NICE National Institute for Health and Care Excellence

Diagnostics Assessment Programme

Crohn's disease: Tests for therapeutic monitoring of TNF inhibitors (LISA-TRACKER ELISA kits, TNFα-Blocker ELISA kits, and Promonitor ELISA kits)

Evaluation Report

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Diagnostics Assessment Programme

Crohn's disease: Tests for therapeutic monitoring of TNF inhibitors (LISA-TRACKER ELISA kits, TNFα-Blocker ELISA kits, and Promonitor ELISA kits)

Contents:

Any information supplied to NICE which has been marked as confidential has been redacted. All personal information has also been redacted.

National Institute for Health and Care Excellence

DIAGNOSTICS ASSESSMENT PROGRAMME

Evidence overview

Therapeutic monitoring of TNF inhibitors in Crohn's disease (LISA-TRACKER ELISA kits, Immundiagnostik TNFα-Blocker ELISA kits, and Promonitor ELISA kits)

This overview summarises the key issues for the Diagnostics Advisory Committee's consideration. This document is intended to be read in conjunction with the final scope issued by NICE for the assessment and the diagnostics assessment report. A glossary of terms can be found in Appendix B.

1 Background

1.1 Introduction

People with Crohn's disease are often given TNF inhibitors, such as infliximab and adalimumab, to inhibit the activity of the cell signalling protein, TNFα, which promotes inflammatory responses. Although TNF inhibitors can bring benefits to many patients with Crohn's disease, there are some patients whose disease does not respond to treatment with TNF inhibitors (primary non-responders). In patients whose disease does respond to treatment with TNF inhibitor, the concentration of drug in the blood immediately before the next dose of TNF inhibitor (trough level) can vary widely between patients even though they received the same initial dose. Furthermore, a large proportion of patients whose disease initially responds to treatment find that their disease stops responding over time (secondary loss of response). Loss of response is thought to be due to the formation of antibodies to TNF inhibitors, and fluctuations in circulating drug levels.

Currently, treatment decisions for patients with Crohn's disease are commonly based on clinical judgement and 'trial and error', so tailoring treatment to the individual patient can be difficult. This can lead to patients experiencing treatment side-effects unnecessarily and a delay in finding optimal therapy. Measuring levels of TNF inhibitors and antibodies against TNF inhibitors in a patient's blood could help a clinician identify when to change the drug dosage or switch to an alternative TNF inhibitor.

The purpose of this assessment is to evaluate the clinical and cost-effectiveness of using ELISA kits (LISA-TRACKER ELISA kits, Immundiagnostik TNFα-Blocker ELISA kits, and Promonitor ELISA kits) to test levels of TNF inhibitors and antibodies to TNF inhibitors in the following 2 populations:

- people with Crohn's disease whose disease responds to treatment with TNF inhibitor
- people with Crohn's disease who experience secondary loss of response during maintenance treatment with TNF inhibitor.

Provisional recommendations on the use of these technologies will be formulated by the Diagnostics Advisory Committee at the Committee meeting on 26 May 2015.

1.2 Scope of the evaluation

Table 1: Scope of the evaluation

Decision questions	Does concurrent testing of TNF inhibitor levels and antibodies to TNF inhibitors represent a clinically and cost-effective use of NHS resources in people with Crohn's disease: 1. whose disease responds to treatment with TNF inhibitor? Testing will be carried out: a. 3 to 4 months after start of treatment

	b. 3 to 4 months and every 12 months from start of treatment 2. who experience secondary loss of response during maintenance treatment with TNF inhibitor? Does testing of TNF inhibitor levels followed by reflex testing of antibodies to TNF inhibitors if drug level is undetectable represent a clinically and cost-effective use of NHS resources in people with Crohn's disease: 3. whose disease responds to treatment with TNF inhibitor? Testing will be carried out: a. 3 to 4 months after start of treatment b. 3 to 4 months and every 12 months from start of treatment 4. who experience secondary loss of response during maintenance treatment with TNF inhibitor?
Populations	People with Crohn's disease who are being treated with infliximab or adalimumab.
Interventions	LISA-TRACKER ELISA kits *(Theradiag/Alpha Labs):* • LISA-TRACKER Adalimumab (LTA002) • LISA-TRACKER Infliximab (LTI002) • LISA-TRACKER anti-Adalimumab (LTA003) • LISA-TRACKER anti-Infliximab (LTI003) • LISA-TRACKER Duo Adalimumab (LTA005) • LISA-TRACKER Duo Infliximab (LTI005) Immundiagnostik TNFα-Blocker ELISA kits

	(Immundiagnosik/BioHit Healthcare): • Immundiagnostik TNFα-Blocker ADA, antibodies against infliximab (e.g. Remicade®) ELISA (K9650) • Immundiagnostik TNFα-Blocker ADA, antibodies against adalimumab (e.g. Humira®) ELISA (K9652) • Immundiagnostik TNFα-Blocker ADA, TOTAL antibodies against infliximab (e.g. Remicade®) ELISA (K9654) • Immundiagnostik TNFα-Blocker ADA, TOTAL antibodies against adalimumab (e.g. Humira®) ELISA (K9651) • Immundiagnostik TNFα-Blocker monitoring, infliximab drug level (e.g. Remicade®) ELISA (K9655) • Immundiagnostik TNFα-Blocker monitoring, adalimumab drug level (e.g. Humira®) ELISA (K9657) Promonitor ELISA kits *(Proteomika):* • Promonitor-ADL ELISA (5080230000) • Promonitor-IFX ELISA (5060230000) • Promonitor-ANTI-ADL ELISA (5090230000) • Promonitor-ANTI-IFX ELISA (5070230000)
Linked-evidence approach	Test methods that are not included as an intervention but have evidence comparing it to an intervention test and evidence reporting clinical outcomes, should be included for the purpose of performing linked evidence modelling only.

Comparator	Treatment decisions made on clinical judgement without measuring levels of TNF inhibitor and antibodies to TNF inhibitors.
Healthcare setting	Secondary and tertiary care
Outcomes	Intermediate measures for consideration may include: • Time to result • Number of inconclusive results • Frequency of dose adjustment • Frequency of treatment switch
	Clinical outcomes for consideration may include: • Measures of disease activity • Rates of response, relapse and remission • Duration of response, relapse and remission • Rates of hospitalisation • Rates of surgical intervention • Time to surgical intervention • Adverse effects of treatment • Health related quality of life
	Costs will be considered from an NHS and Personal Social Services perspective. Costs for consideration may include: • Costs of testing • Costs of treatment • Costs of other resource use – outpatient appointments – hospitalisation – additional tests – surgery
	The cost-effectiveness of interventions should be

	expressed in terms of incremental cost per quality-adjusted life year.
Time horizon	The time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared.

Further details including descriptions of the interventions, comparator, care pathway and outcomes can be found in the final scope.

2 The evidence

This section summarises data from the diagnostics assessment report compiled by the External Assessment Group (EAG).

2.1 Clinical effectiveness

The EAG conducted a systematic review of the evidence on tests to monitor levels of TNF inhibitors and antibodies to TNF inhibitors in people with Crohn's disease treated with infliximab or adalimumab. Details of the review can be found starting on page 54 of the diagnostics assessment report. The review had 4 key objectives:

- To compare the performance of the different tests available
- To compare optimal cut-off levels identified in different studies
- To analyse the correlation between test results and clinical state
- To describe and compare test-informed algorithms used in studies, and to review the clinical effectiveness of these test-informed algorithms compared with standard care (no testing performed).

In total 68 studies were included in the clinical effectiveness review. Some studies addressed more than one of the objectives of the review.

For the purpose of this assessment and to aid understanding, test have been split into 3 groups: index tests, alternative tests, and other tests. The different

tests are summarised in table 2. Because there were no direct clinical outcome data for the index tests (LISA-TRACKER ELISA kits, Immundiagnostik TNFα-Blocker ELISA kits, and Promonitor ELISA kits), the clinical effectiveness review considered alternative tests for which clinical outcome data were available. Evidence on the comparative performance of the index tests and the alternative tests was then sought in order to make a link between the index tests and the clinical outcomes. Other tests are also mentioned in the review as they form an indirect link between the index tests and clinical outcomes via the alternative tests.

Table 2: Summary of the different tests

Index tests	LISA-TRACKER ELISA kits Promonitor ELISA kits Immundiagnostik TNFα-Blocker ELISA kits	Named in the scope and are subject to recommendations by the DAC
Alternative tests	Prometheus ELISA and HMSA Leuven in-house ELISA	Form a link between the index tests and clinical outcomes.
Other tests	Amsterdam Sanquin in-house ELISA and RIA	Form a link between the index tests and the alternative tests

Tests for the therapeutic monitoring of TNF inhibitors may be performed in 2 ways:

- **Concurrent testing** is when tests for TNF inhibitor drug levels and antibodies to TNF inhibitor are performed at the same time.
- **Reflex testing** is when the test for TNF inhibitor drug levels is performed first and the result from this test is used to guide follow-up testing by the laboratory without a further request from the treating clinician. If the drug is undetectable, testing for antibodies to TNF inhibitor would be performed. If TNF inhibitor is present in the sample, testing for antibodies would not be performed.

These 2 different approaches are explored in the modelling.

Test results are used in an algorithm to decide how treatment should be managed. Different algorithms are used for different testing scenarios. For example, in people whose disease responds to treatment with TNF inhibitor:

- if drug trough levels are higher than the target range, the dosing interval of the TNF inhibitor could be increased with the aim of lowering drug trough levels
- if drug trough levels are lower than the target range, the dosing interval of the TNF inhibitor could be decreased, or the dose increased, with the aim of raising drug trough levels.

Alternatively, in people whose disease has stopped responding to TNF inhibitor:

- if drug trough levels are lower than the target range and anti-drug antibodies are detectable, switch to an alternative TNF inhibitor should be considered
- if drug trough levels are lower than the target range and anti-drug antibodies are low or undetectable, increasing the dose of the current TNF inhibitor should be considered
- if drug trough levels are in the target range, a switch to a treatment with a different mechanism of action should be considered.

Evidence on the clinical outcomes

Three studies were identified that implemented a test-informed algorithm in the management of people with Crohn's disease treated with infliximab or adalimumab, and reported clinical outcomes. Table 3 provides an overview of the study design and risk of bias of the studies. A more detailed description of the risk of bias in the studies can be found starting on page 100 of the diagnostics assessment report. A detailed overview of the features and key findings of the studies can be found starting on page 103 of the diagnostics assessment report.

Table 3: Overview of the included studies

Study	Design	Scenario	Population	Test	Risk of bias
Steenholdt et al. (2014, 2015)	RCT	Loss of response to infliximab	CD only (N = 69)	RIA	High (Cochrane risk of bias tool)
Vaughn et al. (2014)	Retrospective	Responders to infliximab	71% CD, 27% UC	ELISA and HMSA (Prometheus)	Adequate quality (Downs and Black checklist)
Vande Casteele et al. (2015)	RCT	Responders to infliximab	68% CD; 31% UC	Leuven in-house ELISA	Unclear (Cochrane risk of bias tool)
CD – Crohn's disease; UC – ulcerative colitis; RCT – randomised controlled trial; RIA – radioimmunoassay; ELISA – Enzyme linked immunosorbent assay; HMSA – homogeneous mobility shift assay.					

Key conclusions from these studies were:

- Steenholdt et al. reported that in patients with Crohn's disease who experienced a loss of response to infliximab, the clinical response in the test-algorithm group was similar to the clinical response in the standard treatment (dose intensification) group.
- Vande Casteele et al. reported that in patients with inflammatory bowel disease with a stable response to infliximab, clinical response was similar in the test-algorithm group and in the clinically-based dosing group.
- Vaughn et al. reported that in patients with inflammatory bowel disease in clinical remission on infliximab, trough concentration monitoring of infliximab resulted in a greater probability of remaining on infliximab compared with no monitoring.

Steenholdt et al. study

This was a single-blind randomised controlled trial of 69 adults with Crohn's disease on maintenance infliximab treatment with loss of response. Full details are provided starting on page 106 of the diagnostics assessment report. Participants were randomised to either an infliximab intensified arm (n=36) or to an algorithm arm (n=33). In the former, the dose frequency of 5

mg/kg infliximab was increased from every 8 weeks to every 4 weeks. In the latter, participants received treatment according to a defined algorithm based on serum concentrations of infliximab and of antibodies to infliximab. Samples were taken immediately before infliximab infusion and were analysed by radioimmunoassay. The algorithm categorised patients into 1 of 4 groups as described in table 4.

Table 4: Treatment algorithm used in the Steendholdt et al. study

Group	Drug levels	Antibody levels	Treatment	ITT population
Group 1	Sub-therapeutic infliximab	Detectable anti-infliximab antibodies	Change to a different TNF-inhibitor (adalimumab)	14 (20%)
Group 2	Sub-therapeutic infliximab	Undetectable anti-infliximab antibodies	Intensify infliximab treatment	3 (4%)
Group 3	Therapeutic infliximab	Undetectable anti-infliximab antibodies	Discontinue treatment with TNF inhibitors. Review of condition.	48 (70%)
Group 4	Therapeutic infliximab	Detectable anti-infliximab antibodies	Repeat testing. If unchanged results then act as for group 3.	4 (6%)

The primary outcomes of the study were the mean cost of treatment over 12 weeks and the proportion of patients with 'clinical response' at 12 weeks. In the dose intensification arm all patients received allocated treatment. In the algorithm arm 14 of 33 patients did not receive treatment according to the algorithm (13 in group 3; 1 in group 4). Most of these 14 patients continued to receive infliximab. There were 2 withdrawals from the algorithm arm and 8 withdrawals from the dose intensification arm.

The study shows that the majority of patients with loss of response to TNF inhibitor had therapeutic levels of TNF inhibitor and undetectable anti-infliximab antibodies (group 3). The authors' interpretation of this is that 'inhibition of TNFα is ineffective due to non-TNFα driven disease'. The External Assessment Group noted that the algorithm treatment for this group

is subject to discretion and requires further investigation and reflection by clinicians.

In the intent to treat population (n = 69), clinical response at week 12 was observed in 53% (19/36) of patients in the dose intensification arm and in 58% (19/33) of patients in the algorithm arm. In the per-protocol population (n = 55) clinical response at week 12 was observed in 53% (19/36) of patients in the dose intensification arm and in 47% (15/33) of patients in the algorithm arm. Mean costs in the intent to treat population were €9178 in the dose intensification arm compared with €6038 in the algorithm arm. Mean costs in the per-protocol population were €9178 in the dose intensification arm compared with €4062 in the algorithm arm.

An extension study reported clinical findings to week 20 and mean costs to week 52. In the intent to treat population, clinical response at 20 weeks was observed in 56% of the dose intensification arm and in 76% of the algorithm arm. In the same population, remission was achieved at week 20 in 39% of patients in the dose intensification arm, and in 55% of patients in the algorithm arm. Mean costs at week 20 and week 52 were $17,236 and $29,072 respectively in the dose intensification arm. In the algorithm arm mean costs at week 20 and week 52 were $11,940 and $22,066 respectively.

Vaughn et al. (2014)

This was a retrospective observational pilot study of patients with inflammatory bowel disease in clinical remission receiving infliximab. Full details are provided starting on page 114 of the diagnostics assessment report. Patients were identified from records and classified into those who received proactive drug monitoring and those who did not (control group). Samples were analysed initially by ELISA (Prometheus) and later with a homogeneous mobility shift assay (Prometheus). In the proactive monitoring group, serum trough levels were used to guide infliximab dose modifications to achieve target drug levels according to the algorithm presented in table 5. Reactive testing was performed in both groups if there was loss of response or there was a concern for side effects due to antibody formation.

Table 5: Treatment algorithm used in the Vaughn et al. study

Test result	Treatment
Undetectable trough levels	Dose of infliximab increased to 7.5 mg/kg and next infusion given after 6 weeks, before returning to every 8 weeks.
Detectable trough level, but < 5 micrograms/mL	Dose of infliximab increased by 50 or 100 mg
Trough drug levels of >10 micrograms/mL on at least 2 occasions	Infliximab dose reduced
Trough drug level between 5 and 10 micrograms/mL	No changes made

There were 48 and 78 patients in the proactive drug monitoring group and the control group, respectively. Key outcomes reported were: time remaining on treatment and reasons for stopping infliximab. In the proactive drug monitoring group infliximab dose was adjusted in 35% (17/48) of patients following initial testing (71% dose escalation, 18% dose decrease, and 12% stopped infliximab). Following subsequent proactive tests dose was adjusted in 25% (12/48) of patients (80% dose escalation and 20% dose decrease).

