Serum CEA trends for Diagnosing Colorectal Cancer Recurrence in the FACS Randomised Clinical Trial

RUNNING HEADER: Monitoring CEA trend to detect colorectal cancer recurrence

Bethany Shinkins, John N Primrose, Siân A Pugh, Brian D Nicholson, Rafael Perera, Tim James, David Mant

University of Leeds, UK
Bethany Shinkins, DPhil, Lecturer in Decision Analytic Modelling

University of Oxford, UK
Brian D Nicholson, MRCGP, Clinical Research Fellow
David Mant, FMedSci, Emeritus Professor of General Practice
Rafael Perera, Head of Statistics

Oxford University Hospitals NHS Foundation Trust, Oxford, UK
Tim James, PhD, Head Biomedical Scientist.

University of Southampton, UK
John Primrose, FMedSci, Professor of Surgery
Sian Pugh, MRCS, MRC Clinical Research Fellow

Correspondence to: Bethany Shinkins
Current address: Leeds Institute of Health Sciences, Worsley Building, University of Leeds, Clarendon Way, LS2 9LJ

e-Mail: b.shinkins@leeds.ac.uk

ABSTRACT

Background

Most guidelines recommend that patients who have undergone curative resection for primary colorectal cancer are followed up for 5 years with regular blood CEA tests to trigger further investigation for recurrence. With this single test approach, recurrence may be missed or patients may have false alarms.

Methods

The diagnostic accuracy of trends in CEA measurements for recurrent colorectal cancer, taken as part of the FACS (Follow-up After Colorectal Surgery) trial (2003-2014), were analysed. Investigation to detect recurrence was triggered by clinical symptoms, scheduled CT or colonoscopy, or a CEA>7μg/L above baseline. Time-dependent ROC analysis compared the diagnostic accuracy of CEA trends to single test results. CEA trends were estimated via linear regression.

Results

The area under the ROC curves (AUC) for CEA trend was consistently over 0.82 across all 5 years of follow-up. In comparison, the AUCs for single measurements ranged from 0.62 – 0.75. Improvement was most marked at the end of the first year of follow-up, with the AUC increasing from 0.62 (95% CI: 0.51-0.74) to 0.88 (95% CI: 0.81-0.95). However, no individual trend threshold achieved a sensitivity above 70% (30% missed cases).

Conclusions

Interpreting trends in CEA measurements instead of single CEA test results improves diagnostic accuracy for recurrence, but not sufficiently to warrant it being used as a single surveillance strategy to trigger further investigation. In the absence of a more accurate biomarker, monitoring trends in CEA should be combined with clinical, endoscopic, and imaging surveillance for improved accuracy.

Key Words

Carcinoembryonic antigen, ROC curve, serial measurements, colorectal cancer, recurrence

INTRODUCTION

After initial treatment for colorectal cancer, most clinicians in Europe and North America follow-up patients to detect asymptomatic recurrence. Guidelines recommend using more than one diagnostic modality (1-5), and most clinicians employ a combination of clinical examination, carcinoembryonic antigen (CEA) testing, some imaging (usually CT-CAP), and colonoscopy to detect intra-luminal recurrence.

Two recent trials, FACS (Follow-up After Colorectal Surgery) and CEA Watch, have confirmed that including CEA testing as part of follow-up advances diagnosis, more than doubling the number of recurrences that can be treated surgically with curative intent (6, 7). However, both trials suggest that the common approach to testing recommended by guidelines (acting on the results of single tests) is sub-optimal. In the FACS trial, recurrences were missed in patients who had experienced a gradual CEA rise and we previously suggested that this might be avoided by considering the trend in a series of results over time(8). In the CEA Watch trial, further investigation was based on CEA trend on the basis of evidence from other studies that it provide better diagnostic yields than a single test (9-14).

Critics of intensive follow-up regimes point to the lack of evidence that pre-symptomatic diagnosis of recurrence improves survival. It is therefore important that the chosen follow-up regime does not cause harm and that it is cost-effective. A potential problem with acting on single raised CEA test results is that it can cause false alarms (leading to unnecessary CT scans). We therefore report here an observational analysis of CEA data from the FACS randomised controlled trial to quantify the advantage of making clinical decisions on the basis of CEA trend rather than the results of a single test, looking at the potential to avoid both missed cases of recurrence and of false alarms.

