1 Appendices

1.1 Appendix: Additional analyses and explanation to inform policy making

Q1: How different are the BCSP and UKFSST detection rates?

Further details have now been provided within the main report section "Screening test characteristics".

Q2 Considering repeated screening with FIT120 or 160 for age ranges 50-74 and 60-74; is the replacement of a FIT screen with a BS at age 58/60 cost effective?

Under the base case analysis (BS sensitivity and uptake from the BCSP)

For FIT120/FIT160 ages 50-74 replacing the FIT age 58 with a bowel scope results in higher QALYs but lower cost effectiveness. i.e. it is not cost effective to replace FIT age 58 with bowel scope. For FIT120/FIT160 ages 60-74 replacing the FIT age 60 with a bowel scope results in higher QALYs and higher cost effectiveness. i.e. it is cost effective to replace FIT age 60 with bowel scope.

Scenario analysis with higher bowel scope detection rates from UKFSST, and higher uptake 55%

This scenario analysis is highlighted in yellow in the table below. For FIT120/FIT160 ages 50-74 or 60-74 it is cost effective to replacing a FIT at age 58/60 with a bowel scope.

Screening strategy	Incremental compared to no screening			Lifetime reduction		Screening referral colonoscopies (year1)	Screening flexible sigmoidoscopy (year 1)
	Costs (£m)	QALYs	NMB (£m)	CRC incidence	CRC mortality		
Current							
gFOBT ages 60-74 biennial	-£37.3m	11,608	£269.4m	10.7%	16.8%	36,406	-
Bowel scope age 55, gFOBT ages 60-74 biennial	£12.9m	19,197	£371.1m	16.5%	22.5%	42,920	329,121
One off screens							
Bowel Scope age 59	£26.2m	9,341	£160.7m	8.4%	9.8%	9,197	289,081
Bowel Scope age 59	£21.7m	14,080	£259.9m	12.0%	14.5%	13,313	357,983
FIT20 age 57	-£39.6m	9,093	£221.5m	7.3%	8.5%	28,995	-
Bowel scope and repeated FIT screening							
2-yearly, age 50-74, FIT120	-£91.1m	27,320	£637.5m	25.0%	31.0%	91,652	-
BS age 58, 2-yearly, age 50-74 (excl 58), FIT120	-£26.8m	29,694	£620.6m	27.1%	33.0%	93,099	300,813
BS age 58, 2-yearly, age 50-74 (excl 58), FIT120	-£17.5m	31,823	£653.9m	28.5%	34.6%	97,162	372,511
2-yearly, age 50-74, FIT160	-£78.5m	24,879	£576.1m	22.6%	28.3%	77,126	-
BS age 58 2-yearly, age 50-74 (excl 58), FIT160	-£17.1m	27,763	£572.4m	25.1%	30.8%	79,829	300,813
BS age 58 2-yearly, age 50-74 (excl 58), FIT160	-£9.0m	30,120	£611.4m	26.7%	32.7%	83,892	372,511
2-yearly, age 60-74, FIT120	-£68.7m	16,034	£389.4m	18.3%	23.7%	51,397	-
BS age 60, 2-yearly, age 60-74, FIT120	-£18.2m	19,530	£408.8m	21.7%	27.0%	53,372	278,432
BS age 60, 2-yearly, age 60-74, FIT120	-£15.5m	22,755	£470.6m	23.8%	29.8%	57,540	344,796
2-yearly, age 60-74, FIT160	-£58.7m	14,466	£348.0m	16.4%	21.4%	43,880	-
BS age 60, 2-yearly, age 60-74, FIT160	-£11.4m	18,479	£381.0m	20.2%	25.3%	47,014	278,432
BS age 60, 2-yearly, age 60-74, FIT160	-£9.5m	21,844	£446.4m	22.5%	28.3%	51,182	344,796

Scenario analysis with higher bowel scope test sensitivity from UKFSST, and higher uptake 55%

For FIT120 ages 50-74 replacement with bowel scope at age 58 reduced the cost effectiveness from £637m to £620m (base case assumption). However, under the analysis with detection rates as in the UKFSST and higher uptake (55%) cost effectiveness increases from £637m to £654m.

With a screening referral colonoscopy capacity of 50,000 (approximately current usage) if bowel scope detection rates and uptake can achieve levels observed in the UKFSST then the most cost effective screening strategy is 2-yearly FIT 160 ages 60-74 and a BS at 60. For BS at age 60 to achieve a similar quality as was observed in the UKFSST target detection rates should be 9.0% (>8.6%), 2.9% (2.6%), 0.34% (>0.26%) for LR adenomas, HR adenomas and CRC respectively. The target uptake rate is 55%. We note that this conclusion assumes that endoscopy capacity cannot be transferred from flexible sigmoidoscopies to screening referral colonoscopies.

With a screening referral colonoscopy capacity of 57,000 (approximately current usage) if bowel scope detection rates and uptake can achieve levels observed in the UKFSST then the most cost effective screening strategy is 2-yearly FIT 160 ages 60-74 and a BS at 58. (Note this is compared to 2-yearly FIT127 ages 58-74). For BS at age 58 to achieve a similar quality as was observed in the UKFSST target detection rates should be 8.8% (>8.4%), 2.6% (2.4%), 0.30% (>0.23%) for LR adenomas, HR adenomas and CRC respectively. The target uptake rate is 55%. We note that this conclusion assumes that endoscopy capacity cannot be transferred from flexible sigmoidoscopies to screening referral colonoscopies.

With a 2-yearly FIT120 ages 60-74 strategy it is not possible to include a bowel scope at a capacity of around 50,000 annual screening referrals.

We note that for 2-yearly FIT ages 50-74 the threshold below which it is not cost-effective to add a bowel scope (with UKFFST detection rates and uptake) is FIT93 which requires a screening referral colonoscopy capacity of 109,000 (which may require conversion of flexible sigmoidoscopy capacity to colonoscopies to achieve).

Q3 If bowel scope capacity could be converted to screening referral colonoscopy capacity does this impact on the conclusions?

We suppose an endoscopy capacity equivalence of *10 bowel scopes = 4 screening referral colonoscopies*. Two screening strategies with equivalent endoscopy capacity (under this assumption) are compared. We consider a one-off bowel scope screen at age 59. This is associated with a high endoscopy capacity which could also be used to undertake repeated FIT screening with a lower test threshold.

Base case analysis (BS sensitivity and uptake from the BCSP)

A one-off bowel scope at age 59 (290k bowel scopes, 9k screening referral colonoscopies) is considerably less effective and a cost effective than a repeated FIT74 screening strategy which is associated with 125k screening referral colonoscopies. Such strategies could be considered to have equivalent endoscopy capacity. Hence, if bowel scope capacity could be converted to screening referral colonoscopy capacity instead, it would result in far higher effectiveness and cost-effectiveness to undertake repeated FIT only screening strategies.

Screening strategy	Incremental compared to no screening			Lifetime reduction		Screening referral colonoscopies (year1)	Screening flexible sigmoidoscopy (year 1)
	Costs (£m)	QALYs	NMB (£m)	CRC incidence	CRC mortality		
Bowel Scope age 59	£26.2m	9,341	£160.7m	8.4%	9.8%	9,197	289,081
2-yearly, age 50-74, FIT74	**-£111.1m**	**31,613**	**£743.3m**	**29.3%**	**35.6%**	125,129	-

Scenario analysis with higher bowel scope test sensitivity from UKFSST, and higher uptake 55%

A one-off bowel scope at age 59 (358k bowel scopes, 13k screening referral colonoscopies) is considerably less effective and a cost effective than a repeated FIT screening strategy associated with 156k screening referral colonoscopies. Such strategies could be considered to have equivalent endoscopy capacity. Hence, if bowel scope capacity could be converted to screening referral colonoscopy capacity instead, it would result in far higher effectiveness and cost-effectiveness to undertake repeated FIT only screening strategies. 2-yearly FIT54 age 50-74 is associated with 2.4 times the effectiveness (QALYs) and 3.1 times the cost effectiveness compared to one off bowel scope at age 59.

Scenario analysis with higher bowel scope test sensitivity from UKFSST, and higher uptake 55%

Screening strategy	Incremental compared to no			Lifetime reduction		Screening referral colonoscopies (year1)	Screening flexible sigmoidoscopy (year 1)
	Costs (£m)	QALYs	NMB (£m)	CRC incidence	CRC mortality		
Bowel Scope age 59	£21.7m	14,080	£259.9m	12.0%	14.5%	13,313	357,983
2-yearly, age 50-74, FIT54	**-£121.9m**	**34,480**	**£811.5m**	**32.1%**	**38.5%**	155,363	-
		2.4	3.1				

The model predictions for the repeated FIT screening strategies are associated with uncertainty as there is a paucity of evidence of how well the test will perform when repeated. However, the repeated FIT screening strategy considered here could be significantly less effective than predicted by this model but still remain more effective and cost effective than a one-off bowel scope.

1.2 Appendix: Scenario Analysis

A series of scenario analyses were undertaken to investigate the impact of the following parameter changes: lower CRC treatment costs; higher Bowel Scope sensitivity from UKFSST, and a higher uptake.

The age at which a one-off bowel scope screen was the most cost effective did not vary under most scenario analyses and changed only slightly (to age 58 rather than age 59) in the case of a higher sensitivity and uptake.

Under the scenario analyses with a higher bowel scope sensitivity and uptake from UKFSST, one-off bowel scope was more cost-effective than one-off FIT20.

Whether it was cost-effective to add a one-off bowel scope screen to a repeated FIT screening strategy varied under several of the scenario analysis. Adding bowel scope was not cost-effective for: lower CRC treatment costs; higher cost for Bowel Scope; increased rate of symptomatic presentation; female-only subjects. Adding bowel scope was costs effective for the scenario analyses: lower discount rate for costs and QALYs; higher Bowel Scope sensitivity from UKFSST, and a higher uptake; lower cost for Bowel Scope; reduced FIT sensitivity in repeated screens; reduced FIT sensitivity and specificity in repeated screens; male-only subjects.

Table 1: Scenario analyses for one-off bowel scope or FIT20 screen

Scenario analysis with lower CRC treatment costs

Screening strategy	Incremental compared to no screening			Lifetime reduction		Screening referral colonoscopies (Yr1)	Screening flexible sigmoidoscopy
	Costs (£m)	QALYs	NMB (£m)	CRC incidence	CRC mortality		
Bowel Scope age 55	£90.0m	9,654	£103.0m	7.3%	8.3%	8,510	329,121
Bowel Scope age 56	£86.1m	**9,655**	£107.0m	7.6%	8.7%	8,626	315,523
Bowel Scope age 57	£82.3m	9,602	£109.8m	7.9%	9.1%	8,879	308,029
Bowel Scope age 58	£78.6m	9,497	£111.3m	8.2%	9.5%	9,120	300,813
Bowel Scope age 59	£75.1m	9,341	**£111.7m**	8.4%	9.8%	9,197	289,081
Bowel Scope age 60	£71.7m	9,139	£111.0m	8.6%	10.1%	9,275	278,432
Bowel Scope age 61	£68.5m	8,892	£109.3m	8.8%	10.3%	8,677	266,846
Bowel Scope age 62	£65.4m	8,605	£106.7m	8.9%	10.5%	9,112	267,185
Bowel Scope age 63	£62.4m	8,283	£103.2m	9.0%	10.7%	8,758	262,564
Bowel Scope age 64	£59.6m	7,928	£98.9m	**9.1%**	10.8%	8,933	254,981
Bowel Scope age 65	£56.9m	7,544	£94.0m	**9.1%**	10.9%	8,761	255,383
Bowel Scope age 66	£54.3m	7,140	£88.5m	9.1%	**11.0%**	9,384	260,232
Bowel Scope age 67	£51.8m	6,716	£82.5m	9.0%	10.9%	9,370	265,506
Bowel Scope age 68	£49.3m	6,277	£76.2m	8.9%	10.8%	10,382	279,845
Bowel Scope age 69	£47.1m	5,833	£69.6m	8.7%	10.7%	10,973	302,066
Bowel Scope age 70	**£44.8m**	5,385	£62.9m	8.5%	10.5%	8,872	232,338

Scenario analysis with higher bowel scope test sensitivity from UKFSST, and higher uptake 55%

Screening strategy	Incremental compared to no screening			Lifetime reduction		Screening referral colonoscopies(Yr1)	Screening flexible sigmoidoscopy
	Costs (£m)	QALYs	NMB (£m)	CRC incidence	CRC mortality		
Bowel Scope age 55	£40.2m	14,440	£248.6m	10.4%	12.3%	12,252	407,567
Bowel Scope age 56	£34.7m	**14,468**	£254.7m	10.9%	12.9%	12,435	390,728
Bowel Scope age 57	£29.8m	14,416	£258.6m	11.3%	13.4%	12,818	381,448
Bowel Scope age 58	£25.4m	14,286	**£260.3m**	11.6%	14.0%	13,184	372,511
Bowel Scope age 59	£21.7m	14,080	£259.9m	12.0%	14.5%	13,313	357,983
Bowel Scope age 60	£17.7m	13,803	£258.3m	12.2%	14.9%	13,443	344,796
Bowel Scope age 61	£14.8m	13,458	£254.4m	12.5%	15.3%	12,563	330,448
Bowel Scope age 62	£12.3m	13,050	£248.8m	12.6%	15.7%	13,214	330,868
Bowel Scope age 63	£10.3m	12,588	£241.4m	12.8%	16.0%	12,690	325,146
Bowel Scope age 64	£8.9m	12,074	£232.6m	**12.8%**	16.2%	12,967	315,755
Bowel Scope age 65	£7.8m	11,516	£222.5m	**12.8%**	16.3%	12,708	316,253
Bowel Scope age 66	£7.3m	10,923	£211.2m	12.8%	**16.4%**	13,637	322,258
Bowel Scope age 67	**£7.1m**	10,298	£198.9m	12.6%	16.4%	13,607	328,789
Bowel Scope age 68	£7.1m	9,648	£185.8m	12.4%	16.3%	15,107	346,546
Bowel Scope age 69	£7.7m	8,987	£172.1m	12.2%	16.2%	15,954	374,064
Bowel Scope age 70	£7.3m	8,316	£159.0m	11.8%	15.9%	12,926	287,716

Scenario analysis with lower CRC treatment costs

Screening strategy	Incremental compared to no screening			Lifetime reduction		Screening referral colonoscopies (Yr1)	Screening flexible sigmoidoscopy
	Costs (£m)	QALYs	NMB (£m)	CRC incidence	CRC mortality		
FIT20 age 55	£7.4m	9,131	£175.2m	6.7%	7.8%	31,210	-
FIT20 age 56	£6.3m	**9,138**	£176.5m	7.0%	8.2%	29,814	-
FIT20 age 57	£5.2m	9,093	**£176.6m**	7.3%	8.5%	28,995	-
FIT20 age 58	£4.3m	8,997	£175.7m	7.5%	8.9%	28,201	-
FIT20 age 59	£3.4m	8,852	£173.7m	7.7%	9.2%	26,985	-
FIT20 age 60	£2.6m	8,662	£170.7m	7.9%	9.4%	25,874	-
FIT20 age 61	£1.9m	8,430	£166.7m	8.1%	9.7%	24,194	-
FIT20 age 62	£1.3m	8,153	£161.8m	8.2%	9.9%	24,104	-
FIT20 age 63	£0.7m	7,845	£156.2m	8.2%	10.0%	23,095	-
FIT20 age 64	£0.3m	7,498	£149.7m	**8.3%**	10.2%	22,320	-
FIT20 age 65	-£0.1m	7,132	£142.8m	**8.3%**	10.2%	21,782	-
FIT20 age 66	-£0.4m	6,745	£135.3m	8.3%	**10.3%**	22,132	-
FIT20 age 67	-£0.6m	6,340	£127.4m	8.2%	10.2%	21,992	-
FIT20 age 68	-£0.9m	5,915	£119.2m	8.0%	10.1%	23,120	-
FIT20 age 69	-£1.0m	5,487	£110.7m	7.9%	10.0%	24,290	-
FIT20 age 70	**-£1.0m**	5,051	£102.0m	7.7%	9.8%	18,644	-

Table 2: Scenario analyses for the key screening strategies

Scenario analysis with lower CRC treatment costs

Screening strategy	Incremental compared to no screening			Lifetime reduction		Screening referral colonoscopies (year1)	Screening flexible sigmoidoscopy (year 1)
	Costs (£m)	QALYs	NMB (£m)	CRC incidence	CRC mortality		
Current							
gFOBT ages 60-74 biennial	£6.1m	11,608	£226.0m	10.7%	16.8%	36,406	-
Bowel scope age 55, gFOBT ages 60-74 biennial	£98.3m	19,197	£285.7m	16.5%	22.5%	42,920	329,121
One-off screens							
Bowel Scope age 59	£75.1m	9,341	£111.7m	8.4%	9.8%	9,197	289,081
FIT20 age 57	£5.2m	9,093	£176.6m	7.3%	8.5%	28,995	-
Repeated FIT screening							
2-yearly, age 51-65, FIT161	£30.8m	19,098	£351.2m	15.1%	18.4%	49,856	-
2-yearly, age 50-70, FIT153	£39.8m	23,600	£432.2m	20.2%	24.9%	69,912	-
2-yearly, age 50-74, FIT124	£43.9m	27,037	£496.8m	24.7%	30.7%	89,822	-
2-yearly, age 50-74, FIT74	£49.3m	31,613	£583.0m	29.3%	35.6%	125,129	-
Bowel scope and repeated FIT screening							
Bowel scope age 59, 2-yearly, age 51-65 (excl.59), FIT161	£105.8m	22,373	£341.7m	18.2%	21.7%	52,805	289,081
Bowel scope age 58, 2-yearly, age 50-70 (excl. 58), FIT153	£118.6m	26,471	£410.9m	22.7%	27.4%	72,438	300,813
Bowel scope age 58, 2-yearly, age 50-74 (excl. 58), FIT124	£122.9m	29,468	£466.5m	26.9%	32.7%	91,428	300,813

Discount rates for future costs and QALYs set to 1.5%

Screening strategy	Incremental compared to no screening			Lifetime reduction		Screening referral colonoscopies (year1)	Screening flexible sigmoidoscopy (year 1)
	Costs (£m)	QALYs	NMB (£m)	CRC incidence	CRC mortality		
no screening	£0.0m	-	£0.0m	-	-	-	-
gFOBT ages 60-74 biennial	-£84.4m	21,906	£522.5m	10.7%	16.8%	36,406	-
Bowel scope age 55, gFOBT ages 60-74 biennial	-£47.7m	34,690	£741.5m	16.5%	22.5%	42,920	329,121
Bowel Scope age 59	£4.5m	16,686	£329.2m	8.4%	9.8%	9,197	289,081
FIT20 age 57	-£69.9m	15,892	£387.7m	7.3%	8.5%	28,995	-
2-yearly, age 51-65, FIT161	-£120.0m	33,372	£787.4m	15.1%	18.4%	49,856	-
2-yearly, age 50-70, FIT153	-£154.2m	41,759	£989.4m	20.2%	24.9%	69,912	-
2-yearly, age 50-74, FIT124	-£188.2m	48,335	£1,154.9m	24.7%	30.7%	89,822	-
2-yearly, age 50-74, FIT74	-£227.5m	56,406	£1,355.6m	29.3%	35.6%	125,129	-
Bowel scope age 59, 2-yearly, age 51-65 (excl.59), FIT161	-£62.9m	39,191	£846.7m	18.2%	21.7%	52,805	289,081
Bowel scope age 58, 2-yearly, age 50-70 (excl. 58), FIT153	-£89.6m	46,747	£1,024.6m	22.7%	27.4%	72,438	300,813
Bowel scope age 58, 2-yearly, age 50-74 (excl. 58), FIT124	-£119.1m	52,530	£1,169.7m	26.9%	32.7%	91,428	300,813

Scenario analysis with bowel scope cost=£450

Screening strategy	Incremental compared to no screening			Lifetime reduction		Screening referral colonoscopies (year1)	Screening flexible sigmoidoscopy (year 1)
	Costs (£m)	QALYs	NMB (£m)	CRC incidence	CRC mortality		
no screening							
gFOBT ages 60-74 biennial	-£37.3m	11,608	£269.4m	10.7%	16.8%	36,406	-
Bowel scope age 55, gFOBT ages 60-74 biennial	£54.3m	19,197	£329.7m	16.5%	22.5%	42,920	329,121
One-off screens							
Bowel Scope age 59	£61.5m	9,341	£125.3m	8.4%	9.8%	9,197	289,081
FIT20 age 57	-£39.6m	9,093	£221.5m	7.3%	8.5%	28,995	-
Repeated FIT screening							
2-yearly, age 51-65, FIT161	-£59.2m	19,098	£441.2m	15.1%	18.4%	49,856	-
2-yearly, age 50-70, FIT153	-£73.7m	23,600	£545.8m	20.2%	24.9%	69,912	-
2-yearly, age 50-74, FIT124	-£89.7m	27,037	£630.4m	24.7%	30.7%	89,822	-
2-yearly, age 50-74, FIT74	-£111.1m	31,613	£743.3m	29.3%	35.6%	125,129	-
Bowel scope and repeated FIT screening							
Bowel scope age 59, 2-yearly, age 51-65 (excl.59), FIT161	£31.4m	22,373	£416.1m	18.2%	21.7%	52,805	289,081
Bowel scope age 58, 2-yearly, age 50-70 (excl. 58), FIT153	£24.1m	26,471	£505.3m	22.7%	27.4%	72,438	300,813
Bowel scope age 58, 2-yearly, age 50-74 (excl. 58), FIT124	£10.7m	29,468	£578.6m	26.9%	32.7%	91,428	300,813

Scenario analysis with bowel scope cost=£150

Screening strategy	Incremental compared to no screening			Lifetime reduction		Screening referral colonoscopies (year1)	Screening flexible sigmoidoscopy (year 1)
	Costs (£m)	QALYs	NMB (£m)	CRC incidence	CRC mortality		
no screening							
gFOBT ages 60-74 biennial	-£37.3m	11,608	£269.4m	10.7%	16.8%	36,406	-
Bowel scope age 55, gFOBT ages 60-74 biennial	-£35.1m	19,197	£419.1m	16.5%	22.5%	42,920	329,121
One-off screens							
Bowel Scope age 59	-£14.8m	9,341	£201.7m	8.4%	9.8%	9,197	289,081
FIT20 age 57	-£39.6m	9,093	£221.5m	7.3%	8.5%	28,995	-
Repeated FIT screening							
2-yearly, age 51-65, FIT161	-£59.2m	19,098	£441.2m	15.1%	18.4%	49,856	-
2-yearly, age 50-70, FIT153	-£73.7m	23,600	£545.8m	20.2%	24.9%	69,912	-
2-yearly, age 50-74, FIT124	-£89.7m	27,037	£630.4m	24.7%	30.7%	89,822	-
2-yearly, age 50-74, FIT74	-£111.1m	31,613	£743.3m	29.3%	35.6%	125,129	-
Bowel scope and repeated FIT screening							
Bowel scope age 59, 2-yearly, age 51-65 (excl.59), FIT161	-£44.1m	22,373	£491.6m	18.2%	21.7%	52,805	289,081
Bowel scope age 58, 2-yearly, age 50-70 (excl. 58), FIT153	-£54.5m	26,471	£583.9m	22.7%	27.4%	72,438	300,813
Bowel scope age 58, 2-yearly, age 50-74 (excl. 58), FIT124	-£67.8m	29,468	£657.2m	26.9%	32.7%	91,428	300,813

Scenario analysis with higher bowel scope test sensitivity from UKFSST, and higher uptake 55%

Screening strategy	Incremental compared to no screening			Lifetime reduction		Screening referral colonoscopies (year1)	Screening flexible sigmoidoscopy (year 1)
	Costs (£m)	QALYs	NMB (£m)	CRC incidence	CRC mortality		
Current							
gFOBT ages 60-74 biennial	-£37.3m	11,608	£269.4m	10.7%	16.8%	36,406	-
Bowel scope age 55, gFOBT ages 60-74 biennial	£15.1m	23,088	£446.6m	18.9%	25.3%	45,762	407,567
One-off screens							
Bowel Scope age 59	£21.7m	14,080	£259.9m	12.0%	14.5%	13,313	357,983
FIT20 age 57	-£39.6m	9,093	£221.5m	7.3%	8.5%	28,995	-
Repeated FIT screening							
2-yearly, age 51-65, FIT161	-£59.2m	19,098	£441.2m	15.1%	18.4%	49,856	-
2-yearly, age 50-70, FIT153	-£73.7m	23,600	£545.8m	20.2%	24.9%	69,912	-
2-yearly, age 50-74, FIT124	-£89.7m	27,037	£630.4m	24.7%	30.7%	89,822	-
2-yearly, age 50-74, FIT74	-£111.1m	31,613	£743.3m	29.3%	35.6%	125,129	-
Bowel scope and repeated FIT screening							
Bowel scope age 59, 2-yearly, age 51-65 (excl.59), FIT161	£2.2m	24,985	£497.5m	20.1%	24.2%	56,920	357,983
Bowel scope age 58, 2-yearly, age 50-70 (excl. 58), FIT153	-£4.3m	28,836	£581.0m	24.3%	29.4%	76,501	372,511
Bowel scope age 58, 2-yearly, age 50-74 (excl. 58), FIT124	-£16.5m	31,623	£649.0m	28.3%	34.4%	95,491	372,511

Scenario analysis with reduced FIT test sensitivity in repeated screens

Screening strategy	Incremental compared to no screening			Lifetime reduction		Screening referral colonoscopies (year1)	Screening flexible sigmoidoscopy (year 1)
	Costs (£m)	QALYs	NMB (£m)	CRC incidence	CRC mortality		
Current							
gFOBT ages 60-74 biennial	-£37.3m	11,608	£269.4m	10.7%	16.8%	36,406	-
Bowel scope age 55, gFOBT ages 60-74 biennial	£12.9m	19,197	£371.1m	16.5%	22.5%	42,920	329,121
One-off screens							
Bowel Scope age 59	£26.2m	9,341	£160.7m	8.4%	9.8%	9,197	289,081
FIT20 age 57	-£39.6m	9,093	£221.5m	7.3%	8.5%	28,995	-
Repeated FIT screening							
2-yearly, age 51-65, FIT161	-£40.7m	15,803	£356.7m	12.4%	15.1%	43,373	-
2-yearly, age 50-70, FIT153	-£50.5m	19,587	£442.3m	16.6%	20.6%	60,138	-
2-yearly, age 50-74, FIT124	-£63.2m	22,615	£515.5m	20.5%	25.6%	77,352	-
2-yearly, age 50-74, FIT74	-£81.8m	26,912	£620.1m	24.7%	30.3%	109,967	-
Bowel scope and repeated FIT screening							
Bowel scope age 59, 2-yearly, age 51-65 (excl.59), FIT161	£9.1m	20,087	£392.6m	16.4%	19.5%	47,292	289,081
Bowel scope age 58, 2-yearly, age 50-70 (excl. 58), FIT153	£5.1m	23,480	£464.5m	20.0%	24.1%	63,646	300,813
Bowel scope age 58, 2-yearly, age 50-74 (excl. 58), FIT124	-£4.9m	26,058	£526.1m	23.5%	28.6%	80,018	300,813

Scenario analysis with reduced FIT test sensitivity and specificity in repeated screens

Screening strategy	Incremental compared to no screening			Lifetime reduction		Screening referral colonoscopies (year1)	Screening flexible sigmoidoscopy (year 1)
	Costs (£m)	QALYs	NMB (£m)	CRC incidence	CRC mortality		
no screening	£0.0m	-	£0.0m	-	-	-	-
gFOBT ages 60-74 biennial	-£37.3m	11,608	£269.4m	10.7%	16.8%	36,406	-
Bowel scope age 55, gFOBT ages 60-74 biennial	£12.9m	19,197	£371.1m	16.5%	22.5%	42,920	329,121
One-off screens							
Bowel Scope age 59	£26.2m	9,341	£160.7m	8.4%	9.8%	9,197	289,081
FIT20 age 57	-£39.6m	9,093	£221.5m	7.3%	8.5%	28,995	-
Repeated FIT screening							
2-yearly, age 51-65, FIT161	£146.6m	15,731	£168.0m	12.4%	15.1%	442,882	-
2-yearly, age 50-70, FIT153	£189.0m	19,497	£201.0m	16.6%	20.6%	590,935	-
2-yearly, age 50-74, FIT124	£192.2m	22,522	£258.2m	20.5%	25.6%	653,768	-
2-yearly, age 50-74, FIT74	£171.6m	26,820	£364.8m	24.7%	30.3%	681,188	-
Bowel scope and repeated FIT screening							
Bowel scope age 59, 2-yearly, age 51-65 (excl.59), FIT161	£173.8m	20,024	£226.7m	16.4%	19.5%	391,563	289,081
Bowel scope age 58, 2-yearly, age 50-70 (excl. 58), FIT153	£221.9m	23,400	£246.1m	20.0%	24.1%	535,251	300,813
Bowel scope age 58, 2-yearly, age 50-74 (excl. 58), FIT124	£228.3m	25,973	£291.2m	23.5%	28.6%	597,393	300,813