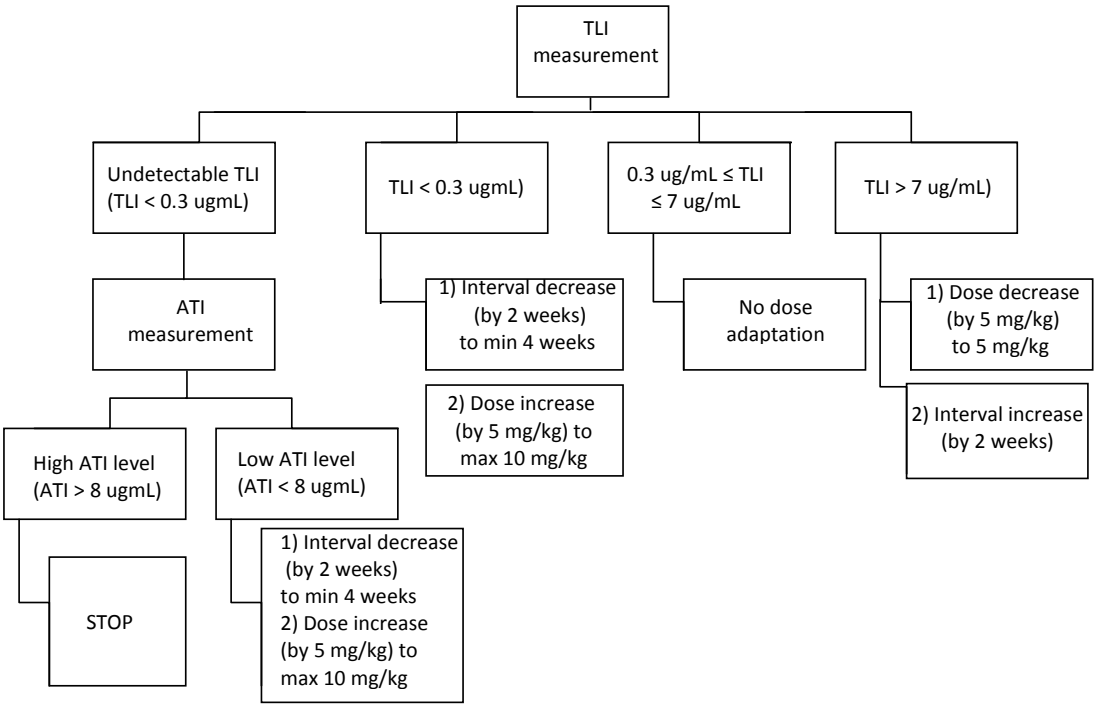
At 5 years, the probability of remaining on treatment was 86% in the proactive drug monitoring group, and 52% in the control group. Regression analysis found that the probability of patients remaining on infliximab treatment was related only to proactive drug monitoring of infliximab. In the control group, the main reasons for stopping infliximab treatment were recurrence of symptoms and acute infusion reactions. In the proactive drug monitoring group, the main reasons from stopping infliximab treatment were adverse events and high antibody levels.

Vande Casteele et al. (2015) – the TAXIT trial

This was a randomised controlled trial of 251 patients with inflammatory bowel disease (173 with Crohn's disease and 78 with ulcerative colitis). Full details are provided starting on page 119 of the diagnostics assessment report. Patients were randomised to clinically-based dosing or to infliximab trough concentration based dosing. Prior to randomisation patients were screened and underwent an optimisation phase. Therefore all randomised patients entered the maintenance phase of the study with trough infliximab levels in

the target range of 3 to 7 micrograms per ml. In the clinically-based dosing arm, all subsequent infliximab dosing was according to clinical symptoms and C-reactive protein levels. In the trough concentration-based dosing arm, all subsequent infliximab dosing was according to the algorithm presented in figure 1.

Figure 1: Treatment algorithm from the TAXIT trial (TLI – trough level infliximab; ATI – antibodies to infliximab)



The primary outcome was the rate of clinical plus biological remission 1 year after randomisation. Secondary outcomes included: durable remission, relapse, trough infliximab levels in target range, anti-drug antibody positivity, quality of life (EQ-5D), and total cost of treatment. Samples were analysed using Leuven in-house ELISAs.

In the optimisation phase, 74% of patients with Crohn's disease were in remission prior to dose optimisation, and 80% were in remission after optimisation. Dose escalation was performed in 43 of 178 patients and the percentage of patients in remission in this group increased from 65% to 88%. Dose reduction was performed in 51 of 178 patients and the percentage of

patients in remission in this group decreased from 80% to 69%. For the dose escalation group an average of 2.1 optimisations were required to reach target trough infliximab levels, and at the end of optimisation the median infusion interval was 6 weeks (range 4 to 8 weeks). For the dose reduction group an average of 1.4 optimisations were required and the median infusion interval was 8 weeks (range 6 to 12 weeks).

In the maintenance phase, similar rates of clinical remission were seen in both arms of the trial: 69% in the concentration based dosing arm, and 66% in the clinically based dosing arm ($p = 0.686$). When restricted to patients with Crohn's disease, rates of clinical remission were 63% in the concentration based dosing arm, and 55% in the clinically based dosing arm ($p = 0.353$).

There was little difference between groups in the probability of maintaining durable remission (26% and 27% in concentration based and clinically based dosing arms, respectively). More patients in the concentration-based dosing arm than in the clinically based dosing arm (74% compared with 57%) had an infliximab trough concentration between 3 and 7 micrograms per ml whereas the risk of patients in the clinically based arm having undetectable trough levels of infliximab was significantly greater, relative risk 3.7 (95% confidence intervals 1.7 to 8.0). None of the patients in the concentration-based dosing arm were positive for anti-drug antibodies but 3 patients in the clinically based arm had anti-drug antibodies.

No deaths occurred in either group, but 2 patients in the clinically based dosing arm required hospital admission, one due to acute appendicitis and another due to ileostomy complications. There were 12 discontinuations in the clinically based dosing arm and 13 discontinuations in the concentration-based dosing arm. More patients in the clinically based arm (17%) than in the concentration-based dosing group (7%) relapsed and needed rescue therapy (relative risk 2.4, 95% CI 1.2 to 5.1).

Evidence on the comparative performance of different tests

The comparative performance of the index tests (LISA-TRACKER ELISA kits, Immundiagnostik TNFα-Blocker ELISA kits or Promonitor ELISA kits) with alternative tests which do have data on clinical outcomes was reviewed. Data comparing the performance of the 3 index tests was also assessed. There were 14 studies which had relevant test comparisons, of which 5 reported concordance as numerical data or Cohen's kappa. In addition an unpublished analysis of data was provided by a company.

Comparisons between the index tests

Comparisons between the index tests are described in detail starting on page 75 of the diagnostics assessment report. In summary:

- Adalimumab levels
  - In an analysis using [redacted], the Immundiagnostik TNFα-Blocker ELISA showed [redacted], the Promonitor ELISA showed [redacted], and the LISA-TRACKER ELISA showed [redacted]. [redacted] between the Promonitor ELISA and the LISA-TRACKER ELISA and [redacted] between the Promonitor ELISA and the Immundiagnostik TNFα-Blocker ELISA.
  - In an analysis using both patient samples and spiked samples, test results were different between the Promonitor ELISA and the LISA-TRACKER ELISA, and Pearson R^2 was 0.83. Authors concluded that the LISA-TRACKER ELISA underestimated adalimumab levels (Nagore et al. 2015).

- Antibodies to adalimumab
  - In an analysis using [redacted], the [redacted] between the Promonitor ELISA and the spiked value.
  - The analysis by Nagore et al. reports a Cohen's Kappa of 0.8 between the Promonitor ELISA and the LISA-TRACKER ELISA.

- Infliximab levels
  - In an analysis using [redacted], the Immundiagnostik TNFα-Blocker ELISA showed [redacted], the Promonitor ELISA showed [redacted], and the LISA-TRACKER ELISA showed [redacted]. For all 3 tests the level of bias appears [redacted]. [redacted] between the Promonitor ELISA and the LISA-TRACKER ELISA and [redacted] between the Promonitor ELISA and the Immundiagnostik TNFα-Blocker ELISA.
  - In an analysis using both patient samples and spiked samples, test results were different between the Promonitor ELISA and the LISA-TRACKER ELISA, and Pearson R^2 was 0.98. Authors concluded that the LISA-TRACKER ELISA overestimated infliximab levels (Nagore et al. 2015).
  - A study of 66 patient samples showed that results from the Immundiagnostik TNFα-Blocker ELISA were on average 1.8 micrograms per ml lower than results from the Promonitor ELISA (Daperno et al. 2013).

- Antibodies to infliximab
  - In an analysis using [redacted], the [redacted] between the Promonitor ELISA and the spiked value.
  - The analysis by Nagore et al. (2015) reports a Cohen's Kappa of 1.0 between the Promonitor ELISA and the LISA-TRACKER ELISA, indicating complete agreement.
  - The study by Daperno et al. (2013) found that test results from the Immundiagnostik TNFα-Blocker ELISA and the Promonitor ELISA were 'identical' in only 6 out of 63 cases.

Based on the limited evidence on the correlation between the 3 index tests, it appears that the LISA-TRACKER ELISAs have the most variation in test

results compared with the Immundiagnostik TNFα-Blocker ELISAs and Promonitor ELISAs. However, it is not clear how this would impact test results at clinically meaningful cut-off points.

Comparisons between the index tests and the alternative tests

Comparisons between the index tests and the alternative tests are described in detail starting on page 79 of the diagnostics assessment report. In summary:

- LISA-TRACKER ELISAs
  - One study was identified which has data on the LISA-TRACKER ELISAs and the Leuven in-house ELISAs for infliximab and antibodies to infliximab (Vande Casteele et al. 2012). This study also included the Amsterdam Sanquin ELISA and radioimmunoassay. A mix of clinical and spiked samples was used. Results suggest that the LISA-TRACKER ELISA may give some false positive results for infliximab levels in the presence of antibodies to infliximab or adalimumab. For detecting antibodies to infliximab, the LISA-TRACKER ELISA gave fewer positive results than the radioimmunoassy, but more positive results than the Leuven in-house ELISA. However it is not clear if these results are true positive.
  - There were no data linking the LISA-TRACKER ELISAs to any of the alternative tests for detection of adalimumab or antibodies to adalimumab.
- Promonitor assays
  - One study compared the Promonitor ELISAs with the Amsterdam Sanquin ELISA and radioimmunoassay (Ruiz-Arguello et al. 2013), and a further study compared the Amsterdam Sanquin ELISA and radioimmunoassay with the Leuven in-house ELISA (Vande Casteele et al. 2012), giving an indirect link between the index test and the alternative test.
  - Ruiz-Arguello et al. (2013) used spiked samples and results suggest that for drug levels, although analytical sensitivity of the Amsterdam Sanquin

ELISA was higher than that of the Promonitor ELISA, the Amersterdam Sanquin ELISA may overestimate drug levels at higher drug concentrations. For anti-drug antibodies, the analytical sensitivity of the Promonitor ELISA was higher than that of the Amsterdam Sanquin radioimmunoassay.

  - Vande Casteele et al. (2012) reported that the Amsterdam Sanquin ELISA and the Leuven in-house ELISA for drug levels perform similarly across all cut-off used. However, the Amsterdam Sanquin radioimmunoassay gave more positive results for anti-drug antibodies than the Leuven in-house ELISA.

- Immundiagnostik TNFα-Blocker ELISAs
  - Two studies compared the Immundiagnostik TNFα-Blocker ELISAs with the Prometheus HMSA (Eser et al. 2013a and 2013b). The Immundiagnostik ELISAs were compared to the Amsterdam Sanquin ELISA and radioimmunoassay in 1 study (Schatz et al. 2013), and Vande Casteele et al. (2012) compared the Amsterdam Sanquin ELISA and radioimmunoassay with the Leuven in-house ELISAs.
  - Eser et al. (2013a and 2013b) used patient samples and report that the Prometheus HMSA was able to detect anti-infliximab antibodies in the presence of infliximab, whereas the Immundiagnostik ELISA returned inconclusive results due to interference from infliximab.
  - Schatz et al. (2013) used patient samples and reported agreement between the Immundiagnostic ELISA and the Amsterdam Sanquin ELISA for infliximab levels as a Cohen's Kappa of 0.792. More positive results were returned by the Amsterdam Sanquin tests than the Immundiagnostik ELISAs for both infliximab levels and antibodies to infliximab.
  - There were no data linking the Immundiagnostik ELISAs to any of the alternative tests for detection of adalimumab or antibodies to adalimumab.

In conclusion, there was insufficient evidence linking any of the index tests (LISA-TRACKER, Immundiagnostik or Promonitor) to any of the alternative

tests with links to clinical outcomes (HMSA, radioimmunoassay, Prometheus ELISA, or Leuven in-house ELISA).

Evidence on optimal cut-off levels

The range of cut-offs reported across the included studies illustrates that no validated threshold has been established to date. Cut-offs strongly depend on the assay used, the drug measured, the clinical marker investigated and the time of testing.

- ROC threshold analyses to determine optimal cut-off levels predictive of clinical response for infliximab, adalimumab or both, were reported in 24 studies. Full details of these studies can be found in the diagnostics assessment report starting on page 89.
- Different studies used different markers to assess clinical response, which is the 'reference standard' to determine the accuracy of the test.
- When identifying optimal cut-off levels, some studies aimed for high sensitivity (0.90) at the expense of specificity (0.37), while other favoured high specificity (1.00) at the expense of sensitivity (0.33).
- Reported cut-offs for infliximab ranged from 0.6 and 7 micrograms per ml.
- Reported cut-offs for adalimumab ranged from 3 to 6.85 micrograms per ml.

Evidence on the correlation between test results and clinical state

The test accuracy of drug level tests and anti-drug antibodies tests as predictors of clinical status was moderate. Positive and negative predictive values across clinical prevalence ranges indicated that 20% to 30% of test results were incorrect.

The review identified 34 studies which reported on the relationship between test results and the clinical status of patients with Crohn's disease or inflammatory bowel disease. Of these, 3 were systematic reviews which included a meta-analysis, and 31 were primary studies. The systematic reviews are described in detail starting on page 126 of the diagnostics

assessment report, and the primary studies are described in detail starting on page 132 of the diagnostics assessment report.

Systematic reviews

The systematic reviews all included meta-analyses which addressed the risk of having a positive or negative test result in patients with a particular clinical state. For example, in patients with loss of response what is the risk of a negative test result compared with the risk of a negative test result in patients without loss of response? Conversely, what is the risk of loss of response in patients who have a negative test result compared with patients who have a positive test result? Data from these studies were also analysed to give test accuracy parameters, with the clinical state used as the reference standard.

Nanda et al. (2013) included 11 studies in the meta-analysis and report a 3-fold greater risk of experiencing loss of response in patients with a positive anti-drug antibodies test result compared with patients with a negative anti-drug antibodies test result (3.16 [95% CI 2.00 to 4.98]). Hierarchical meta-analysis of all studies gave a sensitivity of 0.70 (95% CI 0.55 to 0.82) and specificity of 0.81 (95% CI 0.67 to 0.89) for the anti-drug antibody test in predicting loss of response. At a loss of response prevalence of 34.7%, the positive predictive value was calculated as 65% and the negative predictive value as 84%.

Lee et al. (2012) included 10 studies in the meta-analysis and reported no significant decrease in rates of remission in patients with a positive test result for anti-drug antibodies compared with patients with a negative test result for anti-drug antibodies (0.96 [95%CI 0.77 to 1.19]). Hierarchical meta-analysis of study data gave a sensitivity of 0.42 and specificity of 0.69 for the anti-drug antibody test in predicting remission.

Lee et al. (2012) also examined the association between the development of anti-drug antibodies and the use of immunosuppressant therapies. Meta-analysis of 11 studies indicated a 50% reduction in risk of developing anti-

drug antibodies when immune suppressants are administered (0.50 [95%CI 0.42 to 0.59]).

Paul et al. (2014) included 3 studies in adults and 2 studies in children and reported significantly greater odds of experiencing a lack of clinical response in patients with sub-therapeutic adalimumab levels compared with patients with therapeutic levels of adalimumab (2.60 [95% CI 1.79 to 3.77]). They also reported significantly greater odds of experiencing a lack of clinical response in patients with antibodies to adalimumab present compared with patient with no antibodies to adalimumab identified (10.15 [95% CI 3.90 to 26.40]).

Primary studies

Primary studies were reviewed in order to identify information which could be useful for the economic modelling. Studies were included if they provided dichotomised test results (that is, positive or negative) and related these to dichotomised clinical state (that is, response or lack of response). The review identified 31 relevant studies and full results are presented starting on page 133 of the diagnostics assessment report.

Only 3 studies reported the results of both drug level tests and anti-drug antibody tests for individual patients. These studies allowed estimation of the proportions of patients that would enter each of the treatment categories following concurrent or reflex testing for use in the economic model. Test results from these studies were compared with test results from the meta-analysis of multiple single test studies. The External Assessment Group concluded that results from the patient level studies were sufficiently similar to results from the meta-analysis, and therefore reasonably representative of the patient populations of interest.

2.2 Costs and cost effectiveness

The External Assessment Group conducted a search to identify existing studies investigating the cost effectiveness of LISA-TRACKER ELISA kits, Immundiagnostik TNFα-Blocker ELISA kits, and Promonitor ELISA kits for measuring levels of TNF inhibitors and of anti-drug antibodies.

Systematic review of cost effectiveness evidence

There were 4 studies identified which investigated the cost effectiveness of different assays for measuring levels of TNFα inhibitors and of anti-drug antibodies. These studies are described in detail starting on page 148 of the diagnostics assessment report.

Vande Casteele et al. (2015) conducted a randomised controlled trial to determine whether concentration-based infliximab dosing is more cost effective than clinically-based infliximab dosing in people with moderate-to-severe Crohn's disease or ulcerative colitis (TAXIT trial). The design of the trial and the clinical results are described in section 2.1 of this overview. The time horizon of the model was one year and the perspective was that of the third-party payer. The base-case results demonstrated that concentration-based dosing was slightly less effective (0.8227 versus 0.8421) and less costly (€20,700 versus €21,000) than clinically-based dosing, but overall differences were small.

Steenholdt et al. (2014) assessed the cost-effectiveness of receiving treatment based on serum concentrations of infliximab and infliximab antibodies compared with receiving infliximab at an increased dose frequency of 5 mg/kg every 4 weeks in patients with loss of response to infliximab while on maintenance treatment. The study design and the clinical results are described in section 2.1 of this overview. Authors report that costs at 12 weeks were significantly lower in the algorithm group than in the infliximab intensification group. Mean costs at 12 weeks were €6038 in the algorithm group compared with €9178 in the infliximab intensification group.

Steenholdt et al. conducted a follow-up to the original study which extended the time horizon to 1 year in order to assess the long-term costs and clinical outcomes of treatment of Crohn's disease in people with loss of response to infliximab maintenance therapy (Steenholdt et al. 2015). Costs were assessed at the 20-week scheduled trial visit and again at one year. Clinical outcomes were assessed after 20 weeks. Authors report that the algorithm group had significantly lower costs than the infliximab intensification group at the 20

week follow-up and this was maintained throughout the year. At 20 weeks average costs in the algorithm group were US$11,900 compared with US$17,200 in the infliximab intensification group. At 1 year average costs in the algorithm group were US$22,100 compared with US$29,100 in the infliximab intensification group.