METHODS

Secondary observational analysis of data from the FACS trial, a 2x2 pragmatic randomised factorial controlled trial comparing minimum post-surgery follow-up of colorectal cancer patients for 5 years with 3-6 monthly blood tests for carcinoembryonic antigen (CEA) and 6-12 monthly computerised tomography (CT) imaging (6) (see Appendix Tables A1 and A2 for exact testing schedule). Patients were recruited between 2003 and 2009 from 39 NHS hospitals across England who had undergone curative surgery for primary colorectal cancer and, after extensive testing (histology, imaging and a CEA ≤ 10μg/L), were confirmed to have no residual disease. The CEA analysis was undertaken using a Siemens Centaur XP analyser at a single laboratory with a standard quality control regime to ensure longitudinal stability. If the blood CEA level was 7μg/L or more above the patient's baseline level (which was measured on entry to the trial) after repeat measurement, the GP was asked to refer the patient urgently to the local hospital for further investigation.

The reference standard in the study was clinical diagnosis of colorectal cancer recurrence as determined by the multi-disciplinary team (MDT) at each hospital centre. As this was a pragmatic clinical trial in an NHS setting, all available evidence (which will have varied between hospitals and individual patients) was used to determine a diagnosis of recurrence, including imaging and biopsy results. All participant centres were bowel cancer specialist centres with diagnostic experience and expertise. Investigation for suspected recurrence could be triggered as a result of abnormal results on per-protocol colonoscopy or CT imaging, a CEA>7μg/L above their personal baseline, or suspicious symptoms.

Linear regression was applied to capture the trend in individuals' CEA measurements. The beta coefficient from each model was extracted and compared to using a threshold to interpret single CEA measurements. A selected value was used for the single measurement analysis. The CEA measurement at the time of diagnosing recurrence is often selected for analysis[15]. But this may overestimate the accuracy as in clinical practice all measurements are interpreted prospectively. We therefore looked across all measurements for each individual and identified, based on all possible thresholds, whether their measurements had risen above the threshold at any time point. Time-dependent ROC analysis was utilised to demonstrate the trade-off between sensitivity and sensitivity when clinical decisions are based on different degrees of trend and area under the curves. All analyses were carried out using the statistical package R (http://www.R-project.org/); the dependent ROC analysis was carried using the R package timeROC (16).

RESULTS

Table 1 presents the clinical characteristics of the patients included in the analysis. The median number of CEA measurements available for each participant was 13 (IQR: 10-14), with a median of 6 (IQR: 3–9) measurements in patients who developed a recurrence and 14 (IQR: 13-14) for those who did not develop a recurrence. In this analysis, we included all individuals who had at least two CEA measurements available, but this means that individuals that have a recurrence detected in the first three months had to be excluded. The analysis was based on 582 of the 602 patients allocated to the two arms of the study who received CEA testing (see flow-chart and reasons for exclusions in Appendix Figure A1).

Of the 96/104 patients with recurrence who had undergone at least two CEA tests, 6 had a falling level, 6 had an initial fall before a rise, 13 had stable levels, and 23 had a rising CEA level that did not exceed the 5µg/L threshold. Even among the 52 patients whose CEA did rise above 5µg/L, it took a median of 15.9 months (IQR: 6.8 – 29.9) to go above the threshold. Of the 478 patients without recurrence, 15 (3%) had two or more raised CEA measurements above 5µg/L within the study period. These false alarms occurred in 12% of smokers and 3% of non-smokers (OR=4.43, 95% CI: 1.34-14.7, p=0.01). Smokers should therefore be informed of the false-alarm risk and be followed up with a different modality, such as CT-CAP.

INSERT FIGURE 1 HERE

Figure 1 compares the CEA trend in patients with and without recurrence. CEA levels tended to be very stable in patients free of recurrence but those from patients with disease recurrence had variable trends with a 100 fold standard deviation.

INSERT FIGURE 2 HERE

Figure 2 compares the time-dependent ROC curves at Year 1 for making the decision to investigate on the basis of the trend compared to single measurements. The table shows that across all years of follow-up, the AUC for CEA trend remains over 0.80. The time-dependent AUC estimates for the trend analysis are consistently higher than the single measurement interpretation (although in some cases the 95% confidence intervals overlap). This appeared to be constrained to a maximum sensitivity of 70% without losing substantial test specificity. Making decisions on CEA would not allow CEA alone to be used as a triage test because even at the point of optimal trade-off between sensitivity and specificity, 3 in 10 recurrences would be missed. Additional tests to maintain sensitivity should be

carried out as per local protocol. For details of additional testing and associated timings within the FACS trial, please refer to Tables A1 and A2 in the Appendix.

