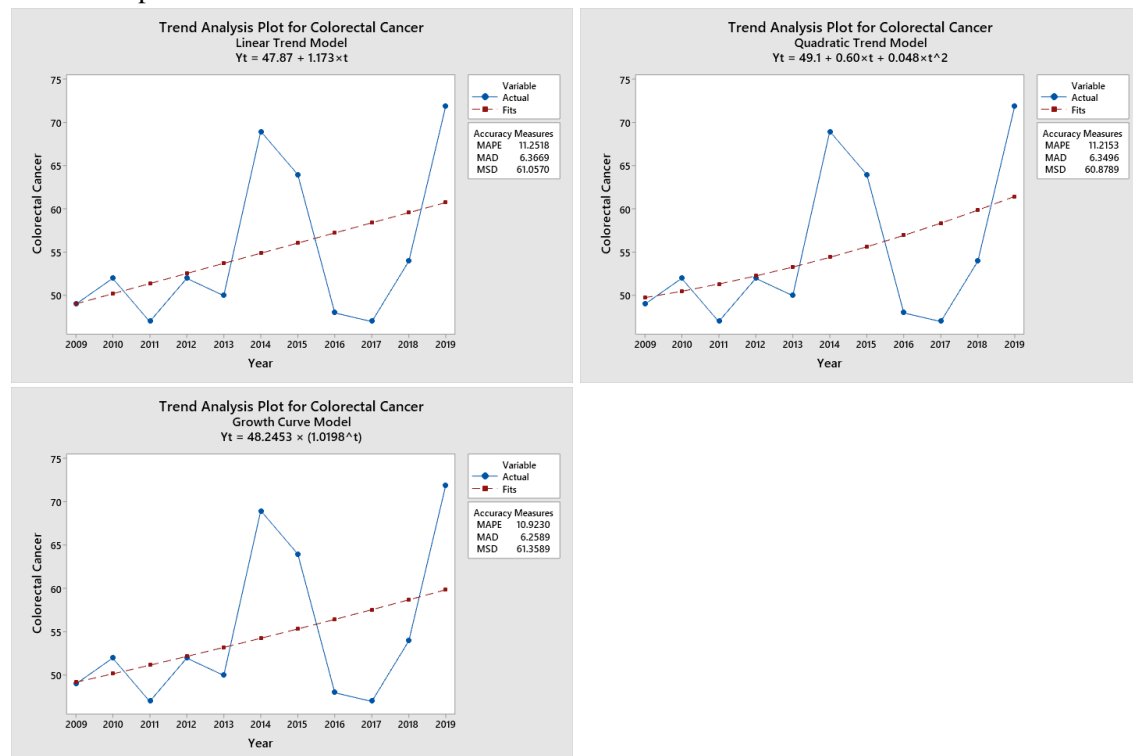
BMJ Open

Supplementary File 5.

Detailed analysis for forecasting future ten-year incidence of colorectal cancer using the best-fitted curve model obtained from regression analysis among young patients based on tumor location and tumor side involvement

a. Regression Model for CRC Cases Among Young Patients

1. Curve Options

Model	Automatic Fitted-Curve	Accuracy Measurements		
		MAPE	**MAD**	**MSD**
Linear	$Yt = 47.87 + 1.173t$	11.2518	6.3669	61.0570
Quadratic	$Yt = 49.1 + 0.60t + 0.049t^2$	11.2153	6.3496	**60.8789**
Exponential Growth*	$Yt = 48.2453 \times (1.0198^t)$	**10.9230**	**6.2589**	61.3589
S-shaped	*Error: Can not fit the model to these data*	*n/a*	*n/a*	*n/a*

*The best-fitted model is the one that has the lowest value for three parameters (MAPE, MAD, and MSD), or at least for two parameters out of three, or at least having the lowest value for MAPE.

2. The significance test results for the slope of curve estimation (i.e., Exponential growth model) employing the ANOVA statistical test in SPSS.

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Regression	.042	1	.042	1.916	.200
Residual	.200	9	.022		
Total	.242	10			

The independent variable is Year.