Velayos et al. (2013) used a decision analytical model to assess the cost-effectiveness of a testing-based strategy with an empiric dose escalation strategy for patients with moderate-to-severe Crohn's disease who experience loss of response to infliximab. The study had a third party payer perspective and a 1 year time horizon. The base-case results showed that that the testing strategy was cheaper and marginally more effective than the empiric dose escalation strategy.

In summary, all of these studies indicated that a testing strategy might be less costly than alternatives with variable small effects on effectiveness, with some indicating small reduced benefits and some small increased benefits.

Economic analysis

The External Assessment Group also constructed two de novo economic models designed to assess the cost-effectiveness of employing TNF inhibitor and anti-drug antibody monitoring with LISA-TRACKER ELISA kits, Immundiagnostik TNFα-Blocker ELISA kits, and Promonitor ELISA kits in patients with Crohn's disease compared with standard care. The first model focuses on patients who respond to infliximab maintenance therapy and the second model focuses on patients who experience a loss of response to infliximab maintenance therapy. The models are described in detail starting on page 158 of the diagnostics assessment report.

Model structure

Illustrative structures of the responder model and the loss of response model are presented in figure 2 and figure 3 respectively. Both models have a 10 year time horizon with a 4-week cycle length. The models start with a

hypothetical cohort of 30-year olds with moderate to severe Crohn's disease. A description of each health state is provided in table 6.

Figure 2: Illustrative structure for the responder model

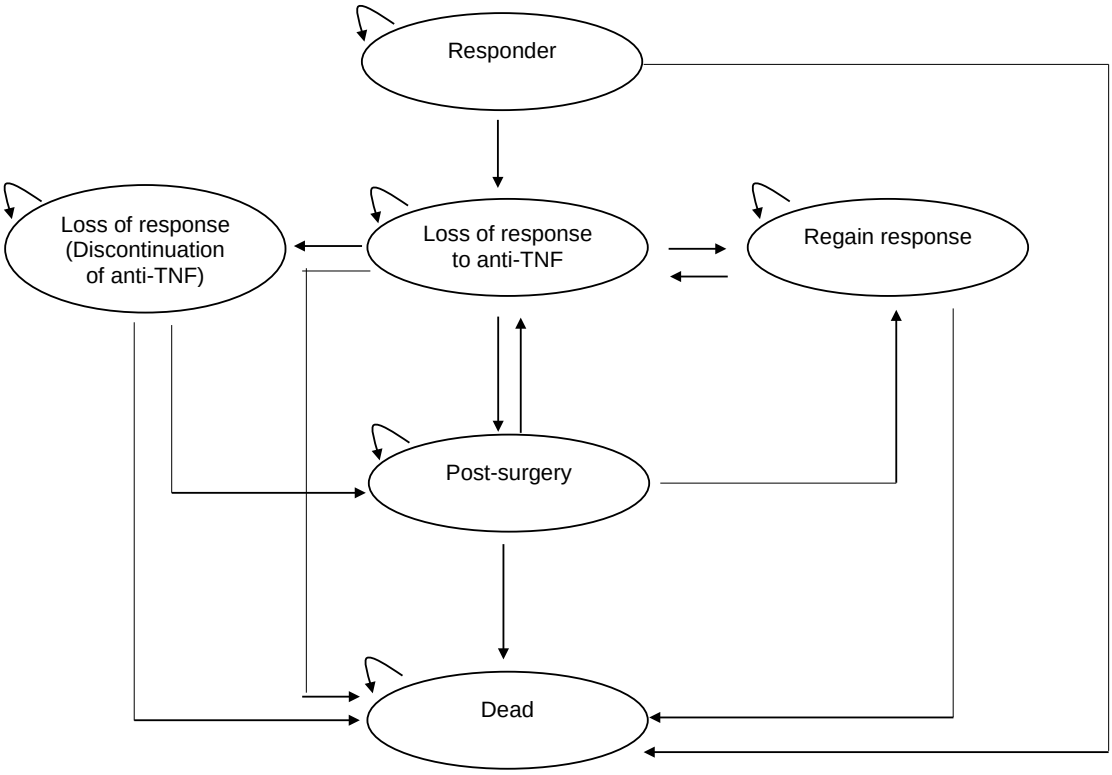


Figure 3: Illustrative structure for the loss of response model

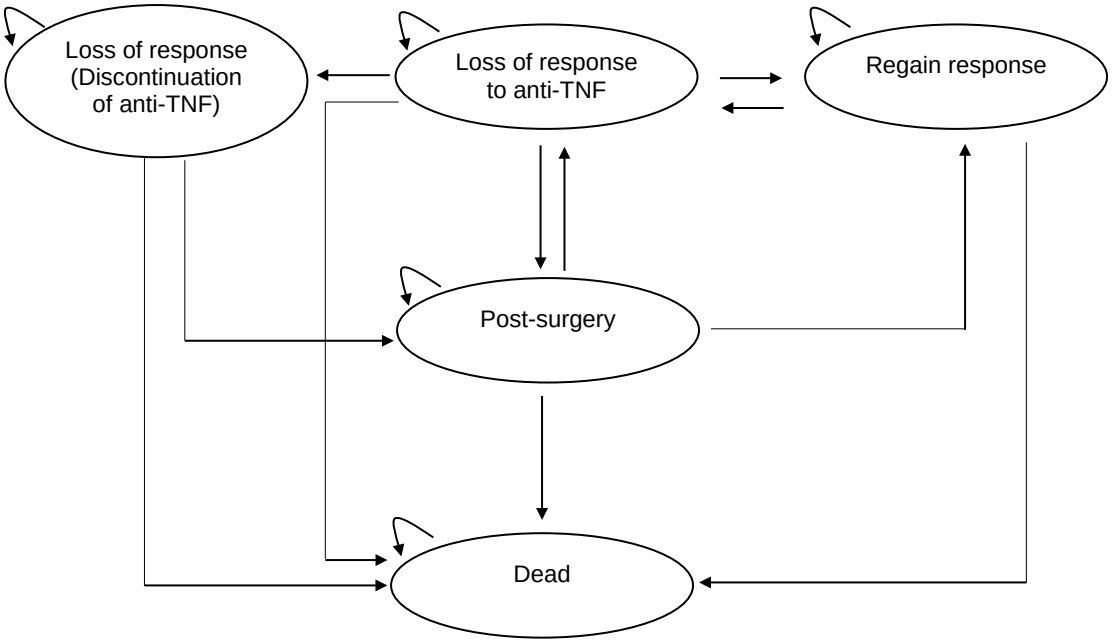


Table 6: Definitions of the health states in the model

Health state	Description
Responder	Maintenance treatment when the patient has supportable active symptoms of abdominal pain, diarrhoea, rectal bleeding or weight loss.
Loss of response	Recurrence of active symptoms while on treatment with maintenance regime, after having responded to treatment.
Loss of response (no anti-TNF)	Recurrence of active symptoms having discontinued anti-TNFα treatment, but receiving best supportive care. People remaining in this state do not require surgery. People who develop active symptoms that require surgery then move to the post-surgery health state or die.
Re-gain response	Maintenance treatment when the patient has no active symptoms having previously lost response.
Post-surgery	After surgery, treatment options are to receive: anti-TNF, immunosuppressant, a combination of anti-TNF and immunosuppressant or no treatment. Patients who receive an anti-TNF alone or in combination will re-enter the model in the re-gain response state or the loss of response state. Patients who receive an immunomodulator or no treatment will remain in the post-surgery state until they require further surgery or they die.
Death	By definition

In each model, patients can receive either:

- standard care
- treatment according to an algorithm based on concurrent testing
- treatment according to an algorithm based on reflex testing.

In the standard care pathway:

- people categorised as responders continue receiving infliximab maintenance therapy every 8 weeks until they lose response
- people who lose response will receive an increase in their dose
- as a result of increasing the dose, people may regain response or continue with loss of response
- people who continue with loss of response will receive another agent in addition to their current treatment

- as a result of adding another agent, people may regain response or continue with loss of response
- people who continue with loss of response will receive a switch to their anti-TNFα treatment
- people whose disease does not respond to the new anti-TNFα will be considered for surgery.

In the concurrent testing pathway:

Tests for infliximab levels and antibodies to infliximab would be performed at the same time. Patients would fall into 1 of 4 categories:

- drug absent and antibodies present
- drug absent and no antibodies
- drug present and antibodies present
- drug present and no antibodies.

For responders, the treatment options for patients falling into each of these categories are based on the algorithm used in the TAXIT trial by Vande Casteele et al. and are listed in table 7. Following treatment in the model, patients may remain responders, lose response or die. For patients with loss of response, treatment options for patients falling into each of the categories are based on the algorithm used in the study by Steenholdt et al. and are listed in table 8.

Table 7: Treatment algorithm for responders (concurrent testing)

Category	Treatment
Drug absent, antibodies present (> 8 mg/mL)	Switch TNF inhibitor
Drug absent, no antibodies (< 8 mg/mL)	Increase dose of current TNF inhibitor
Drug present, antibodies present	If trough level below the target range - decrease in the dosing interval If trough level is within the target range - no dose adaptation If trough level is above the target range - increase in the dosing interval
Drug present, antibodies absent	

Table 8: Treatment algorithm for loss of response (concurrent testing)

Category	Treatment
Drug absent, antibodies present	Switch TNF inhibitor
Drug absent, no antibodies	Increase dose of current TNF inhibitor
Drug present, antibodies present	Treatment with TNF inhibitor discontinued and best supportive care provided
Drug present, antibodies absent	

In the reflex testing pathway:

A test for infliximab levels is performed first. If the drug is absent, a test for antibodies to infliximab would be performed. If the drug is present, no further testing would be done. Patients would fall into 1 of 3 categories:

- drug absent and antibodies present
- drug absent and no antibodies
- drug present.

No studies were identified that used an algorithm based on reflex testing, therefore, the same algorithms were used as for concurrent testing (tables 7 and 8).

Model inputs

The model was populated with data from the clinical effectiveness review and supplemented with information from secondary sources. Where data were lacking, values were obtained from clinical experts. Full details of the data used in the model can be found starting on page 167 of the diagnostics assessment report.

For patients in the responder and loss of response states, the proportions that fall into each of the test result categories are listed in table 9.

Table 9: Proportions of patients with each test result

Health state	Test result category	Proportion (concurrent testing)	Proportion (reflex testing)	Source
Responder	Drug absent, antibodies present	0.17241	0.17241	Imaeda et al. 2012
	Drug absent, antibodies absent	0.12069	0.12069	
	Drug present	0.70690	0.70690	
Loss of response	Drug absent, antibodies present	0.1515	0.2029	Steenholdt et al. 2014
	Drug absent, antibodies absent	0.0303	0.0435	
	Drug present, antibodies present	0.0303	0.7536	
	Drug present, antibodies absent	0.7879		
The Steenholdt study had 33 patients in the intervention arm which were used to generate the proportions in the concurrent testing model as these proportions tie in with the clinical data. In addition, the study had 69 patients in total, which were used to generate the proportions in the reflex testing model (no associated clinical data). The Imaeda study was not a comparative study (there is not an intervention and a control arm), therefore proportions for the concurrent testing model and the reflex testing model are the same.				

For patients with detectable trough drug levels, the proportions with below target range, within target range and above target range were based on the study by Vande Casteele et al. (2015) (table 10).

Table 10: Proportions according to infliximab trough levels

	Threshold	Proportion
Below target range	< 3 micrograms per mL	0.2310
Within target range	3 to 7 micrograms per mL	0.4821
Above target range	> 7 micrograms per mL	0.2869

Patients who have undergone surgery may receive post-operative treatment. The proportions of patients receiving the different treatment options are based on a study by Van der Have et al. (2014) and are listed in table 11.

Table 11: Treatment following surgery

Post-surgery treatment	Proportion
TNF inhibitor	0.1250
Immunosuppressant	0.5000
TNF inhibitor plus immunosuppressant	0.1250
No treatment	0.2500

Costs

Costs were obtains from standard sources such as the British National Formulary and NHS Reference cost database. Costs and resource use are summarised in table 12 and are described in detail starting on page 172 of the diagnostics assessment report and in appendix 18. The test costs used in the model were based on the LISA-TRACKER ELISA kit costs provided by the company, however costs of the other index tests were similar (table 13).

Table 12: Costs and resource use

Variable	Base-case value (£)
Monitoring infliximab	21.74
Monitoring antibodies to infliximab (reflex testing)	41.98
Monitoring infliximab and antibodies to infliximab (concurrent testing)	38.83
Maintenance infliximab	1966.41
Maintenance adalimumab	704.28
Azathioprine	8.40
Mercatopurine	100.94
Predinsolone	14.25
Nutritional therapy (Modulen)	15.06
Laparoscopic ileocolic resection	6908
Responder	725.69
Loss of response	1241.38
Regain response	725.69
Post-surgery	790.69

Table 13: Test costs

Test	Price	Patient samples tested	Cost per patient
LISA-TRACKER Drug level ELISA	£850	42	£20.24
LISA-TRACKER anti-drug antibodies ELISA	£850	42	£20.24
LISA-TRACKER Duo	£1568	2 x 42	£37.33
Immundiagnostik TNFα-blocker Drug level ELISA	£855	40	£21.38
Immundiagnostik TNFα-blocker anti-drug antibodies ELISA	£775	45	£17.22
Immundiagnostik TNFα-blocker ADA, total anti-drug antibodies ELISA	£775	45	£17.22
Promonitor drug level ELISA	£800	40	£20.00
Promonitor anti-drug antibodies ELISA	£800	40	£20.00

Health related quality of life and QALY decrements

Utility weights were taken from published literature and are presented in Table 14. The utility values reported in Velayos et al. (2013) were obtained from the study undertaken by Gregor et al. (1997). Gregor and colleagues compared various elicitation techniques (standard gamble, time trade-off and visual analogue scale) on 180 consecutive Crohn's disease patients. The authors suggested that the standard gamble technique reflected the true value for health states related to people with Crohn's disease.

Table 14: Utility values and sources

Health State	Utility	Source
Responder	0.77	Velayos et al. (2013)
Loss of response	0.62	Derived from Gregor et al. (1997)
Regain response	0.77	Assumption
Surgery	0.60	Marchetti et al. (2014)
Post-surgery	0.86	Velayos et al. (2013)

Assumptions

For the purposes of decision making, the ICERs per QALY gained or lost will be considered. The following assumptions were applied in the base case analysis:

- Patients have received intravenous infusions of infliximab of 5 mg/kg at week 0, 2 and 6.
- Patients weigh more than 70kg
- Patients who regained response have the same utility as those who are considered to be responders
- People with Crohn's disease are not at increased risk of dying from the disease over the lifetime of the model, and there is no difference in mortality between the test-algorithm group and the standard care group.
- For people who have undergone surgery, there is an increased risk of 0.0015 of dying from the procedure.
- The treatment effects for people receiving a dose escalation (from 5mg/kg to 10mg/kg of infliximab) and a decreased interval (from eight week to six week intervals) are the same.
- People who are categorised as a responder and who have trough concentrations within the range that the treatment algorithm suggests receive no dose adaptation.
- Transition probabilities in the test-algorithm group are the same as the transition probabilities in the standard care group for the following transitions:
  - Loss of response to infliximab maintenance therapy (Juillerat et al. 2015)
  - Loss of response in people with dose escalation (Ma et al. 2014)
  - Loss of response to adalimumab maintenance therapy (Karmaris et al. 2009).
- People who remain in the loss of response health state (discontinuation of anti-TNF) have symptoms of Crohn's disease that in time may require surgery. People will receive best supportive care until active symptoms develop that require surgery.

In addition, the testing schedules in the base case models were as follows:

- In the responder model testing was performed every 3 months whilst patients' disease was responding to TNF inhibitor. If patients experienced

loss of response to TNF inhibitor they would also be tested every 3 months until TNF inhibitor was discontinued.

- In the loss of response model, patients experiencing loss of response were tested on entry into the model. If they regained response they would then enter onto the 3 monthly testing regimen. If they continued to experience loss of response to TNF inhibitor they would also be tested every 3 months until TNF inhibitor was discontinued.

It should be noted that the testing schedules in the base case analyses do not match the testing schedules defined in the decision questions set in the scope. Therefore when considering the results from the cost-effectiveness analyses attention should be focused on the alternative scenario analyses which better reflect the decision questions.

Base-case results

The diagnostics assessment report provides 2 sets of base case results. Those presented in the diagnostics assessment report starting on page 176 use time-to event transition probabilities. A revised set of base case results are presented in the diagnostics assessment report addendum which use exponential transition probabilities (which assume the hazard rate does not change over time). These different sets of transition probabilities reflect different assumptions on the stage of Crohn's disease when a patient enters the model. The revised base case results are presented in this overview as the External Assessment Group state that the constant hazard transition probabilities appear to be more appropriate for the model.

Base case results for the responder model are presented in Table 15. Results show that a standard care strategy with no testing is slightly more expensive and more effective than either reflex testing or concurrent testing of drug levels and anti-drug antibodies. ICERs show that if testing strategies were adopted, savings of between £43,700 and £50,800 would be made for each QALY lost.

Base case results for the loss of response model are presented in Table 16. Results show that a standard care strategy with no testing, costs more than reflex testing or concurrent testing of drug levels and anti-drug antibodies, but is more effective. ICERs show that if testing strategies were adopted, savings of between £273,000 and £284,100 would be made for each QALY lost.

Table 15: Base case results for the responder model

Strategy	Mean cost (£)	Incremental costs vs no testing (£)	QALYs	Incremental QALYs vs no testing	ICER (£) (testing vs no testing)
No testing	150,500	-	6.5084	-	-
Reflex testing	138,700	-11,800	6.2761	-0.2323	43,700
Concurrent testing	139,800	-10,700	6.2637	-0.2447	50,800

Table 16: Base case results for the loss of response model

Strategy	Mean cost (£)	Incremental costs vs no testing (£)	QALYs	Incremental QALYs vs no testing	ICER (£) (testing vs no testing)
No testing	215,800	-	6.4961	-	-
Reflex testing	131,000	-84,800	6.1976	-0.2985	284,100
Concurrent testing	129,400	-86,100	6.1807	-0.3154	273,000

Analysis of alternative scenarios

The external assessment group performed a number of scenario analyses to explore the different scenarios presented in the decision questions. These are presented in the diagnostics assessment report starting on page 178, and in the addendum).