The time-dependent ROC analysis allowed us to explore optimal thresholds for interpreting trend (change/yr). These were: Year 1 1.7ug/L; Year 2 1.4ug/L; Year 3 0.8ug/L; Year 4 0.5 ug/L; Year 5 0.3 ug/L.

DISCUSSION

Although debate continues about the benefits of surveillance (17), most clinicians continue to monitor CEA in addition to CT scans at intervals with the aim of detecting pre-symptomatic recurrence. Our results suggest that the diagnostic accuracy of CEA trends is better than interpreting the result of a single test, but a substantial number of recurrences still would be missed if CEA was utilised as a single surveillance method.

In the absence of alternative markers, we can make some suggestions for an improved CEA monitoring schedule based on the FACS data. As per the FACS protocol, 3 monthly measurements in year 2 of follow-up and 6 monthly measurements for the remaining time were sufficient to monitor trend. As almost half the recurrences were identified during the first year of follow-up however, we would suggest a CEA monitoring schedule that consists of monthly measurements in the first 3 months and 2-monthly for the remainder of the first year. This would ensure that enough measurements are taken to observe a trend early on in follow-up (2 months onwards). In the first couple of months, single test measurements would still have to be interpreted. Based on the FACS data, a threshold of 10 ug/L was optimal. Possible thresholds for interpreting trend (change/yr) are reported in the results section, although these require further validation before implementing in clinical practice. To facilitate interpretation, laboratories could report trend over time as well as the absolute CEA level on their report forms.

The cost-effectiveness and feasibility of such an intensive testing schedule would need to be fully evaluated though. The mean direct NHS cost of CEA testing in the FACS trial was estimated as £7.50/test so the cost of the additional 4 tests in year 1 would be approximately £30 (not taking account of subsequent price inflation) (18). A recent study in the Netherlands, the CEAwatch trial, compared an intensive follow-up protocol that included more frequent CEA measurements (bi-monthly), but fewer outpatient visits, detecting more curable recurrences compared to standard practice(19). An economic evaluation of the trial, which included direct medical costs, productivity losses and travel expenses, was conducted and the ICER (incremental cost-effectiveness ratio) per curable recurrence was calculated. It was found that an additional $60,700 (€51,465) is needed to detect one additional patient with curable recurrent disease. Although the authors considered this to be acceptably low (19), until the quality of life and survival benefit of detecting pre-symptomatic recurrence is known, definitive conclusions are not possible (18).

The main strengths of the FACS trial are that the analysis was undertaken on well-staged patients with robust outcome data. The CEA testing was centrally managed with good compliance, and all analyses

were done in one laboratory. The main limitation is that we do not have a reference standard at all time points. The precise time when a recurrence would have been detectable by a trend based analysis is not known. However, the length of follow-up and within-trial surveillance means missed recurrences were not likely. There may have existed a work-up bias as patients with a CEA>7μg/L above their personal baseline were referred for further investigation; however this selected threshold (2002) is higher than current guidelines.

Due to the small number of recurrences it is not possible to use the present analysis to develop a definitive model and cut-off values that clinicians could use now to interpretation of trend. Ideally this would explore other factors that may bring granularity to the interpretation of individual trends (e.g. smoking status, tumour site and stage).

Here we demonstrate that interpreting trends in CEA rather than single tests provides improved diagnostic accuracy. However, CEA is not sufficiently accurate as a single surveillance method and alternative markers are needed.

Acknowledgements

The authors would like to acknowledge the key role of the other FACS Trial investigators in providing the data analysed in this sub-study. We also acknowledge the invaluable contribution of the local NIHR cancer research networks, NHS Trusts and of the patients who agreed to participate in this trial.

Study and Analyses Registration

The original randomised clinical trial was retrospectively registered: ISRCTN41458548.

ClinicalTrials.gov number: NCT00560365

As a secondary analysis, no preregistration exists for the analyses reported in this manuscript.

Contribution of Authors

BS designed and conducted the main analysis under the supervision of RP. BS and DM drafted the manuscript. TJ was responsible for the laboratory analysis of CEA for the FACS trial and gave scientific advice on the interpretation of the data. JP and DM initiated the FACS Trial, drafted the original protocol for this sub-study, and obtained the funding. SP provided clinical support for the trial and advice on colorectal cancer pathology and clinical care of recurrence. BN and SP helped draft the manuscript at each stage. All authors commented on more than one draft of the manuscript and approved the final draft.

Conflicts of interest

Disclosure: The authors have no conflicts of interest to declare

Funding

Both the main FACS project and this sub-study on CEA were funded by the UK National Institute for Health Research Health Technology Assessment (NIHR HTA) Programme (project numbers 99/10/99 and 11/136/81). Bethany Shinkins is currently supported by the NIHR Leeds Diagnostic Evidence Co-operative. The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the HTA programme, NIHR, NHS or the Department of Health. The authors accept full responsibility for the research.