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
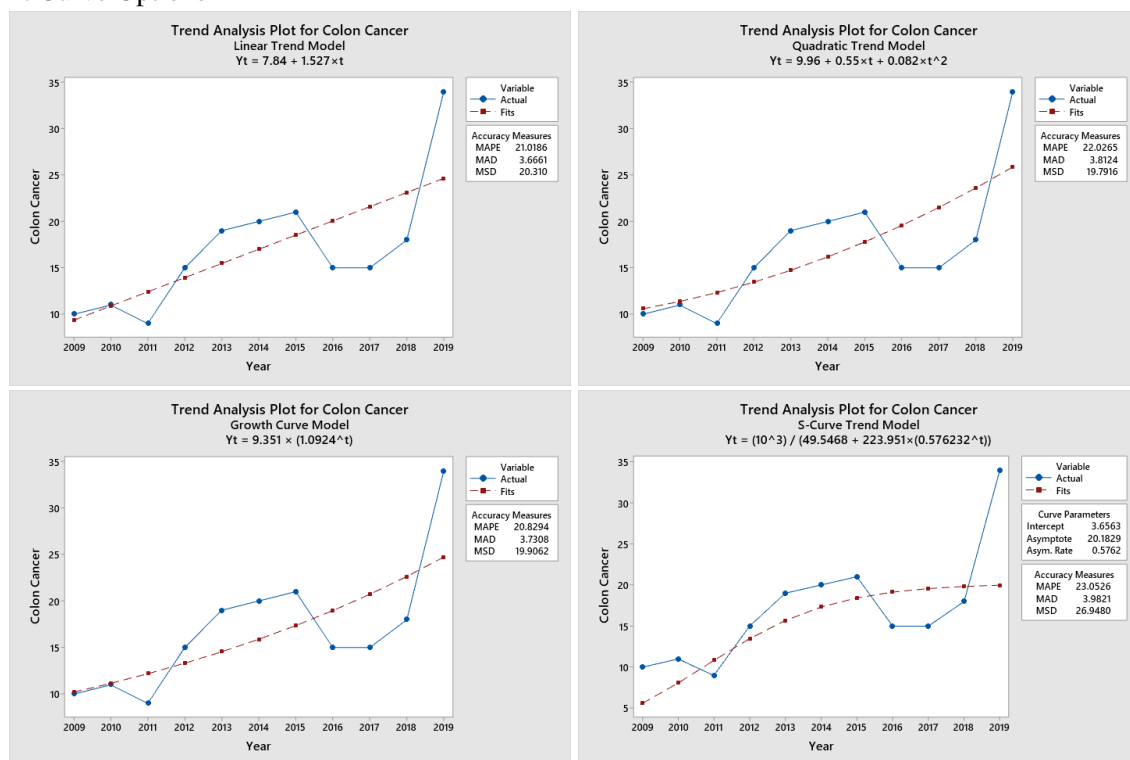
3. The forecast of the number of cases in the following ten-year period using the Exponential growth model.

Year	Forecasted Cases
2020	61.0779
2021	62.2902
2022	63.5266
2023	64.7875
2024	66.0734
2025	67.3849
2026	68.7224
2027	70.0864
2028	71.4776
2029	72.8963
Mean	**66.83232**
Total	**668.3232**

b. Regression Model for Colon Cancer Cases Among Young Patients

1. Curve Options

Model	Automatic Fitted-Curve	Accuracy Measurements		
		MAPE	**MAD**	**MSD**
Linear	$Yt = 7.84 + 1.527t$	21.0186	3.6661	20.3107
Quadratic	$Yt = 9.96 + 0.55t + 0.082t^2$	22.0265	3.8124	**19.7916**
Exponential Growth*	$Yt = 9.351 \times (1.0924^t)$	**20.8294**	**3.7308**	19.9062
S-shaped	$Yt = 10^3 / (49.5468 + (223.951 \times 0.576232^t))$	23.0526	3.9821	26.9480

*The best-fitted model is the one that has the lowest value for three parameters (MAPE, MAD, and MSD), or at least for two parameters out of three, or at least having the lowest value for MAPE.

2. The significance test results for the slope of curve estimation (i.e., Exponential growth model) employing the ANOVA statistical test in SPSS.

ANOVA

|	Sum of Squares	df	Mean Square	F	Sig.
Regression	.859	1	.859	13.320	**.005**
Residual	.581	9	.065		
Total	1.440	10			

The independent variable is Year.

3. The forecast of the number of cases in the following ten-year period using the Exponential growth model.

Year	Forecasted Cases
2020	27.0046
2021	29.4998
2022	32.2256
2023	35.2031
2024	38.4559
2025	42.0091
2026	45.8907
2027	50.1310
2028	54.7630
2029	59.8230
Mean	**41.50058**
Total	**415.0058**