Responder model

The scenario analyses of the responder model included:

- Testing performed annually in patients who disease responds to treatment with TNF inhibitor

- Testing performed initially at 3 months and then annually in patients whose disease responds to treatment with TNF inhibitor
- Testing performed only at 3 months in patients who disease responds to treatment with TNF inhibitor, and in patients who regain response following a loss of response to TNF inhibitor treatment
- Testing performed only at 3 months in patients who disease responds to treatment with TNF inhibitor (no testing of patients who regain response following a loss of response to TNF inhibitor treatment).

In all 4 of the strategies listed above, testing is also performed in patients who experience a loss of response to TNF inhibitor. Testing is performed every 3 months until the patient regains response to TNF inhibitor, or discontinues treatment with TNF inhibitor.

In all these additional scenario analyses, results show that the testing strategies are cheaper and less effective than the standard care strategy of no testing (Table 17 to table 20). ICERs show that if testing strategies were adopted, savings of between £78,900 and £176,300 would be made for each QALY lost.

These additional scenario analyses are based on data from the TAXIT trial, in which patients were tested at each infusion (every 4 to 12 weeks). The strategy of a single test at 3 months followed by annual testing does not correspond with any available data. These results should therefore be treated with caution.

Table 17: Responder model with annual testing

Strategy	Mean cost (£)	Incremental costs vs no testing (£)	QALYs	Incremental QALYs vs no testing	ICER (£) (testing vs no testing)
No testing	150,500	-	6.5084	-	-
Reflex testing	114,100	-36,400	6.2281	-0.2803	129,900
Concurrent testing	114,000	-36,500	6.2201	-0.2883	126,600

Table 18: Responder model with initial testing at 3 months then annual testing for responders and those who regain response

Strategy	Mean cost (£)	Incremental costs vs no testing (£)	QALYs	Incremental QALYs vs no testing	ICER (£) (testing vs no testing)
No testing	150,500	-	6.5084	-	-
Reflex testing	113,400	-37,100	6.2290	-0.2794	132,800
Concurrent testing	113,800	-36,700	6.2244	-0.2840	129,200

Table 19: Responder model with testing at 3 months in responders and people who regain response

Strategy	Mean cost (£)	Incremental costs vs no testing (£)	QALYs	Incremental QALYs vs no testing	ICER (£) (testing vs no testing)
No testing	150,500	-	6.5084	-	-
Reflex testing	103,000	-47,500	6.2390	-0.2694	176,300
Concurrent testing	102,000	-48,500	6.2255	-0.2829	171,400

Table 20: Responder model with testing at 3 months in responders only

Strategy	Mean cost (£)	Incremental costs vs no testing (£)	QALYs	Incremental QALYs vs no testing	ICER (£) (testing vs no testing)
No testing	150,500	-	6.5084	-	-
Reflex testing	102,900	-47,600	6.2255	-0.2829	168,300
Concurrent testing	102,000	-48,500	6.2255	-0.2829	171,400

Loss of response model

The scenario analysis of the loss of response model examined a test schedule where patients experiencing a loss of response to TNF inhibitor receive testing, but patients who regain response to treatment with TNF inhibitor do not receive testing. Testing is performed every 3 months until the patient regains response to TNF inhibitor, or discontinues treatment with TNF

inhibitor. Results show that testing strategies are cheaper than standard care with no testing, but less effective (table 21). ICERs show that if testing of patients who experience a loss of response was adopted, savings of between £340,900 and £354,500 could be made per QALY lost.

This additional scenario analysis is based on data from the study by Steenholdt et al. (2014) where people with loss of response were tested once and followed up for 12 weeks.

Table 21: Loss of response model with testing only in patients with loss of response

Strategy	Mean cost (£)	Incremental costs vs no testing (£)	QALYs	Incremental QALYs vs no testing	ICER (£) (testing vs no testing)
No testing	215,800	-	6.4961	-	-
Reflex testing	97,700	-118,100	6.1630	-0.3331	354,500
Concurrent testing	96,200	-119,600	6.1453	-0.3508	340,900

Sensitivity analyses

In addition to the scenario analyses, a range of univariate sensitivity analyses were performed and results are presented starting on page 178 of the diagnostics assessment report and in the addendum. These included:

- Changing the time horizon from 10 years to 1 year.
- Changing the transition probabilities from exponential transition probabilities (which assume the hazard rate does not change over time) to time to event transition probabilities.
- Transition probabilities from Juillerat Weibull and Vande Casteele at al. (2015) were used.
- Reducing the proportion of people with infliximab and antibodies to infliximab present from 0.7878 to 0.200.
- Patients did not regain response following best supportive care.

Most of these changes had no impact on the direction of the results. However, changing the transition probabilities from exponential transition probabilities to time to event transition probabilities, did impact the responder model, resulting in the no testing strategy becoming cheaper and more effective than the testing strategies (table 22). Also, if patients did not regain response following best supportive care, this resulted in the testing strategies becoming more expensive than the no testing strategy. The difference in QALYs also reduced, but the no testing strategy remained more effective than the testing strategies (table 23).

Table 22: Responder model using time to event transition probabilities

Strategy	Mean cost (£)	Incremental costs vs no testing (£)	QALYs	Incremental QALYs vs no testing	ICER (£) (testing vs no testing)
No testing	137,600	-	6.5146	-	-
Reflex testing	145,900	8300	6.3315	-0.1831	Dominated
Concurrent testing	147,100	9500	6.3215	-0.1931	Dominated

Table 23: Responder model when patients do not regain response following best supportive care

Strategy	Mean cost (£)	Incremental costs vs no testing (£)	QALYs	Incremental QALYs vs no testing	ICER (£) (testing vs no testing)
No testing	150,550		6.5084		
Reflex testing	158,300	7750	6.4813	-0.0271	Dominated
Concurrent testing	160,800	10,250	6.4813	-0.0271	Dominated

In further sensitivity analyses, key model input parameters were varied to determine which inputs influence the ICER. Results are presented on page 180 of the diagnostics assessment report, and show that the models are stable to most changes, but sensitive to a 10% increase in the utility value for people who regain response.

Probabilistic sensitivity analyses were also performed on the base case models and are presented in the addendum to the diagnostics assessment

report. In the responder model, the scatterplot shows considerable uncertainty around both the incremental costs and incremental QALYs. The cost effectiveness acceptability curve suggests that at a willingness to pay of £20,000 per QALY a no testing strategy is 50% likely to be the most cost effective strategy. It should be noted however that this analysis is of the base case model in which testing was performed every 3 months.

In the loss of response model the scatterplot shows less uncertainty in the incremental costs, but considerable uncertainty in the incremental QALYs. The cost effectiveness acceptability curve suggests that at a willingness to pay of £20,000 per QALY there is no preference between a no testing strategy and a testing strategy. However, at higher willingness-to-pay thresholds (greater than 30,000 per QALY) no testing is likely to be the most cost-effective strategy. Again, it should be noted that this analysis is using the base case model in which patients whose disease regained response to TNF inhibitor were tested every 3 months, in addition to testing those with loss of response.

3 Issues for consideration

Clinical effectiveness

- There are no direct clinical outcome data on any of the index test kits (LISA-TRACKER ELISAs, Promonitor ELISAs and Immundiagnostik TNFα Blocker ELISAs). All direct clinical outcome data comes from studies of alternative tests, therefore all of the economic modelling depends on the assumption that the index tests are equivalent to the alternative tests. The External Assessment Group assessed the validity of this assumption by reviewing the comparative performance of the index and the alternative tests. However, a link between the index tests and the alternative tests based on concordance data could not be established. It therefore remains uncertain to what extent the outcomes of the assessment apply to the 3 index test kits.

- All 3 clinical effectiveness studies were based on the use of infliximab; no studies of patients on adalimumab were identified. Therefore the clinical and cost-effectiveness of therapeutic monitoring of adalimumab and antibodies to adalimumab levels is unknown.
- There were no studies in children, it is therefore uncertain if the clinical and cost-effectiveness results would apply to this population.
- There is little difference in the cost-effectiveness results between the concurrent test strategy and the reflex test strategy. It is likely that different test centres will have different preferences for testing strategy. For example, in centres performing a small number of tests, a reflex test strategy could result in delay to test results due to the need to wait for a full batch of anti-drug antibody samples before running the test. Therefore, a concurrent test strategy would be the most effective strategy. However, in large test centres running many tests, a reflex test strategy may be the most effective way of working.
- There is no agreed algorithm for use in the NHS, and none of the clinical effectiveness studies were conducted in the UK. Furthermore, several studies revealed that the algorithm prescribing treatment was not followed by clinicians in many cases. This reflects the fact that the heterogeneity of symptoms in Crohn's disease, the relapsing and remitting disease pattern and the personal preferences of clinicians and individual patients mean that it is difficult to establish a standardised pathway for patients with Crohn's disease. Therefore a successful algorithm would likely incorporate multiple factors rather than just a single test result, to allow a personalised approach to optimal TNF inhibitor treatment.
- There are several issues around the ELISA kits and interpretation of test results which should be considered. Firstly the lack of a validated test threshold; secondly the predictive performance of the assays as tests for clinical status appears to show that a considerable number of patients will have a false positive or false negative result; thirdly there is no gold standard for the assessment of response in Crohn's disease patients, meaning that studies use different definitions for response, remission and

relapse. All these issues impact on the generalisability of the clinical effectiveness studies.

Cost effectiveness

- The clinical outcome data that were available were limited to 3 studies, of which 2 were relevant to the decision questions. This meant that some aspects of the decision questions were not addressed by the available evidence, and therefore assumptions needed to be made in the modelling. For example, the Steenholdt et al. study provided the treatment algorithm and outcome data for the loss of response model. This study used a concurrent testing strategy. There were no studies of therapeutic monitoring of drug and anti-drug antibody levels which used a reflex testing strategy. Therefore in the modelling, the algorithm used in the Steenholdt et al. study was adapted to reflex testing.
- In addition, populating the economic model using outcome and test data from the 2 relevant studies was problematic because of their small population size, short duration, and difficulties in allocating outcomes to categories of patients returning different defined test results. This meant that inputs for the model had to be drawn from many different sources, some of which may have different patient populations and others presented incomplete or ambiguous information regarding drug and anti-drug antibody levels. This may result in uncertainties in the modelling results.
- There is uncertainty in the utility values used in the model, many of which are based on an old study by Gregor et al. (1997). Further, the utility for patients who have regained response is assumed to be the same as the utility for primary responders.
- In the models, the testing strategies are always less effective than the standard care strategy with no testing. Incremental QALY losses for the testing strategies compared with the no testing strategy range from 0.0137 (in the responder model with a 1 year time horizon and testing every 3 months) to 0.3508 (loss of response model with a 10 year time horizon and testing only in people who have loss of response). The scatterplots from the probabilistic sensitivity analyses show that there is considerable

uncertainty in the incremental QALYs in both the responder model and the loss of response model.

- Utility decrements and costs associated with adverse events occurring as a result of drug treatment and complications following surgery were not included in the model. This may have led to an underestimation of costs across both treatment arms.
- One way sensitivity analyses showed that the model is stable to small changes in most of the model input parameters. However, the model is sensitive to changes in the utility value for people who regain response after experiencing a loss of response to TNF inhibitor treatment. A value for this specific utility could not be found in the literature. Therefore it was assumed to be 0.77, which is the same utility value as for people whose disease is responding to treatment with a TNF inhibitor.

4 Equality considerations

NICE is committed to promoting equality of opportunity, eliminating unlawful discrimination and fostering good relations between people with particular protected characteristics and others.

No potential equality issues have been identified.

5 Implementation

Key considerations for the implementation of assays to measure TNF inhibitor levels and anti-drug antibody levels include:

- Assays should only be performed in CPA accredited laboratories.
- If laboratories are testing large numbers of samples they might want to consider using an automated platform, which may be an additional cost.
- If laboratories are only testing a few samples this may impact on the turnaround time of results if samples are collected to run in batches.
- Alternatively, testing could be centralised in a few laboratories. Any laboratory offering a testing service should be able to demonstrate the stability of the sample.

- The ELISA technique does not require additional training to be provided. However, laboratory staff may need to develop skills in interpreting results.
- The results should be reported alongside the test method used to produce the results as different tests require interpretation at different thresholds.
- If a significant number of laboratories offer testing, an external quality assurance scheme may be required.
- TNF inhibitors and the tests to measure levels of TNF inhibitors and anti-drug antibodies may be funded by separate departments. Therefore, if use of the tests results in better allocation of treatment, the department funding the test may incur costs while the department funding the treatment may save costs. Communication and collaboration will be required between departments to resolve this potential issue.

6 Authors

Frances Nixon
Topic Lead

Sarah Byron
Technical Adviser

May 2015

Appendix A: Sources of evidence considered in the preparation of the overview

A. The diagnostics assessment report for this assessment was prepared by Warwick Evidence:

Freeman, K., Connock, M., Taylor-Phillips, S., Auguste, P., Mistry, H., Shyangdan, D., Court, R., Arasaradnam, R., Sutcliffe, P., Clarke, A. Crohn's disease: Tests for therapeutic monitoring of TNFα inhibitors (LISA-TRACKER ELISA kits, TNFα-Blocker ELISA kits, and Promonitor ELISA kits). Diagnostic Assessment Report commissioned by the NIHR HTA Programme on behalf of the National Institute for Health and Care Excellence.

B. The following organisations accepted the invitation to participate in this assessment as stakeholders. They were invited to attend the scoping workshop and to comment on the diagnostics assessment report.

Manufacturer(s) of technologies included in the final scope:

- Alpha Laboratories
- Biohit Healthcare
- Immundiagnostik AG
- Proteomika S.L.U.

Other commercial organisations:

- AbbVie Ltd
- Euro Diagnostica AB
- Matriks Biotek
- Merck Sharp and Dohme
- Viapath

Professional groups and patient/carer groups:

- British Society of Gastroenterology
- British Society of Paediatric Gastroenterology Hepatology and Nutrition
- Crohn's and Colitis UK

- Pelvic Pain Support Network
- Royal College of Nursing
- Royal College of Pathologists
- Royal College of Physicians
- UK Clinical Pharmacy Association (UKCPA)

Research groups:

None

Associated guideline groups:

None

Others:

- Department of Health
- Healthcare Improvement Scotland
- NHS England
- Royal Devon and Exeter Foundation NHS Trust
- Sandwell and West Birmingham Hospitals NHS Trust
- Welsh Government

Appendix B: Glossary of terms

Adalimumab

A recombinant human anti-TNFα IgG1 monoclonal antibody

Anti-drug antibodies

Antibodies produced by the body in an immune response against a therapeutic antigen, for example a monoclonal antibody, which may inactivate the drug and modify the pharmacokinetic characteristics of the drug

Hierarchical meta-analysis

A statistical framework for performing meta-analysis of diagnostic test accuracy data

Immunosuppressants

A class of drugs used to supress of prevent an immune response

Inflammatory bowel disease

A group of inflammatory conditions of the colon and small intestine, the two most common being Crohn’s disease and ulcerative colitis

Infliximab

A chimeric (human-murine) anti-TNFα IgG1 monoclonal antibody

Primary non-response

A lack of improvement of clinical signs and symptoms during induction therapy

Secondary loss of response

Loss of clinical response to therapy in patients whose disease had initially demonstrated clinical response

Severe active Crohn's disease

Very poor general health and one or more symptoms such as weight loss, fever, severe abdominal pain and usually frequent (3 to 4 or more) diarrhoeal stools daily. People with severe active Crohn's disease may or may not develop new fistulae or have extra-intestinal manifestations of the disease. This clinical definition normally, but not exclusively, corresponds to a Crohn's

Disease Activity Index (CDAI) score of 300 or more, or a Harvey-Bradshaw score of 8 to 9 or above.

TNFα

An inflammatory cytokine which helps to regulate the immune system, but when present in high concentrations it is responsible for the destructive inflammatory processes that occur in inflammatory bowel disease

TNF inhibitors (Anti-TNFα drugs)

Biological therapies which target the TNFα protein with the aim of modifying the inflammatory disease process

Trough level

The lowest concentration reached by a drug before the next dose is administered

Diagnostic Assessment Report commissioned by the NIHR HTA Programme on behalf of the National Institute for Health and Care Excellence – Final report

Title of project:

Clinical and cost-effectiveness of use of therapeutic monitoring of TNFα inhibitors (LISA-TRACKER ELISA kits, TNFα-Blocker ELISA kits, and Promonitor ELISA kits) versus standard care in people with Crohn's disease: systematic reviews and economic modelling

Name of External Assessment Group (EAG) and project lead:

Produced by:	Warwick Evidence
Lead author:	Karoline Freeman
Co-authors:	Martin Connock
	Peter Auguste
	Sian Taylor-Phillips
	Hema Mistry
	Deepson Shyangdan
	Rachel Court
	Ramesh Arasaradnam
	Paul Sutcliffe
	Aileen Clarke

Correspondence to: Dr Paul Sutcliffe
Associate Professor
Deputy Director for Warwick Evidence
Populations, Evidence and Technologies
Division of Health Sciences
Warwick Medical School
University of Warwick
Coventry CV4 7AL

Tel: 02476 150189
Fax: 02476 528375
Email: p.a.sutcliffe@warwick.ac.uk

Date completed: 23 April 2015

Source of funding:
This report was commissioned by the NIHR HTA Programme as project number 14/69/03.

Aileen Clarke and Sian Taylor-Phillips are partly supported by the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care West Midlands at the University Hospitals Birmingham NHS Foundation Trust.

Declared competing interests of the authors:
Aileen Clarke is one of the editors of the Health Technology Assessment journal. All other authors have no conflicts of interest.