Role of funding agency

The funding agency (UK NIHR HTA) had no role in: the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; nor decision to submit the manuscript for publication.

REFERENCES

1. Duffy MJ, Lamerz R, Haglund C, Nicolini A, Kalousova M, Holubec L, et al. Tumor markers in colorectal cancer, gastric cancer and gastrointestinal stromal cancers: European Group on Tumour Markers (EGTM) 2014 Guidelines Update. International Journal of Cancer. 2013;134(11):2513-22.
2. Labianca R, Nordlinger B, Beretta GD, Brouquet A, Cervantes A. Primary colon cancer: ESMO Clinical Practice Guidelines for Diagnosis, adjuvant treatment and follow-up. Annals of Oncology. 2010;21(Suppl 5):v70-v7.
3. Locker GY, Hamilton S, Harris J, Jessup JM, Kemeny N, MacDonald JS, et al. ASCO 2006 update of Recommendations for the use of tumour markers in gastrointestinal cancer. Journal of Clinical Oncology. 2006;24(33):5313-27.
4. Nccn. NCCN Guidelines Version 3.2013 Colon Cancer. NCCN Clinical Practice Guidelines in Oncology2013.
5. Nice clinical g. Colorectal Cancer: National Institute for Health and Clinical Excellence; 2011.
6. Primrose JN, Perera R, Gray A, Rose P, Fuller A, Corkhill A, et al. Effect of 3-5 years of scheduled CEA and CT follow-up to detect recurrence of colorectal cancer: FACS randomised clinical trial. JAMA. 2014;311(3):263-70.
7. Verberne C, Doornbos PM, Grossmann I, De Bock GH, Wiggers T, editors. Intensified follow-up in colorectal cancer patients using frequent carcino-embryonic antigen (CEA) measurements and CEA-triggered imaging. European Journal of Cancer; 2013; Amsterdam: European Cancer Congress 2013, ECC 2013.
8. Shinkins B, Nicholson BD, James TJ, Primrose JN, Mant D. Carcinoembryonic antigen monitoring to detect recurrence of colorectal cancer: how should we interpret the results? Clinical Chemistry. 2014;pii: clinchem.2014.228601:[Epub ahead of print]-[Epub ahead of print].
9. Grossmann I, Verberne C, De Bock G, Havenga K, Kema I, Klaase J, et al. The Role of High Frequency Dynamic Threshold (HiDT) Serum Carcinoembryonic Antigen (CEA) Measurements in Colorectal Cancer Surveillance: A (Revisited) Hypothesis Paper. Cancers. 2011;3(2):2302-15.
10. Minton JP, Martin EW, Jr. The use of serial CEA determinations to predict recurrence of colon cancer and when to do a second-look operation. Cancer. 1978;42(3 Suppl):1422-7.
11. Staab HJ, Anderer FA, Stumpf E, Hornung A, Fischer R, Kieninger G. Eighty-four potential second-look operations based on sequential carcinoembryonic antigen determinations and clinical investigations in patients with recurrent gastrointestinal cancer. American Journal of Surgery. 1985;149(2):198-04.
12. Carl J, Bentzen SM, Norgaard-Pedersen B, Kronborg O. Modelling of serial carcinoembryonic antigen changes in colorectal cancer. Scandinavian Journal of Clinical & Laboratory Investigation. 1993;53(7):751-5.
13. Boey J, Cheung HC, Lai CK, Wong J. A Prospective Evaluation of Serum Carcinoembryonic Antigen (CEA) Levelsin the Management of Colorectal Carcinoma. World J Surg. 1984;8(3):279-86.
14. Huang YY, Lee PI, Liu MC, Chen CC, Huang KC, Huang AT. A General Cutoff Level Combined With Personalized Dynamic Change of Serum Carcinoembryonic Antigen Can Suggest Timely Use of FDG PET for Early Detection of Recurrent Colorectal Cancer. Clinical nuclear medicine. 2015;40(10):e465-9.
15. Nicholson BD, Shinkins B, Pathiraja I, Roberts NW, James TJ, Mallett S, et al. Blood CEA levels for detecting recurrent colorectal cancer. The Cochrane database of systematic reviews. 2015(12):CD011134.
16. Blanche P, Dartigues JF, Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. Statistics in medicine. 2013;32(30):5381-97.
17. Treasure T MK, Fiorentino F, Russell C. The CEA Second-Look Trial:a randomised controlled trial of carcinoembryonic antigen prompted reoperation for recurrent colorectal cancer. BMJ Open. 2014;4:e004385-e.