c. Regression Model for Rectal Cancer Cases Among Young Patients

1. Curve Options
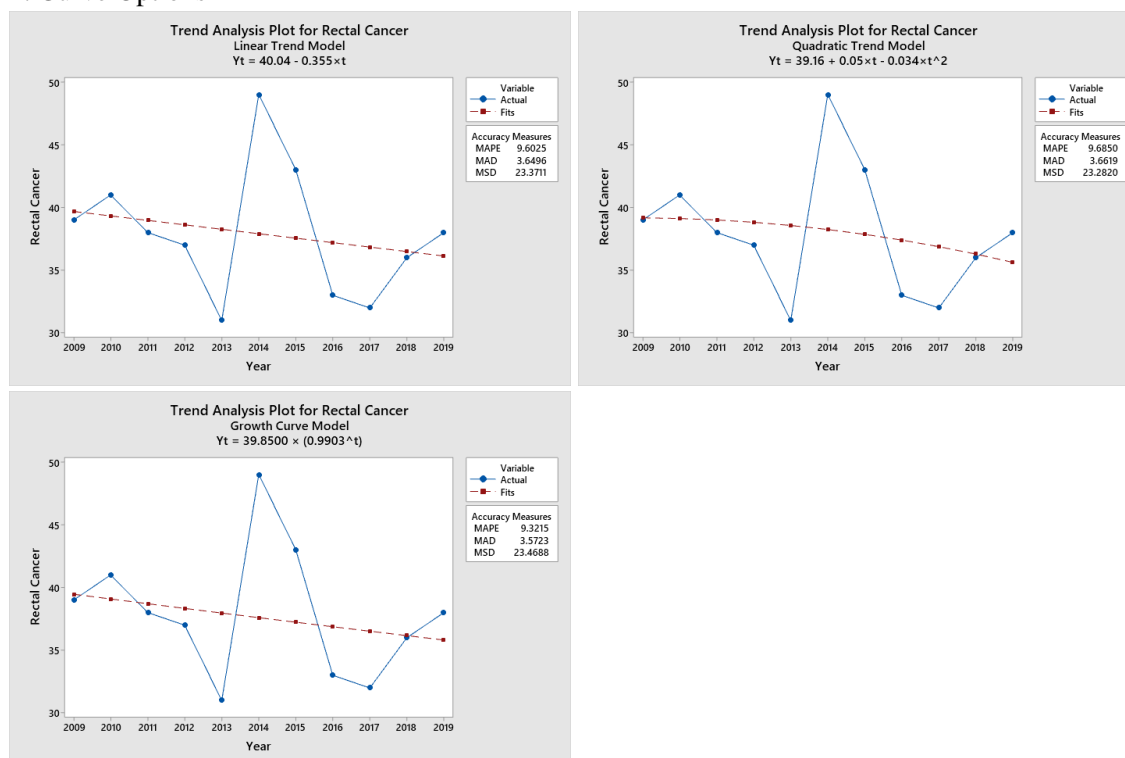
Model	Automatic Fitted-Curve	Accuracy Measurements		
		MAPE	MAD	MSD
Linear	$Yt = 40.04 - 0.355t$	9.6025	3.6496	23.3711
Quadratic	$Yt = 39.16 + 0.05t - 0.034t^2$	9.6850	3.6619	**23.2820**
Exponential Growth*	**$Yt = 39.85 \times (0.9903^t)$**	**9.3215**	**3.5723**	23.4688
S-shaped	*Error: Can not fit the model to these data*	*n/a*	*n/a*	*n/a*

*The best-fitted model is the one that has the lowest value for three parameters (MAPE, MAD, and MSD), or at least for two parameters out of three, or at least having the lowest value for MAPE.

2. The significance test results for the slope of curve estimation (i.e., Exponential growth model) employing the ANOVA statistical test in SPSS.