Acknowledgements:
The authors would like to thank Peter Irving, Anjan Dhar, Joanna Sheldon, Anja St.Clair-Jones, Lisa Younge and Rebecca Harmston for their expert clinical input during the development of the model structure, Tara Gurung for supporting the systematic review process, Joshua Pink for his advice on the health economic modelling, Casper Steenholdt and Niels Vande Casteele for providing helpful additional information on their studies on request and Emma Loveman and Jill Colquitt, Effective Evidence LLP (http://effectiveevidence.org) for their helpful comments on the draft report.

Rider on responsibility for report:
The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:
Freeman, K., Connock, M., Auguste, P, Taylor-Phillips, S, Mistry, H, Shyangdan, D, Court, R., Arasaradnam, R., Sutcliffe, P., Clarke, A. Crohn's disease: Clinical and cost-effectiveness of use of therapeutic monitoring of TNFα inhibitors (LISA-TRACKER ELISA kits, TNFα-Blocker ELISA kits, and Promonitor ELISA kits) versus standard care in people with Crohn's disease: systematic reviews and economic modelling. *Diagnostic Assessment Report commissioned by the NIHR HTA Programme on behalf of the National Institute for Health and Care Excellence.*

Contributions of authors:
Karoline Freeman (Research Fellow) coordinated the review and wrote the introduction and discussion. Rachel Court (Information specialist) developed the search strategy and undertook searches. Martin Connock (Senior Research Fellow), Sian Taylor-Phillips (Assistant Professor), Deepson Shyangdan (Research Fellow), Paul Sutcliffe (Associate Professor) and Karoline Freeman conducted the clinical effectiveness systematic review, this included: screening and retrieving papers,

assessing against the inclusion criteria, appraising the quality of papers and abstracting data from papers for synthesis. Hema Mistry (Assistant Professor) and Peter Auguste (Research Assistant) undertook the health economic work. Martin Connock conducted the data analysis. Paul Sutcliffe (Associate Professor) and Aileen Clarke (Professor of Public Health) provided project management and editing input and Aileen Clarke provided clinical and methodological input and wrote the abstract and summary. Ramesh Arasaradnam (Associate Professor of Gastroenterology) provided clinical comment and guidance. All authors were involved in writing draft versions of the report.

Academic and commercial in confidence information:

Please note that throughout the report academic in confidence (AIC) information is marked yellow and underlined and commercial in confidence (CIC) information is marked blue and underlined.

Table of contents

List of tables

List of figures

List of abbreviations

5-ASA	5-aminosalicylic acid
abs	Antibodies
ACCENT	A Crohn's Disease Clinical Trial Evaluating Infliximab in a New Long-Term Treatment Regimen
ADA	Adalimumab
ADAbs	Anti-Drug Antibodies
AGA	American Gastroenterology Association
AL	Algorithm
ATI	Antibodies to infliximab
AU/mL	Arbitrary Unit
AUC	Area Under the Curve
BNF	British National Formulary
BNFC	British National Formulary For Children
BSG	British Society of Gastroenterology
c	Commercial
CD	Crohn’s Disease
CDAI	Crohn’s Disease Activity Index
CDEIS	Crohn’s Disease Endoscopic Index Of Severity
CENTRAL	Cochrane Central Register of Controlled Trials
CHARM	Crohn’s Trial of the Fully Human Antibody Adalimumab for Remission Maintenance
CHEERS	Consolidated Health Economic Evaluation Reporting Standards
CI	Confidence Interval
CIC	Commercial in Confidence
ClinBD	Clinically-Based Dosing
ConBD	Concentration-Based Dosing
CR	Clinical Remission
CRP	C-Reactive Protein
CT	Computerised Tomography
DARE	Database of Abstracts of Reviews of Effects
DDW	Digestive Diseases Week
dL	Decilitre
DOR	Diagnostic Odds Ratio
ECCO	European Crohn’s And Colitis Organisation
EIA	Enzyme Immunoassay
ELISA	Enzyme Linked Immunosorbent Assay
EMA	European Medicines Agency
EQ-5D	EuroQol five-dimension scale
Fab	Fragment antigen binding
FC	Faecal Calprotectin

Fc	Fragment Crystallising
FDA	Food and Drug Administration
Hb	Hemoglobin
HBI	Harvey Bradshaw Index
HC	Heavy Chains
HMSA	Homogeneous Mobility Shift Assay
HPLC	High Performance Liquid Chromatography
HR	Hazard Ratio
HRQOL	Health Related Quality of Life
HSROC	Hierarchical Summary Receiver Operating Characteristic
HTA	Health Technology Assessment
I^2	Statistical Heterogeneity Unexplained by Chance
IBD	Inflammatory Bowel Disease
IBDQ	Inflammatory Bowel Disease Questionnaire
IC	Indeterminate Colitis
ICER	Incremental Cost-Effectiveness Ratio
IFX	Infliximab
IgG	Immunoglobulin G
II	Infliximab Intensification
INAHTA	International Network Of Agencies For Health Technology Assessment
IPD	Individual Patient Data
IQR	Interquartile Range
ITT	Intention To Treat Analysis
IV	Intravenous therapy
LC	Light Chains
LOQ	Limit Of Quantification
LOR	Loss Of Response
MA	Meta-Analysis
MEDLINE	International Prospective Register Of Systematic Reviews
mg/L	Milligrams per Liter
MH	Mucosal Healing
NA	Not Applicable
nc	Non-Commercial
NHS	National Health Service
NHS EED	National Health Service Economic Evaluation Database
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NPV	Negative Predictive Value
NR	Not Reported
PANTS	Personalised anti-TNF therapy in Crohn’s disease

PDAI	Perianal Disease Activity Index
PMS	Postmarketing Surveillance
Pop	Population
PP	Per Protocol Analysis
PPV	Positive Predictive Value
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PSA	Probabilistic Sensitivity Analysis
PSS	Personal Social Services
QALYs	Quality-Adjusted Life Years
QUADAS-2	Quality Assessment of Diagnostic Accuracy Studies
RCT	Randomised Controlled Trial
RES	Response
RevMan	Review Manager
RFIPC	Rating Form of Inflammatory Bowel Disease Patient Concerns
RIA	Radio-Immunoassay
ROC	Receiver Operating Characteristic
RR	Relative Risk
sc	Subcutaneous
SD	Standard Deviation
SE	Standard Error
SE-HPLC	Size Exclusion High Performance Liquid Chromatography
SEM	Standard Error of Mean
Sens	Sensitivity
SF-6D	Short Form 6D
Spec	Specificity
sROC	Summary Receiver Operating Characteristic
S-S	Disulphide Bonds
TAILORIX	Trial Assigning IndividuaLized Options for Treatment (Rx)
TAXIT	Trough level Adapted infliXImab Treatment
TL	Trough Level
TLI	Trough Level infliximab
TNFα	Tumour Necrosis Factor Alpha
TP	Transition Probability
TRA	Trough Level Adalimumab
UC	Ulcerative Colitis
UEGW	United European Gastroenterology Week
UK	United Kingdom
UKCRN	United Kingdom Clinical Research Network
WBC	White Blood Cells
WHOICTRP	World Health Organization International Clinical Trials Registry Platform

µg/L	Micrograms per Liter
µg/mL	Micrograms per milliliter

ABSTRACT

Background and Objectives

To undertake systematic reviews and economic modelling of clinical and cost effectiveness of use of therapeutic monitoring of TNFα inhibitors (LISA-TRACKER ELISA kits, TNFα-Blocker ELISA kits, and Promonitor ELISA kits) versus standard care in people with Crohn's disease

Methods

Multiple electronic databases were searched from inception to October / November / December 2014. Supplementary searches were run to check for additional studies. We aimed to identify primary studies and systematic reviews with meta-analysis meeting the following inclusion criteria:

- Population: People with moderate to severe active Crohn's disease treated with infliximab or adalimumab
- Intervention: Monitoring of serum anti-TNFα (infliximab or adalimumab) and or of anti-drug antibody levels using intervention tests or other test assays implemented in a test-treatment algorithm
- Comparator: Standard care (no monitoring)
- Outcomes: Any patient related outcome, test agreement, cost-effectiveness estimates

Study quality assessments were undertaken using recognised checklists (QUADAS-2, Cochrane, Philips, CHEERS). Evidence was synthesised using narrative review and meta-analysis methods where appropriate.

A *de novo* Markov model was built in TreeAge Pro 2013. The model had a four week cycle and a ten year time horizon and adopted an NHS and PSS perspective. A linked evidence approach was adopted to populate the model. Costs were adjusted to 2013/14 prices and discounted at 3.5%.

Results

We included 68/2,434 and 4/2,466 studies for the clinical and cost effectiveness reviews, respectively. Twenty three studies comparing test methods were identified. Evidence on test concordance was sparse and contradictory offering scant data for a linked evidence approach. Three studies, (2 RCTs, 1 retrospective observational study), investigated outcomes following implementation of a test-algorithm. None used the intervention tests. Neither of the two RCTs found evidence of clinical benefit for a test-treatment regimen. Thirty one studies were meta-analysed to estimate test accuracy for predicting clinical status indicating that between 20 and 30% of positive and negative test results

are likely to be inaccurate. The 4 cost-effectiveness studies suggested small reductions in costs due to testing.

In the economic analysis the base-case analysis of the model showed that standard practice (no testing/therapeutic monitoring with the intervention tests) dominated other options. Sensitivity and scenario analyses gave similar results. The PSA indicated a 92% likelihood that the 'no-testing' strategy was cost effective at a willingness to pay of £20,000 per QALY.

Conclusions

Our finding that testing is not cost effective should be viewed cautiously in view of the very limited evidence available, the uncertainty about a linked evidence approach in this context and the lack of a gold standard for assay comparison. Clinicians should be mindful of variation in performance of different assays and of the absence of standardised approaches to patient assessment and treatment algorithms.

Research priorities

There is substantial variation in the underlying treatment pathways and uncertainty in the relative effectiveness of assays and test based treatment algorithms that require further investigation. There is very little research evidence on adalimumab or on drug monitoring in children with Crohn's disease.

SCIENTIFIC SUMMARY

Introduction

Crohn's disease (CD) is a serious chronic fluctuating inflammatory condition of the digestive tract. It is uncommon and is currently estimated to affect about 115,000 people in the UK with about 3,000 new cases diagnosed each year. In severe active CD biological therapies are used when other treatment options fail and before surgical removal of the affected bowel is considered. These more recent drugs include monoclonal antibodies that inactivate tumour necrosis factor alpha (TNFα), which is a cytokine identified as having an important role in several inflammatory diseases including CD. The two anti-TNFα agents designated for this report are infliximab and adalimumab.

Response to anti-TNFα agents is variable. Sub-therapeutic drug levels are one cause of loss of response and may often be caused by the development of anti-drug antibodies which neutralise the drugs' action and hasten clearance from the circulation. This idea has led to the development of test kits able to measure circulatory levels of anti-TNF drugs and of the antibodies directed against them, and to the use of test results in treatment algorithms to bring the anti-TNFα agent into the therapeutic range and to prevent continued futile use of ineffective agents.

Decision problem

The decision problem for this assessment is:
Does testing of TNFα inhibitor levels and antibodies to TNFα inhibitors (infliximab or adalimumab) represent a clinically and cost-effective use of NHS resources in people with moderate or severe CD whose disease responds to treatment or who have has lost response to treatment with TNFα inhibitor?

The comparator for testing is standard care with appropriate anti-TNFα.

Three commercially available test kits for estimation of serum anti-TNFα agents and anti-drug antibodies have been identified as the intervention tests for this assessment, these are: LISA-TRACKER ELISA kits, TNFα-Blocker ELISA kits, and Promonitor ELISA kits.

Objectives

Objective A – review of comparative performance of tests

To review and critique studies of tests which compare

- Two or more intervention tests, or an intervention test with another test method which can be used to perform a linked evidence assessment.
- To compare and contrast studies that reported a test threshold analysis to determine the optimal drug level cut-off to predict or diagnose response.

Objective B – description of algorithms

To describe algorithms used in studies which include data on one or more intervention test or on a test which allows a linked evidence approach to be performed (i.e., algorithms used in studies identified in Objective C). The studies were required to provide an algorithm and report clinical outcomes for the management of patients with CD following measurement of serum levels of anti-TNFα drug and anti-drug antibodies and implementation of the algorithm.

Objective C1 – review of clinical effectiveness of test-algorithm combinations

To systematically review the literature comparing the clinical effectiveness of [a] an intervention or other assays for anti-TNFα agents and/ or for anti-drug antibodies used in conjunction with a treatment algorithm in Crohn's patients treated with infliximab or adalimumab; with [b] standard care (no tests performed or test-informed algorithm used) in CD patients treated with the same anti-TNF agent.

Objective C2 – analysis of correlation between test results and clinical outcomes

To analyse correlation studies which investigate the relationship between tests results for anti-TNFα and anti-drug antibody levels and clinical outcome in terms of response in patients with CD. This objective was added post-protocol because of the paucity of management studies which address the decision questions and to generate information of potential use for economic modelling.

Objective D – review of cost effectiveness of test- algorithm combinations versus standard care

To assess the cost-effectiveness of employing anti-TNFα and anti-TNFα antibody monitoring with LISA-TRACKER ELISA kits, TNFα-Blocker ELISA kits, and Promonitor ELISA kits in patients with CD compared with standard care.

In the absence of studies using the intervention tests, to use a linked evidence approach in which evidence of clinical effectiveness is taken from studies using alternative tests to the intervention tests.

Methods

Clinical effectiveness and cost-effectiveness systematic reviews

Searches of multiple electronic databases were undertaken during October, November and December 2014. Databases were searched from inception and no date limits were applied. Several supplementary searches of other online resources were run to check for other published and unpublished studies. Reference lists of included studies and relevant review articles were checked. Citation searches of selected included studies were undertaken. Further information was provided by the manufacturers of the intervention tests.

Two reviewers independently screened and assessed titles and abstracts of all records for inclusion. Studies were included according to the following criteria:

- *Population:* patients (adults and children) with moderate to severe active CD treated with infliximab or adalimumab
- *Intervention:* monitoring of serum anti-TNFα (infliximab or adalimumab) and or of anti-drug antibody levels using intervention tests or other test methods implemented in a test-treatment algorithm
- *Comparator:* standard care (no anti-TNFα or anti-TNFα antibody monitoring)
- *Outcomes:* Any patient related outcome, test agreement, cost-effectiveness estimates
- *Study Design:* Any primary-study design and systematic reviews with meta-analyses

Study quality assessments were undertaken using an adapted QUADAS-2 checklist, the Cochrane risk of bias tool, the Downs and Black checklist and for the economics studies, the Philips and the CHEERS checklists. Data were extracted by one reviewer checked by a second reviewer; disagreements were resolved by consensus or with a third reviewer. Evidence was synthesised using narrative review and statistical methods where appropriate. Individual patient data (IPD) were reconstructed from available Kaplan Meier plots using the method of Guyot et al. 2012. Meta-analyses were undertaken in Stata 11 software or using "MetaAnalyst" software and Review Manager (RevMan) 5.3 (The Cochrane Collaboration, The Nordic Cochrane Centre, Copenhagen, Denmark). The Harbord and Whiting 2009 method of hierarchical meta-analysis was used for diagnostic studies.

Cost-effectiveness model

A *de novo* Markov model was built in TreeAge Pro 2013 to evaluate the cost effectiveness of test-algorithm based treatment strategies versus standard care. Two populations were considered: patients responding to treatment and patients who have lost response to treatment. Two test strategies were assessed: concurrent testing of drugs and antibodies to the drugs and reflex testing (i.e. a drug test followed by an anti-drug antibody test, depending on the results of the drug test). When testing concurrently there are four possible test outcomes: drug + / antibody -; drug + / antibody +; drug - / antibody +; drug - / antibody -. In reflex testing only three test outcomes are possible: drug +; drug – / antibody -; drug - / antibody +. The model structure was informed by studies from the clinical effectiveness review, additional published studies and expert clinical advice. The model had a 4 week cycle and a ten year time horizon and adopted an NHS and PSS perspective. Costs were adjusted to 2013/14 prices and annually discounted at 3.5%. The starting point was a hypothetical cohort of people age 30 years. Outcomes are reported as incremental cost effectiveness ratios (ICER), expressed in terms of cost per quality-adjusted life-year (QALY) gained. A linked-evidence approach was

adopted. (In this approach, evidence from studies using tests other than the designated intervention tests was employed as a proxy for intervention test evidence). A single RCT from the clinical effectiveness review (using a radioimmunoassay) provided proportions of patients for the 4 possible combinations of test results when drug and anti-drug antibody levels were measured concurrently in patients who had lost response to infliximab. This RCT also provided some clinical outcomes for the patients' treatment according to a test-treat algorithm, prescribing treatment options for the 4 drug / antibody test outcome combinations. The TAXIT trial provided equivalent data for responders. Because we have no studies which use reflex testing algorithms for the reflex strategy for responders and people who had lost response had to be adopted using these two studies. A number of sensitivity analyses were undertaken including: a shortened 1 year time horizon with four-week cycle lengths, different transition probabilities for loss of response and changes in proportions of people in the different testing results categories. Probabilistic sensitivity analysis was also undertaken (10,000 model runs).

Results

We identified 2,434 and 2,466 studies for the clinical and cost effectiveness searches respectively of which 68 and 4 studies were included.

Clinical effectiveness

Twenty three studies comparing test methods were identified. Most studies did not investigate any of the three intervention tests. Evidence on concordance between the three intervention assays at a clinically relevant threshold was sparse and sometimes contradictory. Overall there was insufficient evidence to reliably assess comparative performance of the three intervention assays or their performance relative to other assay methods or to any of the comparators with links to clinical outcomes (HMSA, RIA, Prometheus ELISA, or Leuven in-house ELISA).