18. Mant D, Gray A, Pugh S, Campbell H, George S, Fuller A, et al. A randomised controlled trial to assess the cost-effectiveness of intensive versus no scheduled follow-up in patients who have undergone resection for colorectal cancer with curative intent. Health technology assessment (Winchester, England). 2017;21(32):1-86.
19. Verberne CJ, Wiggers T, Grossmann I, de Bock GH, Vermeulen KM. Cost-effectiveness of a carcinoembryonic antigen (CEA) based follow-up programme for colorectal cancer (the CEA Watch trial). Colorectal Disease. 2016;18(3):O91-O6.

Table 1. Patient characteristics

	Total (n=582)	Recurrence (n=104)	No Recurrence (n=478)
Age (years)			
<65	185 (32%)	42 (40%)	143 (30%)
65-79	339 (58%)	53 (51%)	286 (60%)
80+	58 (10%)	9 (9%)	49 (10%)
Smoking status			
Never Smoked	232 (40%)	37 (36%)	195 (41%)
Ex-smoker	297(51%)	49 (47%)	248 (52%)
Current smoker	36 (6%)	11 (11%)	25 (5%)
Unknown	17 (3%)	7 (7%)	10 (2%)
Stage of primary cancer			
I	110 (19%)	10 (10%)	100 (21%)
II	282 (48%)	46 (44%)	236 (49%)
III	166 (29%)	45 (43%)	121 (25%)
Unknown	24 (4%)	3 (3%)	21 (4%)
Site of primary cancer			
Right colon	181 (31%)	28 (27%)	153 (32%)
Left colon	219 (38%)	34 (32%)	185 (39%)
Rectum	170 (29%)	41 (39%)	129 (27%)
Unknown	12 (2%)	1 (2%)	11 (2%)

APPENDIX

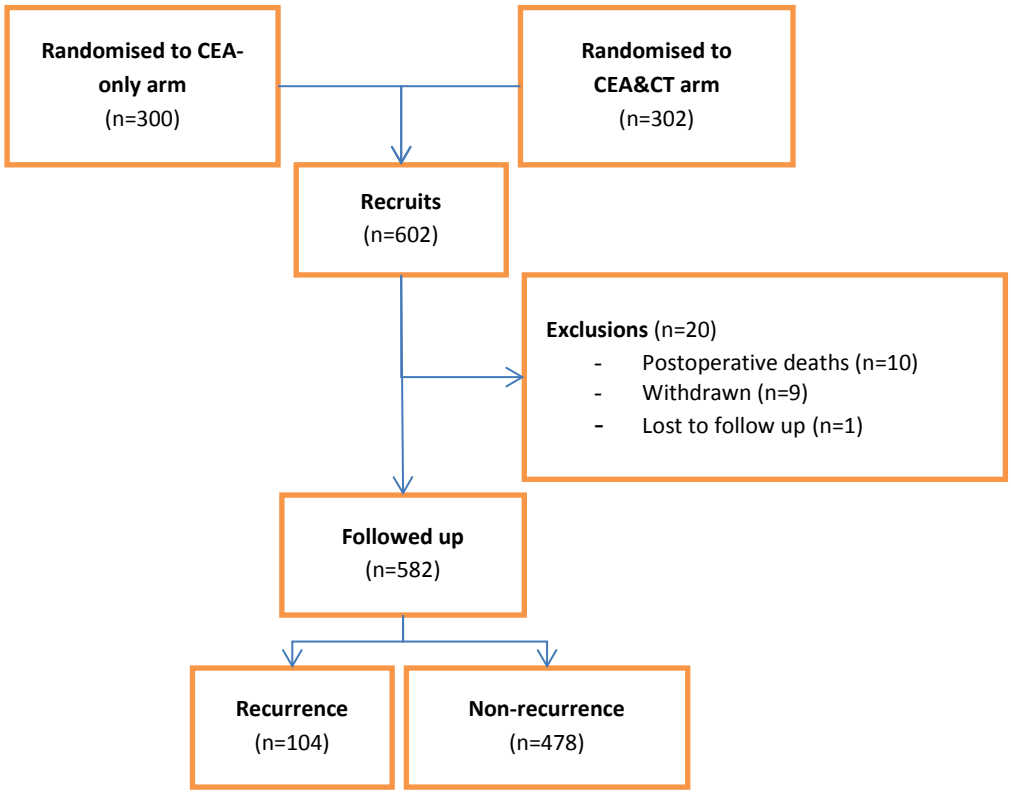
Table A1 Testing schedule in the FACS trial for the CEA & CT arm included in this analysis