Scenario analysis with increased rates of symptomatic presentation (+10%)

Screening strategy	Incremental compared to no screening			Lifetime reduction		Screening referral colonoscopies (year1)	Screening flexible sigmoidoscopy (year 1)
	Costs (£m)	QALYs	NMB (£m)	CRC incidence	CRC mortality		
no screening	£0.0m	-	£0.0m	-	-	-	-
gFOBT ages 60-74 biennial	-£45.4m	10,807	£261.5m	11.2%	16.8%	36,309	-
Bowel scope age 55, gFOBT ages 60-74 biennial	£2.0m	17,938	£356.7m	16.9%	22.6%	42,830	329,054
One-off screens							
Bowel Scope age 59	£21.7m	8,763	£153.6m	8.5%	9.8%	9,180	288,987
FIT20 age 57	-£43.7m	8,514	£214.0m	7.3%	8.5%	28,976	-
Repeated FIT screening							
2-yearly, age 51-65, FIT161	-£68.2m	17,856	£425.4m	15.3%	18.4%	49,793	-
2-yearly, age 50-70, FIT153	-£85.6m	22,077	£527.1m	20.4%	24.9%	69,809	-
2-yearly, age 50-74, FIT124	-£104.0m	25,308	£610.1m	25.1%	30.7%	89,681	-
2-yearly, age 50-74, FIT74	-£127.4m	29,612	£719.6m	29.7%	35.5%	124,953	-
Bowel scope and repeated FIT screening							
Bowel scope age 59, 2-yearly, age 51-65 (excl.59), FIT161	-£14.1m	20,941	£432.9m	18.4%	21.7%	52,734	288,987
Bowel scope age 58, 2-yearly, age 50-70 (excl. 58), FIT153	-£25.2m	24,783	£520.9m	23.0%	27.4%	72,328	300,725
Bowel scope age 58, 2-yearly, age 50-74 (excl. 58), FIT124	-£40.8m	27,601	£592.8m	27.2%	32.7%	91,280	300,725

Males

Screening strategy	Incremental compared to no screening			Lifetime reduction		Screening referral colonoscopies (year1)	Screening flexible sigmoidoscopy (year 1)
	Costs (£m)	QALYs	NMB (£m)	CRC incidence	CRC mortality		
Current							
gFOBT ages 60-74 biennial	-£50.2m	14,049	£331.2m	10.4%	16.2%	36,596	-
Bowel scope age 55, gFOBT ages 60-74 biennial	-£14.0m	24,257	£499.1m	16.7%	22.5%	45,315	337,645
One-off screens							
Bowel Scope age 59	£9.2m	12,443	£239.7m	9.1%	10.5%	12,212	296,399
FIT20 age 57	-£55.4m	11,340	£282.2m	7.3%	8.6%	26,116	-
Repeated FIT screening							
2-yearly, age 51-65, FIT161	-£87.3m	23,668	£560.6m	15.1%	18.3%	51,348	-
2-yearly, age 50-70, FIT153	-£108.2m	29,082	£689.8m	19.9%	24.5%	71,815	-
2-yearly, age 50-74, FIT124	-£129.8m	33,165	£793.0m	24.3%	29.9%	90,997	-
2-yearly, age 50-74, FIT74	-£160.8m	38,740	£935.6m	28.6%	34.6%	122,425	-
Bowel scope and repeated FIT screening							
Bowel scope age 59, 2-yearly, age 51-65 (excl.59), FIT161	-£37.4m	28,159	£600.6m	18.5%	22.0%	57,056	296,399
Bowel scope age 58, 2-yearly, age 50-70 (excl. 58), FIT153	-£51.4m	33,036	£712.2m	22.8%	27.3%	77,140	308,483
Bowel scope age 58, 2-yearly, age 50-74 (excl. 58), FIT124	-£69.4m	36,533	£800.0m	26.6%	32.2%	95,478	308,483

Female

Screening strategy	Incremental compared to no screening			Lifetime reduction		Screening referral colonoscopies (year1)	Screening flexible sigmoidoscopy (year 1)
	Costs (£m)	QALYs	NMB (£m)	CRC incidence	CRC mortality		
Current							
gFOBT ages 60-74 biennial	-£25.7m	9,478	£215.3m	11.0%	17.2%	36,493	-
Bowel scope age 55, gFOBT ages 60-74 biennial	£34.1m	15,117	£268.3m	16.3%	22.6%	41,346	320,565
One-off screens							
Bowel Scope age 59	£38.4m	7,007	£101.8m	7.9%	9.2%	6,912	281,647
FIT20 age 57	-£26.6m	7,277	£172.1m	7.3%	8.6%	31,791	-
Repeated FIT screening							
2-yearly, age 51-65, FIT161	-£35.7m	15,314	£341.9m	15.1%	18.5%	48,988	-
2-yearly, age 50-70, FIT153	-£44.3m	18,995	£424.2m	20.3%	25.2%	68,806	-
2-yearly, age 50-74, FIT124	-£54.9m	21,825	£491.4m	25.1%	31.3%	89,520	-
2-yearly, age 50-74, FIT74	-£67.5m	25,515	£577.8m	29.7%	36.2%	128,742	-
Bowel scope and repeated FIT screening							
Bowel scope age 59, 2-yearly, age 51-65 (excl.59), FIT161	£23.7m	17,669	£329.7m	18.0%	21.5%	49,817	281,647
Bowel scope age 58, 2-yearly, age 50-70 (excl. 58), FIT153	£20.0m	21,049	£400.9m	22.6%	27.5%	69,187	293,049
Bowel scope age 58, 2-yearly, age 50-74 (excl. 58), FIT124	£11.2m	23,551	£459.8m	27.0%	33.1%	88,902	293,049

1.3 Appendix: Detailed results for repeated FIT screening strategies

Table 1: Optimal repeated FIT screening strategy for different referral colonoscopy capacities

Screening referral colonoscopy capacity	Screening strategy
50,000 (similar to current capacity)	2-yearly, age 51-65, FIT161 (8 screens)
70000	2-yearly, age 50-70, FIT153 (11 screens)
90,000 (optimistic future capacity)	2-yearly, age 50-74, FIT124 (13 screens)
110,000	2-yearly, age 50-74, FIT90 (13 screens)
130,000	2-yearly, age 50-74, FIT70 (13 screens)
150,000	1-yearly, age 50-74, FIT159 (25 screens)

Table 2: Repeated FIT screening strategy results for referral colonoscopy capacity 50,000

Strategy	Costs (discounted, incremental compared to no screening)	QALYs (discounted, incremental compared to no screening)	Cancer incidence	Cancer mortality	Number of screens	NMB	Screening referral colonoscopys
Screening referral colonoscopys < 50000							
1-yearly, age 56-63, FIT156	-£60.5m	18,372	15.4%	18.9%	8	£427.9m	49,895
1-yearly, age 55-62, FIT161	-£57.9m	18,464	14.8%	18.1%	8	£427.2m	49,942
1-yearly, age 56-63, FIT157	-£60.3m	18,333	15.3%	18.8%	8	£427.0m	49,705
1-yearly, age 55-62, FIT162	-£57.8m	18,426	14.8%	18.0%	8	£426.3m	49,757
1-yearly, age 57-64, FIT151	-£62.5m	18,187	15.9%	19.6%	8	£426.3m	49,938
1-yearly, age 56-63, FIT158	-£60.1m	18,295	15.3%	18.8%	8	£426.0m	49,518
1-yearly, age 55-62, FIT163	-£57.6m	18,389	14.8%	18.0%	8	£425.4m	49,573
1-yearly, age 57-64, FIT152	-£62.3m	18,148	15.9%	19.6%	8	£425.3m	49,744
2-yearly, age 51-65, FIT161	-£59.2m	19,098	15.1%	18.4%	8	**£441.2m**	49,856
2-yearly, age 53-67, FIT155	-£64.1m	18,824	16.1%	19.9%	8	**£440.6m**	49,945
2-yearly, age 51-65, FIT162	-£59.0m	19,056	15.1%	18.4%	8	**£440.1m**	49,668
2-yearly, age 52-66, FIT160	-£61.5m	18,915	15.5%	19.1%	8	**£439.8m**	49,991
2-yearly, age 53-67, FIT156	-£63.8m	18,780	16.1%	19.8%	8	**£439.5m**	49,753
2-yearly, age 51-65, FIT163	-£58.8m	19,014	15.0%	18.3%	8	**£439.1m**	49,482
2-yearly, age 52-66, FIT161	-£61.3m	18,872	15.5%	19.0%	8	**£438.8m**	49,804
2-yearly, age 53-67, FIT157	-£63.6m	18,737	16.0%	19.8%	8	**£438.4m**	49,564
3-yearly, age 52-70, FIT122	-£67.8m	18,274	16.3%	20.3%	7	**£433.3m**	49,837
3-yearly, age 50-71, FIT154	-£62.3m	18,539	16.2%	20.2%	8	**£433.1m**	49,951
3-yearly, age 50-68, FIT130	-£62.9m	18,485	15.3%	18.9%	7	£432.6m	49,913
3-yearly, age 50-71, FIT155	-£62.1m	18,493	16.1%	20.2%	8	£431.9m	49,759
3-yearly, age 52-70, FIT123	-£67.5m	18,218	16.3%	20.2%	7	£431.8m	49,586
3-yearly, age 50-68, FIT131	-£62.7m	18,431	15.3%	18.8%	7	£431.3m	49,674
3-yearly, age 52-67, FIT100	-£68.0m	18,157	15.4%	18.9%	6	£431.1m	49,841
3-yearly, age 50-71, FIT156	-£61.8m	18,447	16.1%	20.1%	8	£430.8m	49,568
4-yearly, age 50-70, FIT98	-£66.8m	17,933	15.6%	19.3%	6	£425.5m	49,921
4-yearly, age 51-71, FIT95	-£68.9m	17,807	16.0%	19.9%	6	£425.1m	49,936
4-yearly, age 50-74, FIT115	-£66.5m	17,886	16.4%	20.7%	7	£424.2m	49,832
4-yearly, age 50-70, FIT99	-£66.5m	17,861	15.5%	19.2%	6	£423.7m	49,590
4-yearly, age 51-71, FIT96	-£68.6m	17,734	16.0%	19.8%	6	£423.2m	49,597
4-yearly, age 50-74, FIT116	-£66.2m	17,825	16.4%	20.7%	7	£422.7m	49,564
4-yearly, age 50-70, FIT100	-£66.1m	17,791	15.5%	19.1%	6	£421.9m	49,266
4-yearly, age 51-71, FIT97	-£68.2m	17,662	15.9%	19.7%	6	£421.4m	49,264
5-yearly, age 50-70, FIT74	-£68.1m	17,379	15.2%	18.7%	5	£415.7m	49,620
5-yearly, age 51-71, FIT72	-£70.1m	17,239	15.7%	19.3%	5	£414.9m	49,529
5-yearly, age 50-70, FIT75	-£67.7m	17,284	15.2%	18.6%	5	£413.4m	49,165
5-yearly, age 51-71, FIT73	-£69.6m	17,143	15.6%	19.2%	5	£412.5m	49,065
5-yearly, age 50-70, FIT76	-£67.2m	17,192	15.1%	18.5%	5	£411.1m	48,721
5-yearly, age 51-71, FIT74	-£69.1m	17,048	15.5%	19.1%	5	£410.1m	48,614
5-yearly, age 52-72, FIT71	-£70.9m	16,937	15.9%	19.8%	5	£409.7m	49,933
5-yearly, age 50-70, FIT77	-£66.8m	17,101	15.0%	18.5%	5	£408.8m	48,288
6-yearly, age 50-74, FIT70	-£67.7m	16,646	15.5%	19.4%	5	£400.6m	49,624
6-yearly, age 52-70, FIT53	-£70.5m	16,496	15.1%	18.6%	4	£400.4m	49,534
6-yearly, age 53-71, FIT51	-£72.2m	16,364	15.6%	19.2%	4	£399.5m	49,344
6-yearly, age 50-74, FIT71	-£67.2m	16,548	15.4%	19.3%	5	£398.2m	49,148
6-yearly, age 52-70, FIT54	-£69.9m	16,369	15.0%	18.4%	4	£397.3m	48,880
6-yearly, age 53-71, FIT52	-£71.6m	16,233	15.5%	19.0%	4	£396.2m	48,671
6-yearly, age 50-74, FIT72	-£66.7m	16,452	15.3%	19.2%	5	£395.8m	48,683
6-yearly, age 50-68, FIT58	-£65.5m	16,472	14.1%	17.1%	4	£394.9m	49,494

Table 3: Repeated FIT screening strategy results for referral colonoscopy capacity 90,000

Strategy	Costs (discounted, incremental compared to no screening)	QALYs (discounted, incremental compared to no screening)	Cancer incidence	Cancer mortality	Number of screens	NMB	Screening referral colonoscopys
Screening referral colonoscopys < 90000							
1-yearly, age 52-66, FIT170	-£71.4m	27,397	22.4%	27.1%	15	**£619.3m**	89,905
1-yearly, age 52-66, FIT171	-£71.2m	27,354	22.4%	27.1%	15	**£618.3m**	89,592
1-yearly, age 51-65, FIT173	-£66.7m	27,566	21.7%	26.1%	15	£618.1m	89,914
1-yearly, age 53-67, FIT167	-£75.2m	27,116	23.0%	28.1%	15	£617.5m	89,842
1-yearly, age 52-66, FIT172	-£71.0m	27,311	22.3%	27.0%	15	£617.2m	89,283
1-yearly, age 51-65, FIT174	-£66.6m	27,523	21.6%	26.1%	15	£617.0m	89,603
1-yearly, age 53-67, FIT168	-£75.0m	27,072	23.0%	28.0%	15	£616.4m	89,528
1-yearly, age 52-66, FIT173	-£70.8m	27,268	22.3%	27.0%	15	£616.1m	88,977
2-yearly, age 50-74, FIT124	-£89.7m	27,037	24.7%	30.7%	13	**£630.4m**	89,822
2-yearly, age 50-72, FIT114	-£89.9m	26,989	24.0%	29.5%	12	**£629.7m**	89,778
2-yearly, age 50-74, FIT125	-£89.4m	26,968	24.7%	30.6%	13	**£628.7m**	89,382
2-yearly, age 50-72, FIT115	-£89.5m	26,914	23.9%	29.4%	12	**£627.8m**	89,287
2-yearly, age 50-74, FIT126	-£89.0m	26,899	24.6%	30.5%	13	**£627.0m**	88,947
2-yearly, age 50-70, FIT103	-£89.6m	26,853	23.1%	28.2%	11	**£626.6m**	89,505
2-yearly, age 50-72, FIT116	-£89.1m	26,840	23.9%	29.3%	12	**£626.0m**	88,803
2-yearly, age 51-73, FIT111	-£92.6m	26,634	24.5%	30.3%	12	**£625.3m**	89,546
3-yearly, age 50-74, FIT70	-£98.9m	25,777	24.0%	29.4%	9	£614.4m	89,258
3-yearly, age 50-74, FIT71	-£98.3m	25,653	23.8%	29.3%	9	£611.3m	88,404
3-yearly, age 50-71, FIT63	-£97.1m	25,606	22.8%	27.6%	8	£609.2m	89,247
3-yearly, age 50-74, FIT72	-£97.7m	25,531	23.7%	29.2%	9	£608.3m	87,571
3-yearly, age 52-73, FIT58	-£102.9m	25,186	24.1%	29.4%	8	£606.6m	89,251
3-yearly, age 50-71, FIT64	-£96.5m	25,470	22.7%	27.5%	8	£605.9m	88,272
3-yearly, age 50-74, FIT73	-£97.1m	25,411	23.6%	29.0%	9	£605.4m	86,760
3-yearly, age 52-73, FIT59	-£102.2m	25,041	23.9%	29.2%	8	£603.0m	88,204
4-yearly, age 50-74, FIT48	-£101.1m	24,869	23.3%	28.5%	7	£598.4m	89,699
4-yearly, age 50-74, FIT49	-£100.3m	24,686	23.2%	28.3%	7	£594.0m	88,383
4-yearly, age 50-74, FIT50	-£99.5m	24,508	23.0%	28.1%	7	£589.6m	87,116
4-yearly, age 51-71, FIT41	-£100.5m	24,414	22.4%	27.0%	6	£588.7m	89,575
4-yearly, age 50-70, FIT43	-£96.9m	24,454	21.6%	26.0%	6	£586.0m	88,735
4-yearly, age 50-74, FIT51	-£98.7m	24,334	22.8%	27.9%	7	£585.3m	85,895
4-yearly, age 51-71, FIT42	-£99.6m	24,205	22.2%	26.8%	6	£583.7m	88,001
4-yearly, age 50-74, FIT52	-£97.9m	24,164	22.7%	27.7%	7	£581.2m	84,718
5-yearly, age 50-70, FIT33	-£96.7m	23,802	21.2%	25.5%	5	£572.7m	89,214
5-yearly, age 51-71, FIT32	-£99.7m	23,634	21.8%	26.3%	5	£572.4m	88,948
5-yearly, age 50-70, FIT34	-£95.7m	23,539	21.0%	25.2%	5	£566.4m	87,208
5-yearly, age 51-71, FIT33	-£98.6m	23,365	21.6%	26.0%	5	£565.9m	86,898
5-yearly, age 52-72, FIT32	-£100.7m	23,107	22.1%	26.7%	5	£562.8m	88,157
5-yearly, age 50-70, FIT35	-£94.7m	23,285	20.7%	24.9%	5	£560.4m	85,309
5-yearly, age 51-71, FIT34	-£97.5m	23,106	21.3%	25.7%	5	£559.6m	84,961
5-yearly, age 53-73, FIT31	-£102.2m	22,762	22.6%	27.4%	5	£557.5m	88,201
6-yearly, age 50-74, FIT31	-£98.0m	23,093	21.9%	26.7%	5	£559.8m	89,253
6-yearly, age 50-74, FIT32	-£96.8m	22,811	21.6%	26.3%	5	£553.0m	87,135
6-yearly, age 52-70, FIT24	-£98.2m	22,663	21.1%	25.3%	4	£551.5m	89,515
6-yearly, age 53-71, FIT23	-£100.9m	22,508	21.8%	26.1%	4	£551.1m	89,237
6-yearly, age 50-74, FIT33	-£95.6m	22,540	21.4%	26.0%	5	£546.4m	85,137
6-yearly, age 55-73, FIT21	-£104.3m	21,978	22.9%	27.6%	4	£543.9m	89,486
6-yearly, age 52-70, FIT25	-£96.8m	22,309	20.8%	24.9%	4	£543.0m	86,715
6-yearly, age 53-71, FIT24	-£99.4m	22,142	21.4%	25.7%	4	£542.2m	86,338

Table 4: Repeated FIT screening strategy results for referral colonoscopy capacity 110,000

Strategy	Costs (discounted, incremental compared to no screening)	QALYs (discounted, incremental compared to no screening)	Cancer incidence	Cancer mortality	Number of screens	NMB	Screening referral colonoscopys
Screening referral colonoscopys < 110000							
2-yearly, age 50-74, FIT90	-£103.3m	29,852	27.5%	33.7%	13	£700.4m	109,964
2-yearly, age 50-74, FIT91	-£102.9m	29,754	27.4%	33.6%	13	£698.0m	109,178
2-yearly, age 50-74, FIT92	-£102.4m	29,656	27.3%	33.5%	13	£695.6m	108,408
2-yearly, age 50-72, FIT84	-£102.1m	29,644	26.5%	32.2%	12	£695.0m	109,348
2-yearly, age 50-74, FIT93	-£102.0m	29,560	27.2%	33.4%	13	£693.2m	107,654
2-yearly, age 50-72, FIT85	-£101.7m	29,540	26.4%	32.1%	12	£692.5m	108,495
2-yearly, age 51-73, FIT81	-£105.6m	29,335	27.2%	33.2%	12	£692.3m	109,629
1-yearly, age 50-67, FIT167	-£70.8m	31,027	24.9%	30.0%	18	£691.3m	109,605
2-yearly, age 50-74, FIT94	-£101.5m	29,465	27.1%	33.3%	13	£690.8m	106,913
1-yearly, age 51-68, FIT165	-£75.9m	30,741	25.6%	31.0%	18	£690.7m	109,997
1-yearly, age 50-67, FIT168	-£70.6m	30,980	24.9%	29.9%	18	£690.2m	109,212
2-yearly, age 50-70, FIT76	-£100.9m	29,460	25.5%	30.8%	11	£690.1m	109,521
2-yearly, age 50-72, FIT86	-£101.2m	29,437	26.3%	32.0%	12	£690.0m	107,659
2-yearly, age 51-73, FIT82	-£105.1m	29,228	27.1%	33.1%	12	£689.6m	108,746
1-yearly, age 51-68, FIT166	-£75.7m	30,694	25.6%	31.0%	18	£689.6m	109,602
1-yearly, age 50-67, FIT169	-£70.4m	30,933	24.8%	29.9%	18	£689.1m	108,822
2-yearly, age 50-74, FIT95	-£101.1m	29,371	27.0%	33.2%	13	£688.5m	106,187
1-yearly, age 51-68, FIT167	-£75.5m	30,647	25.5%	30.9%	18	£688.4m	109,212
1-yearly, age 50-67, FIT170	-£70.2m	30,886	24.8%	29.9%	18	£687.9m	108,437
2-yearly, age 51-71, FIT74	-£104.5m	29,162	26.2%	31.7%	11	£687.7m	109,656

Table 5: Repeated FIT screening strategy results for referral colonoscopy capacity 130,000

Strategy	Costs (discounted, incremental compared to no screening)	QALYs (discounted, incremental compared to no screening)	Cancer incidence	Cancer mortality	Number of screens	NMB	Screening referral colonoscopys
Screening referral colonoscopys < 130000							
2-yearly, age 50-74, FIT70	-£113.2m	32,116	29.8%	36.1%	13	£755.5m	129,900
1-yearly, age 50-71, FIT172	-£78.1m	33,827	29.2%	35.5%	22	£754.6m	129,765
1-yearly, age 50-71, FIT173	-£77.9m	33,779	29.2%	35.5%	22	£753.5m	129,328
2-yearly, age 50-74, FIT71	-£112.6m	31,988	29.6%	35.9%	13	£752.4m	128,661
1-yearly, age 50-71, FIT174	-£77.7m	33,731	29.1%	35.4%	22	£752.3m	128,895
1-yearly, age 50-70, FIT162	-£78.4m	33,659	28.6%	34.6%	21	£751.6m	129,526
1-yearly, age 50-71, FIT175	-£77.4m	33,684	29.1%	35.4%	22	£751.1m	128,466
1-yearly, age 51-72, FIT168	-£83.0m	33,386	29.9%	36.5%	22	£750.7m	129,970
1-yearly, age 50-70, FIT163	-£78.2m	33,608	28.5%	34.6%	21	£750.3m	129,054
1-yearly, age 50-71, FIT176	-£77.2m	33,637	29.0%	35.3%	22	£750.0m	128,042
1-yearly, age 51-73, FIT179	-£81.9m	33,396	30.3%	37.2%	23	£749.8m	129,595
1-yearly, age 51-72, FIT169	-£82.7m	33,338	29.8%	36.4%	22	£749.5m	129,526
2-yearly, age 50-74, FIT72	-£112.1m	31,861	29.5%	35.8%	13	£749.3m	127,454
1-yearly, age 51-71, FIT158	-£83.6m	33,284	29.3%	35.6%	21	£749.3m	129,680
1-yearly, age 50-70, FIT164	-£78.0m	33,558	28.5%	34.5%	21	£749.1m	128,588
1-yearly, age 50-71, FIT177	-£77.0m	33,589	29.0%	35.3%	22	£748.8m	127,623
1-yearly, age 50-69, FIT151	-£78.7m	33,496	28.0%	33.7%	20	£748.6m	129,746
1-yearly, age 51-73, FIT180	-£81.6m	33,350	30.3%	37.2%	23	£748.6m	129,187
1-yearly, age 51-72, FIT170	-£82.5m	33,289	29.8%	36.4%	22	£748.3m	129,087
1-yearly, age 51-71, FIT159	-£83.4m	33,233	29.3%	35.6%	21	£748.0m	129,200

Table 6: Repeated FIT screening strategy results for referral colonoscopy capacity 150,000

Strategy	Costs (discounted, incremental compared to no screening)	QALYs (discounted, incremental compared to no screening)	Cancer incidence	Cancer mortality	Number of screens	NMB	Screening referral colonoscopys
Screening referral colonoscopys < 150000							
1-yearly, age 50-74, FIT159	-£85.4m	36,088	32.8%	40.0%	25	£807.1m	149,500
1-yearly, age 50-73, FIT151	-£86.0m	36,048	32.3%	39.2%	24	£807.0m	149,889
1-yearly, age 50-74, FIT160	-£85.2m	36,034	32.7%	39.9%	25	£805.8m	148,955
1-yearly, age 50-73, FIT152	-£85.8m	35,993	32.2%	39.1%	24	£805.6m	149,305
1-yearly, age 50-74, FIT161	-£84.9m	35,982	32.7%	39.9%	25	£804.6m	148,417
1-yearly, age 50-73, FIT153	-£85.5m	35,938	32.2%	39.1%	24	£804.3m	148,728
1-yearly, age 50-72, FIT144	-£86.2m	35,901	31.7%	38.3%	23	£804.2m	149,441
1-yearly, age 50-74, FIT162	-£84.7m	35,929	32.6%	39.8%	25	£803.3m	147,884
1-yearly, age 50-73, FIT154	-£85.3m	35,883	32.1%	39.0%	24	£803.0m	148,158
1-yearly, age 50-71, FIT135	-£86.5m	35,821	31.1%	37.4%	22	£802.9m	149,970
1-yearly, age 50-72, FIT145	-£85.9m	35,844	31.6%	38.2%	23	£802.8m	148,821
1-yearly, age 50-74, FIT163	-£84.4m	35,877	32.6%	39.8%	25	£802.0m	147,357
1-yearly, age 50-73, FIT155	-£85.1m	35,829	32.1%	39.0%	24	£801.6m	147,594
1-yearly, age 50-71, FIT136	-£86.3m	35,760	31.0%	37.4%	22	£801.5m	149,295
1-yearly, age 50-72, FIT146	-£85.7m	35,787	31.6%	38.2%	23	£801.4m	148,209
1-yearly, age 50-74, FIT164	-£84.2m	35,825	32.5%	39.7%	25	£800.7m	146,836
1-yearly, age 51-74, FIT147	-£90.8m	35,493	32.9%	40.1%	24	£800.7m	149,479
1-yearly, age 50-73, FIT156	-£84.8m	35,775	32.0%	38.9%	24	£800.3m	147,036
1-yearly, age 50-72, FIT147	-£85.4m	35,731	31.5%	38.1%	23	£800.1m	147,605
1-yearly, age 50-71, FIT137	-£86.0m	35,701	31.0%	37.3%	22	£800.0m	148,629

1.4 Appendix: Probabilistic Sensitivity Analyses (PSA)

We ran probabilistic sensitivity analysis for the nine key screening strategies presented in the main results. The analysis was run for 1000 runs, varying parameters using the distributions presented in the model parameters table. We looked at the impact of the PSA for the following pairwise comparisons:

i) comparing one-off Bowel Scope with one-off FIT20;
ii) replacing FIT at 59 with Bowel Scope for a 2-yearly FIT 161 screening strategy;
iii) replacing FIT at 58 with Bowel Scope for a 2-yearly FIT 161 screening strategy;
iv) replacing FIT at 58 with Bowel scope for a 2-yearly FIT 161 screening strategy;

The mean cost and QALYs from the PSA differed only slightly to those for the deterministic results. This is illustrated on the PSA scatterplot. The CEACs illustrate that the probability that it is cost effective to replace a FIT with a bowel scope varies according to the willingness to pay threshold. At higher willingness to pay thresholds the probability that it is cost effective to replace a FIT with a bowel scope is higher.