Three studies, two RCTs and one retrospective observational study provided comparative evidence on clinical outcomes following implementation of a test-algorithm versus a non-algorithm strategy. None of these studies used the intervention tests. All investigated infliximab treated patients. Neither of the RCTs found evidence of clinical benefit for a test-algorithm treatment regimen. In the TAXIT trial which investigated the effectiveness of drug monitoring following dose optimisation in patients with response to infliximab treatment 131/178 (73.59%) CD patients were in clinical remission before dose optimisation and 138/173 (79.77%) after dose optimisation using a test-treatment algorithm; at 52 weeks post randomisation there was likewise no difference in clinical and biological remission between the intervention test-treatment group and controls (P = 0.353). Both RCTs estimated cost savings in drug expenditure with a test-treatment algorithm compared to normal care. The retrospective observational study compared a proactive test-treatment algorithm versus normal care

reporting greater retention on infliximab treatment for the intervention group. However the algorithm was ill defined. Much of the evidence including this retrospective study investigated mixed groups of IBD (CD and ulcerative colitis) patients.

Thirty one studies reported on the correlation between test results and subsequent clinical state (response / no response). The studies were meta-analysed to estimate test accuracy for predicting clinical status. Meta-analyses indicated moderate test accuracy; positive and negative predictive value estimates derived from meta-analyses indicated that between 20 and 30% of positive and negative test results are likely to be inaccurate. This was confirmed by re-analysis of three meta-analyses of the ability to predict response / loss of response using drug and or anti-drug antibody levels.

Three studies reported results of both drug and anti-drug antibody tests for individual patients (one for infliximab treated responders, one for infliximab treated patients with loss of response, and one for adalimumab treated responders). These studies allowed estimation of the proportion of patients who might enter each of the treatment categories following concurrent or reflex testing strategies to enter into the economic model. However, the patients did not receive treatment according to a test-treat algorithm, therefore no outcome data from these studies was available, and outcomes data from the trials had to be used in the economic modelling.

Cost-effectiveness

The systematic review of cost effectiveness studies identified four studies. All of these indicated that a testing strategy might be less costly than alternatives with variable small effects on effectiveness. Use of standard checklists suggested that all the studies are subject to some limitations. There was insufficient published information to model an adalimumab test-based treatment strategy. The model therefore addressed infliximab therapy only.

In the base-case, the *de novo* Markov model results show that standard practice was less costly and produced more QALYs, hence dominating both the reflex testing and the concurrent testing strategy (Table 1). Sensitivity analyses indicated that change in testing frequency from 3 monthly to annually or reducing the time horizon to 1 year changed the most cost effective option to a concurrent testing strategy. The PSA indicated a 92% likelihood that the 'no-testing' strategy was cost effective at a willingness to pay of £20,000 per QALY.

Table 1 Base-case results for the analysis cost per QALY (2013/14 prices)

Strategy	Mean cost per strategy (£)	Difference in costs £	Effectiveness (QALYs)	Incremental QALYs	Incremental cost-effectiveness ratio (£)

					(ICER)
No testing	137,600	-	6.5146	-	-
Reflex testing	145,900	8,300	6.3315	0.1831	Dominated
Concurrent testing	147,100	9,500	6.3215	0.1931	Dominated

Discussion and Conclusions

Main findings

Meta-analysis indicates that tests have only moderate predictive accuracy for clinical status. There was insufficient evidence to assess the performance of the intervention tests properly relative to one another or to tests using alternative methodology. The literature indicates a lack of clinical consensus about what are the best and most appropriate tests to employ in clinical practice.

The limited RCT evidence from short term studies indicated that there is little or no benefit from a test-algorithm strategy although there may be some cost savings.

The base case cost-effectiveness analysis indicated that standard care – the no-testing strategy accumulates slightly greater QALYs, at a lower cost. This strategy is 92% likely to be cost effective at standard levels of willingness to pay.

Strengths and limitations

Strengths of the work include a robust and comprehensive systematic review (literature search, data extraction and analysis) strategy and the building of a *de novo* Markov model for the cost effectiveness assessment.

Although we undertook extensive systematic searches for relevant evidence and screened more than 30,000 titles, the findings of the systematic review warrant a cautious interpretation. Definitions of severity of disease (including response and loss of response) lack standardisation which impacts on the classification of patients in different studies. Consensus on a treatment algorithm is missing, possibly impacting on clinicians' confidence in using them. The evidence on assay performance was sparse and sometimes conflicting with lack of an agreed gold or reference standard for tests. There was very limited concordance data from studies comparing test performance of different assays.

Populating the economic model with information from the literature was problematical because of the small size, short duration, and the subjective methods for outcomes measurement. None of the studies used an appropriate standard care arm for economic modelling and many external sources of data and assumptions were required to populate the model. Inputs for the economic model needed to be drawn from disparate studies so that conclusions need to be tested with data from further research. Several studies sourced for model inputs included a proportion of patients with ulcerative colitis; the impact of this on model outputs is difficult to gauge. Variation in clinical practice in the management of patients with CD further complicated assumptions for model structure and inputs. We were unable to include adverse events and their treatment costs and this may have underestimated the costs.

Implications

Our findings that testing anti-TNFα drugs and their antibodies are not cost effective should be viewed cautiously by clinicians and policy makers, in view of the linked-evidence approach required and the poor quality of the evidence available to us. Clinicians should be mindful of the potential variation in performance of the different testing methods and strategies in their day to day practice.

Research priorities

We found that there is uncertainty about underlying treatment pathways, about the relative effectiveness of assays in the absence of a gold standard or agreed reference test, about which assays to use under which circumstances and which clinical algorithms to follow as a result of testing. There is very little research on adalimumab or the use of testing strategies and algorithms in children. The key questions for future research consideration are:

- What is the relative performance of methods of measuring anti-TNFα drug and their antibodies by ELISA kits compared to other methods such as RIA and HMSA and are these clinically significant? For example is there a validated drug threshold that is a useful predictor of clinical outcome?
- What are the best criteria for estimating response, non-response and loss of response in CD?
- At what time should assessments of drug and antibody take place?
- What is the effectiveness of clinical algorithms for disease management in response to testing in the UK?
- What is the effectiveness and cost-effectiveness of monitoring CD patients on adalimumab and for paediatric patients with CD?
- What is the relevance of co-treatment with immunomodulators in the monitoring of anti-TNFα agents and their antibodies?
- Is there a benefit of measuring total drug / antibodies as compared to measurements of free drug / antibody alone?

PLAIN ENGLISH SUMMARY

Crohn's disease (CD) is a serious chronic inflammatory condition of the digestive tract. It is uncommon and is currently estimated to affect about 115,000 people in the UK[1] with about 3000 new cases diagnosed each year. The causes of CD are not known but environmental, genetic and immunity related factors are believed to play a role as are previous infections and smoking. Crohn's disease often occurs in young adults and in women more than men.

In severe active CD in people who do not respond to first line treatments, surgical removal of parts of the bowel is an option. More recently a newer, type of drug has been used in these severely ill patients. These new drugs are known as antitumor necrosis factor agents (anti-TNFα agents) and the two most commonly used in CD are infliximab and adalimumab.

These are expensive drugs for the NHS. Some patients – but not all – respond to the drugs, that is experience improvement, and some patients respond and then lose their response. One cause of loss of response is that the patient develops an immune reaction and makes their own antibodies to cancel out the effect of infliximab and adalimumab. In the immune reaction the body registers the drugs as foreign proteins and eliminates them even though they are actually being used to treat the patient.

Tests have been developed to measure both the level of drug (infliximab and adalimumab) in the patient's blood and the level of antibodies that the patient has produced against these drugs. The idea is that drug levels and treatment options can be changed and improved in response to the test outcome to ensure that the patient is on the best treatment for them.

In this review we looked at the clinical and cost effectiveness of new tests (LISA-TRACKER ELISA kits, TNFα-Blocker ELISA kits, and Promonitor ELISA kits) to help work out how best to manage patients with moderate to severe CD taking these new drugs. The report will help NICE to make recommendations about how well the tests work and whether the benefits are worth the cost of the tests for use in the NHS in England and Wales.

We undertook systematic reviews of the clinical and cost effectiveness of the new tests and built our own economic model to assess cost-effectiveness.

For the clinical effectiveness review we found 72 relevant studies. We found that there is no 'gold standard' test for the levels of drugs or of antibodies to drugs and that current tests disagree, which means that it is difficult to assess the effectiveness of new tests. The cost-effectiveness model drew mainly on evidence from two randomised controlled trials. We made a number of assumptions in the

model. Even so, the model showed that over a 10 year period and compared to standard care, testing appeared to be not cost effective- it was more costly and offered less benefit.

We conclude that further research and assessments of testing schemes is needed, particularly with regard to frequency of testing and to design of plans for treatment options after testing, to ensure best care for patients with severe active Crohn's disease.

1 INTRODUCTION

1.1 Overview

Anti-TNFαs including infliximab [Remicade®, Merck Sharp & Dohme Ltd.] and adalimumab [Humira®, AbbVie]) are given to people with inflammatory bowel disease (IBD) including Crohn's disease (CD) as second or third line therapy. Response to anti-TNFα treatment varies among patients treated for inflammatory chronic conditions. While some stay in response over a long period of time, others are weaned off the drug because it is no longer needed and others may lose response at some stage during treatment. Loss of response can occur for various reasons, the most prominent being 1) formation of antibodies against the drug which neutralise the drug rendering it ineffective and 2) ongoing illness due to inflammation which is not meditated by TNFα. It has been proposed that measurement of serum levels of anti-TNFα and its antibodies can aid the management of patients with chronic diseases on anti-TNFα drugs.

Measurement of anti-TNFα levels and its antibodies can either be carried out concurrently (concurrent testing strategy), or antibody testing can be carried out conditional on the absence of measurable drug levels (reflex testing strategy).

The linked evidence approach adopted in this review is a methodology to handle shortcomings in the evidence for medical test evaluations.[2] The idea is to link evidence from other relevant research to the anticipated benefits of the test in question when direct evidence from the test and its effects on patient outcomes is absent. The decision analytic model is informed by systematically identified indirect evidence to predict the impact of the test under evaluation on patient outcome. The validity of this approach is dependent on the similarity between the populations, tests and outcomes across the linkages.[2]

1.2 Descriptions of the health problem – Crohn's disease

CD is a chronic fluctuating episodic inflammatory condition of the digestive tract; it is uncommon and is currently estimated to affect about 115,000 people in the UK[1] with about 3000 new cases diagnosed each year.[3] The aetiology of CD is still largely unknown but environmental, genetic and immunological factors are believed to play a role as are previous infections and smoking.[4]

1.2.1 Aetiology and pathology

CD can affect adults, adolescents or children. CD manifests itself mainly during late adolescence or early adulthood. The first onset most commonly occurs between the ages of 16 and 30 with a second peak between the ages of 60 and 80. Women are slightly more frequently affected than men but in

children it is seen more often in boys than in girls. CD is most common in white people in westernised countries and has its highest prevalence among Jewish people with European descent.[4]

CD follows a pattern of acute disease (relapse) interspersed with periods of remission (lack of symptoms). CD causes inflammation of the lining of the digestive tract which, depending on the individual, occurs at any location from the mouth to the rectum, but most commonly affects the end of the small intestine (terminal ileum) (35%) or the connection between the small intestine and large intestine (ileocaecal region) (40%).[5] Within individuals the disease location is fairly stable.

Fistulising CD describes the condition in patients who have developed complications in the form of abnormal connections between the bowel and other organs known as fistulae. Fistulae develop in between 17% and 43% of people with CD.[6] Active luminal Crohn's disease describes the condition in patients who have inflammation in the tube of the intestine.

The main symptoms of CD depend on the location of disease. They include abdominal pain, chronic or nocturnal diarrhoea, anal lesions, rectal bleeding, weight loss and swelling of the abdomen with tenderness. Complications include strictures, perforations, abdominal obstructions and development of fistulae. Extra-intestinal symptoms related to intestinal inflammation include inflammation of the joints, skin, liver and the eyes.[7] CD in children is often noticed because of growth failure.[8] Symptoms range in severity and the assessment of severity is used to classify CD into mild, moderate or severe disease according to disease activity scales. Reponses is defined as a reduction in symptoms.

1.2.2 Measurement of disease activity

CD can be difficult to diagnose because of overlapping symptoms with those of other gastrointestinal disorders such as UC and irritable bowel syndrome. Investigations to aid diagnosis include taking of the patient's medical history, physical examination, blood and stool tests and finally endoscopy to confirm diagnosis. Because the treatment for CD depends on the location and severity of disease, assessment of disease activity once disease is confirmed is important. However, disease activity is difficult to assess, and a global measure which includes clinical, endoscopic, biochemical and pathological features to define the heterogeneous disease pattern of CD is not available.[9] This means that there is no 'gold standard' for the assessment of disease severity which has important implications for this assessment. For instance there is no standardised definition for when remission has been achieved. The two most commonly used measures of disease activity are the Crohn's Disease Activity Index (CDAI) and the Harvey-Bradshaw Index (HBI) (a simplified version of the CDAI) which are based on the patient's history, physical features and laboratory data. A paediatric CDAI has been developed which emphasises the less subjective laboratory parameters.[10] Additional measures include the Perianal Disease Activity Index (PDAI), the Inflammatory Bowel Disease

Questionnaire (IBDQ) and the Crohn's Disease Endoscopic Index of Severity (CDEIS).[5] However, these tools have been primarily developed for clinical trials rather than clinical practice. In clinical practice, assessment of mucosal healing by endoscopy to assess response and remission is becoming increasingly important and the potential of objective laboratory markers such as C-reactive protein and faecal calprotectin for the assessment of disease activity, risk of complications, prediction of relapse, and monitoring the effect of therapy have been recognised.[11]

The CDAI measures variables including number of liquid stools, abdominal pain, general well-being, extra-intestinal complications, use of anti-diarrhoeal drugs, abdominal mass, haematocrit and body weight.[12] These are weighted according to their ability to predict disease activity leading to an individual score ranging from 0 to 600. The CDAI has been criticised for giving too much weight to relatively subjective items,[12] however more objective measures such as mucosal healing on endoscopy are not infallible either, because of the patchy distribution of inflammation in CD. Samples taken for examination may not necessarily be representative of the whole bowel.[1]

While the CDAI uses a symptom diary of the patient over 7 days for the assessment, the HBI uses only a single day's diary entry for assessment. Furthermore, the HBI does not take into consideration body weight, haematocrit and use of drugs for diarrhoea for the measurement of disease activity. Scores from the HBI range from 0 to 20.[13]

In the absence of standardised definitions in which scores correspond to the different disease severity stages, this review adopts the definitions from the NICE guidance TA187:[6]

- Remission is defined as a CDAI score <150
- Moderate to severe disease is defined as a CDAI score >220
- Severe disease is defined as a CDAI score >300

Response (i.e., relief of symptoms) has often been defined as a reduction in the CDAI score of at least 70 points from baseline.[14]

Severe active Crohn's disease was defined for the purpose of the indication of infliximab or adalimumab treatment as:[6]

> *"Very poor general health and one or more symptoms such as weight loss, fever, severe abdominal pain and usually frequent (3–4 or more) diarrhoeal stools daily. People with severe active Crohn's disease may or may not develop new fistulae or have extra-intestinal manifestations of the disease. This clinical definition normally, but not exclusively,*

corresponds to a Crohn's Disease Activity Index (CDAI) score of 300 or more, or a Harvey-Bradshaw score of 8 to 9 or above."

Furthermore, the Practice Parameter Committee of the American College of Gastroenterology have produced definitions of disease severity.[7] These are:

Mild-moderate disease:

- *"Mild-moderate disease applies to ambulatory patients able to tolerate oral alimentation without manifestations of dehydration, toxicity (high fevers, rigors, prostration), abdominal tenderness, painful mass, obstruction, or >10% weight loss"*

Moderate-severe disease:

- *"Moderate-severe disease applies to patients who have failed to respond to treatment for mild-moderate disease or those with more prominent symptoms of fever, significant weight loss, abdominal pain or tenderness, intermittent nausea or vomiting (without obstructive findings), or significant anaemia."*

Severe-fulminant disease:

- *"Severe-fulminant disease refers to patients with persisting symptoms despite the introduction of steroids as outpatients, or individuals presenting with high fever, persistent vomiting, evidence of intestinal obstruction, rebound tenderness, cachexia, or evidence of an abscess."*

Remission:

- *"Remission" refers to patients who are asymptomatic or without inflammatory sequelae and includes patients who have responded to acute medical intervention or have undergone surgical resection without gross evidence of residual disease. Patients requiring steroids to maintain well-being are considered to be 'steroid-dependent' and are usually not considered to be 'in remission'."*

1.2.3 Management and Care pathway

The treatment of CD is complex, in general it aims at: a) reducing symptoms through induction and maintenance of remission, b) minimising drug-related toxicity, and c) reducing the risk of surgery.[15] The management options for CD include drug therapy (e.g., glucocorticosteroids, 5-aminosalicylate, antibiotics, immunomodulators, TNFα inhibitors), enteral nutrition, smoking cessation and, in severe or chronic active disease, surgery. The choice of treatment amongst the available drugs is influenced by patient age, site and activity of disease, previous drug tolerance and response to treatment, and the presence of extra-intestinal manifestations.[16, 17] Enteral nutrition is widely used as a first line treatment to facilitate growth and development in children and young people. Adjuvant therapy commonly coexists and includes management of extra-intestinal manifestations, antibiotics, corticosteroids or

immunomodulator therapy. Between 50% and 80% of people with CD require surgery due to complications such as strictures causing symptoms of obstruction, fistula formation, perforation or failure of medical therapy.[1]

Once remission has been achieved, maintenance therapy can be considered following assessment of the course and extent of CD, effectiveness and tolerance of previous treatments, presence of biological or endoscopic signs of inflammation, and potential for complications.[15]

1.2.3.1 Induction of remission according to the NICE Clinical Guideline 152[1]

Usually, at first presentation, people with active CD are recommended monotherapy with conventional steroid therapy (i.e. glucocorticoids including prednisolone, methylprednisolone or intravenous hydrocortisone), which is aimed at inducing remission as a first line treatment. Alternatively, treatment with budesonide, 5-aminosalicylic acid (5-ASA), or enteral nutrition may be offered for people who do not choose to take or who are intolerant of glucocorticosteroid therapy.