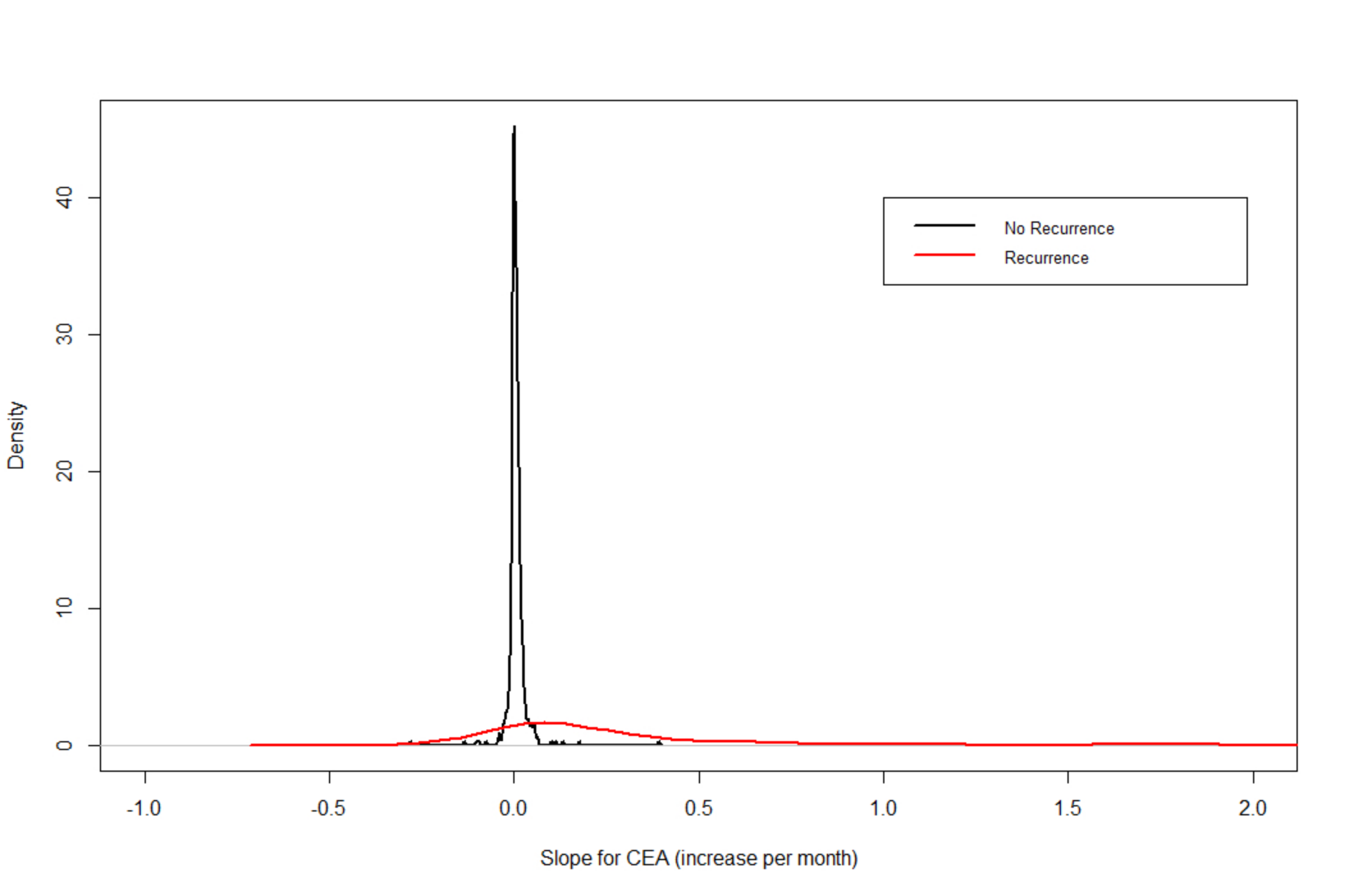
	Month of Follow-up													
Test	**3**	**6**	**9**	**12**	**15**	**18**	**21**	**24**	**30**	**36**	**42**	**48**	**54**	**60**
CEA	x	x	x	x	x	x	x	x	x	x	x	x	x	x
CT CAP		x		x		x		x	x	x	x	x	x	x
OPA Review		x		x		x		x		x		x		x
Colonoscopy														x