Figure 1: Cost-effectiveness for the key strategy comparisons, assessed at £20,000 per QALY

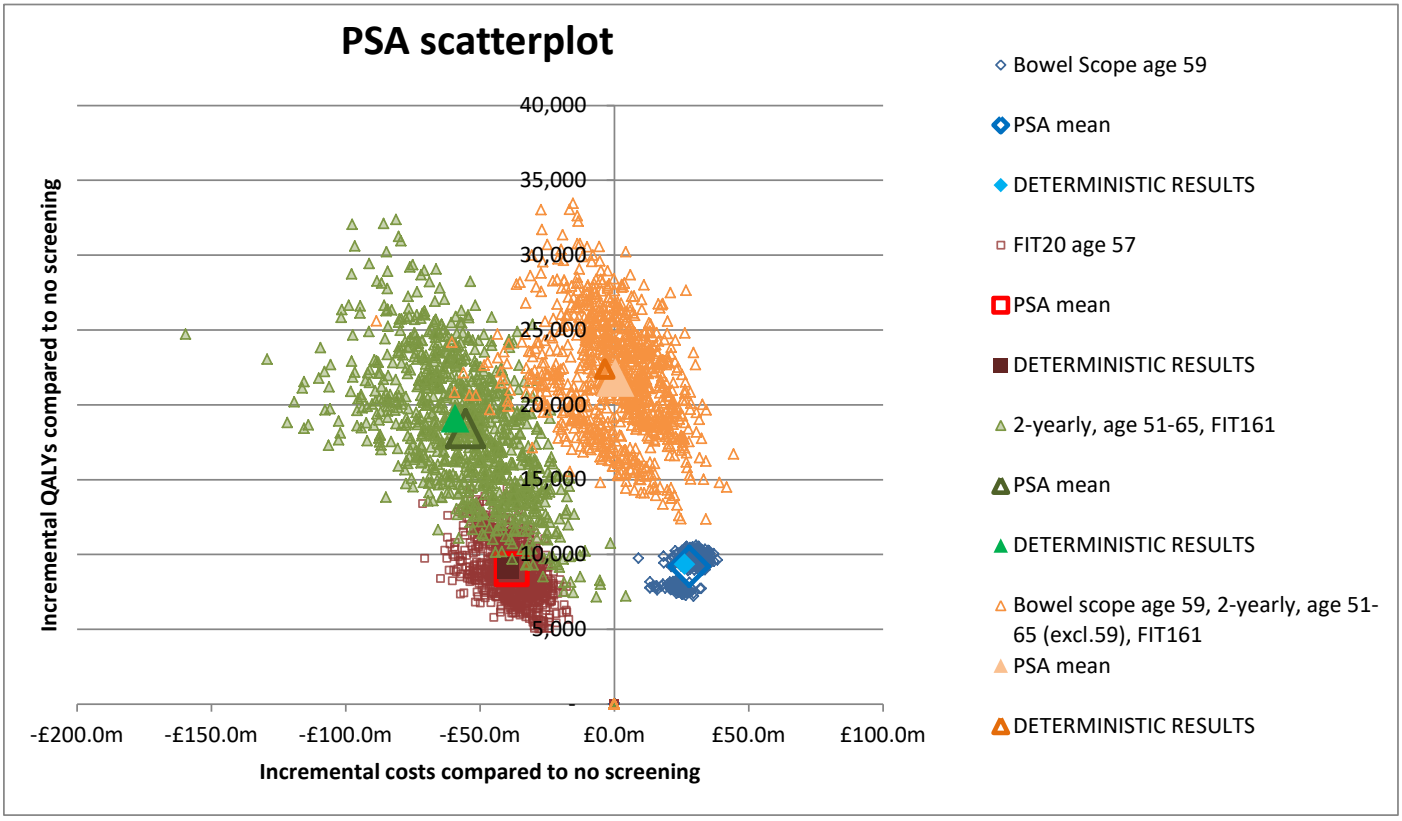
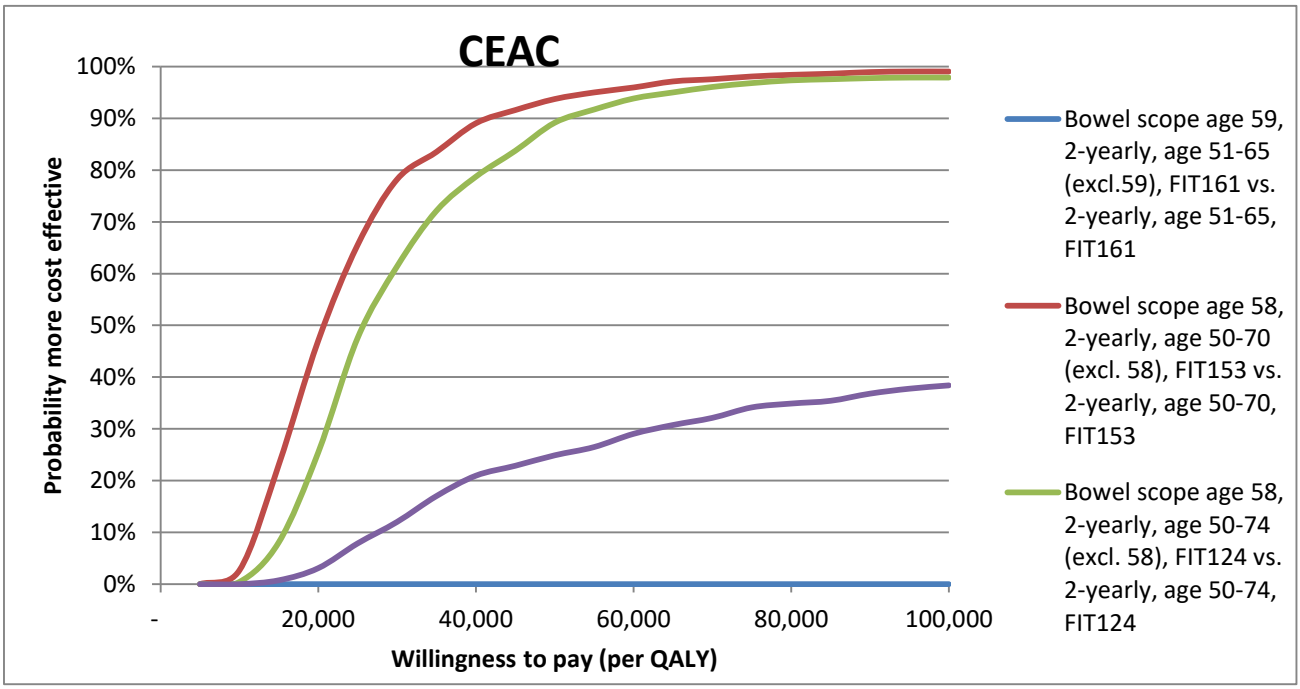


Figure 2: Cost effectiveness acceptability curves (CEACs) for the key strategy comparisons



WTP	Bowel scope age 59, 2-yearly, age 51-65 (excl.59), FIT161 vs. 2-yearly, age 51-65, FIT161	Bowel scope age 58, 2-yearly, age 50-70 (excl. 58), FIT153 vs. 2-yearly, age 50-70, FIT153	Bowel scope age 58, 2-yearly, age 50-74 (excl. 58), FIT124 vs. 2-yearly, age 50-74, FIT124	Bowel Scope age 59 vs. FIT20 age 57
£20,000	0%	47%	26%	3%
£30,000	0%	78%	61%	12%

1.5 Appendix: Table of model parameter values

Table 7: Table of parameters used in the model

Test Characteristics	*mean*	*Distribution used in PSA and 95% CI*	*source*
gFOBT Sensitivity for LR adenomas	0.01	Beta(422,72,457) (0.01-0.01)	BCSP data
gFOBT Sensitivity for HR adenomas	0.10	Beta(732,6,777) (0.09-0.10)	BCSP data
gFOBT Sensitivity for CRC	0.17	Beta(139,691) (0.14-0.19)	BCSP data
gFOBT Specificity age 50	0.99	Beta(128,370,1,428) (0.99-0.99)	BCSP data
gFOBT Specificity age 70	0.99	Beta(128,370,1,428) (0.99-0.99)	BCSP data
FS Sensitivity for LR adenomas	0.24	Beta(1,535,4,839) (0.23-0.25)	BCSP data
FS Sensitivity for HR adenomas	0.68	Beta(1,967,940) (0.66-0.69)	BCSP data
FS Sensitivity for CRC	0.44	Beta(294,378) (0.40-0.48)	BCSP data
FS Specificity	1.00	NA	Assumption due to nature of the test
FIT Sensitivity for LR adenomas	0.01	Beta(8,1,454) (0.00-0.01)	FIT pilot, Moss et al 2016.
FIT Sensitivity for HR adenomas	0.17	Beta(20,99) (0.11-0.24)	FIT pilot, Moss et al 2016.
FIT Sensitivity for CRC	0.17	Beta(3,13) (0.03-0.38)	FIT pilot, Moss et al 2016.
FIT Specificity	0.99	Beta(2,254,32) (0.98-0.99)	FIT pilot, Moss et al 2016.
COL Sensitivity for LR adenomas	0.77	Beta(544,167) (0.73-0.80)	Van Rijn et al 2006
COL Sensitivity for HR adenomas	0.98	Beta(94,2) (0.94-1.00)	Bressler et al 2007
COL Sensitivity for CRC	0.98	Beta(94,2) (0.94-1.00)	Bressler et al 2007
COL Specificity	-	NA	Assumption due to nature of the test
CTC Sensitivity for LR adenomas	0.63	Beta(446,265) (0.40-0.98)	Assumption based on detection rates relative to COL in Atkin 2013
CTC Sensitivity for HR adenomas	0.80	Beta(77,19) (0.51-1.25)	Assumption based on detection rates relative to COL in Atkin 2013
CTC Sensitivity for CRC	0.96	Beta(92,4) (0.62-1.49)	Assumption based on detection rates relative to COL in Atkin 2013
CTC Specificity	0.88	Beta(5,267,710) (0.87-0.89)	Lin et al 2015 review
Proportion CTC of all referrals Age 60-61	0.03	Beta(255,7,832) (0.03-0.04)	NHS BCSP data 2014/15.
Proportion CTC of all referrals Age 62-63	0.04	Beta(304,7,918) (0.03-0.04)	NHS BCSP data 2014/15.
Proportion CTC of all referrals Age 64-65	0.04	Beta(399,8,478) (0.04-0.05)	NHS BCSP data 2014/15.
Proportion CTC of all referrals Age 66-67	0.04	Beta(415,8,965) (0.04-0.05)	NHS BCSP data 2014/15.
Proportion CTC of all referrals Age 72-74	0.06	Beta(414,6,405) (0.06-0.07)	NHS BCSP data 2014/15.
Proportion CTC of all referrals Age 60-61	0.03	Beta(9,29,991) (0.00-0.00)	NHS BCSP data 2014/15.
COL (without polypectomy) perforation rate	0.00	Beta(28,30,853) (0.00-0.00)	Rutter et al 2014
COL (with polypectomy) perforation rate	0.00	Beta(4,73) (0.01-0.11)	Rutter et al 2014
COL Probability of death following perforation	0.05	N/A	Gatto et al 2003
FS (without polypectomy) perforation rate	-	Beta(1,9,498) (0.00-0.00)	FS UK screening trial data, Atkin et al 2002
FS (with polypectomy) perforation rate	0.00	Beta(2,29) (0.01-0.17)	FS UK screening trial data, Atkin et al 2002
FS Probability of death following perforation	0.06	Beta(12,40,609) (0.00-0.00)	Gatto et al 2003
FS probability of hospitalisation for bleeding	0.00	Beta(52,130,779) (0.00-0.00)	FS UK screening trial data, Atkin et al 2002
COL probability of hospitalisation for bleeding	0.00	Beta(6,29,042) (0.00-0.00)	Rutter et al 2014
CTC perforation rate	0.00	N/A	Bellini et al 2014 metanalysis
CTC Probability of death following perforation	-	N/A	Bellini et al 2014 metanalysis
gFOBT mean number of tests completed	1.08	N/A	Assumption based on number of gFOBTs returned within 7 days
Repeat rates	*mean*	*Distribution used in PSA and 95% CI*	*source*
gFOBT mean number of tests completed	1.08	N/A	Assumption based on number of gFOBTs returned within 7 days
iFOBT mean number of tests completed	1.01	N/A	NHS BCSP data, Italian iFOBT screening programme Zorzi et al 2009
FS Probability test repeated on a later day	0.02	Beta(839,39,782) (0.02-0.02)	FS UK screening trial data, Atkin et al 2002
COL repeat test rate	0.07	Beta(5,453,72,858) (0.07-0.07)	NHS BCSP data
CTC additional investigation rate	0.89	Beta(911,116)	Plumb et al 2013
Mean gFOBT uptake over all screening rounds	0.58	Beta(2,398,418,1,719,460) (0.58-0.58)	NHS BCSP data 2014/15
Mean iFOBT uptake over all screening rounds	0.65	Beta(26,674,14,256) (0.65-0.66)	Calculated compared with gFOBT data from RR derived from Moss et al 20

Natural history parameters	*mean*	*Distribution used in PSA and 95% CI*	*source*
Normal epithelium to LR adenomas - age 30	0.021	Correlated parameter set (0.018-0.022)	Model calibration
Normal epithelium to LR adenomas - age 50	0.020	Correlated parameter set (0.019-0.022)	Model calibration
Normal epithelium to LR adenomas - age 70	0.045	Correlated parameter set (0.029-0.042)	Model calibration
Normal epithelium to LR adenomas - age 100	0.011	Correlated parameter set (0.005-0.031)	Model calibration
LR adenomas to high risk adenomas - age 30	0.009	Correlated parameter set (0.008-0.014)	Model calibration
LR adenomas to high risk adenomas - age 50	0.008	Correlated parameter set (0.006-0.008)	Model calibration
LR adenomas to high risk adenomas - age 70	0.008	Correlated parameter set (0.008-0.010)	Model calibration
LR adenomas to high risk adenomas - age 100	0.004	Correlated parameter set (0.004-0.010)	Model calibration
HR adenomas to Dukes A CRC - age 30	0.029	Correlated parameter set (0.004-0.036)	Model calibration
HR adenomas to Dukes A CRC - age 50	0.025	Correlated parameter set (0.022-0.026)	Model calibration
HR adenomas to Dukes A CRC - age 70	0.054	Correlated parameter set (0.050-0.058)	Model calibration
HR adenomas to Dukes A CRC - age 100	0.115	Correlated parameter set (0.084-0.115)	Model calibration
Normal epithelium to CRC Dukes A	0.00004	Correlated parameter set (0.000-0.000)	Model calibration
Preclinical CRC: Dukes Stage A to B	0.51	Correlated parameter set (0.504-0.886)	Model calibration
Preclinical CRC: Dukes Stage B to C	0.69	Correlated parameter set (0.499-0.797)	Model calibration
Preclinical CRC: Dukes Stage C to D	0.71	Correlated parameter set (0.594-0.762)	Model calibration
Symptomatic presentation with CRC Dukes A	0.04	Correlated parameter set (0.043-0.070)	Model calibration
Symptomatic presentation with CRC Dukes B	0.18	Correlated parameter set (0.124-0.195)	Model calibration
Symptomatic presentation with CRC Dukes C	0.37	Correlated parameter set (0.303-0.394)	Model calibration
Symptomatic presentation with CRC Dukes D	0.74	Correlated parameter set (0.647-0.924)	Model calibration
Proportion of cancer incidence classified as proximal	0.38	Beta(422,72,457) (0.01-0.01)	Cancer Registrations 2007, England
Average number of adenomas present in patient with at least one	2.30	Beta(732,6,777) (0.09-0.10)	Rutter et al 2014
Proportion of advanced adenomas classified as HR adenomas	0.75	Beta(139,691) (0.14-0.19)	FS UK screening trial data, Atkin et al 2002
Surveillance parameters	*mean*	*Distribution used in PSA and 95% CI*	*source*
Proportion of HR polypectomy requiring annual surveillance	0.29		NHS BCSP data
LR polypectomy, transition probability LR	0.10		
LR polypectomy, transition probability HR	0.04		
LR polypectomy, transition probability CRC	0.00		
IR polypectomy, transition probability LR	0.16		England BCSP data, Martinez et al 2009 LR=low risk, IR-intermediate
IR polypectomy, transition probability HR	0.09		risk, HR=high risk
IR polypectomy, transition probability CRC	0.00		
HR polypectomy, transition probability LR	0.19		
HR polypectomy, transition probability HR	0.57		

Harm/complications parameters	*mean*	*Distribution used in PSA and 95% CI*	*source*
COL (without polypectomy) perforation rate	0.000	Beta(9,29,991) (0.00-0.00)	Rutter et al 2014
COL (with polypectomy) perforation rate	0.001	Beta(28,30,853) (0.00-0.00)	Rutter et al 2014
COL Probability of death following perforation	0.052	Beta(4,73) (0.01-0.11)	Gatto et al 2003
FS (without polypectomy) perforation rate	-	N/A	FS UK screening trial data, Atkin et al 2002
FS (with polypectomy) perforation rate	0.000	Beta(1,9,498) (0.00-0.00)	FS UK screening trial data, Atkin et al 2002
FS Probability of death following perforation	0.065	Beta(2,29) (0.01-0.17)	Gatto et al 2003
FS probability of hospitalisation for bleeding	0.000	Beta(12,40,609) (0.00-0.00)	FS UK screening trial data, Atkin et al 2002
COL probability of hospitalisation for bleeding	0.000	Beta(52,130,779) (0.00-0.00)	Rutter et al 2014
CTC perforation rate	0.000	Beta(6,29,042) (0.00-0.00)	Bellini et al 2014 metanalysis
CTC Probability of death following perforation	-	N/A	Bellini et al 2014 metanalysis
gFOBT mean number of tests completed	1.080	N/A	Assumption based on number of gFOBTs returned within 7 days
iFOBT mean number of tests completed	1.010	N/A	NHS BCSP data, Italian iFOBT screening programme Zorzi et al 2009
FS Probability test repeated on a later day	0.021	Beta(839,39,782) (0.02-0.02)	FS UK screening trial data, Atkin et al 2002
CTC Probability of death following perforation	-	Beta(5,453,72,858) (0.07-0.07)	Bellini et al 2014 metanalysis

Resource Use parameters	*mean*	*Distribution used in PSA and 95% CI*	*source*
Cost of gFOBT screen (non-compliers)	£2.23	Uniform(2.01,2.45)	Southern Hub screening costings model inflated
Cost of gFOBT screen (normal result)	£3.69	Uniform(3.32,4.06)	Southern Hub screening costings model inflated
Cost of gFOBT screen (positive result)	£13.11	Uniform(11.80,14.42)	Southern Hub screening costings model inflated
Cost of iFOBT screen (non-compliers)	£7.06	Uniform(6.35,7.77)	Southern Hub screening costings model inflated
Cost of iFOBT screen (normal result)	£8.09	Uniform(7.28,8.90)	Southern Hub screening costings model inflated
Cost of iFOBT screen (positive result)	£17.78	Uniform(16.00,19.56)	Southern Hub screening costings model inflated
Cost of FS screen excl. FS exam (non-compliers)	£5.51	Uniform(4.96,6.07)	Southern Hub screening costings model inflated
Cost of FS screen excl. FS exam (not referred to COL)	£6.60	Uniform(5.94,7.26)	Southern Hub screening costings model inflated
Cost of FS screen excl. FS exam (referred to COL)	£16.29	Uniform(14.66,17.92)	Southern Hub screening costings model inflated
Cost of FS (without polypectomy)	£311.02	Uniform(280,342)	NHS reference costs 14/15 for colonoscopy * 0.6 on advice of Wendy Atki
Cost of FS (with polypectomy)	£360.10	Uniform(324,396)	NHS reference costs 14/15 for colonoscopy * 0.6 on advice of Wendy Atki
Proportion of LR adenomas being referred for COL following FS	£0.03	Uniform(0.03,0.03)	FS UK screening trial data, Atkin et al 2002
Cost of specialised screening nurse post FS	£16.25	Uniform(14.63,17.88)	PSSRU Unit Costs 14/15
Cost of specialised screening nurse post FOBT	£32.50	Uniform(29.25,35.75)	PSSRU Unit Costs 14/16
Cost of COL (without polypectomy)	£518.36	Uniform(467,570)	NHS reference costs 14/15
Cost of COL (with polypectomy)	£600.16	Uniform(540,660)	NHS reference costs 14/15
Cost of treating bowel perforation (major surgery)	£1,272.78	Gamma(100,13)	NHS reference costs 14/15
Cost of admittance for bleeding (overnight stay on medical ward)	£474.54	Gamma(100,5)	NHS reference costs 14/16
Pathology cost for adenoma	£28.82	Gamma(100,0)	NHS reference costs 14/15, histopathology
Pathology cost for cancer	£28.82	Gamma(100,0)	NHS reference costs 14/15, histopathology
Cost of CTC	£136.21	Uniform(76,163)	NHS reference costs 14/15
Cost treatment Dukes' A Age 40-49	£31,218	Gamma(100,312)	Laudicella et al. 2016
Cost treatment Dukes' A Age 50-59	£31,218	Gamma(100,312)	Laudicella et al. 2016
Cost treatment Dukes' A Age 60-69	£31,798	Gamma(100,318)	Laudicella et al. 2016
Cost treatment Dukes' A Age 70-79	£32,377	Gamma(100,324)	Laudicella et al. 2016
Cost treatment Dukes' A Age 80-100	£32,377	Gamma(100,324)	Laudicella et al. 2016
Cost treatment Dukes' B Age 40-49	£31,218	Gamma(100,312)	Laudicella et al. 2016
Cost treatment Dukes' B Age 50-59	£31,218	Gamma(100,312)	Laudicella et al. 2016
Cost treatment Dukes' B Age 60-69	£31,798	Gamma(100,318)	Laudicella et al. 2016
Cost treatment Dukes' B Age 70-79	£32,377	Gamma(100,324)	Laudicella et al. 2016
Cost treatment Dukes' B Age 80-100	£32,377	Gamma(100,324)	Laudicella et al. 2016
Cost treatment Dukes' C Age 40-49	£44,086	Gamma(100,441)	Laudicella et al. 2016
Cost treatment Dukes' C Age 50-59	£44,086	Gamma(100,441)	Laudicella et al. 2016
Cost treatment Dukes' C Age 60-69	£40,729	Gamma(100,407)	Laudicella et al. 2016
Cost treatment Dukes' C Age 70-79	£37,371	Gamma(100,374)	Laudicella et al. 2016
Cost treatment Dukes' C Age 80-100	£37,371	Gamma(100,374)	Laudicella et al. 2016
Cost treatment Dukes' D Age 40-49	£44,086	Gamma(100,441)	Laudicella et al. 2016
Cost treatment Dukes' D Age 50-59	£44,086	Gamma(100,441)	Laudicella et al. 2016
Cost treatment Dukes' D Age 60-69	£40,729	Gamma(100,407)	Laudicella et al. 2016
Cost treatment Dukes' D Age 70-79	£37,371	Gamma(100,374)	Laudicella et al. 2016
Cost treatment Dukes' D Age 80-100	£37,371	Gamma(100,374)	Laudicella et al. 2016
Discount rate for costs	3.5%	N/A	NICE methods of techonology appraisal 2008
Discount rate for health outcomes	3.5%	N/A	NICE methods of techonology appraisal 2008
Willingness to pay threshold	£20,000	N/A	NICE methods of techonology appraisal 2008
Screening participation parameters	*mean*	*Distribution used in PSA and 95% CI*	*source*
Mean gFOBT uptake over all screening rounds	0.58	Beta(2,398,418,1,719,460) (0.58-0.58)	NHS BCSP data 2014/15
Mean iFOBT uptake over all screening rounds	0.65	Beta(26,674,14,256) (0.65-0.66)	Calculated compared with gFOBT data from RR derived from Moss et al 20
gFOBT participation for a round for those who comply with at least one gFOBT test (incidenct uptake)	0.85	Beta(1,934,059,331,839) (0.85-0.85)	NHS BCSP data 2014/15
iFOBT participation for a round for those who comply with at least one iFOBT test (incident uptake)	0.90	Beta(20,287,2,360) (0.89-0.90)	Calculated compared with gFOBT data from RR derived from Moss et al 20
Follow-up compliance FOBT screening	0.87	Beta(37,517,5,509) (0.87-0.88)	NHS BCSP data 2014/15
Follow-up compliance FS screening	0.96	Beta(2,047,79) (0.95-0.97)	FS UK screening trial data, Atkin et al 2002
COL surveillance compliance	-	N/A	NHS BCSP data
FS screening compliance	0.44	Beta(34,265,42,884) (0.44-0.45)	NHS BCSP data 2014/15
CTC follow-up compliance	0.99	Beta(2,731,22) (0.99-0.99)	Plumb et al 2013
Health-related quality of life parameters	*mean*	*Distribution used in PSA and 95% CI*	*source*
Utility value cancer free Age 30-34	0.91	Beta(28,945,2,823) (0.91-0.92)	Ara et al 2010.
Utility value cancer free Age 35-39	0.90	Beta(26,499,3,028) (0.90-0.90)	Ara et al 2010.
Utility value cancer free Age 40-44	0.88	Beta(22,447,2,961) (0.88-0.89)	Ara et al 2010.
Utility value cancer free Age 45-49	0.87	Beta(20,271,3,097) (0.87-0.87)	Ara et al 2010.
Utility value cancer free Age 50-54	0.85	Beta(19,605,3,439) (0.85-0.86)	Ara et al 2010.
Utility value cancer free Age 55-59	0.83	Beta(20,411,4,104) (0.83-0.84)	Ara et al 2010.
Utility value cancer free Age 60-64	0.81	Beta(21,437,4,932) (0.81-0.82)	Ara et al 2010.
Utility value cancer free Age 65-69	0.79	Beta(20,011,5,258) (0.79-0.80)	Ara et al 2010.
Utility value cancer free Age 70-74	0.77	Beta(15,018,4,496) (0.76-0.78)	Ara et al 2010.
Utility value cancer free Age 75-79	0.74	Beta(9,155,3,148) (0.74-0.75)	Ara et al 2010.
Utility value cancer free Age 80-84	0.72	Beta(5,144,2,008) (0.71-0.73)	Ara et al 2010.
Utility value cancer free Age 85+	0.68	Beta(2,896,1,334) (0.68-0.70)	Ara et al 2010.
Utility value CRC Age 30-34 Stage A-C	0.87	Beta(364,55) (0.84-0.90)	Ara et al 2010.
Utility value CRC Age 35-39 Stage A-C	0.85	Beta(391,67) (0.82-0.89)	Ara et al 2010.
Utility value CRC Age 40-44 Stage A-C	0.84	Beta(415,80) (0.81-0.87)	Ara et al 2010.
Utility value CRC Age 45-49 Stage A-C	0.82	Beta(440,96) (0.79-0.86)	Ara et al 2010.
Utility value CRC Age 50-54 Stage A-C	0.80	Beta(465,113) (0.77-0.84)	Ara et al 2010.
Utility value CRC Age 55-59 Stage A-C	0.79	Beta(490,134) (0.75-0.82)	Ara et al 2010.
Utility value CRC Age 60-64 Stage A-C	0.77	Beta(514,158) (0.73-0.80)	Ara et al 2010.
Utility value CRC Age 65-69 Stage A-C	0.74	Beta(535,184) (0.71-0.78)	Ara et al 2010.
Utility value CRC Age 70-74 Stage A-C	0.72	Beta(549,214) (0.69-0.75)	Ara et al 2010.
Utility value CRC Age 75-79 Stage A-C	0.69	Beta(552,243) (0.66-0.73)	Ara et al 2010.
Utility value CRC Age 80-84 Stage A-C	0.67	Beta(536,268) (0.63-0.70)	Ara et al 2010.
Utility value CRC Age 85+ Stage A-C	0.63	Beta(494,291) (0.60-0.67)	Ara et al 2010.
Utility value CRC Age 30-34 Stage D	0.67	Beta(913,451) (0.65-0.70)	Ara et al 2010.
Utility value CRC Age 35-39 Stage D	0.66	Beta(912,474) (0.63-0.68)	Ara et al 2010.
Utility value CRC Age 40-44 Stage D	0.65	Beta(909,497) (0.62-0.67)	Ara et al 2010.
Utility value CRC Age 45-49 Stage D	0.63	Beta(906,524) (0.61-0.66)	Ara et al 2010.
Utility value CRC Age 50-54 Stage D	0.62	Beta(903,553) (0.60-0.65)	Ara et al 2010.
Utility value CRC Age 55-59 Stage D	0.61	Beta(901,586) (0.58-0.63)	Ara et al 2010.
Utility value CRC Age 60-64 Stage D	0.59	Beta(898,624) (0.57-0.62)	Ara et al 2010.
Utility value CRC Age 65-69 Stage D	0.57	Beta(890,662) (0.55-0.60)	Ara et al 2010.
Utility value CRC Age 70-74 Stage D	0.56	Beta(873,700) (0.53-0.58)	Ara et al 2010.
Utility value CRC Age 75-79 Stage D	0.54	Beta(839,728) (0.51-0.56)	Ara et al 2010.
Utility value CRC Age 80-84 Stage D	0.51	Beta(781,738) (0.49-0.54)	Ara et al 2010.
Utility value CRC Age 85+ Stage D	0.49	Beta(687,728) (0.47-0.52)	Ara et al 2010.

1.6 Appendix: Model Natural History

Evidence suggests that most CRC develops from adenomas in the lining of the bowel which is known as the adenoma-carcinoma sequence [1]. Various approaches can be taken to model the development of adenomas and CRC. These include modelling: the growth of individual adenomas; the number/size/type/location of adenomas; an individual's progression from non-advanced to advanced adenomas; an individual's progression from low-risk to high-risk adenomas.