The addition of an immunomodulator (azathioprine, mercaptopurine or methotrexate) to a conventional glucocorticosteroid or budesonide is recommended as an add-on therapy for inducing remission for people who have active CD, who have experienced two or more inflammatory exacerbations in a 12-month period, or for whom glucocorticosteroid doses cannot be tapered. As advised in the current online version of the British National Formulary (BNF)[18] or British National Formulary for Children (BNFC),[18] the effects of azathioprine, mercaptopurine, and methotrexate as well as levels of neutropenia (in people on azathioprine or mercaptopurine) should be monitored.[15]

In adults and people aged 6–17 years with severe active CD who fail to respond to the first line of treatment with conventional therapy (e.g., immunomodulators, corticosteroids), or who are intolerant to or who have contraindications to conventional therapy, anti-TNFα agents (infliximab and adalimumab) are recommended as treatment options within their licensed indications. The administration of anti-TNFα agents is recommended until 12 months after the start of treatment or until treatment failure (including the need for surgery), depending on whichever occurs first. Reassessment and monitoring of disease activity (at least every 12 months) is advised to ascertain the clinical appropriateness of ongoing treatment. Usually, treatment is initiated with the less expensive drug (i.e. infliximab), considering drug administration costs, dose, and product price per dose. The use of anti-TNFα drugs for the treatment of CD is covered in the 2010 NICE technology appraisal guidance 187 (Infliximab (review) and adalimumab for the treatment of Crohn's disease) which is summarised in section 1.2.3.3.[6]

Surgery should be considered early in the course of the disease for people whose disease is limited to the distal ileum or for children and young people who have growth impairment despite optimal medical treatment and/or who have refractory disease.[1]

1.2.3.2 Maintenance of remission according to the NICE Clinical Guideline 152[1]

People with CD in remission can be managed with or without maintenance treatment. The options for maintenance (including treatment or no treatment) need to be discussed with patients and parents or carers. The discussion should include risk of relapse and the potential side effects of drug treatments. People who decide not to use maintenance treatment should agree follow-up plans (e.g., frequency and duration of visits) and should receive information on markers and symptoms of relapse (e.g., unintended weight loss, abdominal pain, diarrhoea, general ill-health) to ensure that they keep their disease appropriately under review with their healthcare professionals.

People with CD in remission who choose to receive maintenance therapy may be offered a single drug such as azathioprine or mercaptopurine if remission has been induced using a conventional glucocorticosteroid or budesonide. Methotrexate can be offered if remission was induced by methotrexate or to people who are not able to tolerate, or who have contra-indications to azathioprine or mercaptopurine. Treatment with 5-ASA can be used to maintain remission after surgery.

If remission has been achieved with anti-TNFα medication, then maintenance with anti-TNFα with or without an immunomodulator can be used. Continuation of treatment with infliximab or adalimumab during remission is advised only if there is evidence of ongoing active disease assessed by clinical symptoms, biological markers, and endoscopy if necessary. The balance between harms and benefits of ongoing treatment should be taken into account. The guideline states that people who relapse after anti-TNFα treatment may start it again.[15]

1.2.3.3 NICE guidelines

The NICE guideline TA187 describes when infliximab or adalimumab should be used to treat people with severe active or fistulising Crohn's disease in the NHS in England and Wales. The guideline states:[6]

1.2.3.3.1 Infliximab

"Infliximab has a UK marketing authorisation for the treatment of:

- *severe, active Crohn's disease in people whose disease has not responded despite a full and adequate course of therapy with a corticosteroid and/or an immunosuppressant, or who are intolerant to or have medical contraindications for such therapies*

- *fistulising, active Crohn's disease in people whose disease has not responded despite a full and adequate course of therapy with conventional treatment (including antibiotics, drainage and immunosuppressive therapy)*
- *severe, active Crohn's disease in people aged 6–17 years whose disease has not responded to conventional therapy, including a corticosteroid, an immunomodulator and primary nutrition therapy, or who are intolerant to or have contraindications for such therapies."*

Administration of infliximab should follow this pattern:[6]

> *"5-mg/kg intravenous infusion over a 2-hour period followed by another 5-mg/kg infusion 2 weeks after the first. If a person's disease does not respond after two doses, no additional treatment with infliximab should be given. In people whose disease responds, infliximab regimens include maintenance treatment (another 5-mg/kg infusion at 6 weeks after the initial dose, followed by infusions every 8 weeks) or re-administration, otherwise known as episodic treatment (an infusion of 5-mg/kg if signs and symptoms of the disease recur)."*

For fistulising disease the first three doses at weeks 0, 2, and 6 are considered as induction therapy and additional infliximab therapy should not be given if the first three doses have not induced a response. The patient pathway for people responding to infliximab induction therapy and moving onto maintenance therapy is given in Figure 1.

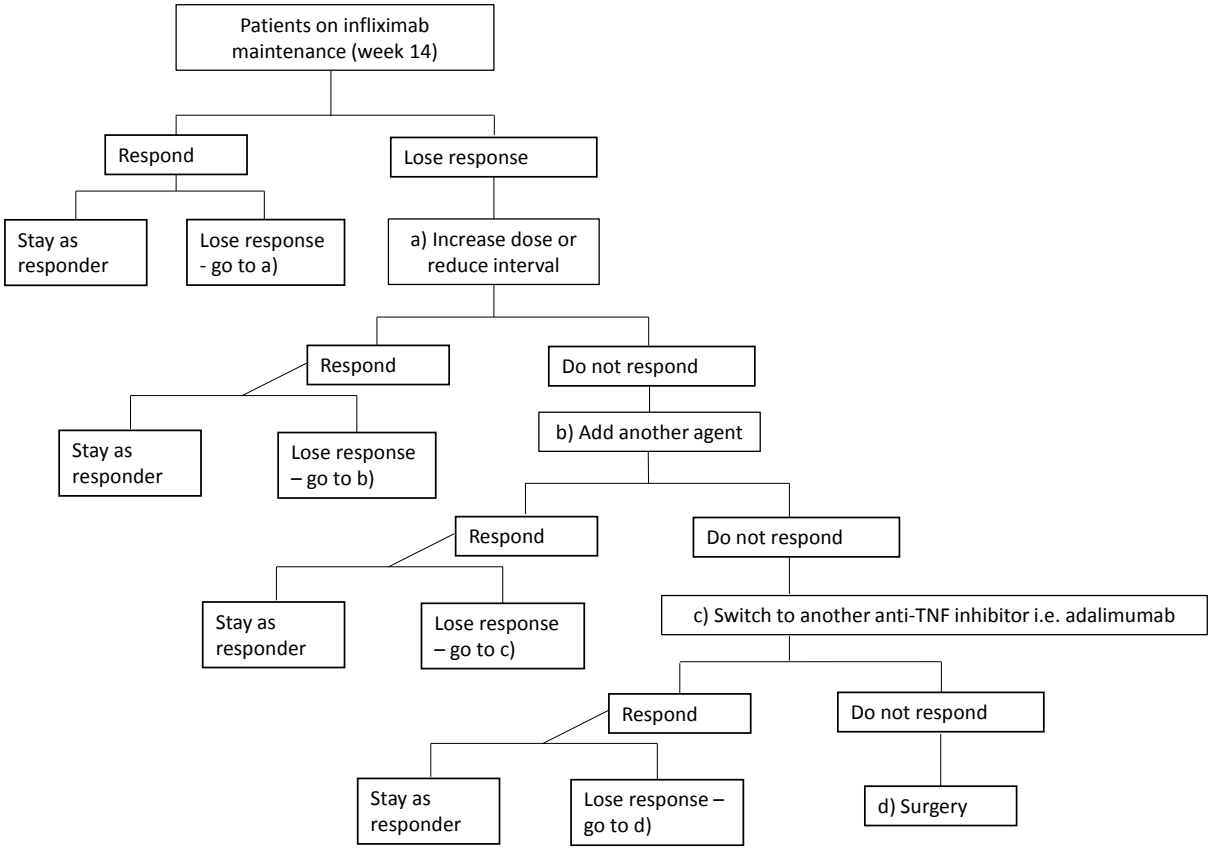


Figure 1 Patient pathway of Crohn's disease patients on infliximab therapy

1.2.3.3.2 Adalimumab

Adalimumab can be used to treat severe, active CD in adults whose disease has not responded to treatment with an immunosuppressant and/or corticosteroid, or who are intolerant to or have contraindications to such therapies.

Administration of adalimumab should follow this pattern:[6]

> *"The adalimumab induction treatment dose regimen for adults with severe Crohn's disease is 80 mg via subcutaneous injection, followed by 40 mg 2 weeks later. After induction treatment the recommended dose is 40 mg every other week. This can be increased to 40 mg every week in people whose disease shows a decrease in response to treatment."*

1.2.3.4 Anti-tumour necrosis factor alpha (anti-TNFα) agents

CD is associated with elevated levels of the immune-regulatory protein TNFα. The reasons for this elevation in CD is still largely unknown. TNFα is a small cell-signalling protein (cytokine) involved in inflammatory responses primarily by influencing regulation of various effector cells of the immune system. TNFα has been shown to have a role in several inflammatory diseases including CD, ulcerative colitis, rheumatoid arthritis and ankylosing spondylitis. Anti-tumour necrosis factor alpha (anti-TNFα) agents bind to cell surface TNFα and free TNFα and block their activity. Blocking of

TNFα with anti-TNFα drugs has been shown to be successful for some patients with inflammatory diseases including CD. Anti-TNFα agents recommended by NICE for the treatment of CD are infliximab and adalimumab. These monoclonal antibodies are introduced into the human body to bind and block TNFα. They are classed as monoclonal antibodies because they are derived from genetically engineered immune cells, which are all daughters of a single parent cell, so that in culture they generate and secrete antibodies which are all of identical structure and affinity for TNFα.[15]

1.2.3.4.1 Infliximab

Infliximab is a chimeric (mouse-human) monoclonal antibody. It is said to be chimeric because the genetic code determining its amino acid sequences is partly derived from the mouse genome and partly from the human genome. Infliximab belongs to the IgG1 (immunoglobulin gamma type 1) group of antibody molecules (Figure 2). It should be borne in mind that IgG1 molecules are globular (not linear as in the diagram) and that they are glycoproteins which have carbohydrate chains attached (not shown in Figure 2). As infliximab is generated from cultured mouse cells, the carbohydrate part of the molecules corresponds to that of mouse rather than human glycoproteins.[15]

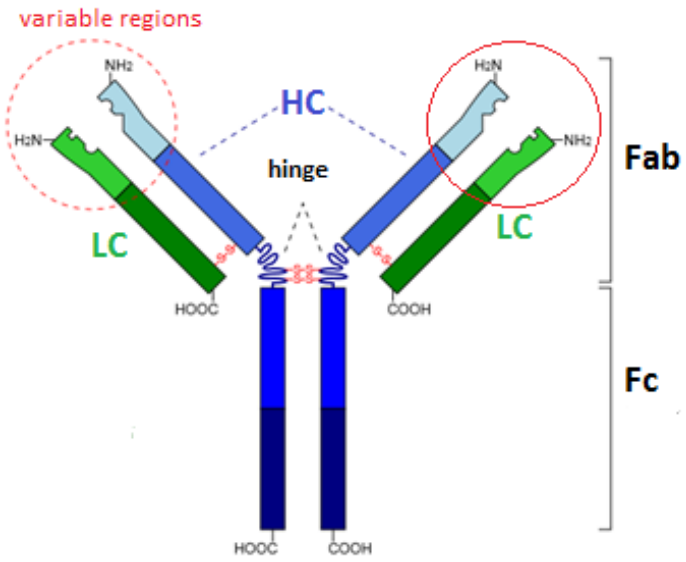


Figure 2 Diagrammatic representation of the structure of an IgG1 antibody molecule
The molecule comprises two heavy chains (HC) and two light chains (LC); the HCs are joined together across disulphide bonds (S-S) and each LC is joined to a HC by S-S bonding. The LC and HC have a variable region (different from all other antibodies) at the amino (NH_2) end of the chain; these variable regions are responsible for binding antigen. The rest of the HC and LC are identical to other IgG1 antibodies and are called constant regions. Proteolytic enzymes papain and pepsin cut the molecule just above or below the S-S bonds holding the HC together. When below the HC S-S bond this generates an Fc (Fragment crystallising) and an Fab (Fragment antigen binding) product. When the split is above the HC S-S bond two antigen binding fragments are formed ($F(ab)_2$)

Infliximab is composed of human IgG1 heavy chain constant regions and human Kappa light chain constant regions (together representing 70% of the genetic makeup of the molecule), plus mouse-derived heavy chain and light chain variable regions (30% of the genetic makeup, 4 out of 12 domains) which carry the binding sites with high affinity and specificity to TNFα (Figure 1). Infliximab was the first anti-TNFα agent that was approved and licensed for treating severe active Crohn's disease and active fistulising Crohn's disease in adults and children over the age of six. It is administered intravenously over 1–2 hours.

Side effects of infliximab include:

- Allergic reaction to the infusion (or infliximab) apparent by:
    - hives (red, raised, itchy patches of skin) or other skin rashes
    - difficulty swallowing or breathing
    - pains in the chest or muscle or joint pain fever or chills
    - swelling of the face or hands
    - headaches or a sore throat
- Serious viral or bacterial infections including tuberculosis, especially in people over 65
- Skin reactions including psoriasis (red scaly patches), rashes, skin lesions, ulcers and hives, and swollen face and lips
- Worsening of heart problems
- Increased risk of cancer or lymphoma
- Liver inflammation

Many of the side effects are reversible if the drug is stopped.[15]

1.2.3.4.2 Adalimumab

Adalimumab is a human IgG1 monoclonal antibody with Kappa light chains. It consists of purely human antibody polypeptide domains (Figure 2). However, as adalimumab is generated from cultured Chinese hamster ovary cells, the carbohydrate part of the molecules corresponds to that of hamster rather than human glycoprotein. Adalimumab is a more recent anti-TNFα therapy that was approved for treating CD in adults only. It is administered as a subcutaneous injection by a doctor or nurse or can be self-injected by the patient or a family member.[15]

Side effects of adalimumab include:

- Reactions to the injection including pain, swelling, redness, bruising and itching
- Allergic reaction to adalimumab including:
    - rashes or hives

    - swollen face, hands and feet
    - trouble breathing
- Greater susceptibility to infections such as colds, flu, pneumonia, sepsis and tuberculosis
- Skin reactions including psoriasis (scaly patches), eczema, other skin rashes and ulcers
- Skin cancer, lymphoma or leukaemia
- Damage to nerves (demyelination)
- Lupus

Many of the side effects are reversible if the drug is stopped.[15]

1.2.4 Significance to the NHS and current service cost

The aim of successful therapies in CD is to prolong remission and to minimise relapse. Patients' quality of life fluctuates through time and unsurprisingly has been found to be better during remission. Studies using various disease-specific health-related quality of life measures (such as McMaster IBDQ, IBDQ-36, short IBDQ, rating form of IBD patient concerns (RFIPC), Cleveland Clinic questionnaire, Gastrointestinal Quality of Life Index) show a clear correlation between HRQOL and symptoms. These measures in Crohn's patients have allowed utility estimates to be developed for patients in various clinical states.[19]). QOL has been found to be somewhat worse in Crohn's disease than in ulcerative colitis, and substantially worse in relapse than in healthy matched individuals. It has also been found to be similar or worse to that experienced in many other medical conditions.[20] Gastroenterologists tend to rely on global clinical judgement, which tends to be less reproducible than QOL assessment tools, but is of course, simpler for decision-making in everyday clinical practice.[9]

CD patients can be cared for in primary or secondary care depending on symptom severity. While GPs manage patients in remission or with mild symptoms, patients with more severe active disease are managed in secondary care. These are patients who are likely to be steroid-dependent, on immunomodulators or anti-TNFαs, or patients requiring surgery. It has been estimated that about 50% of CD patients experience at least one flare per year. Of these, 20% of patients will require hospitalisation.[21] Disease flares have been found to be associated with a 2-3 fold increase in hospitalisation and a 20-fold increase in cost compared to managing people in remission.[22] Audit data shows that anti-TNFα agents are potentially cost-saving by successfully maintaining patients in remission and reducing hospital admissions. A cost reduction of £138 per patient at 6 months[23] and £2750 per patient at 12 months (excluding infliximab costs) has been demonstrated in a before- and after study of infliximab therapy.[24] However, in the latter study both non-responders (£3608) as well as responders (£1656) incurred a considerably higher annual cost than the previously costs estimated using decision modelling of long term care in CD (£631). [25] Using 2008 prevalence figures the total

annual cost to the NHS for the ~60,000 patients with CD was estimated at £38 million. Updating this figure with more recent prevalence data (115, 000) but still using 2008 prices would double that cost to £73 million. This however, might be a modest estimate when considering the wide range of measured 6 month costs for individual CD patients (£73 - £33,254).[22] The main drivers of costs were hospital admission, surgery and anti-TNFα treatment.[25] In the HTA by Clark et al. (2003)[3] the average cost of a single 5mg/kg infliximab infusion for a 70kg patient was reported to be £1804.80. No comparable data for adalimumab are available.

The significance of CD to the NHS is increased by the fact that the prevalence of CD is increasing and the disease affects many people at a young age. The life time care costs for CD patients can now be compared to other major chronic diseases such as diabetes and cancer.[21] This argues not only clinically but also economically for interventions which keep patients in remission and out of hospital. This review focuses on whether monitoring anti-TNFα agents and their antibodies with ELISA tests could potentially contribute to this aim.

1.3 Rationale for measuring anti-TNFα drug and anti-drug antibody levels

1.3.1 Responders and non-responders definitions and incidence rates

Like other treatment regimens for CD, anti-TNFα treatment aims to induce remission (known as induction therapy defined as <150 points on CDAI and no draining fistulae) and to prevent relapse (maintenance therapy). However failure to induce a response and loss of response to anti-TNFα are common problems in clinical practice. The lack of consensus regarding clear definitions for response and remission result in inconsistencies in reported incidence rates of non-response and loss of response (LOR) covered in this section.