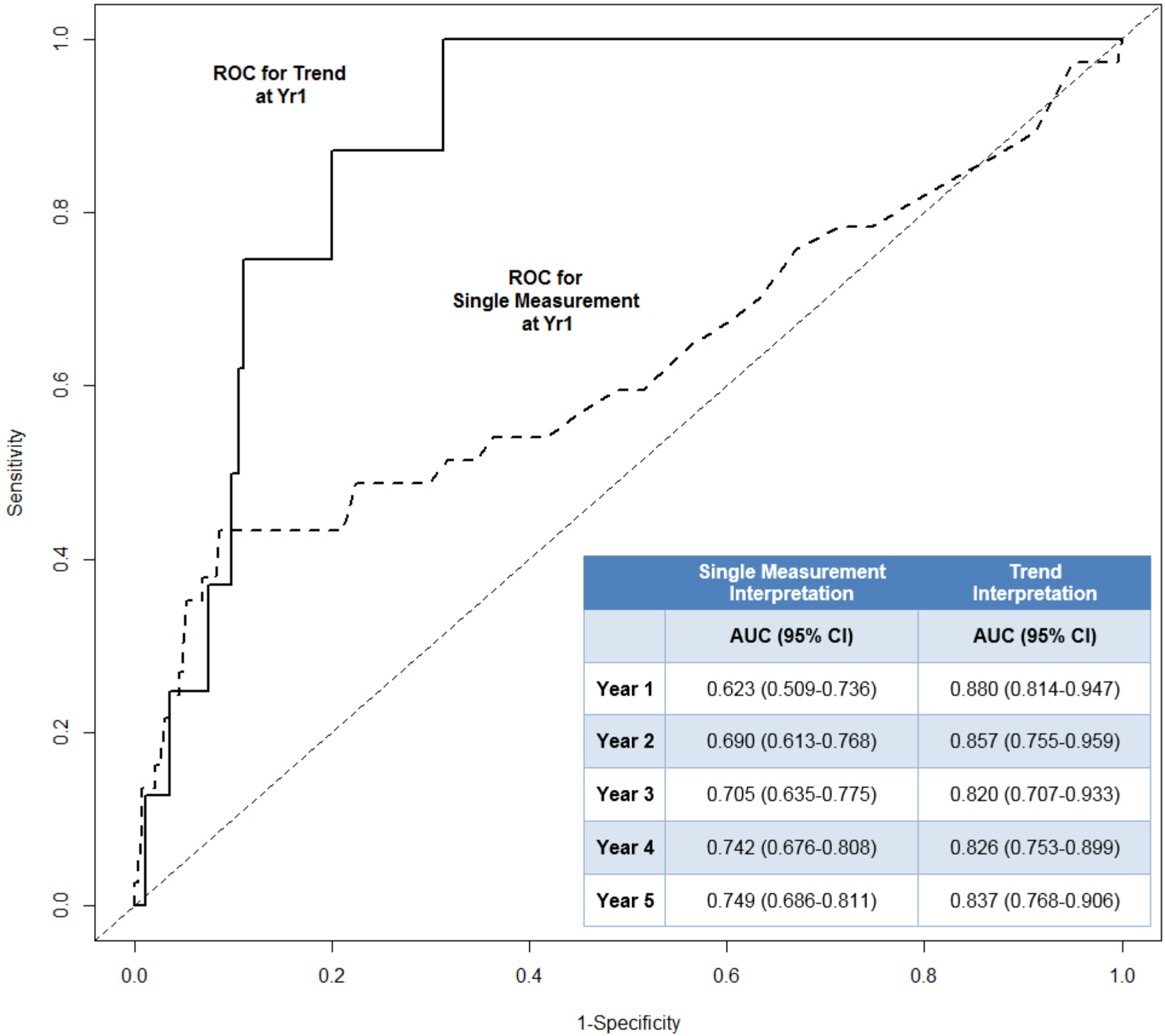
Table A2 Testing schedule in the FACS trial for the CEA only arm included in this analysis

	Month of Follow-up													
Test	**3**	**6**	**9**	**12**	**15**	**18**	**21**	**24**	**30**	**36**	**42**	**48**	**54**	**60**
CEA	x	x	x	x	x	x	x	x	x	x	x	x	x	X
CT CAP					x									
OPA Review		x		x		x		x		x		x		X
Colonoscopy								x						x