The natural history of CRC can be modelled using a patient-level or a cohort model [2, 3]. A patient-level simulation gives greater flexibility in modelling disease natural history and management, allowing, for instance, easier implementation of surveillance colonoscopy (as a patient's pathways will depend on their past surveillance results). A patient-level modelling approach will generally require more parameters and distributional assumptions than a cohort model. For example, a cohort modelling approach requires information on the average rate at which an adenoma would develop into a CRC, but a patient-level modelling approach would also require knowledge of the between-patient variation in this rate.

There is considerable uncertainty surrounding several of the natural history parameters such as adenoma growth rates. A cohort modelling approach was used in preference to a patient-level model in this instance to reduce the number of assumptions required and to ensure that there was sufficient data available to inform the model parameters. This choice was based on previous experience with both methods in modelling CRC. A state transition model was used to simulate the life experience of a cohort of 30 year old individuals in the general population of England with normal epithelium through to the development of adenomas and CRC and subsequent death.

Definition of health states

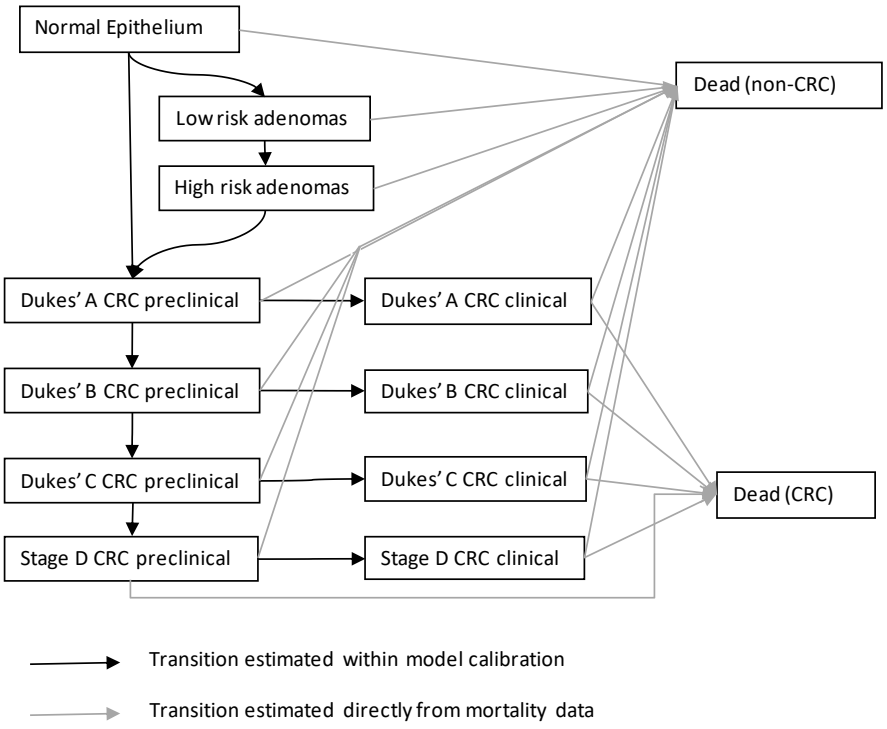
Health states were defined according to an individual's true underlying histological state. CRC was divided into eight health states which describe the Dukes' stages A-D and whether or not the CRC has been clinically diagnosed: preclinical/clinical.

Individuals with adenomas can be classified in many different ways to reflect the size, type, number and location of adenomas present, but it is important that the choice of adenoma health states reflects the data available to inform the model. The gFOBT screening programme in England records detection rates for "low-risk" and "intermediate/high-risk" adenomas as defined by the current British Society of Gastroenterology (BSG) guidelines for endoscopic surveillance following adenoma removal [4]. Detection rates from the FS screening trial which use this classification into "low-risk" and "intermediate/high-risk" adenomas were also obtained. The modelling uses this classification of adenomas to define two health states to describe individuals with adenomas. The "high risk adenomas" health state includes persons with at least 3 small adenomas or at least one adenoma of size >1cm (this includes the BSG intermediate and high risk surveillance categories). The "low-risk adenomas" health state includes persons with 1-2 small (<1cm) adenomas. These health states correspond to those used to determine an individual's surveillance strategy, so this approach eases the modelling of surveillance.

The model health states are: normal epithelium, low risk adenomas, high risk adenomas, preclinical CRC Dukes' stages A-D, clinical CRC Dukes' stages A-D, and dead. The health states and transitions included within the natural history model are shown in

Figure 3.

Figure 3: Diagram of Model Structure



Transition between health states

We define a sequence of annual transition probabilities between these states relating to CRC developing through the adenoma–carcinoma sequence, as this is thought to be the natural history of most CRC. In addition, we define a transition probability from normal epithelium to Dukes' A CRC to allow for the hypothesis that a proportion of cancers do not arise from adenomas (*de novo* cancers). For each cancer state we define the probability of being diagnosed through symptomatic presentation or chance detection, and this transition corresponds to moving from a preclinical to a clinical health state.

There is evidence to suggest that adenoma growth rate varies with age. Brenner et al. examined the results of 840,149 screening colonoscopies and found that the age gradient is much stronger for CRC incidence than for advanced adenoma prevalence, hence projected annual transition rates from advanced adenomas to CRC strongly increase with age [5]. The probability of developing a low risk adenoma, the transition probability from low to high risk adenoma, and the transition probability from high risk adenoma to Dukes' stage A CRC were allowed to vary by age using a piecewise linear model whose parameter values were the transition probabilities at ages 30, 50,70,100.

Transitions between the preclinical CRC states and from preclinical to clinical CRC are assumed to be independent of age. All persons may die of non-CRC causes, and this is modelled using age-specific mortality rates. Once a person is diagnosed with CRC, the transitions between Dukes' stages are no longer modelled and a stage-specific CRC relative survival rate is applied. In addition, preclinical stage D CRC may be fatal. Survival rates for clinical CRC stages A-D and preclinical stage D CRC is assumed to be dependent on the CRC stage at diagnosis and patient age.

Location of adenomas and cancer

Adenomas and CRC may develop in various locations within the colon and rectum. Little data was identified describing CRC/adenoma prevalence by location and age. A study by Yamaji et al. analysed the records of a colonoscopic follow-up study on 2900 subjects after polypectomy [6]. They describe the change in adenoma location by age: "Although there may be individual predilection for right-side or left-side location of colorectal adenomas, aging tends to increase the number of adenomas in the right-side colon, while only modestly affecting those in the left-side colon." We observed that the proportion of persons who only had adenomas in the proximal colon did not vary significantly by age; see Table 8.

Table 8: Location of adenomas by age as reported by Yamaji et al. 2007

Age group	Adenomas located only in the left side colon and rectum	Adenomas located in both the left side and the right-side colon	Adenomas located only in the right-side colon
<40	59%	12%	30%
40-49	56%	15%	29%
50-59	43%	24%	34%
>=60	37%	34%	29%

Table 9 shows incidence of cancer in the proximal and distal colon by age for newly diagnosed cases in England in 2007. Of diagnosed cases of CRC with known location, 62% are located in the distal colon and 38% in the proximal colon. Distal and proximal CRC may be associated with different likelihoods of displaying symptoms and receiving a diagnosis. Hence the difference in incidence between the proximal and distal colon is unlikely to accurately reflect the difference in prevalence between the distal and proximal colon.

Table 9: CRC by age and location, registrations of newly diagnosed cases 2007

	Incidence Rates per 100,000 population						CRC with known location	
Age range	Proximal Colon (C18.0-C18.6)		Distal Colon (C18.7,C18.8,C19,C20)		Unknown location (C18.9)		Proximal	Distal
30-34	1.2	41%	1.4	51%	0.2	8%	45%	55%
35-39	2.1	39%	2.9	53%	0.4	8%	42%	58%
40-44	3.6	36%	5.6	56%	0.8	8%	39%	61%
45-49	5.5	29%	12.2	64%	1.4	7%	31%	69%
50-54	10.2	27%	25.6	67%	2.6	7%	29%	71%
55-59	18.2	27%	44.7	66%	5.3	8%	29%	71%
60-64	36.5	31%	70.8	61%	9.0	8%	34%	66%
65-69	57.9	31%	112.0	61%	15.1	8%	34%	66%
70-74	79.0	33%	143.4	59%	20.0	8%	36%	64%
75-79	115.8	37%	166.8	54%	28.8	9%	41%	59%
80-84	149.9	40%	181.1	49%	40.8	11%	45%	55%
85 and ove	140.4	39%	165.6	46%	55.4	15%	46%	54%
All ages	20.7	34%	33.7	56%	5.7	10%	38%	62%

Screening test sensitivity by location

The sensitivity of a screening test may vary between the distal and the proximal colon. This gives two important considerations for the modelling of screening. Firstly, as CRC/adenoma location distributions vary by age, it follows that the overall sensitivity of a screening test may vary by age. Secondly, a screening test with significantly different proximal and distal sensitivity will impact the location distribution for remaining undetected CRC and adenomas. This in turn will impact on the detection rates seen at subsequent screens. Hence, adenoma/CRC location distribution and screening test sensitivity by location may be important considerations when modelling combined or repeated screening strategies.

The extent to which the CRC sensitivity of a screening test varies between the distal/proximal colon can be estimated by comparing the location distribution of screen detected CRC with that of prevalent CRC. As no data on the location distribution of prevalent CRC was available, data on the location distribution of CRC incidence was used. The use of incidence as a proxy for prevalence will introduce errors, as symptoms and diagnosis rates will vary by location. Hence this calculation is simply a crude estimate for illustrative purposes. Location specific sensitivities for CRC are estimated in

Table 10.

Table 10: Screening test CRC detection by location

Screening	Screen detected CRC Distal	Screen detected CRC Proximal	Age group screened	Proportion of CRC incidence in distal colon for age group	Sensitivity to CRC Overall	Sensitivity to CRC Distal*	Sensitivity to CRC Proximal*
gFOBT BCSP data	72%	28%	60-69	66%	0.24	0.26	0.20
FS trial data	90%	10%	55-64	69%	0.62	0.81	0.20

*Formulae used in calculation: overall sensitivity = proportion distal * distal sensitivity + proportion proximal * proximal sensitivity

In the England gFOBT screening programme, 72% of CRC detected (with a known location) was found in the distal colon, compared to 66% of CRC incidence which is distal for this age group [7]. Using this data we estimate that gFOBT has very similar sensitivity in the distal and proximal colon.

Flexible sigmoidoscopy examines the distal colon only; however, a participant may be referred to colonoscopy following FS and colonoscopy may find lesions in both the proximal and distal colon. In the UK flexible sigmoidoscopy trial, 90% of all CRC detected at screening was found in the distal colon, compared to 69% of CRC incidence which is distal for this age group. This implies a significant difference between distal and proximal sensitivity which corresponds with the nature of the test. A FS CRC sensitivity of 20% for the proximal colon implies that 20% of proximal CRC was associated with a distal adenoma which required referral to colonoscopy.

Sensitivity at repeat screens

The estimated location specific test sensitivities were used to examine the degree to which the overall sensitivity to CRC may vary between a first and a repeat screen. An initial distal:proximal CRC split of 70:30 was assumed, and calculation details are presented in Table 11. This calculation estimated the maximum possible change in overall sensitivity, as it assumes that the CRC location distribution does not change in the time after the first screen to before the repeat screen. The gFOBT overall sensitivity to CRC did not vary significantly by first/repeat screen; however, FS overall sensitivity to CRC may be reduced to as little at 0.42 for a repeat screen. Hence modelling varying FS sensitivity by first/repeat screen is important for a strategy involving two or more FS screens. This estimate of minimum FS overall sensitivity to CRC for a repeat FS screen is used within a sensitivity analysis.

Table 11: Estimated overall sensitivity at first/repeat screen incorporating location-specific sensitivities

		CRC location distribution Distal	CRC location distribution Proximal	Sensitivity to CRC Distal	Sensitivity to CRC Proximal	Sensitivity to CRC Overall
gFOBT	First screen	0.70	0.30	0.26	0.20	0.24
	Repeat screen	0.68	0.32	0.26	0.20	0.24
FS	First screen	0.70	0.30	0.81	0.20	0.63
	Repeat screen	0.36	0.64	0.81	0.20	0.42

Data on detection rates in the distal/proximal colon for FIT is not available, so no conclusions can be reached on the sensitivity in the proximal and distal locations.

Location-specific sensitivity to adenomas

Data on the location of adenomas is very complex to report. The definition used for high risk adenomas (or advanced adenomas) refers to the whole colon. An individual will often have adenomas in both the proximal and distal colon, and it may be the combination of these that determines the risk level.

Yamaji et al. found that the proportion of persons who only had adenomas in the proximal colon did not vary significantly by age; see Table 8 [6]. Hence, even though the sensitivity of FS varies significantly between the proximal and the distal colon, this suggests that the overall sensitivity of FS may not significantly vary by age.

Data from gFOBT screening showed a significantly lower HR adenoma detection rate at the repeat screen. This may suggest that the location specific variation in gFOBT HR adenoma sensitivity is significant. However, data on HR adenoma prevalence by location is not available, so this remains an area requiring further research.

Metachronous adenomas – adenoma recurrence rates post-polypectomy

The model uses data on the risk of recurrence of adenomas in persons who have had adenomas removed by polypectomy and are undergoing surveillance. To ensure consistency between the model parameters, it is important that the post-polypectomy transition probabilities used align with the other natural history transition probabilities in the model. We assume that persons who are undergoing surveillance post-polypectomy are at higher risk of developing adenomas than persons with a normal epithelium. We also assume that polypectomy reduces the risk of developing CRC. Hence we place restrictions on the post-polypectomy transition probabilities as described in Table 12.

Table 12: Restrictions on transition probabilities post-polypectomy

Restrictions on transition probabilities post polypectomy
Post polypectomy(LR) to LR adenoma **>** Normal epithelium to LR adenoma
Post polypectomy(HR) to LR adenoma **>** Normal epithelium to LR adenoma
Post polypectomy(LR) to HR adenoma **<** LR adenoma to HR adenoma **>** Normal epithelium to HR adenoma
Post polypectomy(HR) to HR adenoma **>** Normal epithelium to HR adenoma
Post polypectomy(LR) to CRC **<** LR adenoma to CRC **>** Normal epithelium to CRC
Post polypectomy(HR) to CRC **<** HR adenoma to CRC **>** Normal epithelium to CRC
Post polypectomy(LR) to LR adenoma< Post polypectomy(HR) to LR adenoma
Post polypectomy(LR) to HR adenoma< Post polypectomy(HR) to HR adenoma
Post polypectomy(LR) to CRC adenoma< Post polypectomy(HR) to CRC adenoma

Data on the surveillance results from the England gFOBT BCSP details over 4000 surveillance colonoscopies (Table 13) [7]. Unfortunately, data which details the results of 1 and 3 year (IR/HR) surveillance separately is not currently available, so some assumptions had to be made.

Table 13: Detection rates at surveillance in the England gFOBT screening programme

	Detection rates at surveillance		Estimated annual reccurrence rate *	
Find	Surveillance undertaken in 2008 assumed to be 1-year surveillance)	Surveillance (undertaken in 2010 assumed to be mainly 3-year surveillance)	Persons undergoing 3-yearly surveillance	Persons undergoing 1-year surveillance
CRC	1.3%	0.3%	0.1%	1.3%
HR adenomas	55.7%	24.4%	9.1%	56.8%
LR adenomas	14.5%	31.9%	16.3%	18.8%

*Estimated annual recurrence rates were calculated by adjusting for the number of years until surveillance and colonoscopy miss rates.

There is currently no data available of recurrence rates for persons with LR adenomas who do not receive surveillance in the English BCSP.

Martinez et al. report a pooled analysis of individual data from 8 prospective studies comprising 9167 men and women aged 22 to 80 with previously resected colorectal adenomas to quantify their risk of developing subsequent advanced adenoma or cancer, as well as identify factors associated with the development of advanced colorectal neoplasms during surveillance [8]. Risk of new neoplasia at follow-up evaluation is estimated according to baseline adenoma

characteristics. Data from the Martinez study was converted into annual transition probabilities assuming a follow-up period of 4 years; see Table 14. It should be noted that the definitions of low and high risk used in the Martinez study differs slightly from the definitions used in the BSG surveillance guidelines; however, the Martinez study was still deemed to be the best available data source.

Table 14: Data from Martinez et al. 2009

	Risk of new neoplasia at follow-up evaluation (median duration of follow up 47.2 months)			Annual transition probabilities (assuming a follow-up of 4 years)		
Adenoma history*	Non advanced adenoma	Advanced adenoma**	Colorectal cancer	Non advanced adenoma	Advanced adenoma	Colorectal cancer
Low-risk	0.345 (0.331,0.358)	0.069 (0.062,0.076)	0.005 (0.003,0.007)	10.0%	1.8%	0.1%
High-risk	0.353 (0.339,0.367)	0.155 (0.145,0.166)	0.008 (0.005,0.01)	10.3%	4.1%	0.2%

*The low-risk group includes patients with 1–2 small (<1 cm), tubular adenoma(s) with low-grade dysplasia.

**Advanced adenoma are defined as those with a diameter 10mm or larger, having greater than 25% villous

The model uses recurrence rates from the English BCSP for persons with HR adenomas and data from Martinez et al. for persons with LR adenomas. This data on recurrence rates post-polypectomy has several limitations. The transition probabilities reported are not age-dependent; however, the transition probabilities used in the model are age-dependent. The study populations do not reflect the English screening population, are quite small in size, do not use the BSG surveillance guidelines to categorise adenomas, and report highly varying recurrence rates. It is very important that detailed data on outcomes at surveillance in the English gFOBT screening programme is collected and available for future modelling work to improve the accuracy of decision support for the screening programmes.

Classification of adenomas

Adenomas can be categorised in the following ways: by size: <5mm, 5-10mm, 10-20mm, 20+mm and by type: tubulovillous/villous (>25% villous features), advanced/non-advanced, high grade dysplasia. In addition, persons can be classified by number of adenomas present or by BSG surveillance guidelines risk level: low/intermediate/high.

The majority of the colonoscopy studies identified in the systematic review classify adenomas as advanced or non-advanced. As the definition of “advanced adenoma” includes tubulovillous or villous adenomas, it will include some individuals who would be classified as low-risk according the BSG guidelines. There will also be some individuals with 3-4 small adenomas who are classified as intermediate risk according the BSG guidelines but who do not have advanced adenoma. Out of persons found to have an advanced adenoma in the FS trial, 74% were classified as intermediate or high risk according to the BSG guidelines. Hence it was assumed that 74% of persons with advanced adenoma had high-risk adenomas.

Table 15: Classification of persons with adenomas

	BSG surveillence guidelines			Definition used in Brenner et al.	Model health states	
	low risk	intermediate risk	high risk	Advanced adenoma	low risk adenomas	high risk adenomas
1-2 small (<10mm) adenomas	X				X	
3-4 small (<10mm) adenomas		X				X
large (<=10mm) adenoma		X		X		X
5+ small (<10mm) adenomas			X			X
3+ adenomas at least one of which is >=10mm			X			X
high grade dysplasia				X		X
1-2 small (<10mm) tubulovillous or villous adenoma	X			X	X	
3-4 small (<10mm) tubulovillous or villous adenoma		X		X		X
5+ small (<10mm) tubulovillous or villous adenoma			X	X		X

Data from the gFOBT screening programme in England reports detection rates of low/intermediate/high-risk adenomas (according to BSG guidelines), and this classification is used to determine an individual’s surveillance [7]. Data from FIT screening in Italy and colonoscopy screening in Germany reports detection rates for “advanced adenomas”. There is great value to be had in using all of these data sources, as they provide valuable information regarding the different screening modalities. The differences in the reporting of adenoma detection rates are problematic and introduce great uncertainty into the modelling. An internationally consistent way of reporting adenoma findings from screening programmes and trials should be a priority for the future.

Adenoma and CRC prevalence in an asymptomatic population

Data on the prevalence of CRC and adenomas by age in a screening population (asymptomatic) was required to inform the CRC natural history model. Such data are available from autopsy studies and can also be estimated from colonoscopy screening studies. A systematic review of data from colonoscopy studies in an asymptomatic population and autopsy studies was undertaken. Studies which report adenoma detection/prevalence rates by age were identified. Full details of the systematic review are included in Appendix 1 of the 2011 reappraisal report [9].

Colonoscopy studies provide data on adenoma prevalence but as colonoscopy is not a perfect test some adenomas (in particular small adenomas) may be missed. Adenoma prevalence estimates from colonoscopy screening studies may also be biased as they consist of a population who attend screening which is likely to differ slightly to the general population. The systematic review identified eight colonoscopy studies which are described in Table 16; the largest of which described the results of over 2 million colonoscopies from the German screening programme [10]. For the model calibration data the study by Brenner et al. was selected due to the large sample sizes, broad age range, and the expected similarity between the German and English screening populations. To incorporate some data on LR adenomas (not reported by Brenner et al.) and some information for persons aged under 60, data from Chung et al. 2010 was also included [11]. Figure 4 and Figure 5 present data on advanced adenoma prevalence by age from colonoscopy studies identified by the systematic review.

Table 16: Summary of colonoscopy study characteristics

Study	Data Collected (Time-Interval)	Country of study	Sample Size	Age Range	Included within study definition of advanced adenoma: adenoma >=1cm in size/diameter	adenoma containing villous features />= 25% villous features	adenoma with high grade dysplasia	adenoma with malignant features	adenoma with carcinoma in situ
Rundle et al (2008)	2004 - 2006	United States	905	40-59	Y	Y	Y		
Lin et al (2006)	2002 - 2005	United States	1244	>= 50	Y	Y			
Strull et al (2006)	1996 - 2003	Israel	1177	40-80	Y	Y	Y		
Soon et al (2005)	2002 - 2004	United States Taiwan	3403 1456	40-70 40-70	Y	Y	Y	Y	
Yamaji et al (2004)	1988 - 2002	Japan	4084	all ages	Y		Y		Y
Chung et al (2010)	2004 - 2007	Korea	5254	30-59	Y	Y	Y		
Brenner et al (2007)	2003 - 2004	Germany	840,149	50-80+	Y	Y	Y		
Brenner et al (2010)	2003 - 2007	Germany	2,185,153	50-75	Y	Y	Y		
Choe et al (2007)	1998 - 2004	Korea	5086	>=20	Y	Y	Y		Y

Figure 4: Advanced adenoma detection rates in colonoscopy studies identified by systematic review

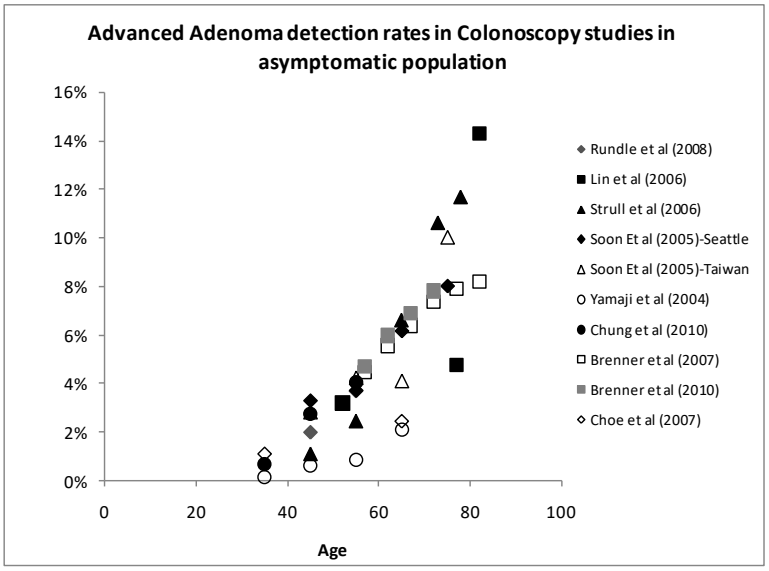
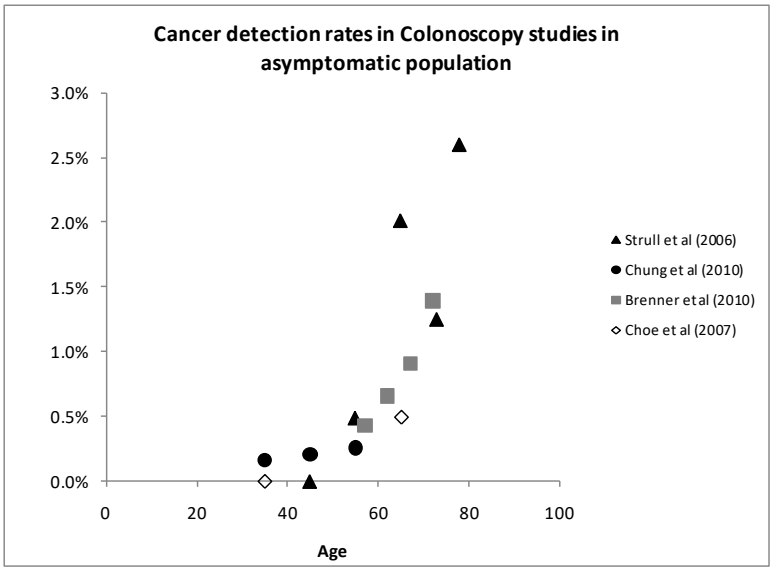


Figure 5: Advanced adenoma detection rates in colonoscopy studies identified by systematic review

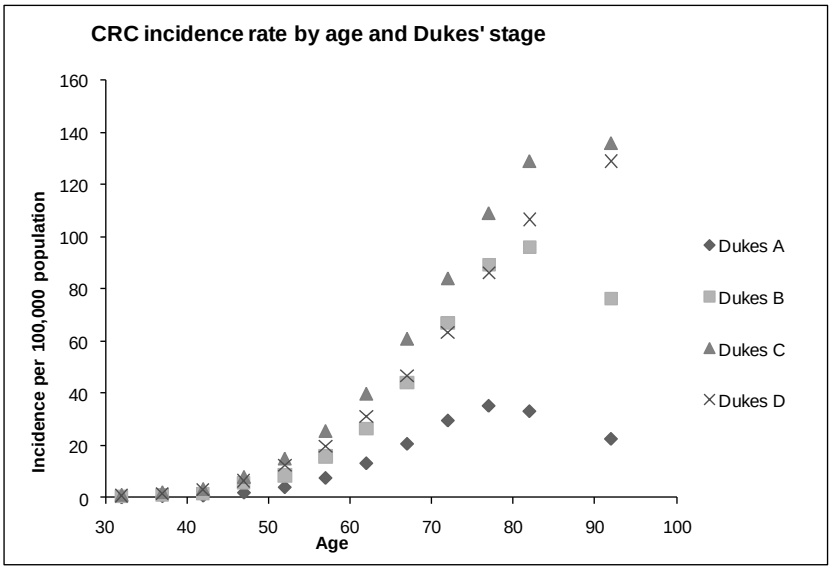


Autopsy studies allow a complete and thorough examination of the colon and rectum; however, data from autopsy studies may be biased, as autopsied individuals represent a biased sample of deaths. In addition, autopsy studies do not always include an equal cross-section of ages. Due to the large amount of heterogeneity in the autopsy studies and the small sample sizes when compared to colonoscopy studies, the autopsy study data was not used within the model calibration.

Colorectal cancer incidence in the absence of screening by age and stage

Data on CRC incidence in the absence of screening categorised by age and Dukes' stage at diagnosis was taken from England cancer registry data for Oxford, Northern and Yorkshire, and Eastern regions from 2004 – 2006 (personal communication from Northern & Yorkshire Cancer Registry & Information Service and the National Cancer Intelligence Network).

Figure 6: CRC incidence rates in the absence of screening by age and Duke' stage



Screening programme data

Observed data from existing screening programmes and screening trials was used within the calibration of the model. The screening detection rates are essential to estimate the sensitivities of the screening tests while the false positive rates inform screening test specificity. Note that we define the false positive rate to be the proportion of persons undergoing colonoscopy following FOBT in whom no CRC or adenomas were found at colonoscopy. The change in screening positivity and detection rates by age provide important information for the natural history model, i.e. the change in underlying adenoma and CRC prevalence by age.

Table 17 provides a summary of the screening data used within the model calibration. The current gFOBT BCSP in England reported numbers of persons with positive gFOBT result and the detection rates of low and high risk adenomas and CRC at screening. Data from the FS trial consisted of detection rates of CRC, low/high risk adenomas and non-advanced/advanced adenomas at screening [12]. As UK data is only available for the gFOBT and FS, screening test data from Italy was used for FIT screening.

The population of the FS trial differed slightly from a screening population, as all persons had indicated that they were interested in attending screening in the questionnaire. The screening data used in the calibration relates to persons who attended screening. Screening attenders in the FS trial may be slightly healthier than those undergoing gFOBT screening, hence they may have slightly lower detection rates at FS screening leading to a slightly lower estimate of FS sensitivity, thus biasing the result slightly in the favour of FOBT. This slight difference between the screening populations is not expected to significantly bias the model results. In fact, an analysis demonstrated that the FS trial control population had lower mortality rates than Norwegian control but incidence was the same.