ANOVA

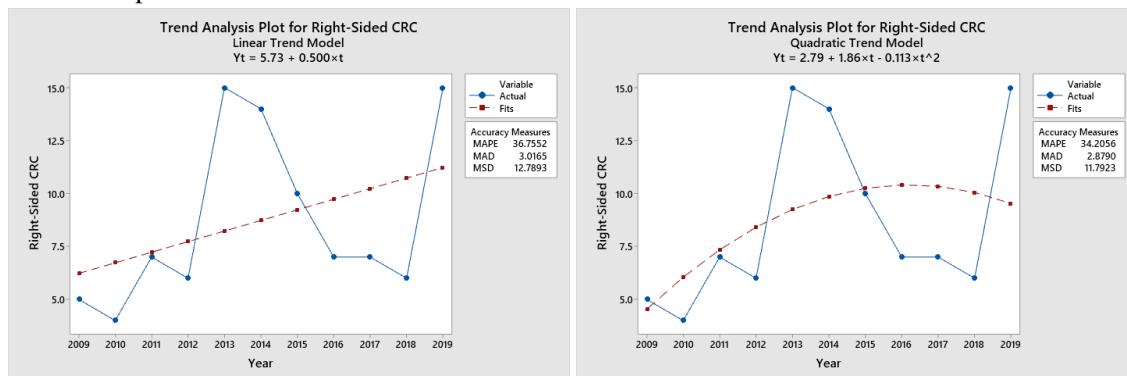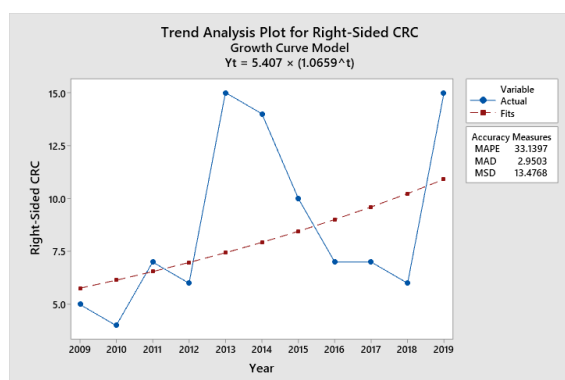|	Sum of Squares	df	Mean Square	F	Sig.
Regression	.010	1	.010	.551	.477
Residual	.169	9	.019		
Total	.180	10			

The independent variable is Year.

3. The forecast of the number of cases in the following ten-year period using the Exponential growth model.

Year	Forecasted Cases
2020	35.4703
2021	35.1278
2022	34.7886
2023	34.4527
2024	34.1201
2025	33.7906
2026	33.4644
2027	33.1413
2028	32.8213
2029	32.5044
Mean	**33.96815**
Total	**339.6815**

d. Regression Model for Right-Sided CRC Cases Among Young Patients

1. Curve Options

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

Trend Analysis Plot for Right-Sided CRC
Growth Curve Model
$Yt = 5.407 \times (1.0659^t)$

Model	Automatic Fitted-Curve	Accuracy Measurements		
		MAPE	**MAD**	**MSD**
Linear	$Yt = 5.73 + 0.5t$	36.7552	3.0165	12.7893
Quadratic*	**$Yt = 2.79 + 1.86t - 0.113t^2$**	34.2056	**2.8790**	**11.7923**
Exponential Growth	$Yt = 5.407 \times (1.0659^t)$	**33.1397**	2.9503	13.4768
S-shaped	*Error: Can not fit the model to these data*	*n/a*	*n/a*	*n/a*

*The best-fitted model is the one that has the lowest value for three parameters (MAPE, MAD, and MSD), or at least for two parameters out of three, or at least having the lowest value for MAPE.

2. The significance test results for the slope of curve estimation (i.e., Quadratic model) employing the ANOVA statistical test in SPSS.

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Regression	27.500	1	27.500	1.759	.217
Residual	140.682	9	15.631		
Total	168.182	10			

The independent variable is Year.

3. The forecast of the number of cases in the following ten-year period using the Quadratic model.

Year	Forecasted Cases
2020	8.78788
2021	7.81818
2022	6.62238
2023	5.20047
2024	3.55245
2025	1.67832
2026	-0.42191
2027	-2.74825
2028	-5.30070
2029	-8.07925
Mean	**1.710957**
Total	**17.10957**

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

e. Regression Model for Left-Sided CRC Cases Among Young Patients

1. Curve Options

Model	Automatic Fitted-Curve	Accuracy Measurements		
		MAPE	MAD	MSD
Linear	$Yt = 42.15 + 0.673t$	12.3297	5.5041	40.5322
Quadratic*	$Yt = 46.33 - 1.26t + 0.161t^2$	**11.9922**	**5.4062**	**38.5144**
Exponential Growth	$Yt = 42.108 \times (1.0137^t)$	12.1815	5.4906	40.6512
S-shaped	*Error: Can not fit the model to these data*	*n/a*	*n/a*	*n/a*

*The best-fitted model is the one that has the lowest value for three parameters (MAPE, MAD, and MSD), or at least for two parameters out of three, or at least having the lowest value for MAPE.

2. The significance test results for the slope of curve estimation (i.e., Quadratic model) employing the ANOVA statistical test in SPSS.

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Regression	49.828	1	49.828	1.006	.342
Residual	445.808	9	49.534		
Total	495.636	10			

The independent variable is Year.

3. The forecast of the number of cases in the following ten-year period using the Quadratic model.

Year	Forecasted Cases
2020	54.4000
2021	57.1636
2022	60.2490
2023	63.6559
2024	67.3846
2025	71.4350
2026	75.8070
2027	80.5007
2028	85.5161
2029	90.8531
Mean	**70.6965**
Total	**706.965**