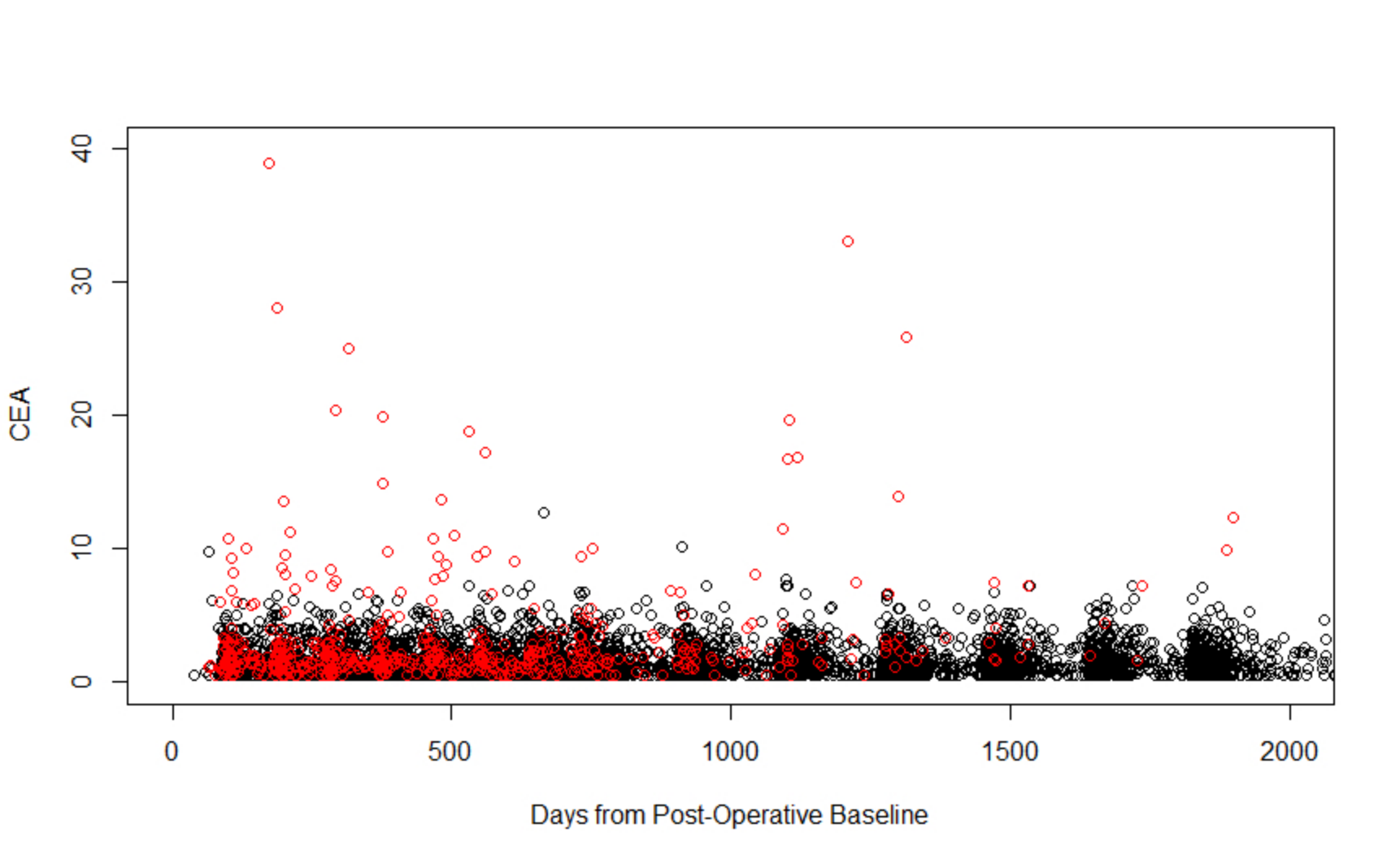
Figure A1 Flow-chart of patients allocated to CEA testing within the FACS cohort to show origin of the data analysed here




No Recurrence
Recurrence
Density
0
10
20
30
40
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
Slope for CEA (increase per month)



	Single Measurement Interpretation	Trend Interpretation
	AUC (95% CI)	AUC (95% CI)
Year 1	0.623 (0.509-0.736)	0.880 (0.814-0.947)
Year 2	0.690 (0.613-0.768)	0.857 (0.755-0.959)
Year 3	0.705 (0.635-0.775)	0.820 (0.707-0.933)
Year 4	0.742 (0.676-0.808)	0.826 (0.753-0.899)
Year 5	0.749 (0.686-0.811)	0.837 (0.768-0.906)


CEA
0
10
20
30
40
0
500
1000
1500
2000
Days from Post-Operative Baseline