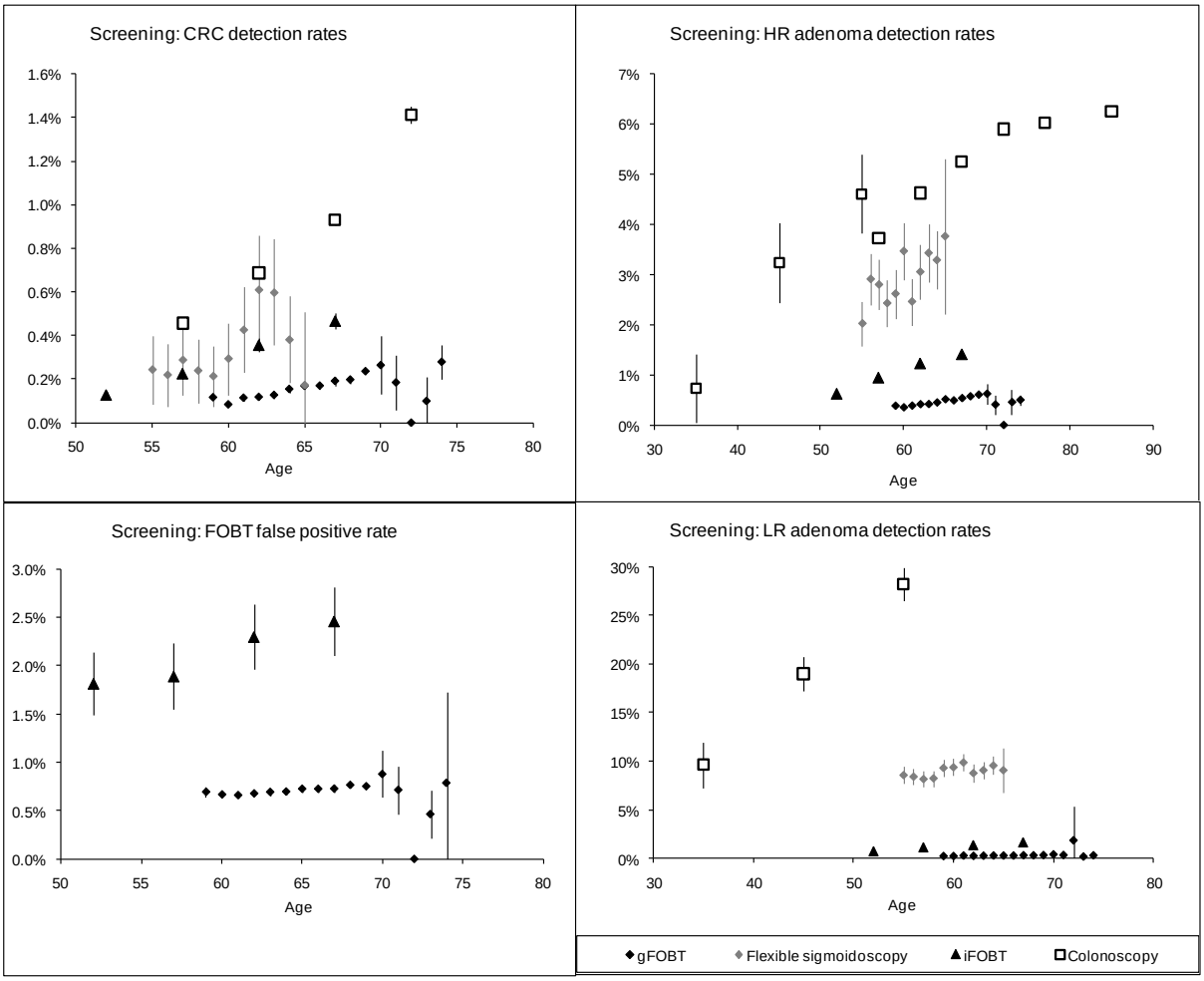
Colonoscopy screening is not considered in this evaluation; however, data from screening colonoscopies is of particular use for calibrating the model because of the accuracy of colonoscopy. As mentioned earlier, colonoscopy screening data was used in preference to autopsy study data as the sample sizes are much larger.

Figure 7 presents the screening data which was used within the calibration process. The higher detection rates seen at FS screening indicate that FS is much more sensitive than gFOBT.

Table 17: Screening data used within model calibration

Screening test	Source	Country	Time period screening undertaken	Number of particpants undergoing screening	Age range of participants	Data reported
gFOBT	England BCSP	England	2006-2010	2,889,925	59-74	false positive rate; detection rates for LR adenomas, HR adenomas and CRC
iFOBT	Zorzi et al	Italy	2006-2010	591,152	50-69	false positive rate; detection rates for non-advanced adenomas, advanced adenomas and CRC
FS	Atkin et al	England	2005-2008	40,621	55-65	detection rates for LR adenomas, HR adenomas and CRC
Colonoscopy	Brenner et al	Germany	2003-2007	2,185,153	55-75	detection rates for advanced adenomas and CRC
Colonoscopy	Brenner et al	Germany	2003-2004	840,149	50-80+	detection rates for advanced adenomas
Colonoscopy	Chung et al		2003-2007	5,254	30-59	detection rates non-advanced adenomas, advanced adenomas and CRC

Figure 7: CRC and adenoma detection rates at screening and FOBT false positive rates with 95% confidence intervals presented as vertical lines



International variation in CRC and adenoma prevalence

There exists data describing the international differences in the incidence of CRC, however, there is little evidence describing the difference in the prevalence of CRC and adenomas. Soon et al. undertook a study in which a cohort of patients in both Taiwan and Seattle received colonoscopy [13]. They concluded that "compared to Westerners, Chinese patients have a slightly lower prevalence of colon neoplasia (but not advanced neoplasia), more distal distribution of neoplasia, and higher likelihood of concomitant proximal advanced neoplasia and distal neoplasia."

Differences in adenoma and CRC prevalence between England, Germany and Italy may exist; however, the extent of these differences is unknown. The value of using data from more than one country is that it allows the use of large datasets from several different screening modalities. The benefit of including data on different screening modalities was considered to outweigh the uncertainty introduced by using datasets from different countries.

Natural history model calibration method

Model calibration used the methods described by Whyte et al. [3]. For a given parameter set, the model can be run to produce predictions of CRC incidence, adenoma prevalence and screening outcomes. The aim of the calibration is to obtain parameter sets whose predictions are close to the observed data. For each data set, the sum squared error (SSE) was calculated by comparing the observed number of observations to the predicted number of observations for each age. The total SSE is a measure of how well the model fits to all the observed data sets. The aim of the calibration is to obtain multiple parameter sets which each produces a model that has a good fit to the observed data sets (determined by consideration of total SSE).

The Metropolis Hastings (MH) algorithm was used for the calibration process to generate multiple sets of parameters [14]. These parameter sets form the posterior distribution which is compatible with the observed data, accurately representing parameter uncertainty. This approach embeds the problem in the framework of Bayesian inference and produces correlated parameter sets which can be used for probabilistic sensitivity analyses (PSA). Correct representation of the joint uncertainty in these parameters is particularly important because of the potential for correlation between several of these parameters.

The model calibration was run eight times using different sets (randomly generated) of initial parameter values to ensure that the best fitting parameter set was obtained. Each run consisted of 50,000 iterations of the MH algorithm and could be run overnight on a standard PC. A sample of 250 parameter sets from after convergence from four of the runs were combined to form 1000 parameters sets to be used to run the PSA.

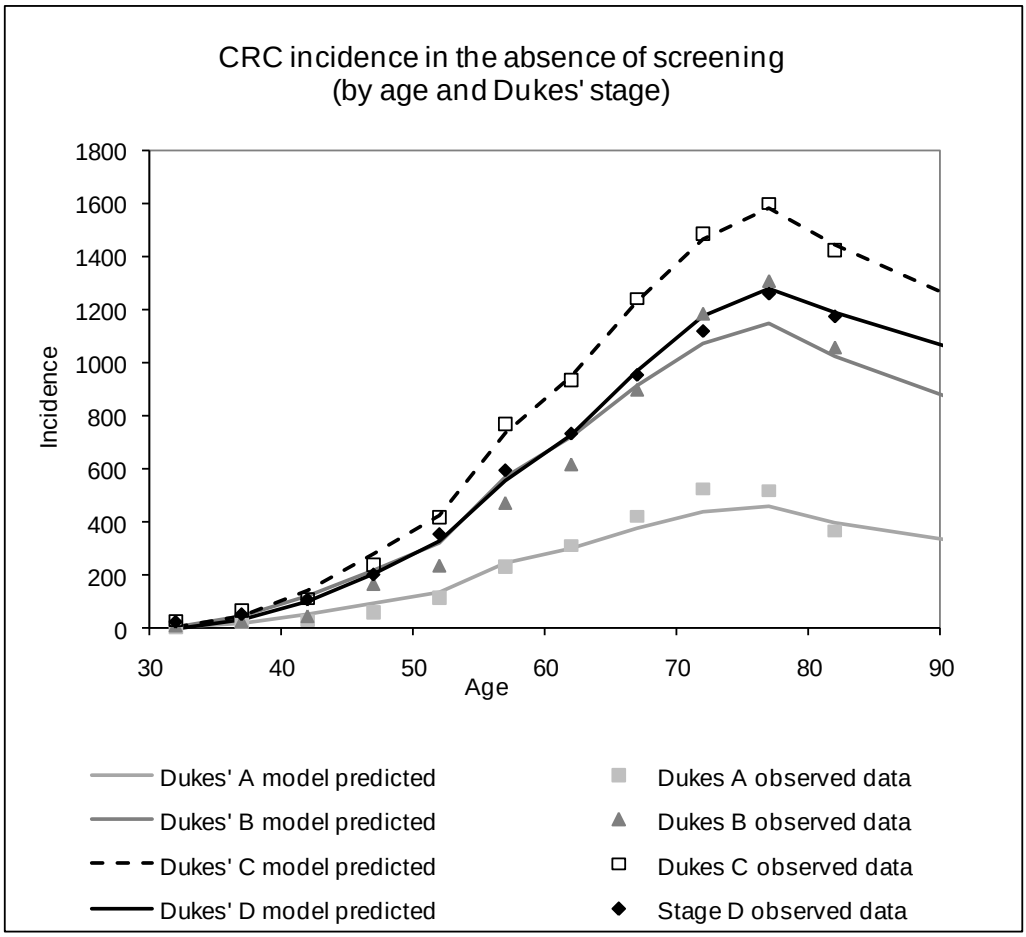
A large number of parameters were being estimated within the calibration process, which can lead to low acceptance rates and slow convergence. Hence an approach was implemented in which there was a random 30% probability that a given parameter was varied on each run, and this increased acceptance rates and time to convergence.

Model calibration results

Figure 8 shows the model predictions compared to the observed data for the best fitting parameter set resulting from the calibration process. The model obtained a good fit to the observed data on CRC incidence in the absence of screening.

The best fitting parameter set and 95% percentiles are presented in Figure 8. The 95% percentiles demonstrate that there are varying degrees of uncertainty surrounding the different parameter values. For example, there is considerable uncertainty surrounding the FS CRC sensitivity value, as the sample sizes are quite small for the CRC detection rates at FS screening. We note that although the CRC sensitivity estimates for FS and FIT were similar, FS has higher detection rates because it is associated with a higher rate of compliance with follow-up colonoscopy.

Figure 8: Model predictions compared to observed data for CRC incidence in the absence of screening



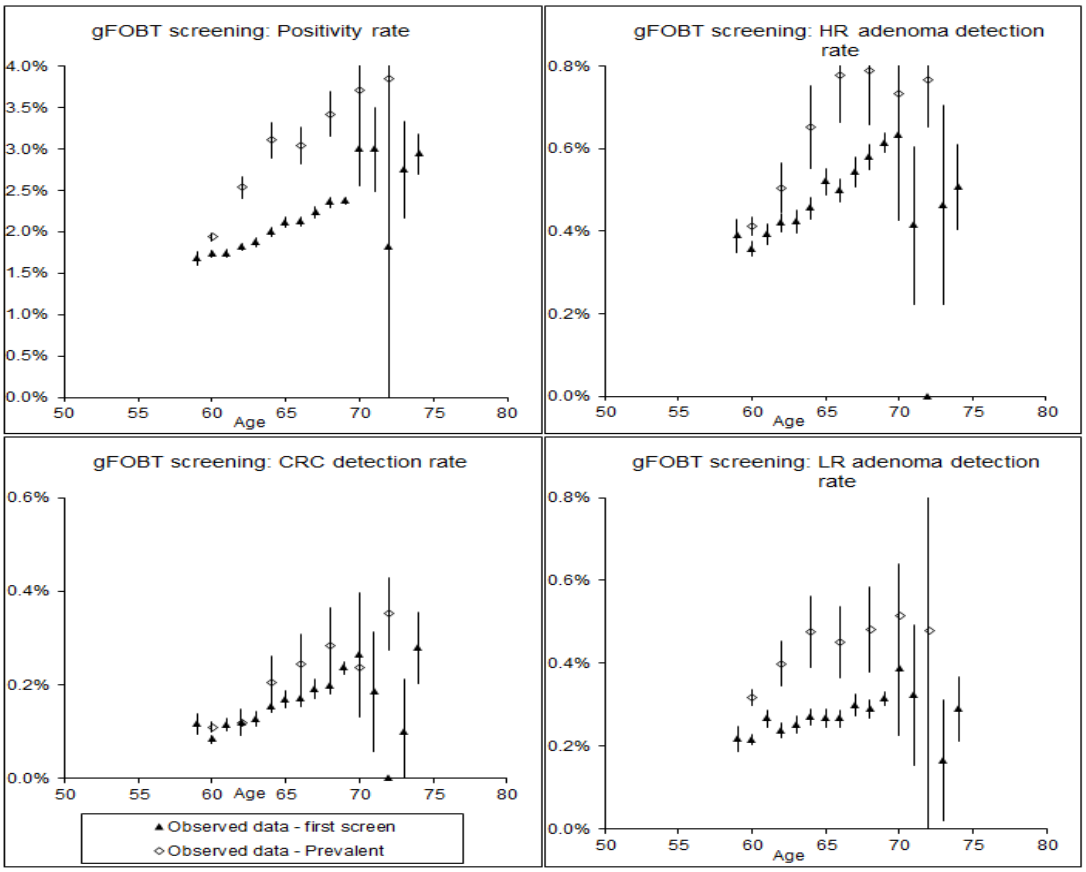
1.7 Appendix: BCSP gFOBT data

The existing model calibration uses BCSP data from the first and second rounds of gFOBT screening (2006-2010) to inform gFOBT screening characteristics. However, although BCSP gFOBT data is available for 2014/15 (presented in section 3.1) there is an issue with updating the model to use the more recent screening data. Current BCSP data is segregated by prevalent and incident screening episodes and is based on a slightly different screening population from the first and second round data used in previous model calibrations. The first round of BCSP screening represents the first time that anyone has been invited to screening (no matter what age), whilst prevalent screening represents invitations to individuals who have never before been screened, but who may have previously been invited. These populations are identical aged 60, but for older age groups, prevalent screening predominantly identifies those who have previously been invited and refused. As discussed previously, those who refuse screening are likely to be at higher risk of CRC than those who have never previously been invited.

This is illustrated in

Figure 9, which compares first round and prevalent data. Higher positivity and detection rates are observed in the recent prevalent data than in the old first round data, and whilst the difference is small at age 60, it increases with age. For example, the positivity rate in 68 year olds invited for the first time was 2.4% compared to 3.4% for 68 year olds who were invited for screening at ages 60, 62, 64, and 66 but did not attend. Individual screening history cannot currently be incorporated due to the cohort nature of the model, but this is something that could be included in phase II of the project, therefore allowing natural history modelling to be based on more recent prevalent data. The data for age 60 does indicate that positivity and detection rates have increased slightly between 2006/7 and 2014/15, perhaps due to improved colonoscopy quality. This indicates that the current version of the model may slightly underestimate the benefits of gFOBT screening.

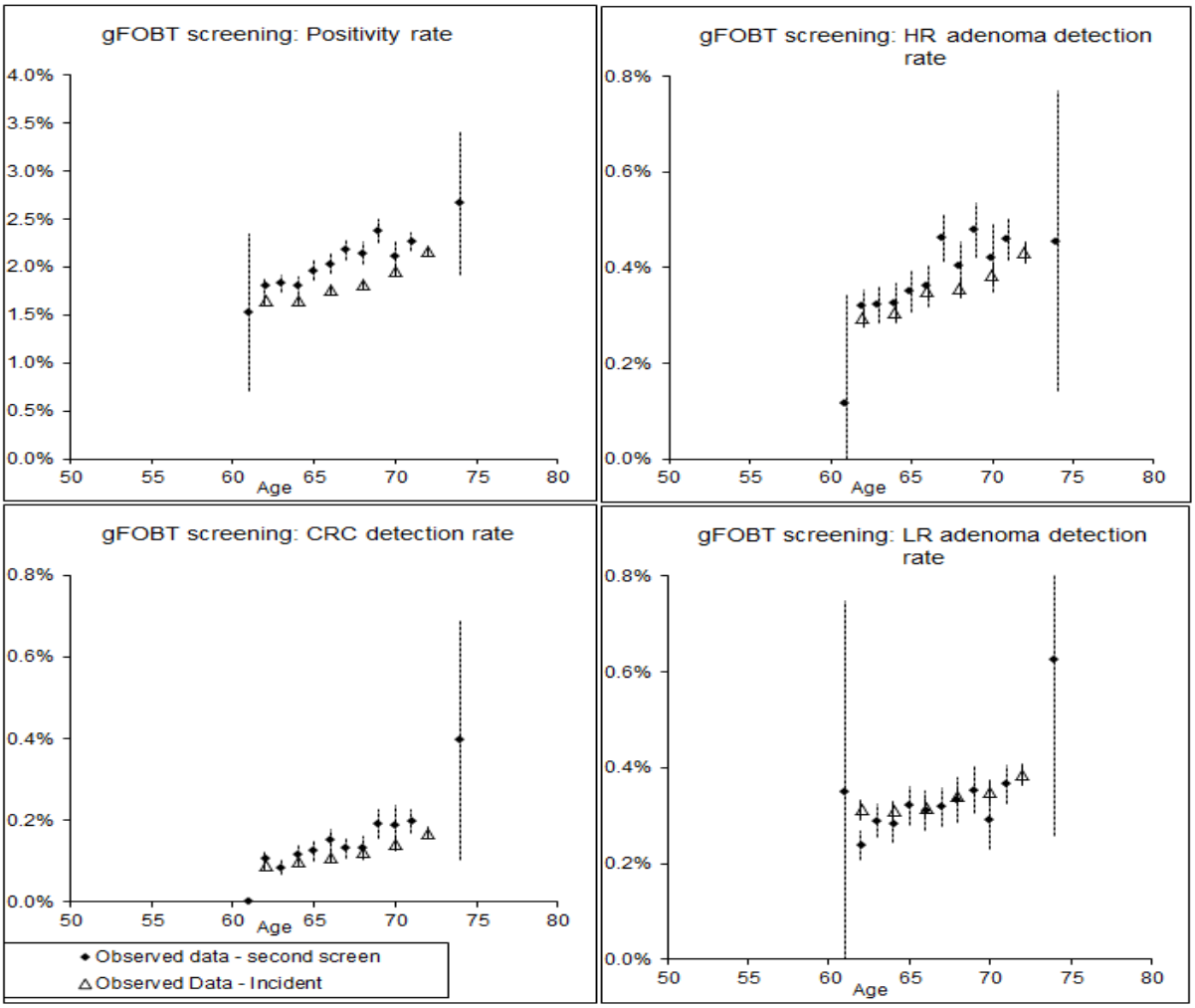
Figure 9: Comparison of BCSP first screen (2006/7) and prevalent (2014/15) data. Vertical lines represent 95% confidence intervals.



Similarly, second round and incident data do not represent the same population, with second round data including some prevalent screening (those who were invited but refused the first round) plus second invitation incident

screening, whereas the 2014/15 incident data includes all incident screening no matter how many times individuals have been previously screened. Given that the more frequently someone is screened the less likely they are to have CRC or adenomas (as those with CRC or adenomas are removed from the eligible screening pool in the previous rounds), it is unsurprising that the data indicates that incident positivity and detection rates (with the exception of detection of low risk adenomas) are lower than second round values (Figure 10).

Figure 10: Comparison of BCSP second round (2008/9) and incident (2014/15) data. Vertical lines represent 95% confidence intervals.



Given that screening is now underway, it is not possible to update first and second round data with equivalent but more recent data. The calibration model was therefore not updated with more recent BCSP data.

FIT and bowel scope test characteristics

The estimation of these has been updated using FIT pilot data and NHSBCSP data on bowel scope. These are now estimated outside of the calibration process.

1.8 Appendix: Utility values

A utility value is a preference weight reflecting the relative value that individuals place on different health states. Here different utility values are used for persons with CRC and for persons without CRC. NICE recommends that utilities should be based upon public preferences (e.g. EQ-5D values) and valued by patients [15]. Given that the focus of the model is comparison of screening strategies at different ages, and that screening may result in earlier detection of CRC (which may also be at an earlier stage) it is also particularly important that age and stage dependent utilities are included in the model. As one of the screening strategies considered is that of not screening, it is also important that a comparison with the health-related quality of life of the general population is considered.

A recent systematic review and meta-analysis of utilities in CRC used data from 26 articles to construct a linear fixed effects model to predict patient utility depending upon cancer type (colon, rectal or colorectal), stage, time from surgery, utility measurement instrument and survey administration method [16]. The fixed effects model estimates a utility of 0.83 for stage I-III CRC, where time from surgery is not indicated using EQ-5D as the utility measurement instrument; a utility of 0.64 for the equivalent stage IV CRC, and a utility of 0.77 for mixed stages. However, the analysis does not give a value for persons without CRC for comparison, nor does it state the mean age of individuals with CRC in the included studies to allow comparison against general population age-related EQ-5D estimates. If it is assumed that the mean age of individuals in the included studies corresponds to 71, which is the mean age of individuals at diagnosis of CRC [17] then the utility of 0.83 for stage I-III CRC is actually considerably higher than the mean utility of the general population at this age (0.78).

Another study measuring EQ-5D in CRC patients has also found health related quality of life to be slightly higher in CRC patients undergoing rehabilitation or remission than in the age standardised population [18]. Small but non-significant utility decrements were found for patients undergoing primary treatment (-0.033) or with metastatic disease (-0.005), with the only significant negative effects found in patients undergoing palliative care (-0.119), confirming that utility is lower in patients suffering from the most severe stages of the disease.

For this study, pooled data from the annual Health Survey for England (HSE) was used to estimate the impact of age and having cancer on utility, as measured using the EQ-5D-3L (hereafter the EQ-5D) [19]. This data is limited by the fact that the health survey for England does not include persons in hospital or in nursing home, and relies on self-report of long-standing disease. Cancer is one of the health conditions represented, but there is no subdivision by cancer type, stage or whether individuals are in remission.

Data used

Data on EQ-5D were collected in the HSEs from 2003 to 2014 inclusive, with the exception of the surveys conducted in 2007, 2009 and 2013. Hence a total of nine years of HSE data were available for analysis. The data were obtained from the UK Data Service [19]. Of the nine surveys with EQ-5D data, mean (index) EQ-5D scores were available for all of the surveys apart from 2010 and 2011, for which EQ-5D responses were only available for individual domains. These were combined to estimate the mean EQ-5D using UK population preferences [20]. This method was validated using the 2012 HSE data as this included both EQ-5D scores and responses for individual domains

Statistical analyses

Both descriptive analyses and linear regression were used to estimate the impact of age and cancer status on mean EQ-5D score. Gender was also included in the analyses, as this is known to be associated with EQ-5D. All analyses were performed using STATA version 14.1.

The descriptive analyses undertaken included univariate analyses of the differences in EQ-5D score, age, and gender by cancer status. For continuous variables (EQ-5D score and age) associations were tested for statistical significance using t-tests, for gender Fisher's exact test was used. Values of less than 0.05 were used to indicate statistical significance, although findings were treated with caution due to the univariate, exploratory aspect of the descriptive analyses. The association of EQ-5D score with both cancer status and age was also visually displayed using on local polynomial smoothing.

Multivariable linear regression modelling was using to estimate the impact of age and cancer status on EQ-5D score, whilst controlling for differences in gender. Interactions between age and cancer status were also considered. The functional form to use for age was based on the results of multivariable fractional polynomial modelling. This considers up to two fractional polynomial terms for age, for each term the potential powers considered are (−2, -1, -½, 0, ½, 1, 2,

3), or the logarithm of age may be chosen. The number of terms, and the powers for each term, are based on minimising the deviance. For comparison, the model which used age untransformed was also considered.

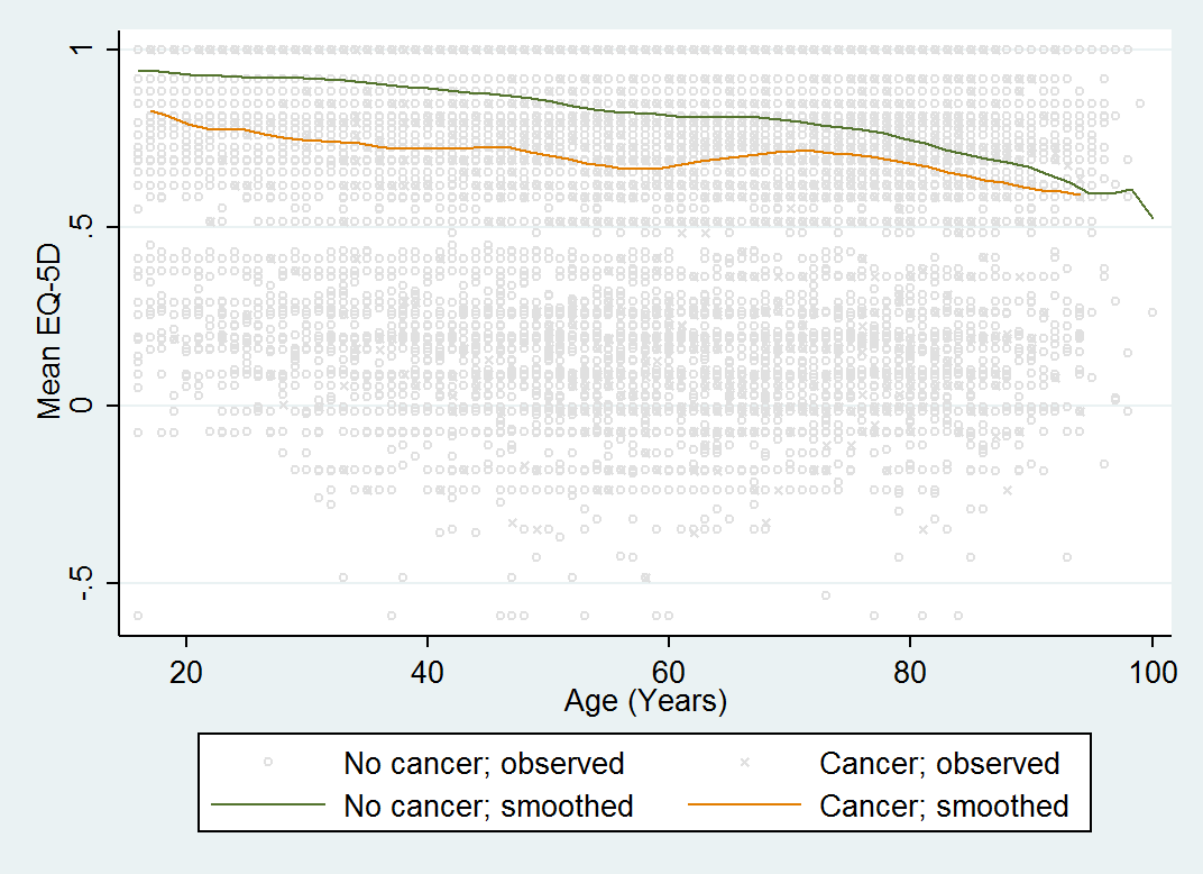
Results

The combined dataset of nine surveys contained 129,364 individuals. Of these, 42,446 (32.8%) had missing EQ-5D data and so were dropped from the analysis. This included all individuals aged less than 16 (n = 34,884; 82.2% of all the missing data). The resulting dataset used in the analysis was 86,918. A breakdown of the number of individuals, including those with cancer, per survey year is provided in Table 18. The association of mean EQ-5D with both cancer status and age is displayed in Figure 11; individual EQ-5D values are shown as scatterplots, along with smoothed averages.

Table 18: Health Survey for England data; individuals with EQ-5D data by year.

Survey Year	Count	Count with cancer	Percent with cancer
2003	13,753	266	1.93%
2004	6,114	130	2.13%
2005	9,211	243	2.64%
2006	12,926	264	2.04%
2008	14,113	292	2.07%
2010	7,332	180	2.45%
2011	7,517	171	2.27%
2012	8,060	162	2.01%
2014	7,892	170	2.15%
Total	86,918	1,878	2.16%

Figure 11: The association between mean EQ-5D and age, by cancer status.



A comparison of age, EQ-5D score and gender by cancer status is provided in Table 19. All of the differences are statistically significant; compared to those without cancer, those with cancer are older, have lower quality of life and are less likely to be female.

Table 19: Sample characteristics of individuals with and without cancer.

	Have cancer (n = 1,878)	No cancer (n = 85,040)	p-value
Age (mean, 95% CI)	64.78 (64.12 to 65.44)	49.17 (49.05 to 49.30)	<0.001
EQ-5D score (mean, 95% CI)	0.694 (0.680 to 0.708)	0.855 (0.854 to 0.857)	<0.001
Female (count, %)	1,003 (53.41%)	47,469 (55.82%)	0.039

CI: Confidence interval

Results from the multivariate fractional polynomial suggest that two age terms should be included; age and age^3. Results from the linear regression model using these two terms are displayed in Table 20, whilst results from the model just using age are displayed in Table 21. Interactions between age and cancer status were not significant for either model (p-values all greater than 0.7), so are not displayed.

Table 20: Parameter estimates for Model 1.

Covariate	N	Coef.	Std. Err.	P>t	95% CI
Age	84,576	-0.0216	0.0022	<0.001	(0.0259 to -0.0172)
Age^3	84,576	-0.0002	<0.001	<0.001	(0.0002 to -0.0001)
Female	47,174	-0.0197	0.0030	<0.001	(0.0256 to -0.0138)
Have cancer	1,768	-0.1065	0.0153	<0.001	(0.1364 to -0.0766)
Coefficient	84,576	0.9972	0.0063	<0.001	(0.9848 to 1.0095)

CI: Confidence interval. Age terms are divided by 10.

Table 21: Parameter estimates for Model 2.

Covariate	N	Coef.	Std. Err.	P>t	95% CI
Age	84,576	-0.0034	<0.001	<0.001	(-0.0035 to -0.0032)
Female	47,174	-0.0202	0.0030	<0.001	(-0.0261 to -0.0143)
Have cancer	1,768	-0.1084	0.0153	<0.001	(-0.1384 to -0.0785)
Coefficient	84,576	1.0304	0.0039	<0.001	(1.0228 to 1.0379)

CI: Confidence interval.

To validate the two-models, model predictions were compared with observed EQ-5D scores, by age-group. Results are displayed in Figure 12, with squared differences displayed in

Figure 13. These show close agreements for both models, although model 2 (which includes one age term) provides a relatively poor fit for the oldest age-group.

Figure 12: Comparison of observed and predicted mean EQ-5D scores.

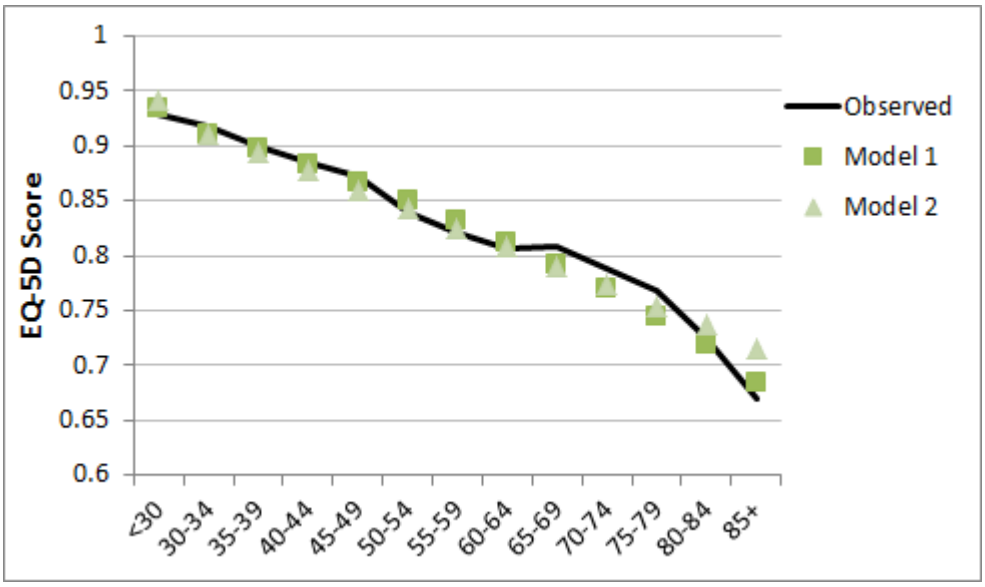
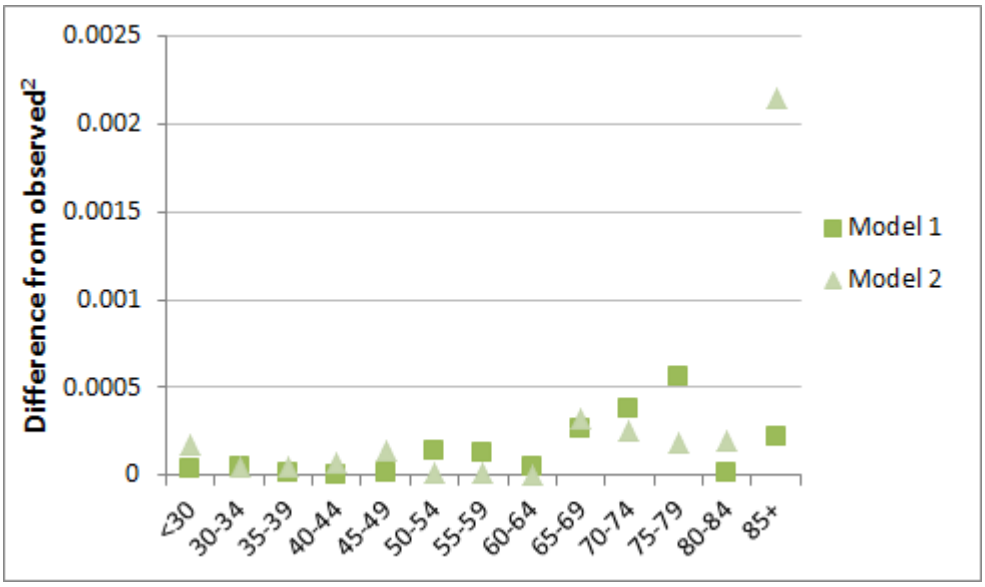


Figure 13: Squared residuals for both models.



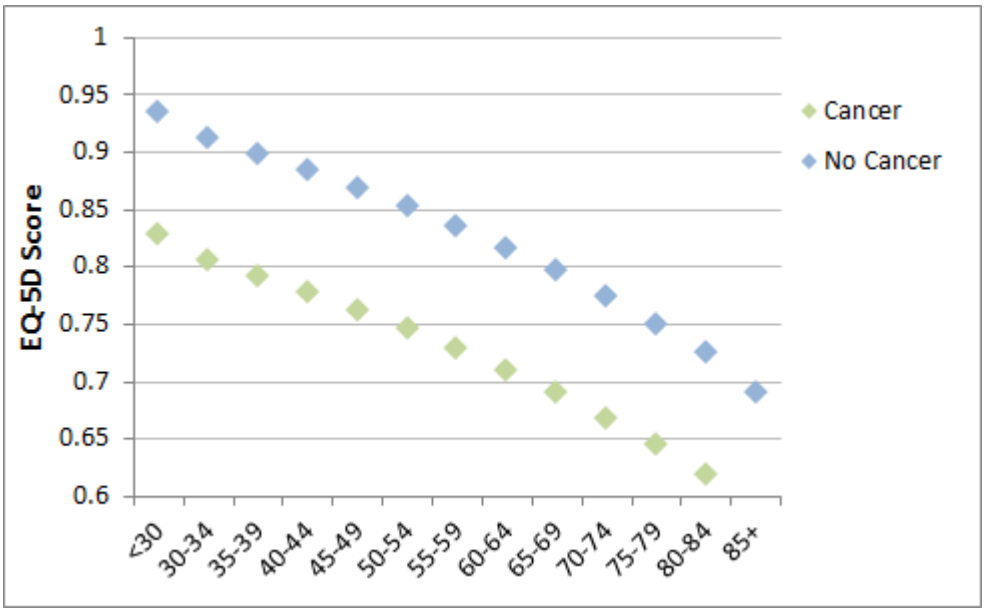
For the health-economic model, utility values for individuals with and without cancer are required by age-group. The predictions from Model 1 are provided in Table 22 and Figure 14.

Table 22: Model-predictions of mean EQ-5D score by age-group.

Age-group	Mean EQ-5D; cancer	Mean EQ-5D; no cancer	Mean EQ-5D; general population
<30	0.828947	0.935461	0.935155
30-34	0.805172	0.911686	0.911139
35-39	0.791586	0.898100	0.897436
40-44	0.777783	0.884297	0.883453
45-49	0.762204	0.868718	0.867453
50-54	0.746368	0.852882	0.850767
55-59	0.728722	0.835236	0.832586
60-64	0.709954	0.816469	0.812973
65-69	0.689847	0.796361	0.791923

70-74	0.667858	0.774372	0.769612
75-79	0.644148	0.750663	0.744147
80-84	0.61865	0.725164	0.719288
85+	0.584025	0.690539	0.684607

Figure 14: Average utilities by age for individuals with and without cancer.



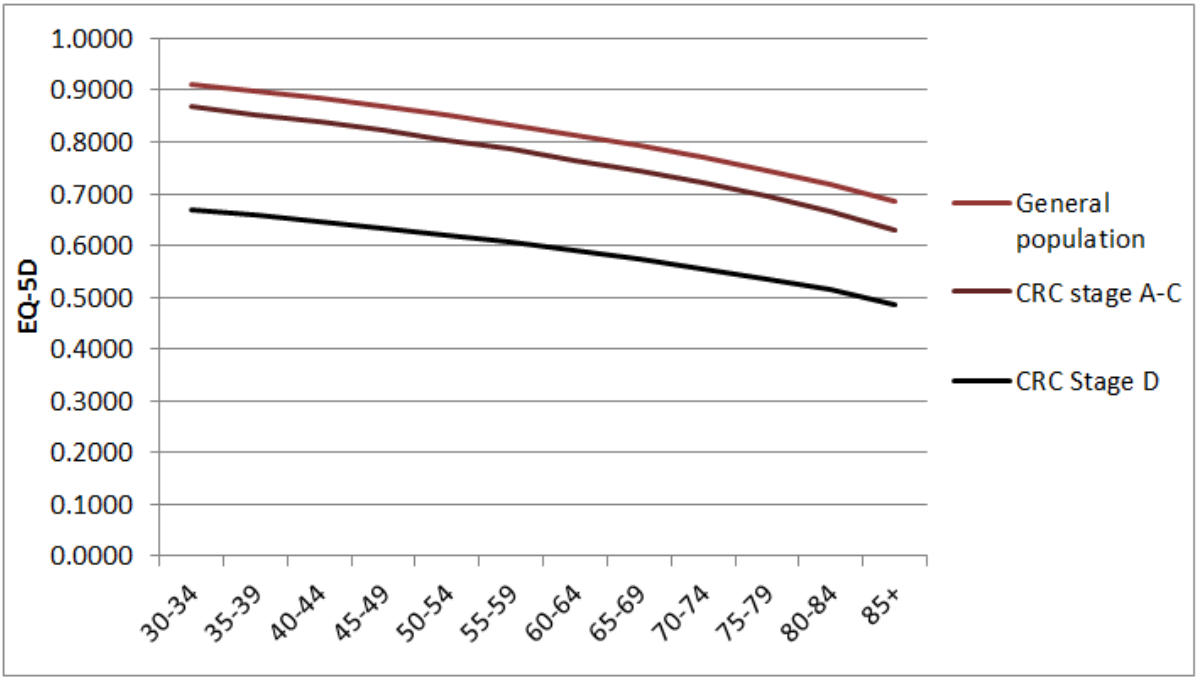
The final step in calculating utilities used in the model was to separate them by stage. The utility values for mixed stage, stage I-III and stage IV from Djalalov et al. 2014 [16] were used to calculate a further two multipliers for converting mixed stage utilities into either stage I-III utilities or stage IV utilities. These utilities are used in the model in the base case scenario, assuming that cancer stages I-III equate to CRC Duke's stages A-C and cancer stage IV equates to CRC Duke's stage D (Table 23 and

Figure 15). It is assumed that once CRC is diagnosed then individuals stay with age-specific CRC utilities for the rest of their life and never return to general population utilities.

Table 23: Utilities used in the model for the base case scenario

Age Group	Without CRC	With CRC – stage A-C	With CRC – stage D
30-34	0.9111	0.8679	0.6692
35-39	0.8974	0.8533	0.6579
40-44	0.8835	0.8384	0.6465
45-49	0.8675	0.8216	0.6335
50-54	0.8508	0.8045	0.6204
55-59	0.8326	0.7855	0.6057
60-64	0.8130	0.7653	0.5901
65-69	0.7919	0.7436	0.5734
70-74	0.7696	0.7199	0.5551
75-79	0.7441	0.6943	0.5354
80-84	0.7193	0.6669	0.5142
85+	0.6846	0.6295	0.4854

Figure 15: Utilities used in the model for the base case scenario



Given that these utilities are for cancer in general and not specifically CRC, and given that both the Djalalov et al. 2014 meta-analysis [16] and the Farkkila et al. 2013 study [18] indicate that utilities may be quite a bit higher for individuals with non-terminal cancer, a sensitivity analysis was carried out based on the utility decrements reported in Farkkila et al. These were age standardised by calculating a utility multiplier in comparison with the general population values calculated above (Table 24). It was assumed that individuals in remission had the same utility as the general population as the increase in utility reported in the study did not seem plausible.

Table 24: Utility multipliers for sensitivity analysis derived from Farkkila et al. 2013 [18]

	Palliative Care	Treatment	Remission
Utility Decrement	-0.119	-0.033	0
Mean Age	69	65	68
Utility of general population at age	0.804	0.804	0.804
Utility Multiplier	0.847	0.958	1

It was assumed that utility of the terminal cancer cases in the model would use the utility multiplier for palliative care, whilst the utility of non-terminal cancer cases would be a composite of five years for the treatment utility multiplier and the remaining life expectancy for the remission utility multiplier. This is because the model cannot incorporate time spent within a health state (e.g. five years in treatment before moving to remission) without a large increase in complexity. The remaining life expectancy for each age group was calculated from the 2012-2014 interim life tables for the UK, taking an unweighted mean of values for male and female of all ages within the group. For the 85+ age group, life expectancy was under five years and therefore only the treatment utility multiplier was used. Final utilities for each age group for the sensitivity analysis are shown in Table 25.

Table 25: Utilities used in the model for the sensitivity analysis

Age Group	Without CRC	Terminal CRC	Non-terminal CRC
30-34	0.9145	0.7748	0.9106
35-39	0.9069	0.7684	0.9026
40-44	0.8824	0.7476	0.8777
45-49	0.8639	0.7319	0.8587
50-54	0.8344	0.7069	0.8287
55-59	0.8222	0.6966	0.8156
60-64	0.8072	0.6839	0.7995

65-69	0.8041	0.6813	0.7947
70-74	0.7790	0.6600	0.7674
75-79	0.7533	0.6382	0.7385
80-84	0.6985	0.5918	0.6796
85+	0.6497	0.5505	0.6222

1.9 Appendix: Model Validation

Several model validations were carried out to check for errors in the model and investigate discrepancies between the ScHARR model, other existing models and the data.

Validation against recent incidence data

The ScHARR model was validated against the most recent UK CRC incidence data (2011-2013) available from CRUK [17]. This reports average number of CRC cases by year and age-specific incidence rates per 100,000 population, by gender and age. Total incidence by age was calculated and compared against model estimates of CRC incidence either with no screening, or with gFOBT screening (the 2012 situation where some individuals could have had up to 3-4 screening rounds would be expected to be somewhere in between these two extremes). FS screening was not modelled as it had not yet begun in 2012. To accurately estimate number of cases in the population, the model used UK population data from 2012 from the ONS [21].

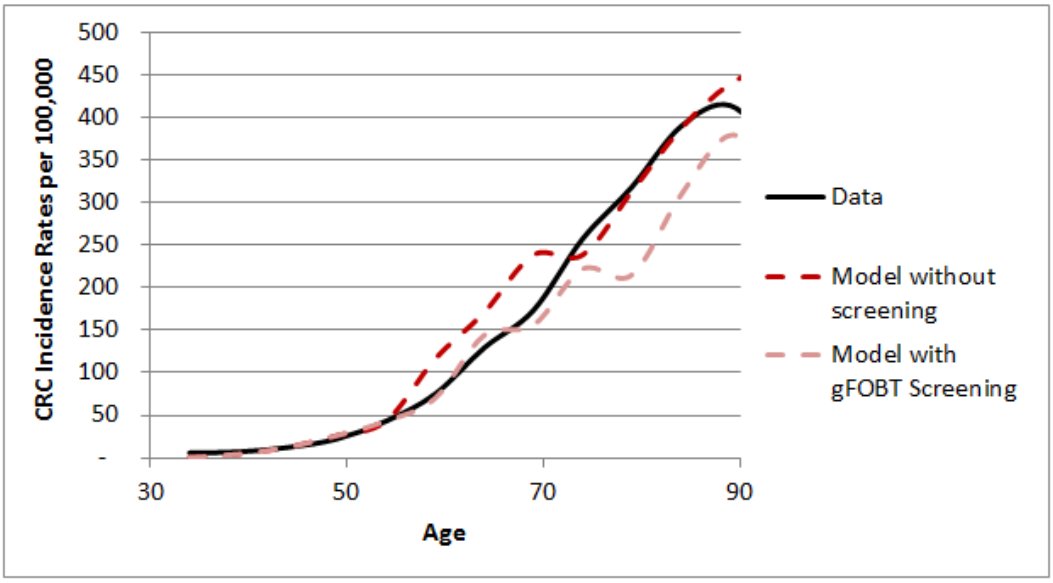
Results are shown in Table 26 and Figure 16. The no screening model predictions are fairly accurate. Slight model under-prediction in the 30-39 age range is due to the model structure in which the entire cohort starts the simulation with normal epithelium at the age of 30. This means that very few individuals develop CRC within the following 10 years. Overall this makes very little difference to the incidence statistics.

Table 26: Model predictions of incidence data from CRUK (2011-2013)

Age Range	**CRUK Data [17]**		**Model (no screening)**		**Model (gFOBT) screening**	
	Annual Cases	*Rates**	Annual Cases	*Rates**	Annual Cases	*Rates**
30-34	243	*6*	18	*0*	18	*0*
35 to 39	284	*7*	178	*4*	178	*4*
40 to 44	559	*12*	567	*12*	567	*12*
45 to 49	1,043	*22*	1,227	*26*	1,227	*26*
50 to 54	1,828	*43*	1,829	*43*	1,829	*43*
55 to 59	2,769	*75*	2,584	*116*	2,584	*70*
60 to 64	4,693	*129*	4,215	*171*	5,228	*144*
65 to 69	5,806	*174*	5,722	*239*	5,217	*156*
70 to 74	6,375	*257*	5,905	*239*	5,495	*222*
75 to 79	6,512	*318*	6,405	*313*	4,399	*215*
80 to 84	5,975	*390*	5,933	*387*	4,739	*309*
85 to 89	3,836	*414*	4,070	*440*	3,511	*379*
90+	1,832	*357*	2,381	*464*	1,725	*336*
TOTAL	**41,755**	***66***	**41,035**	***64***	**36,716**	***58***
*Rates per 100,000 population						

The model also under-estimates incidence in individuals aged between 69 and 74. This could be due to the model using incidence data from the pre-screening era (i.e. pre 2006), which is necessary in order to be able to capture the natural history of CRC without screening. According to CRUK, incidence rates have increased slightly since 2005 [17], likely due to the introduction of screening as this change in incidence is particularly noticeable in the screening-eligible age group. The introduction of screening is not accurately modelled by the screening model results as these reflect a steady state situation where the benefits of screening in reducing actual CRC disease burden outweigh the effect of screening in detecting more cases and therefore increasing incidence temporarily.

Figure 16: Comparison of incidence data from CRUK (2011-2013) and the ScHARR model



Validation against recent death certificate data

The ScHARR model was validated against the most recent death certificate registrations (2014) available from the ONS for England and Wales [22]. This reports certified deaths due to a variety of different causes by age group. Deaths due to C18, Malignant neoplasm of colon, and C19-C21, Malignant neoplasm of rectosigmoid junction, rectum and anus were included as CRC deaths. This slightly over-estimates the total number of CRC deaths as it includes a small proportion of anal cancers. CRC mortality by age was calculated and compared against model estimates of CRC mortality either with no screening, or with gFOBT screening (the 2012 situation where some individuals could have had up to 3-4 screening rounds would be expected to be somewhere in between these two extremes). FS screening was not modelled as it had not yet begun in 2012. To accurately estimate number of cases in the population, the model used ONS population data for England and Wales from 2012 [21].

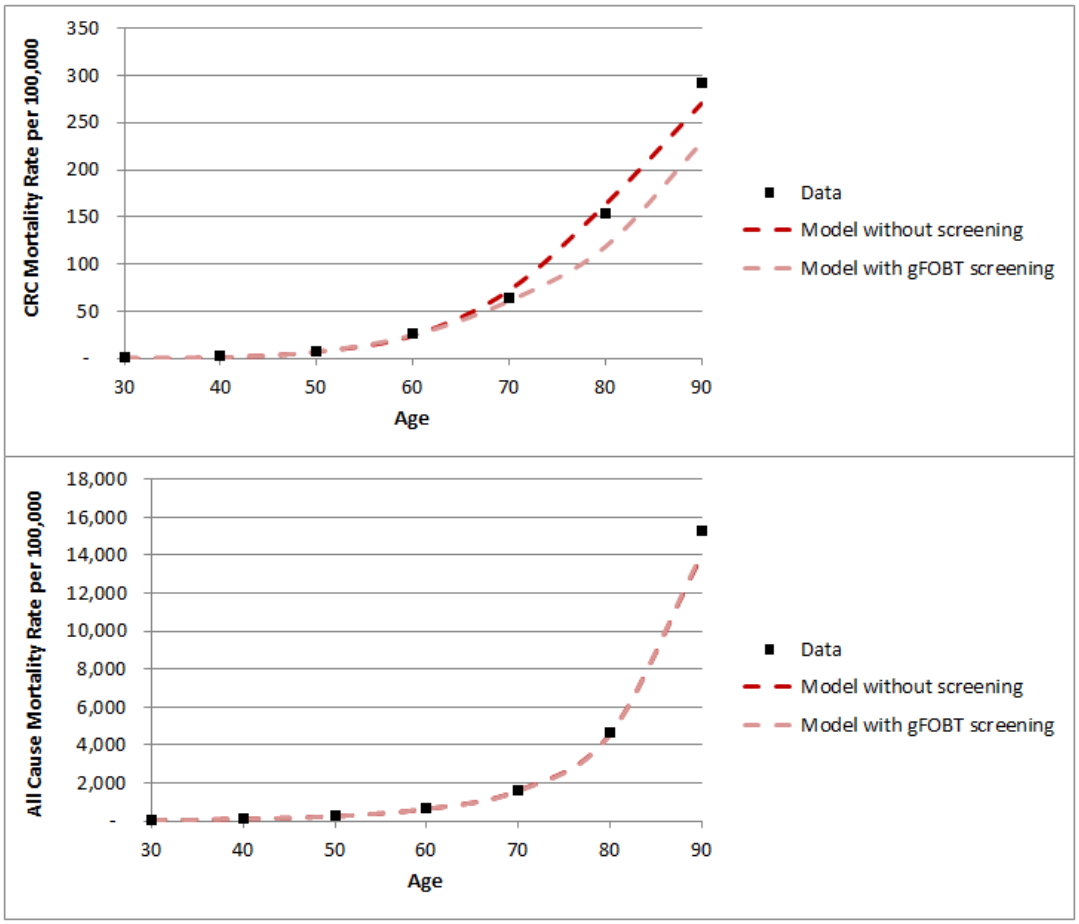
Results are shown in

Table 27 and Figure 17. The model with no screening slightly over-predicts CRC mortality, whilst the model with screening included under-predicts CRC mortality. 'All cause' mortality is fairly accurately predicted in both cases. The differences between the models occur in particular between the ages of 65 and 90, suggesting that they are due to screening. Current screening practice would be expected to lie somewhere in between the screening and no screening model results, which is exactly what the data suggests.

Table 27: Model predictions of ONS death certificate data from 2012. Mortality in rates per 100,000

Age Range	ONS Death Certificate Data		Model (no screening)		Model (gFOBT screening)	
	CRC Mortality	All Cause Mortality	CRC Mortality	All Cause Mortality	CRC Mortality	All Cause Mortality
25–34	1	56	0	62	0	62
35–44	3	121	1	117	1	117
45–54	8	266	7	265	7	265
55–64	27	658	24	645	25	646
65–74	63	1,596	72	1,608	61	1,590
75–84	153	4,690	163	4,567	118	4,535
85+	292	15,245	270	13,961	229	14,061

Figure 17: Comparison of ONS death certificate data from 2014 and mortality data from the ScHARR model



Validation against the previous FIT cost-effectiveness model

The model was validated by comparing results against those obtained in the FIT versus gFOBT cost-effectiveness analysis performed by Jacqueline Murphy and Alastair Gray [23]. Murphy & Gray based their model structure upon an older version of the ScHARR Bowel Cancer Screening Model, but updated parameters relating to unit costs, colonoscopy complications and FIT screening characteristics. Since 2011, and for this project specifically, many updates to the ScHARR model have been made. However, it was anticipated that it should be possible to compare results given by the two models and explain any differences in terms of updated parameters in the ScHARR model.

The following modifications to the ScHARR model were made in order to aid comparison of results from the two models:

- The ScHARR model is able to compare a variety of different screening strategies. To model those strategies compared in the Murphy & Gray report, results were derived from a comparison of gFOBT versus FIT, using a FIT cut-off of 180µg/g. No FS screening was incorporated in the comparison.
- Murphy & Gray use a different set of screening costs from those used as default in the ScHARR model. For the purposes of this validation, the Murphy & Gray screening costs were used instead.
- The Murphy & Gray analysis is based upon a cohort of individuals aged 60 at baseline. The ScHARR model can model either a cohort of individuals or the entire population. To model a cohort representative of the Murphy & Gray cohort in the ScHARR model, it was necessary to model natural history from the age of 30 in sufficient individuals to result in 701,809 individuals aged 60 without diagnosed CRC and hence eligible for screening. The numbers modelled are shown in Table 28. Results were collected only from the age of 60 onwards, and discounting started from age 60.

Table 28: Comparison of cohort size in the ScHARR and Murphy Gray models

	ScHARR Model	Murphy & Gray Model
Cohort aged 30	756,002	N/A
Cohort aged 60	706,962	711,228
Eligible cohort aged 60	701,809	701,809

Each of the results tables from the Murphy and Gray report was compared against the results obtained from the ScHARR model. This enabled us to find and correct some minor errors in the ScHARR model in addition to documenting the differences between the two models. The major differences found were as follows:

FIT Sensitivity Assumptions

The Murphy & Gray analysis used the FIT pilot to estimate FIT screening characteristics relative to gFOBT, whilst FIT screening characteristics were estimated for the ScHARR model comparing detection rates at age 60 with the expected number of underlying adenomas in the population. The overall sensitivity of FIT180 and gFOBT for cancer and adenomas is much lower in the ScHARR model, resulting in the lower positivity rate. However, the number of responders is higher in the ScHARR model which uses new BCSP data (whereas Murphy & Gray use the FIT pilot response rate which is lower). This means that overall, screening costs are estimated as being very similar in the two analyses as shown in

Table 29; total incremental costs per person associated with the first year of FIT screening compared with gFOBT screening are estimated at £2.27 in the ScHARR model and £2.36 in the Murphy & Gray model.

Table 29: Comparison of resource use and costs associated with screening kits in the first year of screening (age 60) in the ScHARR (black text) and Murphy Gray (red underlined text) models

	Resource Use			Cost		
	gFOBT	FIT 180	Inc.	gFOBT	FIT 180	Inc.
Total Invites	701,809	701,809	0			
	701,809	701,809	0			
Total People Normal Returns	429,981	450,880	20,898			
	375,329	439,745	64,416			
Total People Positive Returns	5,817	6,494	678			
	7,133	7,366	233			
Positivity Rate	1.33%	1.42%	0.09%			
	1.86%	1.65%	-0.22%			
Non Returners	266,011	244,435	-21,576			
	319,347	254,698	-64,649			
Number Normal Kits Used	464,380	455,388	-8,991	£942,691	£2,340,696	£1,398,006
	402,351	450,166	47,815	£810,183	£2,291,569	£1,481,386
Number Positive Kits Used	6,282	6,559	277	£12,753	£33,715	£20,962
	7,646	7,540	-106	£15,397	£38,383	£22,986
Number Unreturned Kits Used	287,292	246,879	-40,413	£238,452	£409,820	£171,367
	342,339	260,734	-81,605	£281,998	£431,072	£149,075
TOTAL COSTS				£1,193,896	£2,784,231	£1,590,335
				£1,107,578	£2,761,024	£1,653,446
TOTAL COSTS PER PERSON				**£1.70**	**£3.97**	**£2.27**
				£1.58	**£3.93**	**£2.36**

Follow-up Colonoscopy

The number of follow-up colonoscopies estimated in the ScHARR model is roughly the same as that seen in the Murphy & Gray model. This occurs because whilst the absolute number of individuals eligible for follow-up is lower in the ScHARR model (due to the lower FIT180 and gFOBT positivity), the ScHARR model has a higher uptake of follow-up colonoscopy (87% from recent BCSP data) than Murphy & Gray, and also implements a colonoscopy repeat test rate of 7% that is not implemented in the Murphy & Gray model. However, the number of diagnostic colonoscopies is much lower and the number of therapeutic colonoscopies is slightly higher in the ScHARR model, which is likely to be a consequence of the differences in the modelling of screening sensitivity (as described above). For surveillance colonoscopy, overall numbers are fairly similar between the two models, but the proportion of diagnostic to therapeutic colonoscopies is much higher in the ScHARR model and lower in the Murphy & Gray model. Other differences include a small percentage of individuals undergoing CTC rather than colonoscopy in the ScHARR model, which reduces the total number of colonoscopies undertaken slightly. The result of these differences is that the ScHARR model estimates slightly lower incremental costs per person for colonoscopy use than the Murphy & Gray model (

Table 30).

Table 30: Comparison of colonoscopy use between the ScHARR (black text) and Murphy Gray (red underlined text) models. Lifetime outcomes for a cohort aged 60. All costs are discounted by 3.5%.

	Resource Use			Cost		
	gFOBT	FIT 180	Inc.	gFOBT	FIT 180	Inc.
Follow-up Diagnostic Colonoscopy	15,003 24,315	13,659 21,379	-1,343 -2,935			
Follow-up Therapeutic Colonoscopy	28,741 22,419	33,949 26,157	5,208 3,738			
Follow-up CTC	1,985 0	2,155 0	169 0			
Total Follow-up	45,729 46,736	49,763 47,538	4,034 802	£20,298,629 £21,007,230	£22,320,550 £21,661,237	£2,021,921 £654,006
Surveillance Diagnostic Colonoscopy	17,637 8,094	24,400 10,982	6,762 2,889			
Surveillance Therapeutic Colonoscopy	9,210 20,996	12,760 28,468	3,550 7,472			
Total Surveillance Colonoscopy	26,847 29,090	37,160 39,451	10,313 10,361	£10,136,457 £12,167,682	£14,107,646 £16,594,752	£3,971,189 £4,427,069
Total Bleeds	17 31	18 35	1 4			
Total Perforation	45 43	55 47	10 3			
Total Deaths due to Colonoscopy	2 2	3 2	1 0			
TOTAL PROCEDURES	72,576 75,826	86,923 86,990	14,347 11,164	£30,485,293 £33,271,225	£36,488,313 £38,360,061	£6,003,020 £5,088,836
TOTAL PER PERSON	**0.10** **0.11**	**0.12** **0.12**	**0.02** **0.00**	**£43.44** **£47.41**	**£51.99** **£54.66**	**£8.55** **£7.25**

CRC Stage Distribution

Although the two models use the same set of natural history parameters, the ScHARR model contains several updates that alter CRC stage distribution at diagnosis and survival rates. Updates to the incidence by age and stage data mean that there are now a higher proportion of individuals diagnosed with Dukes' C than Dukes' D CRC in the ScHARR model. Also, the survival curves used in the ScHARR model to estimate cancer mortality now incorporate differential survival by age group. Furthermore, overall survival has improved in recent years meaning that mortality rates are lower (Table 31).

The predicted stage D incidence with gFOBT was double in the Murphy and Gray model compared to the ScHARR model. The Murphy and Gray model predicted proportion of incidence which is stage D was 45% which is much higher than both 2004-6 incidence data (29%), current incidence data (22%) and the ScHARR model predicted incidence data (26%).

The increase in survival and stage shift impact upon the results in two ways. Firstly, CRC stage A and B are cheaper to treat than stage C and D, which makes treatment costs slightly lower in the ScHARR model than they would otherwise be. Secondly, changing to FIT means that the absolute numbers and incremental percentage of lives saved is higher in the Murphy & Gray model than in the ScHARR model due to the CRC incidence updates in the ScHARR model, which impacts quite considerably upon life years and hence QALYs gained. Because of this, incremental life years and incremental QALYs are much smaller in the ScHARR analysis compared with the Murphy & Gray analysis (Table 31).

Table 31: Comparison of CRC incidence and mortality between the ScHARR (black text) and Murphy Gray (red underlined text) models. Lifetime outcomes for a cohort aged 60.

	Number			Proportion		
	gFOBT	FIT 180	Inc.	gFOBT	FIT 180	Inc.
Number diagnosed (CRC stage A)	6,392	5,849	-542	14%	14%	
	4,201	4,057	-144	9%	10%	
Number diagnosed (CRC stage B)	11,861	11,208	-653	26%	26%	
	9,021	8,613	-408	20%	20%	
Number diagnosed (CRC stage C)	15,157	14,460	-697	33%	33%	
	11,402	10,823	-578	26%	26%	
Number diagnosed (CRC stage D)	11,752	11,262	-490	26%	26%	
	19,881	18,722	-1,159	45%	44%	
CRC Incidence (any stage)	45,161	42,779	-2,382			-5.3%
	44,504	42,215	-2,289			-5.1%
CRC Mortality (any stage)	32,050	30,637	-1,413			-4.4%
	37,730	35,694	-2,036			-5.4%

Cost of Cancer Treatment

The cost of cancer treatment produces the largest difference between the two models. Murphy & Gray use treatment costs by stage that are inflated from Pilgrim 2009 [24], whereas the ScHARR model uses a new set of treatment costs that are inflated from Laudicella and include the indirect healthcare costs of cancer treatment so are much higher than those used by Murphy & Gray. The ScHARR costs are higher than the Murphy & Gray costs, particularly for the younger age groups (Table 32). The outcome of this is that the incremental cost savings of using FIT rather than gFOBT screening are almost double in the ScHARR model compared to the Murphy & Gray model (Table 33).

Table 32: CRC Treatment costs used in the ScHARR Model and the Murphy & Gray Model

CRC Treatment Costs	ScHARR Model (from Laudicella)	Murphy & Gray Model (from Pilgrim 2009 [24])
Dukes' A: <65	£31,218	£13,469
65+	£32,377	
Dukes' B: <65	£31,218	£18,532
65+	£32,377	
Dukes' C: <65	£44,086	£25,416
65+	£37,371	
Dukes' D: <65	£44,086	£27,796
65+	£37,371	

Table 33: Comparison of overall results between the ScHARR (black text) and Murphy & Gray (red underlined text) models. Lifetime outcomes for a cohort aged 60. All costs are discounted by 3.5%.

	Costs and Benefits		
	gFOBT	FIT 180	Inc.
Total Screening & Follow-up Costs	£37,522,161 £39,984,076	£52,854,593 £54,866,451	£15,332,431 £14,882,375
Total Cancer Treatment Costs	£835,425,281 £592,879,328	£788,111,130 £558,018,891	-£47,314,151 -£34,860,437
TOTAL COSTS	£872,947,442 £632,863,404	£840,965,722 £612,885,342	-£31,981,720 -£19,978,062
TOTAL COSTS PER PERSON	**£1,243** **£902**	**£1,198** **£873**	**-£46** **-£28**
Total Life Years	11,338,996 11,263,240	11,342,231 11,276,575	3,236 13,335
TOTAL QALYs	8,743,528 8,962,563	8,746,499 8,972,325	2,971 9,762
TOTAL QALYs PER PERSON	**12.46** **12.77**	**12.46** **12.78**	**0.004** **0.014**
ICER	**Cost saving, QALY gain** **Cost saving, QALY gain**		**-£10,764** **-£2,047**

The comparison of the two models indicates that despite being based on many of the same parameters and structural assumptions, they give very different results for some outcomes. Cancer treatment costs are one of the key differences; if the ScHARR model is run using the same CRC treatment costs as those used in the Murphy & Gray model, then the cost-savings of using FIT compared to gFOBT are slightly lower than those found in the Murphy & Gray analysis, rather than almost twice as high. These remaining differences are likely to be due to the differences in stage distribution, CRC survival and estimates of FIT sensitivity used in the two models.

Table 34: Comparison of overall results using different cancer treatment costs

	ScHARR Model Laudicella costs	ScHARR Model Pilgrim costs	Murphy & Gray Pilgrim costs
Total Cancer Treatment Costs	-£47,314,151	-£28,081,461	-£34,860,437
TOTAL COSTS	-£31,981,720	-£12,749,029	-£19,978,062
TOTAL COSTS PER PERSON	**-£46**	**-£18**	**-£28**
TOTAL QALYs PER PERSON	**0.004**	**0.004**	**0.014**
ICER	**-£10,764**	**-£4,291**	**-£2,047**

Validation against the Nottingham Study

A model validation was carried out to compare model results against the findings of the gFOBT randomised controlled trial conducted in Nottingham from 1981 [25]. This randomised 152,850 individuals to biennial gFOBT screening or control (no screening) groups, and has recently published 20 year follow-up results [26].

The model was adapted as much as possible to reflect the trial setting; however, as the model is a cohort model it was not possible to incorporate the patient variability within the trial in terms of age, number of screening rounds and follow-up. The validation exercise therefore had the following limitations:

- The trial included individuals aged between 45 and 75, and reports the number of individuals recruited in each five year age group. To model this as closely as possible, six age cohorts were modelled with mean ages of 47; 52; 57; 62; 67 and 73 and results for these added together.
- The trial delivered screening biennially to individuals, averaging 3-5 rounds of screening per person. The model is only able to simulate screening between the ages of 50 and 75. Therefore for most cohorts it was assumed that individuals received 4 rounds of biennial screening, starting in the first year, but for the cohort aged 47, screening was assumed to not start until age 50, and for the cohort aged 73, only 2 rounds of biennial screening were modelled.
- The trial has a median of 19.5 years of follow-up per person, ranging from 0 to 28.4 years. In the model, a 20 year follow-up was assumed for everyone.
- The screening part of the trial finished over 20 years ago, so some aspects such as CRC survival and all-cause mortality may be out of date compared to the model. To test this, an analysis was carried out in which life table data and CRC mortality data from 1996 (halfway through the trial follow-up) was used instead of current data.

The trial follow-up analysis reported a significant reduction in certified CRC mortality with gFOBT screening compared to no screening with a rate ratio of 0.91 (0.84 to 0.99) and adjusted rate ratio (to take account of incomplete uptake) of 0.82 (0.70 to 0.98), but reduction in CRC incidence was not significant. If current day life tables and CRC mortality rates are used, then the model is able to replicate this reduction in CRC mortality due to gFOBT screening very accurately, but is not able to estimates a higher CRC incidence reduction than expected, although results are well within the reported 95% confidence interval. Rate ratios for CRC mortality are made worse if historical mortality data from 1996 is used.

Table 35: Comparison of adjusted rate ratios for gFOBT screening vs no screening in the gFOBT trial and in the model

	CRC Incidence	**CRC Mortality**	**All Cause Mortality**
Trial	**0.94 (0.85-1.05)**	**0.82 (0.70-0.98)**	**1.00 (0.99-1.02)**
Model current mortality data	0.89	0.82	1.00
Model 1996 mortality data	0.89	0.87	1.00

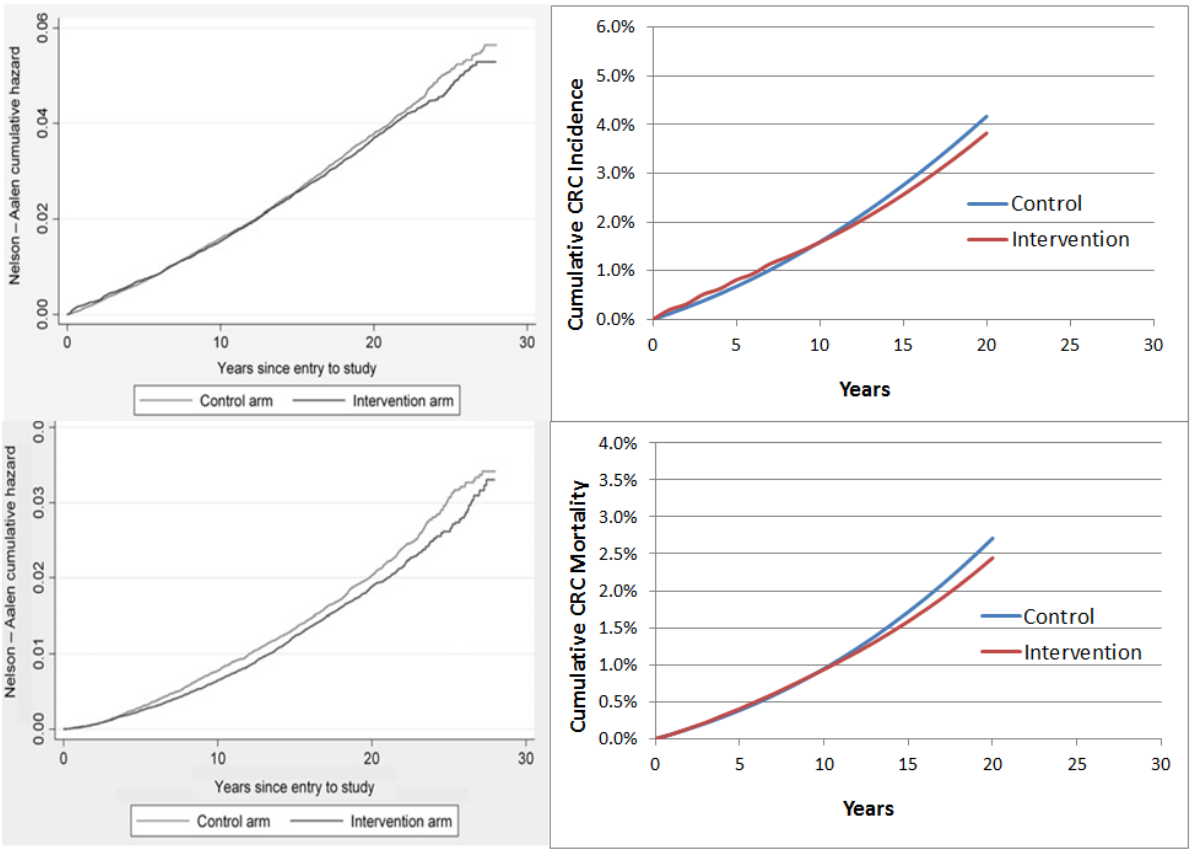
The absolute numbers of individuals dying with CRC or from other causes is underestimated using current mortality data, but is estimated more accurately using historical mortality data from 1996. Absolute CRC incidence however is slightly overestimated by the model using current mortality data. This makes sense as the reduction in 'all cause' mortality over the past 35 years has contributed to a larger number of people living long enough to develop CRC, and indeed CRC incidence has increased over that period according to CRUK statistics [17].

Table 36: Comparison of absolute values for gFOBT screening and no screening in the gFOBT trial and in the model

	Control – no screening				**gFOBT Screening**			
	CRC incidence	**CRC mortality**	**Other cause mortality**	**All cause mortality**	**CRC incidence**	**CRC mortality**	**Other cause mortality**	**All cause mortality**
Trial	**2,354**	**1,300**	**39,250**	**40,550**	**2,279**	**1,176**	**39,505**	**40,681**
Model current	2,655	1,073	26,076	27,150	2,503	978	26,129	27,079
Model 1996	2,380	1,193	34,396	35,589	2,253	1,114	34,485	35,599

Figure 18 shows that CRC incidence does not reduce in the model until year 10-11, and perhaps even later in the trial. This indicates that the benefits of screening in reducing cancer incidence are only seen after screening has stopped: whilst screening is ongoing, cases detected through screening outweigh the number of cases prevented by screening. Mortality reductions in contrast should occur early after screening begins as is seen in the trial. The model is unable to simulate these early mortality reductions as the cohort nature of the model means that reduced mortality in screen detected cancer cases cannot be incorporated.

Figure 18: Estimates of cumulative CRC incidence (top) and mortality (bottom) from the study (left) and the model using 1996 mortality data (right) with 20 years of follow-up



There are several potential explanations for the differences between the trial results and model predictions for cancer incidence. The first relates to differences between the trial population and a general screening population. The trial selected only those individuals who did not have serious illness, meaning that the trial population was healthier than the general population and therefore may have benefitted less from screening (this is known as the healthy volunteer effect [27]). Also, for the first ten years of the trial, individuals were not re-invited for subsequent screening rounds if they had not attended the first screening round. However, we now know from BCSP data presented in the data section of this report that those who have not previously attended screening have a higher incidence of CRC and polyps than those who have previously attended screening [28] and therefore screening has more potential to detect pre-cancerous abnormality and thereby prevent cancer in these individuals than in previous screenees.

Secondly, in the past 35 years, colonoscopy quality is likely to have improved, potentially leading to greater detection and more successful removal of adenomas. Finally, there was no surveillance programme operating at the time of the trial, whereas surveillance in the model has a small effect in reducing cancer incidence and mortality. All these reasons could lead to the trial having underestimated the benefits of BCSP screening in a current day population. However, the model may also overestimate the benefits of screening. Model natural history is based on general population data and the cohort model assumes that screen attendees are representative of the general

population. However, this isn't the case - those who attend screening have fewer abnormalities than those who do not attend screening and therefore may benefit less from screening.

Validation against the FS Trial

A model validation was carried out to compare model results against the 11 year follow-up findings of the FS randomised controlled trial conducted between 1994 and 1999 by Wendy Atkin and others [12, 29]. The trial randomised 57,237 individuals to FS screening (of which 71% took it up), and 113,195 individuals to the control group (no screening).

The following assumptions were used when adapting the model to simulate trial results:

- The trial enrolled individuals aged between 55 and 65. The model simulates the correct number of individuals in each age-year cohort and adds results together to produce a total.
- Mean follow-up in the trial was 11.2 years, whereas the model uses an 11 year follow-up for the entire cohort.
- Uptake for FS screening was changed from 44% (BCSP data) to 71% to reflect the trial.
- The trial was carried out between 1996 and 1999 with follow-up data published in 2010, so some aspects such as CRC survival and all cause mortality may be slightly out of date compared to the model. To test this, life table data and CRC mortality data from 2003 (halfway between 1996 and 2010) was used instead of current data.

The model slightly underestimates absolute CRC incidence but estimates the benefits of screening on CRC incidence very accurately (HR = 0.78 for model and 0.77 for study) (Table 37). Absolute CRC mortality is overestimated by the model, but the benefits of screening on CRC mortality are underestimated compared to the study although this does lie just within the study 95% confidence intervals (HR = 0.82 for the model and 0.69 for the study). Note that other cause mortality is also slightly over-estimated by the model. Interestingly, if current day life table and CRC mortality data is used instead of historic (2003) data, the model is much better at estimating absolute CRC mortality in the control arm and slightly better at estimating CRC mortality hazard ratios (HR = 0.78), but then underestimates absolute other cause mortality quite considerably (data not shown in table).

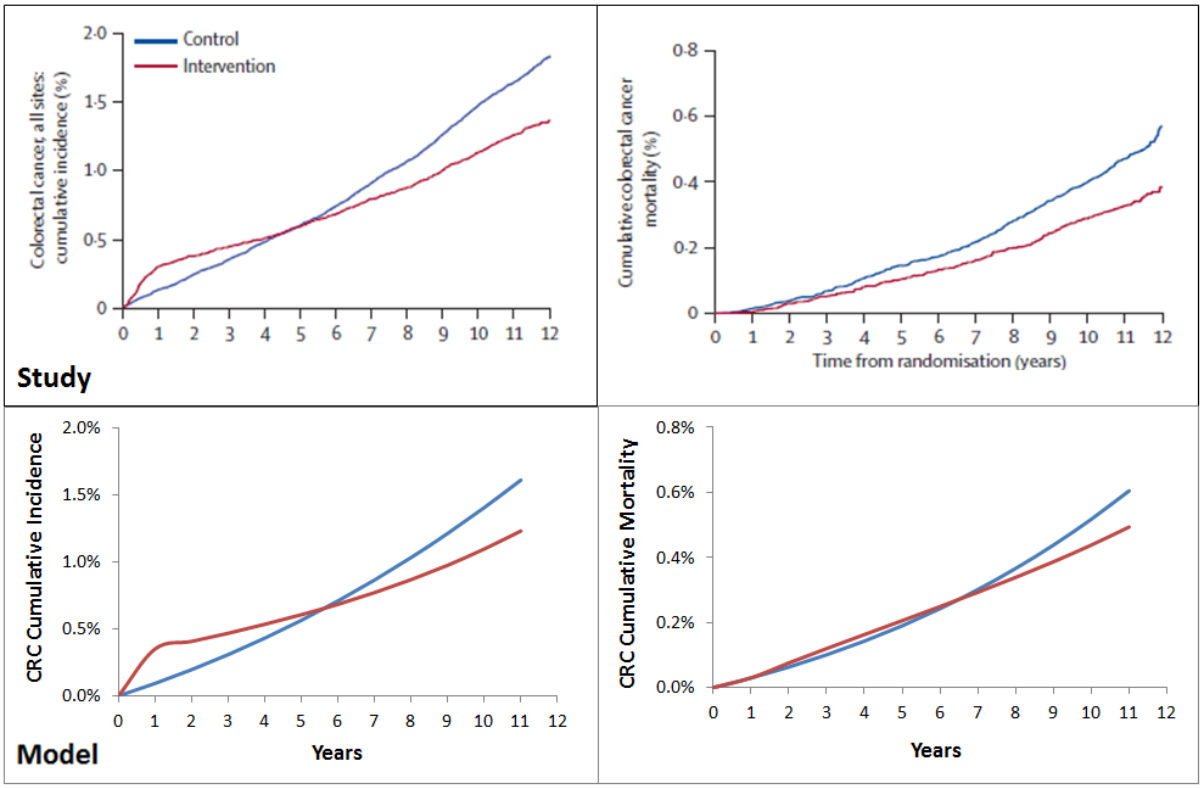
Table 37: CRC incidence and mortality in control and intervention groups in the study and the model after 11 years follow-up

	Control		**Intervention**		**Hazard Ratio**	
	Study	Model	Study	Model	Study	Model
CRC Incidence	1,818	1,701	706	667	0.77 (0.70 -0.84)	0.78
*CRC Incidence Rates**	*149 (143-156)*	*145*	*114 (106-123)*	*112*		
CRC Mortality	538	636	189	265	0.69 (0.59 -0.82)	0.82
*CRC Mortality Rates**	*44 (40-48)*	*54*	*30 (26-35)*	*45*		
Other Cause Mortality	13,230	13,439	6,554	6,774	0.99 (0.96 -1.02)	1.00
*Other Cause Mortality Rates**	*1,080(1,062-1,099)*	*1,144*	*1,057 (1,032-1,083)*	*1,141*		
All Cause Mortality	13,768	14,075	6,775	7,040	0.97 (0.94 -1.00)	0.99
*All Cause Mortality Rates**	*1,124 (1,106-1,143)*	*1,198*	*1,093 (1,067-1,119)*	*1,185*		
*Rates per 100,000 person years						

Comparison of CRC incidence over time indicates that the model seems to be fairly accurate at estimating the initial increase in incidence seen due to screening and the point at which a net reduction in CRC incidence becomes apparent (at about 5-6 years following screening) (Table 37). However, estimates of CRC mortality over time are slightly anomalous, as the screened group shows slightly higher CRC mortality in the early years. This is an

unavoidable limitation of the model which cannot distinguish between screen detected cancer cases and undetected cases which in reality have different mortality rates.

Figure 19: Estimates of cumulative CRC incidence (left) and mortality (right) from the study (top) and the model (bottom) with 11 years of follow-up



There are several possible reasons for the differences between the trial and model results. Firstly, whilst the model uses natural history data from the general population, the trial selected a healthier population for analysis, which excluded individuals with poor health, family history of CRC or current symptoms of CRC. This healthy volunteer effect has been noted previously [27], and could partially explain the differences in absolute CRC mortality observed, particularly given that there is also some difference in 'other cause' mortality. It was reported by Atkin et al. (2010) that CRC incidence in the study control group was almost exactly as expected in the general population which may argue against a healthy volunteer effect, but no mention of a similar CRC mortality comparison was made. Furthermore it is not clear whether a difference in population composition that impact upon absolute mortality would also impact upon relative mortality between control and screening trial arms.

Secondly, the model assumes that individuals receive no screening other than the single FS. However, gFOBT screening was initiated during the trial follow-up period, meaning that it is likely that some individuals in both trial arms received gFOBT screening, known to impact upon CRC mortality and potentially incidence (see Nottingham trial validation above). Thirdly, the model uses a fixed natural history model which does not vary over time whereas in reality presentation rates and incidence may change over time.

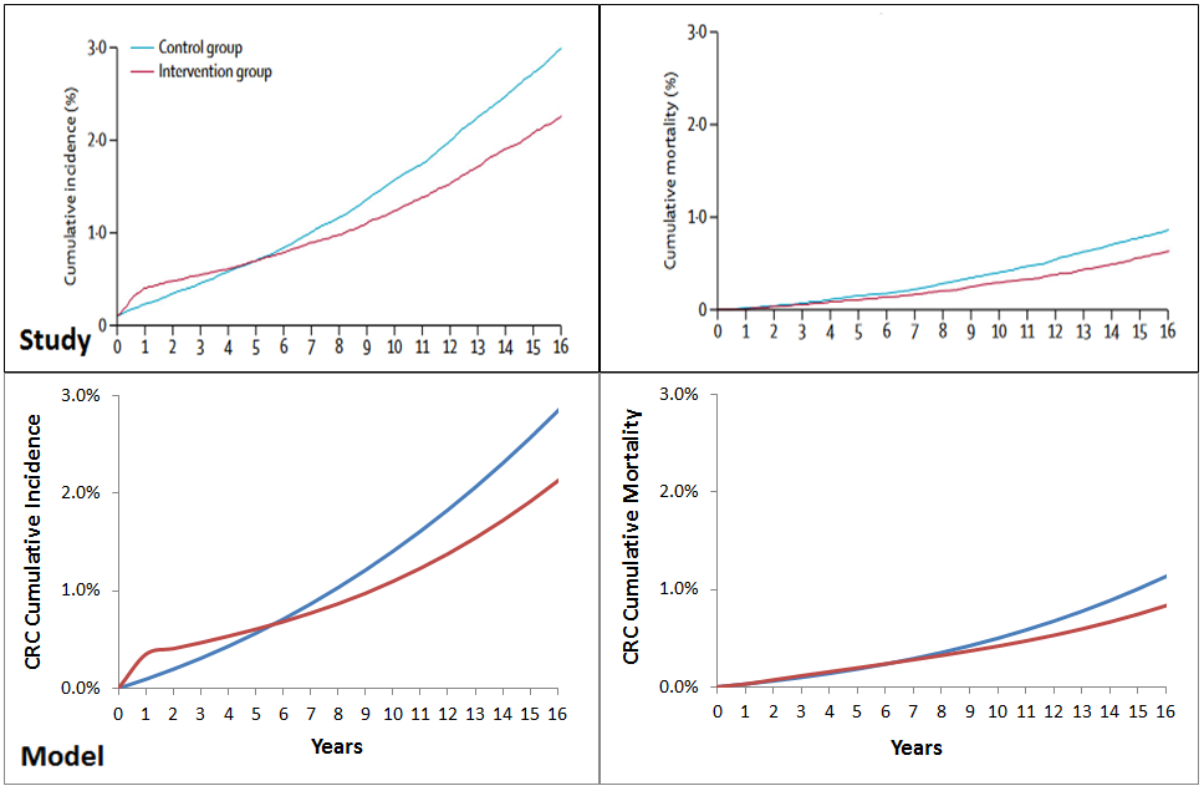
A further validation was carried out against the unpublished FS trial 17 year follow-up data. All model parameters were kept as described above for the 11 year follow-up. However, to reflect the longer follow-up time, life tables and CRC mortality data from 2006 were used instead of 2003.

The model predicts that the benefits of FS screening compared to no screening will continue to increase slightly over this period, indicating that the benefits of FS screening are long-lasting.

Table 38: CRC incidence and mortality in control and intervention groups the model after 17 years follow-up

	Control		Intervention		Hazard Ratio	
	Study	Model	Study	Model	Study	Model
CRC Incidence	3,253	3,105	1,230	1,190	0.74 (0.70-0.80)	0.76
*CRC Incidence Rates**	*184 (178-191)*	*179*	*137 (130-145)*	*136*		
CRC Mortality	996	1,184	353	468	0.70 (0.62-0.79)	0.78
*CRC Mortality Rates**	*56 (53-60)*	*68*	*39 (35-43)*	*53*		
Other Cause Mortality	25,413	26,196	12,926	13,196	1.00 (0.98-1.03)	1.00
*Other Cause Mortality Rates**	*1,427 (1,410-1,445)*	*1,512*	*1,433 (1,408-1,458)*	*1,505*		
All Cause Mortality	27,379	27,390	13,279	13,664	0.99 (0.97-1.01)	0.99
*All Cause Mortality Rates**	*1,483 (1,465-1,501)*	*1,580*	*1,472 (1,447-1,497)*	*1,558*		
*Rates per 100,000 person years						

Figure 20: Estimates of cumulative CRC incidence (left) and mortality (right) from the study (top) and the model (bottom) with 17 years of follow-up



Validation against FIT estimates in the literature

At the lowest threshold of FIT modelled (cut-off of 20 µg/ml) a sensitivity for CRC of 54.4% is used in the model. This value was derived from the UK pilot [30], and is based on individuals receiving both an initial and subsequent screens. This value is similar to that of 57.5% reported by Murphy and Gray, who used a similar methodology. Both this study and the study by Murphy and Gray used data from the UK pilot; it is expected that this is the most relevant evidence source as it related to the UK population, and hence will reflect the natural history of cancer in the UK. It is unclear if the natural history of CRC in different countries is generalisable to the UK setting. Despite this potential limitation Murphy and Gray noted that their derived value (57.5%) is lower than that reported in other studies. For example, the recent systematic review of screening for CRC prepared for the U.S. preventative services task force

[31], identified four studies for which comparable estimates of sensitivity were available [30, 32-34], which reported sensitivities of 73% (n = 2,220), 75% (n = 1,256), 74% (n = 9,989) and 100% (n = 779). To investigate this issue further the estimate from the largest study by Imperiale et al. (2014), which used the same FIT test as the UK pilot [30] was compared with the estimate used for this study. Three reasons for the different sensitivity estimates were identified:

1. Colonoscopy is used by the authors to identify all instances of CRC, with the assumption that this is 100% specific. However, for this report it is estimated that colonoscopy only identifies 98% of cancers, based on published evidence [35]. Adjusting the Imperiale *et al.* sensitivity estimate to account for imperfect colonoscopy adjusts the sensitivity from 73.8% to 72.4%.
2. The authors only consider sensitivity for the first screen. However, it has been shown that sensitivity is higher in the first screen than in subsequent screens [36], with sensitivity from the first screen over-estimating the overall sensitivity by about 8.6%. This results in an adjusted sensitivity from 72.4% to 66.6%.

This adjusted sensitivity of 66.6% is closer to the sensitivities used in both this report and by Murphy & Gray (2015) [23], suggesting that much of the difference was due to differences in how sensitivity was defined.

Validation of surveillance parameters

Model estimate of surveillance colonoscopy usage in year 1 was compared with the actual number of surveillance colonoscopies from the BCSP, in order to validate the surveillance part of the model. Two scenarios (gFOBT screening alone and gFOBT screening plus FS aged 55) were run as the current situation of partial FS roll-out should be intermediate between them. Table 39 indicates that the model estimates are far larger than the observed BCSP data, indicating a problem with the surveillance model. Note that the FIT cost-effectiveness model developed independently by Murphy & Gray does estimate similar number of surveillance colonoscopies to the ScHARR model, indicating that it is unlikely that the difference with BCSP data is due to model coding errors.

Table 39: Number of surveillance colonoscopies observed in the BCSP data and estimated in a single year in the model

BCSP data	Model estimate (gFOBT screening 8 years into roll-out)	Model estimate (gFOBT screening steady state)	Model estimate (FS age 55 plus gFOBT screening steady state)
12,642	27,518	22,469	35,381

There are several reasons why the model could be inaccurately estimating the number of surveillance colonoscopies. Firstly, the model estimates an increased risk of adenoma development in the surveillance population compared with the general population. If the risk has been set too high, then individuals will stay in surveillance for longer as they will be less likely to have consecutive surveillance colonoscopies with negative results. This was tested by altering the model so that adenoma risk in surveillance was unchanged from adenoma risk in the general population. It can be seen in

Table 40 that using an unchanged adenoma risk has only a small effect on the total number of surveillance colonoscopies in year 1 of the model. This seems to be fairly consistent when considering different screening strategies.

Table 40: A comparison of the number of surveillance detected cancer diagnoses and annual surveillance colonoscopies for several different screening strategies and different surveillance assumptions.

	Year 1 Surveillance Colonoscopies	Number of cancer diagnoses: population lifetime			
		Dukes A	Dukes B	Dukes C	Dukes D
gFOBT biannual age 60-74 steady state					
Base case	22,469	1,403	429	120	33
Unchanged adenoma risk in surveillance	21,996	146	55	20	4
60% surveillance attendance rate	19,486	1,694	648	261	85
10% annual surveillance	20,506	1,399	432	122	34
50% annual surveillance	24,652	1,406	425	118	33
gFOBT biannual age 60-74 8 years into roll-out					
Base case	27,518	1,454	443	123	34
Unchanged adenoma risk in surveillance	27,352	151	57	21	4
60% surveillance attendance rate	20,753	1,725	665	261	85
10% annual surveillance	24,658	1,450	447	125	35
50% annual surveillance	30,669	1,457	439	121	33
FIT40 age 55, 60, 65 & 70					
Base case	36,467	1,804	589	187	55
Unchanged adenoma risk in surveillance	35,415	199	79	30	6
60% surveillance attendance rate	34,756	2,766	1,149	512	179
10% annual surveillance	33,489	1,778	584	186	54
50% annual surveillance	39,766	1,832	595	189	55
FS age 50, FIT20 annual 51-74					
Base case	102,276	4,826	1,529	461	131
Unchanged adenoma risk in surveillance	92,711	474	188	71	14
60% surveillance attendance rate	96,460	7,074	2,883	1,260	437
10% annual surveillance	93,811	4,804	1,533	463	132
50% annual surveillance	111,639	4,851	1,525	457	130

Secondly, the model could be assuming a higher attendance rate for surveillance than currently occurs (the model assumes 83% attend). This was tested by reducing attendance rate to 60% (

Table 40), which reduces the number of surveillance colonoscopies carried out variably depending upon screening strategy, but increases the number of cancer cases detected, presumably due to a reduction in individuals leaving surveillance because those that do not turn up do not have the opportunity to be tested negative and leave surveillance. Thirdly, the model could be assuming that a higher proportion of individuals are being referred for annual surveillance versus 3 year surveillance than occurs in practice (the model assumes 29% are referred to annual surveillance). The Murphy& Gray FIT cost-effectiveness analysis reported different numbers of therapeutic versus diagnostic surveillance colonoscopies than the ScHARR model (see validation against Murphy & Gray data above) [23]; however it is unclear what assumptions they used to obtain these results. To test the ScHARR model assumptions, a sensitivity analysis altering the proportion going to annual surveillance to either 10% or 50% was carried out (

Table 40). This again had little effect on the results.

Finally, the model could be overestimating the number of individuals undergoing surveillance if, for example the guidelines pathway, used in the model to determine who is referred for surveillance and who should come off surveillance, is not being followed in practice. This could affect either numbers going into surveillance or coming out of surveillance. Data about surveillance is not available directly from the BCSP, but obtaining this in future will allow the model assumptions to be updated and the surveillance model to be improved.

1.10 Appendix: Evidence on changing sensitivity of FOB-tests over repeat screening rounds.

The evidence on the change in gFOBT/FIT sensitivity by round due to changes in the proportion of disease which is detectable by round is presented below.

Kearns et al. 2014, Guaiac faecal occult blood test performance at initial and repeat screens in the English Bowel Cancer Screening Programme [36]

The gFOBT sensitivity for CRC was estimated to decrease from 27.35% at the initial screen to 20.22% at the repeat screen (a relative reduction of 26.1%). Decreases were also observed for the positive predictive value (8.4–7.2%) and detection rate for CRC (0.19–0.14%). Assuming equal performance measures for both the initial and repeat screens led to an overestimate of the cost effectiveness of gFOBT screening compared with the other screening modalities.

van der Meulen et al. 2016, Nonbleeding Adenomas: Evidence of Systematic False-Negative Faecal Immunochemical Test Results and Their Implications for Screening Effectiveness—A Modeling Study [37]

The model without systematic false-negativity simulated higher detection rates in the second screening round than observed. These observed rates could be reproduced when assuming that FIT systematically missed 26% of advanced and 73% of nonadvanced adenomas. To reduce the false-positive rate in the second round to the observed level, the authors also had to assume that 30% of false-positive findings were systematically false-positive. Systematic false-negative FIT testing limits the long-term reduction of biennial FIT screening in the incidence of colorectal cancer (35.6% vs 40.9%) and its mortality (55.2% vs 59.0%) in participants.

Uri Ladabaum et al. 2016, Sensitivity of Repeated Faecal Immunochemical Testing (FIT) Over Time: Does Each "Bite at the Apple" Stand the Same Chance? [38]

The authors considered two different scenarios: one with independent FIT sensitivities for each round, and one with a diminishing FIT sensitivity for each round (sensitivity at round 2 = ¾ that at round 1, sensitivity at round 3 = ½ that at round 1. The authors stated that modelling independent FIT sensitivities provided a better fit to the observed data, and concluded that “Our results support the hypothesis that the diminishing yield of FIT observed in subsequent screening cycles is attributed primarily to the progressive removal of people with neoplasia from the screen-eligible population, instead of the presence of a substantial fraction of lesions that are "silent" with respect to FIT (i.e. never-bleeding).”

The usefulness of this study is limited as it is only available as an abstract. In addition, the authors’ conclusion that “Our simulation suggests that FIT sensitivity can be considered independent through the initial 4 cycles of a screening program” appears to be based on a visual comparison of modelled and observed data, with no systematic quantitative method for making this judgement. Of the three cohorts modelled, the assumption of a diminishing FIT sensitivity appears (visually) to fit as well as the assumption of independent FIT sensitivities for two (Dutch and USA). Modelling independent FIT sensitivities gives a better fit for the Italian data. It is unclear what these differences are, although difference in FIT sensitivity may be a cause. In addition, there are also other scenarios for a diminishing FIT that the authors could have explored, which may have provided a better fit to the data. For example, the authors did not model a scenario of just two different FITs: one for the initial screen, and one for subsequent screens. In addition, the estimate of the magnitude of the change in FIT sensitivity could have been data-driven to enable a better fit to the observed data.

Steele et al. 2012, Interval cancers in a FOBT-based colorectal cancer population screening programme: implications for stage, gender and tumour site [39]

This study used data from the Scottish BCSP and interval cancers to estimate gFOBT sensitivity at repeat screens. Sensitivity for a given round was defined as: 'screen detected cancers / (screen detected + interval cancers)'. As screen detected cancers will be affected by uptake rates, it is important to adjust for this when making comparisons across rounds. This was achieved by adjusting the screen detected cancers for a given round by the difference in uptake between that round and the first round. The number of interval cancers was also adjusted so that the overall number of cancers per round was the same. For example, for round two there were 208 screen-detected cancers and 213 interval cancers (total cancers = 421). Uptake for round 2 was 96.2% that of round 1 (53.0% compared with 55.0%). Hence if uptake in round 2 had been 55.0%, the number of screen detected cancers would have increased to approximately 216 (= 208*1/0.962). In other words, 8 extra cancers would have been screen detected instead of being interval cancers, so the number of interval cancers would have reduced from 213 to 205. Results from this adjustment are provided in Table 41.

Table 41: Screening results for uptake adjustment

Round	**1**	**2**	**3**
Uptake	55.0%	53.0%	55.3%
Screen-detected	351	208	139
Interval	193	213	229
Screen-detected adjusted	351	216.1	138.3
Interval adjusted	193	204.9	229.7
Sensitivity	64.5%	51.3%	37.6%
Relative change		20.4%	26.8%

However, this data is limited in making comparisons. Whilst round 1 will contain only initial screens, subsequent rounds will contain a mixture of initial and repeat screens. Sensitivity should not change for initial screens (case-mix should not change), what is required is an estimate of how gFOBT sensitivity changes for repeat screens. Data on the proportion of screens that were repeats was not presented in the publication, so was instead taken from Steele et al. [39], and is presented in Table 42.

Table 42: Repeat screening proportions from Steele et al. [39]

Round	2	3
Total invitations	309,803	317,864
Non-responder in all previous rounds	111,763	105,115
Responder in any previous round	147,469	158,006
Total Uptake (%)	53%	55%
Non-responder in all previous rounds	0.138	0.139
Responder in any previous round	0.854	0.852
Total Uptake (n; derived)	164,196	175,779
Non-responder in all previous rounds	15,423	14,611
Responder in any previous round	125,939	134,621
% Repeats*:	76.7%	76.6%

* Defined as number of responders in any previous round / total uptake

Adjusting the relative change for the number of repeats provides a relative change from initial screens to repeat round 2 of 26.7% (20.4%/76.7%). This value is very similar to the relative change found in the UK BCSP by Kearns et al. (26.1%), supporting the conclusions of this study.

Garcia et al. 2015, Interval Cancers in a Population-Based Screening Program for Colorectal Cancer in Catalonia, Spain [40]

This study used data from the biennial screening programme in Barcelona. Interval cancers (defined as cancers diagnosed within 30 months from the last screen) were used to derive estimates of sensitivity, using the same formula as Steel et al. 2012. Data for up to four rounds of screening were available. For the first three screens gFOBT was used, for the fourth a mixture of gFOBT and FIT were used.

No details on uptake were provided, which limit comparisons across screening rounds as differences in uptake will affect estimates of sensitivity when using interval cancers. The study is also limited by relatively low number of identified cancer, which leads to uncertainty in the results. To emphasis this uncertainty, some of the results from Table 3 have been replicated below in Table 43, but including 95% confidence intervals (based on the Wilson score method, as described in Newcombe [41].

Table 43: Screening round sensitivities with 95% confidence intervals

Last screening round	Sensitivity and 95% CI	Number of screens	Sensitivity and 95% CI
1	67.65% (50.84% to 80.87%)	1	67.09% (56.15% to 76.45%)
2	48.15% (30.74% to 66.01%)	2	38.64% (25.72% to 53.38%)
3	57.45% (43.28% to 70.49%)	3	54.29% (38.19% to 69.53%)
4	53.97% (41.79% to 65.69%)	4	61.54% (35.52% to 82.29%)
Rounds 2 to 4	54.01% (45.67% to 62.14%)	Rounds 2 to 4	47.83% (37.91% to 57.91%)

When comparing results by screening round, there is little difference in sensitivities between rounds two and four. However, these results will be limited as the populations will include a mixture of initial and repeat screens. When comparing initial screens (number of screens = 1) with repeat screens (number of screens = 2 to 4), the relative change in sensitivity is 28.7%. This number is similar to, although slightly larger than, the changes reported in the Kearns (26.1%) and Steele (26.7%) studies. It is unclear what impact differential uptake rates would have on the estimate of 28.7%.
Another notable result from the Garcia study is that there is no evidence that sensitivity continues to decrease with repeat screening. In fact, the opposite occurs: amongst people receiving repeat screens, sensitivity increases with the number of repeat screens received. Some of this may be an artefact of natural variation due to small numbers, or differences in uptake (for which there is no data).

A limitation of both the Steele and the Garcia papers is their reliance on using interval cancers. This limitation has already been noted in the Kearns study: "The use of interval cancers to estimate sensitivity has two main limitations; first, some interval cancers would not have been cancers at the time of the screen and second, not all undetected cancers will be diagnosed within any given interval." However, the results from these two studies provide similar estimates of the relative change in sensitivity from the initial to repeat screen.

Conclusion

In conclusion, for gFOBT there is evidence to suggest that sensitivity for repeat screens is lower than for the initial screen. This includes evidence from both the NHS BCSP and the Scottish BCSP. There is currently no evidence that sensitivity continues to diminish for recurring repeat screens, although this is more due to a lack of evidence rather than evidence for lack of an effect.

The evidence base for repeat FIT screens is small, and there is no UK-based evidence. The Ladabaum study, using Dutch data, suggests that FIT sensitivity will decrease for repeat screens. This conclusion is contradicted by the authors of the van der Meulen study. However, the conclusions of this study are limited as the results are only available as an abstract and it is unclear if the presented results actually support the conclusions.

Hence, due to a paucity of data, for the base-case analysis no differences in FIT sensitivity (beyond that due to differences in case-mix as captured by the natural history model) are modelled. The robustness of results to this assumption are explored in sensitivity analyses.

1.11 Appendix: Comparison with studies which consider 'a longer screening interval, with a more sensitive FIT test'

The potential of a longer screening interval, with a more sensitive FIT test, has previously been investigated by Haug et al. (2017) [42] and Digby et. al. (2016) [43]. Both based their findings on a trial on a standard screening interval of 2 years, and used the results to model outcomes for alternative regimes. They concluded, respectively, that a shorter interval leads to performance comparable to, or worse than, the standard scenario. This agrees with our findings in this report, that an interval of 2 years is optimal. However, they only looked at the effects of alternative screening strategies for the first two screening rounds, rather than over a whole screening regime.

Haug et al. [42] analysed data from an ongoing CRC screening study in the southwest Netherlands, with a standard threshold of FIT50. The interval between the first and second screening rounds varied between 1-3 years, and the interval between the second and third rounds was fixed at 2 years. Subjects that tested positive were referred to a surveillance programme. Haug et al. followed subjects up to the time of the third round, exclusive. Alternative scenarios had different thresholds, with the second round omitted, resulting in an interval of 3-5 years. Alternative scenario outcomes were estimated comparing subjects' baseline Haemoglobin (Hb) readings to the scenario's threshold. The results found FIT11 to have a number of subjects diagnosed with advanced adenomas similar to that of the standard scenario, and a higher positivity rate. FIT22 had a similar positivity rate to the standard scenario, with 10% less subject diagnosed with advanced adenomas. All alternative scenarios used less FIT tests, but thresholds below FIT22 had more colonoscopies than the standard scenario. Haug et al. concluded that the alternative scenarios did not markedly differ with respect to diagnostic yield and cumulative positivity rate. They also suggested that an alternative screening strategy, with a higher sensitivity for a single screening, would be advantageous for subjects that attend screenings irregularly, or in the case of lack of a regular invitation system, such as in a decentralised health system. Further, they refer to another study (check which), which found that varying the interval between first and second screenings showed no change in the second-round detection rate for advanced neoplasia. They expect similar results for programmes with higher thresholds.

Digby et al. [43] analysed data from a demonstration pilot study involving 30893 subjects, between 50 and 74 years of age, in Scotland, with a standard threshold of 80 μg/g (FIT200). Alternative scenario results were estimated on the assumption that all cancers detected in the second screening round were present and detectable during the first screening round. Digby et al. found that a threshold of 27.6 μg/g (FIT138) doubled the positivity rate, and had the same number of colonoscopies as the standard scenario. However, there was a substantial decrease in the number of screen-detected cancers, and a substantial increase in the number of interval cancers. There were similar trade-offs for other thresholds. Digby et al. therefore concluded that the alternative regimes seemed unattractive. However, based on an earlier study, which found that the first-screen Hb measurement predicts the subsequent risk of incident colorectal neoplasia, they suggested that a possible improvement would be to make the screening interval dependent on the first-round Hb results, on an individual basis.

1.12 Appendix: Obtaining data on endoscopy capacity

The following document was circulated to the HEE endoscopy working group for comment. Unfortunately no additional data/information was provided on which to base the modelling assumptions.

Estimates of current and future endoscopy usage within the BCSP

Chloe Thomas, Sophie Whyte 24th October 2016, ScHARR, University of Sheffield

These data form part of the Optimising Bowel Cancer Screening research project being undertaken by ScHARR for the National Screening Committee.

The current estimates are based on data from the BCSP and information about current BS roll-out from John Davy.

Estimates future endoscopy capacity were generated via estimates of maximum extra capacity based on information about endoscopy training from Neil Hawkes (via Matt Rutter).

There is some uncertainty in the estimates of future capacity due to limited information on which to base assumptions. Please provide any feedback or additional data/information to inform modelling assumptions to ScHARR ASAP via c.thomas@sheffield.ac.uk or sophie.whyte@sheffield.ac.uk.

Table 1: Estimate of the number and type of endoscopy procedures currently performed annually within the BCSP. Surveillance data from 2015/16, Claire Nickerson at the BCSP; gFOBT data from 2015, Claire Nickerson at the BCSP; BS data extrapolated from September 2016 data from John Davy.

Type of Procedures	Number of Procedures
gFOBT follow-up colonoscopy	39,783
gFOBT follow-up BS	1,896
gFOBT follow-up partial colon.	475
Total gFOBT follow-up	**42,154**
BS screening	106,020
BS follow-up colonoscopy	4,530
Total BS	**110,550**
Surveillance colonoscopy	12,642
Surveillance BS	208
Total Surveillance	**12,850**
TOTAL ENDOSCOPY PROCEDURES	**165,554**

Table 2: Maximum estimates of increase in endoscopy capacity over the next two years, based on information about a training programme for non-medical endoscopists from Neil Hawkes.

	By end 2016	By end 2018
Max. Number new trainees	40	200
Max. Number additional BS procedures	80,000	400,000
If one third trainees go on to train further in colonoscopy:		
Max. Number additional BS procedures	53,333	266,667
Max. Number additional diagnostic colonoscopies	13,333	66,667

1.13 Appendix References

1. Cotton, S., L. Sharp, and J. Little, *The adenoma-carcinaoma sequence and prospects for the prevention of colorectal neoplasia.* Crit Rev Oncog, 1996. **7**(5-6): p. 293-342.
2. Tappenden, P., et al., *Using whole disease modeling to inform resource allocation decisions: economic evaluation of a clinical guideline for colorectal cancer using a single model.* Value Health, 2013. **16**(4): p. 542-53.
3. Whyte, S., C. Walsh, and J. Chilcott, *Bayesian calibration of a natural history model with application to a population model for colorectal cancer.* Med Decis Making, 2011. **31**(4): p. 625-41.
4. Atkin, W.S. and B.P. Saunders, *Surveillance guidelines after removal of colorectal adenomatous polyps.* Gut, 2002. **51**(Suppl V): p. v6-v9.
5. Brenner, H., et al., *Risk of progression of advanced adenomas to colorectal cancer by age and sex: estimates based on 840,149 screening colonoscopies.* Gut, 2007. **56**(11): p. 1585-9.
6. Yamaji, Y., et al., *Right-side shift of metachronous colorectal adenomas after polypectomy.* Scand J Gastroenterol, 2007. **42**(12): p. 1466-72.
7. Nickerson, C., *Bowel Cancer Screening Programme Data.* 2010.
8. Martinez, M.E., et al., *A pooled analysis of advanced colorectal neoplasia diagnoses after colonoscopic polypectomy.* Gastroenterology, 2009. **136**(3): p. 832-41.
9. Whyte, S., et al., *Re-appraisal of the options for colorectal cancer screening: Full report for the NHS Bowel Cancer Screening Programme*. 2011, ScHARR, University of Sheffield.
10. Brenner, H., L. Altenhofen, and M. Hoffmeister, *Sex, Age, and Birth Cohort Effects in Colorectal Neoplasms: A Cohort Analysis.* Ann Intern Med, 2010. **152**(11): p. 697-703.
11. Chung, S.J., et al., *Prevalence and risk of colorectal adenoma in asymptomatic Koreans aged 40–49 years undergoing screening colonoscopy.* Journal of Gastroenterology and Hepatology, 2009. **25**(3): p. 519-525.
12. UK Flexible Sigmoidoscopy Screening Trial Investigators, *Single flexible sigmoidoscopy screening to prevent colorectal cancer: baseline findings of a UK multicentre randomised trial.* Lancet, 2002. **359**: p. 1291-1300.
13. Soon, M.S., et al., *Screening colonoscopy in Chinese and Western patients: a comparative study.* Am J Gastroenterol, 2005. **100**(12): p. 2749-55.
14. Chib, S. and E.R. Greenberg, *Understanding the Metropolis-Hastings Algorithm.* The American Statistician, 1995. **49**(4): p. 327-335.
15. National Institute for Health and Care Excellence, *Guide to the methods of technology appraisal*. 2013.
16. Djalalov, S., et al., *A Review and Meta-analysis of Colorectal Cancer Utilities.* Med Decis Making, 2014. **34**(6): p. 809-18.
17. Cancer Research UK. *Bowel Cancer Statistics*. 2015 [cited 2016 27th September]; Available from: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer.
18. Farkkila, N., et al., *Health-related quality of life in colorectal cancer.* Colorectal Dis, 2013. **15**(5): p. e215-22.
19. Centre, H.a.S.C.I., *Health Survey for England*. 2014.
20. Dolan, P., et al., *A social tariff for EuroQol: Results from a UK general population survey*. 1995, University of York.
21. Statistics, O.f.N., *Population estimates for UK, England and Wales, Scotland and Northern Ireland*. 2012.
22. Statistics, O.f.N., *Deaths Registered in England and Wales*. 2015.
23. Murphy, J. and A. Gray, *The cost-effectiveness of immunochemical faecal occult blood testing vs. guaiac faecal occult blood testing for colorectal cancer screening in the NHS Bowel Cancer Screening Programme*. 2015, Health Economics Research Centre, University of Oxford.
24. Pilgrim, H., et al., *The costs and benefits of bowel cancer service developments using discrete event simulation.* Journal of the Operational Research Society, 2009. **60**(10): p. 1305-1314.

25. Hardcastle, J.D., et al., *Randomised controlled trial of faecal-occult-blood screening for colorectal cancer.* Lancet, 1996. **348**(9040): p. 1472-1477.
26. Scholefield, J.H., et al., *Nottingham trial of faecal occult blood testing for colorectal cancer: a 20-year follow-up.* Gut, 2012. **61**(7): p. 1036-40.
27. Croswell, J.M., D.F. Ransohoff, and B.S. Kramer, *Principles of Cancer Screening: Lessons from History and Study Design Issues.* Semin. Oncol., 2010. **37**(3): p. 202-215.
28. Nickerson, C., *Bowel Cancer Screening Programme Data*. 2016.
29. Atkin, W.S., et al., *Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised controlled trial.* The Lancet, 2010. **375**(9726): p. 1624-1633.
30. Imperiale, T.F., et al., *Multitarget stool DNA testing for colorectal-cancer screening.* N Engl J Med, 2014. **370**(14): p. 1287-97.
31. Lin, J.S., et al., *Screening for colorectal cancer: updated evidence report and systematic review for the US preventative services task force.* JAMA, 2016. **315**(23): p. 2576-2594.
32. Brenner, H. and S. Tao, *Superior diagnostic performance of faecal immunochemical tests for haemoglobin in a head-to-head comparison with guaiac based faecal occult blood test among 2235 participants of screening colonoscopy.* Eur J Cancer, 2013. **49**(14): p. 3049-54.
33. de Wijkerslooth, T.R., et al., *Immunochemical fecal occult blood testing is equally sensitive for proximal and distal advanced neoplasia.* Am J Gastroenterol, 2012. **107**(10): p. 1570-8.
34. Hernandez, V., et al., *Fecal immunochemical test accuracy in average-risk colorectal cancer screening.* World J Gastroenterol, 2014. **20**(4): p. 1038-47.
35. Bressler, B., et al., *Rates of new or missed colorectal cancers after colonoscopy and their risk factors: a population-based analysis.* Gastroenterology, 2007. **132**(1): p. 96-102.
36. Kearns, B., et al., *Guaiac faecal occult blood test performance at initial and repeat screens in the English Bowel Cancer Screening Programme.* Br J Cancer, 2014. **111**(9): p. 1734-41.
37. van der Meulen, M.P., et al., *Nonbleeding adenomas: Evidence of systematic false-negative fecal immunochemical test results and their implications for screening effectiveness—A modeling study.* Cancer, 2016. **122**(11): p. 1680-1688.
38. Ladabaum, U., F.W. Chen, and A. Mannalithara, *Mo1711 Sensitivity of Repeated Fecal Immunochemical Testing (FIT) Over Time: Does Each "Bite at the Apple" Stand the Same Chance?* Gastroenterology, 2016. **150**(4): p. S759.
39. Steele, R.J.C., et al., *Results from the first three rounds of the Scottish demonstration pilot of FOBT screening for colorectal cancer.* Gut, 2009. **58**: p. 530-535.
40. Garcia, M., et al., *Interval Cancers in a Population-Based Screening Program for Colorectal Cancer in Catalonia, Spain*. Vol. 2015. 2015. 9.
41. Newcombe, R.G., *Two-sided confidence intervals for the single proportion: comparison of seven methods.* Statistics in Medicine, 1998. **17**(8): p. 857-872.
42. Haug, U., et al., *Immunochemical faecal occult blood testing to screen for colorectal cancer: can the screening interval be extended?* Gut, 2016. **66**(7): p. 6.
43. Digby, J., et al., *Can the performance of a quantitative FIT-based colorectal cancer screening programme be enhanced by lowering the threshold and increasing the interval?* Gut - Online First, 2017.