1.3.1.1 Primary non-response

Patients not achieving at least a 70 point reduction on the CDAI during induction therapy are classed as primary non-responders. Incidence rates of primary non-response vary greatly depending on the clinical outcome measured (response / remission) and on the time point the assessment of response is undertaken. For example, the ACCENT I study, an RCT investigating the benefit of maintenance infliximab therapy in 573 active CD patients, assessed response after a single infliximab infusion at 2 weeks.[26] The ACCENT II study was a post hoc analysis to determine the efficacy and safety of infliximab therapy in patients with fistulising CD in which patients were assessed at 10 weeks.[27] The CHARM RCT assessed response to adalimumab at 4 weeks to evaluate the drug's efficacy and safety in the maintenance of response and remission in 854 patients with moderate to severe CD.[28] However Ben-Horin et al. (2014)[29] stated that in clinical practice non-response should not be assessed before 8-12 weeks as remission might still be induced at this time. It is therefore not surprising that a review of

incidence rates of non-response in clinical trials reported a range of between 20-40% and 10-20% from 'real-life' series.[29] In contrast, lack of remission at week 4 in patients with luminal CD was reported to be as high as 67% for infliximab and 64% for adalimumab.[30, 31] The true magnitude of the rate of primary non-response is therefore difficult to determine.

Factors associated with non-response are believed to include:[29, 32, 33]

- Severity of disease
- Duration of disease
- Smoking
- Drug elimination
- Drug binding
- Anti-drug antibodies
- Alternative non-TNFα mediated disease pathways
- Concomitant treatment with immunomodulators
- Prior failure of other anti-TNFα

1.3.1.2 Loss of response (LOR)

Patients with an initial response to anti-TNFα treatment can lose response (loss of response: LOR) at any time during induction or maintenance therapy despite intensification of treatment (i.e. increase in dose or decrease in dosing interval). Again lack of a clear definition, assessment at different time points, different outcome measures and different drug doses mean that reported incidence rates of secondary LOR vary considerably across studies. The true extent of this problem is largely unknown. Gisbert and Panes (2009)[34] in their review reported a range of LOR to infliximab of 11-48% (mean 37%) for varying length of follow up and deBoer et al. (2014)[32] reported a range of 21% to 46% for LOR to adalimumab. For this reason the incidence of LOR is better expressed as the annual risk for LOR per patient year (13% for infliximab[34] and 20.3% for adalimumab[35]). LOR to adalimumab and infliximab did not differ significantly in a retrospective study of 375 patients who had lost response to either infliximab or adalimumab, however, patients treated with adalimumab required more dose optimisation intervention than patients on infliximab.[36]

The following factors are believed to prevent LOR:[14]

- Pre-medication with steroids
- Concomitant immunomodulators
- Maintenance therapy as opposed to episodic treatment

Mechanisms of loss of response to anti-TNFα agents are still unclear. The next section describes some of the possible mechanisms in more detail.

1.3.2 Anti-drug antibodies

Anti-drug antibodies can be elicited by infliximab or adalimumab during therapy as a response by the human immune system to these foreign proteins. This is termed immunogenicity of anti-TNFα agents. These anti-drug antibodies bind to the anti-TNFα agent and neutralise its action. If sufficient amounts of antibodies are present, the individual loses response to the drug treatment. During scheduled maintenance therapy the incidence of anti-drug antibodies is 5-18%[26, 37, 38] and 3-17%[38] for infliximab and adalimumab, respectively. The similar rates for infliximab and adalimumab might initially appear counterintuitive as adalimumab is a fully human recombinant protein while infliximab is partly human and partly mouse protein and therefore 'more' foreign. However, adalimumab, like infliximab, is a foreign protein which will prompt a response when coming into contact with the immune system. This indicates that the degree of 'human-ness' is not the main determinant of immunogenicity (formation of antibodies to a foreign protein).[38]

Levels of antibodies have been found to be higher during episodic treatment at 36-61%[38-40] as compared to levels found during maintenance therapy. This indicates that other factors may influence immunogenicity. The true incidence of antibodies in anti-TNFα treated patients is therefore unknown. The ability to mount an immune response which depends on a number of factors including the method of measuring antibody levels and on age, also depends on concomitant treatment with immunomodulators. [41, 14, 26, 39, 40, 42, 43] For that reason concomitant immunomodulators might be given to patients to prevent or reduce the formation of antibodies. This effect was not observed in one study for adalimumab,[14] and Vermeire et al. (2007)[44] reported that increasing anti-TNFα above antibody binding capacity might have similar effects to immunomodulators, by neutralising free antibodies. Vande Casteele et al. (2012)[45] made a similar observation for transient antibodies, (antibodies detectable for a short period during a series of follow up test assays conducted during a course of infusions), while sustained antibodies did not disappear after dose optimisation and were associated with LOR.

The clinical importance of antibodies can be presented as:[38]

- Positive / negative / inconclusive
- High / low
- Above / below a threshold in arbitrary units or in µg/ml
- Drug concentration
- Clinical effect (duration of response, need for dose intensification, switch drug)

- Impact on safety (infusion reactions)

The clinical relevance of antibodies has been debated. However, numerous studies report the correlation between presence of antibodies with low or absent drug levels and consequent response.[26, 37, 39, 42, 44, 46, 47] This can be explained by antibodies binding to the epitope of anti-TNFα and neutralising the drug (i.e. making it unable to bind to TNFα and inhibiting the working mechanism of anti-TNF) or by forming immune complexes with the drug (non-neutralising antibodies) which are subsequently cleared from the circulation (reducing the drug's bioavailability).[48]

While the importance of the neutralising antibodies has been universally acknowledged,[14, 33, 48, 49] reviewers seem to disagree in their conclusion about the importance of non-neutralising antibodies.[14, 33] Over 90% of antibodies to infliximab and adalimumab are neutralising.[50] In a meta-analysis, Garces et al. (2013)[48] estimated that detectable antibodies can decrease response to anti-TNFα by as much as 80%.

An interesting additional observation was made by Steenholdt et al. (2013)[51] who showed that IgG antibodies reacting with the fragment antigen binding portion (Fab) (see Figure 2) of infliximab exist in infliximab naïve IBD patients prior to treatment. The presence of pre-existing antibodies affected response and safety of infliximab treatment in CD patients and the study concluded that the clinical utility of measuring pre-treatment antibodies should be assessed.

1.3.3 Drug levels

While anti-drug antibodies have been shown to reduce anti-TNFα drug levels, there are other known mechanisms which affect drug levels. These include dose and dosing interval, body-mass index, gender, serum albumin levels (serum albumin transports drugs and can affect the half-life of drugs), concomitant immunomodulators, severity of inflammation, mode of administration (intravenous versus subcutaneous) and drug half-life.[14, 43]

As a consequence drug levels vary considerably between patients and within individuals over time.[33] Following administration of the anti-TNFα agent, circulating drug concentration will be at its peak level; the concentration just before the next round of treatment is classed as trough level. The optimal time of testing drug levels within this cycle has been debated[43] and it is largely unknown what the optimal drug levels would be at the different time points. While a threshold trough level is thought to be needed for effectiveness it is also known that supra-therapeutic levels can cause infections and other adverse events.[14]

1.3.4 Anti-TNFα and antibody level monitoring in Crohn's disease

One of the key studies to demonstrate and quantify the link between drug and anti-drug antibody levels, immunomodulator therapy and response in CD patients on anti-TNFα agents was the study by Baert et al. (2003).[42]

Baert et al. (2003)[42] was an early and influential study of the development of anti-drug antibodies to infliximab in patients with CD; this study stimulated numerous subsequent investigations. The study enrolled 125 consecutive patients (38 fistulising and 87 with luminal disease) who received 5mg/kg infliximab at weeks 0, 2 and 6 weeks. Responders (89/125, 71%) were re-treated with this regimen should they require re-start of infliximab therapy according to clinical judgement. Mean treatment period was 10 months and median follow up 36 months. Anti-drug antibodies and infliximab serum levels were measured before and at 4, 8, and 12 weeks after each infusion, using ELISAs (Prometheus Laboratories). After 5 infusions 76/125 (61%) patients were classified as positive for anti-drug antibodies.

When a level of 8µg anti-drug antibodies/mL serum was selected it was found that 24/56 (43%) of patients taking immunomodulators had >8ug/mL anti-drug antibodies compared with 52/69 (75%) of those not taking immune-suppressive agents. The relative risk of anti-drug antibodies >8µg/mL in patients taking versus those not taking suppressive therapy was 2.40 (95% CI: 1.65 to 3.66; $p < 0.001$). Infusion reactions had occurred in 27% of patients by the fifth infusion. The median anti-drug antibody level for patients with reactions was higher than for those with no reactions ($p < 0.001$). Reactions were significantly more common in patients not taking immunomodulator therapy than in those who were taking it. When time to next infusion was taken as a measure of response duration it was found that response duration was reduced in those with anti-drug antibodies >8 µg/mL relative to those with levels <8µg/mL (median: 35 days versus 71 days; $p < 0.001$).

The level of infliximab at 4 weeks after an infusion was correlated with the level of anti-drug antibodies prior to the infusion ($R2 = 0.34$, $p < 0.001$) and was positively correlated with duration of response. Infliximab level and anti-drug antibodies level were independent variables influencing response duration. Logistic regression indicated that the only variable that influenced a 4 week level of infliximab >12µg/mL was the use of immunomodulator therapy. Infliximab level was higher in those without an infusion reaction than those with one.

In summary the results of this study suggest that production of anti-drug antibodies is common during infliximab therapy, that anti-drug antibodies are associated with reduced infliximab levels, that duration of response is reduced by the presence of anti-drug antibodies, and that production of anti-drug antibodies may be reduced when concomitant immunomodulator therapy is employed.

Further evidence steadily accumulated from retrospective analyses of multiple clinical trials and case series[37, 52-54] and the observation that detectable drug at trough is associated with greater clinical efficacy is now well established.[43]

This accumulating evidence has formed the basis of investigations into drug and anti-drug antibody monitoring in anti-TNFα treated CD patients.

Without monitoring, the options for a clinician if the anti-TNFα agent fails are to wait and see, intensify drug treatment, switch drug within its class, or switch to a different class of drugs.[33] Measuring drug and anti-drug antibodies, however, could enable clinicians and patients to make informed choices on management. A number of studies have investigated the clinical utility of measuring drug and anti-drug antibody levels in sera by translating the clinical management decision following a test outcome into a treatment algorithm stipulating the management pathways for patients with a specific test outcome in clinical practice.[55-58]

Drug and anti-drug antibody monitoring could be undertaken in good responders (i.e., those responding to initial induction course of anti-TNFα treatment) as well as in patients with loss of response (i.e., those initially responding to anti-TNFα treatment but losing this response over time). The use of these technologies provides a clinician with potentially useful information that may guide individual patient's future treatment. Such information may aid in anticipating the loss of response in responders or allow drug optimisation, while for non-responders such analyses may help in estimating the likelihood of various candidate reasons for loss of response. For example in non-responders with low levels of drug and high levels of anti-drug antibodies the loss or lack of response may be surmised to be due to rapid clearance of the drug due to the action of anti-drug antibodies; on the other hand a low level of anti-TNFα in the absence of anti-drug antibodies may be suggestive of non-immune mechanisms of rapid drug clearance, while high levels of drug in the absence of antibodies in non-responders may be suggestive of a pathology for the condition independent of TNFα in a particular patient. Algorithms for future treatment based on anti-TNFα and anti-drug antibody estimates have been published.[55-58]

In theory the application of the tests in conjunction with an appropriate algorithm for treatment based on test results might:

- improve quality of life and other outcomes (e.g., faster healing of flare-ups, reduced abdominal pain and associated diarrhoea)
- optimise the treatment plan (facilitate adoption of the most suitable future treatment for individual patients; this might involve a switch to an alternative anti-TNFα or a biologic with an alternative mechanism of action)

- minimise the risk of drug overdose and associated adverse events
- allow earlier de-escalation of therapy, leading to a reduction in the overall drug used
- help to reduce the amount of drugs used inappropriately, unnecessary hospital visits, risk of surgery, and associated costs[15]

1.4 Description of technology under assessment

1.4.1 Intervention technologies

Various assay procedures for anti-TNFα agents and for anti-drug antibodies have been developed in the belief that the levels of circulating anti-TNFα and of anti-drug antibodies can provide information useful to clinicians in indicating potential reasons for treatment failure, and for dosage or treatment adjustment.

Commercially available ELISA kits: the LISA-TRACKER ELISA kits (Theradiag / Alpha Laboratories), the TNFα-Blocker ELISA kits (Immundiagnostik AG), and the Promonitor ELISA kits (Proteomika) are the intervention technologies designed to measure infliximab and adalimumab levels and their antibodies investigated in this review.

These are all particular examples of solid phase Enzyme Linked Immunosorbent Assays (ELISA assays). They estimate the following molecules in patient blood sera:

- Infliximab
- Adalimumab
- Anti-infliximab antibodies
- Anti-adalimumab antibodies

Details of the ELISA kits available from these companies are summarised in Appendix 1.

Other ELISA assays commercially available for measuring these molecules in sera include: the Shikari ELISA kits (Matriks Biotek). These are not included as index tests in the NICE scope.

Other methodologies based on alternative principles of detection and measurement include: [a] radioimmunoassays (RAI); liquid phase assays [b] cell reporter assays based on genetically engineered cells incubated in culture medium; [c] mobility shift assays; liquid phase assays using size-exclusion HPLC and fluorescent dye detection. The differences in these assays may have an effect on their individual performance and describing and contrasting them will help the reader to understand the abilities and limitations of the assays.

1.4.1.1 ELISAs for infliximab and adalimumab

For details of the ELISA kits please refer to Appendix 1. All three specified ELISA methods employ similar principles in which, typically, micro-titre plates with 96 wells coated with reagent receive the patient serum samples or various standards and calibrators. Reagents are added with wash steps between additions. The final step involves quantifying the amount of a peroxidase label in the titre well, this amount being proportional to the amount of anti-TNFα or anti-drug antibody in the patient's sample or in the calibrator standard.[15]

The amount of peroxidase present in the well is quantified using a timed incubation with excess substrates (hydrogen peroxide + 3,3',5,5'-tetramethylbenzidine). Peroxidase catalyses the following reaction: Tetramethylbenzidine + hydrogen peroxide → chromogen + water
The incubation is stopped after an appropriate time by the addition of acid and the accumulated chromogen quantified by measuring optical density with a spectrophotometer.

The reagents used for coating the microtitre plate wells and the reagents used in subsequent steps of the assay procedure differ in detail according to manufacturer. The LISA-TRACKER assays for Infliximab and for adalimumab are illustrated in Figure 3.

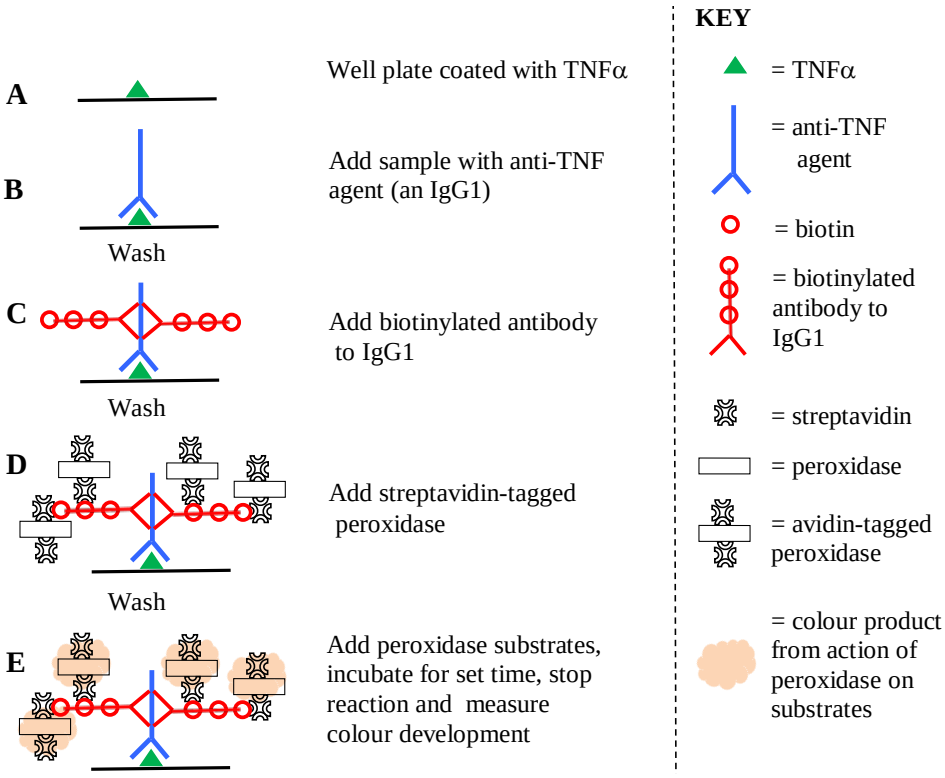


Figure 3 Diagrammatic representation of the LISA-TRACKER assay for infliximab and adalimumab
Procedural steps C and D are detection steps that function to detect the anti-TNFα that is bound to the well surface via TNFα, ensuring a quantitative relationship between anti-TNFα and peroxidase. Step E quantifies the amount of peroxidase (and therefore anti-TNFα) in the titre well (Streptavidin has four very high affinity binding sites for biotin).

Serum samples from patients may contain soluble TNFα receptors; these could compete with anti-TNFα for the immobilised TNFα on the well plate and may potentially interfere with the assay. The assay quantifies free anti-TNFα. Samples may contain anti-TNFα bound to antibodies to anti-TNFα, especially in patients who have lost a response to treatment. These anti-TNFα-antibody complexes will be washed away at the first wash step leaving only free anti-TNFα bound to immobilised TNFα. The amount of anti-TNFα lost at the wash step is likely to vary between patients and is unknown; the practical implications of this are uncertain.[15]

TNFα-Blocker and Promonitor assays for anti-TNFα drugs differ from the LISA-TRACKER assay in that the well-coat is not TNFα but rather a reagent (antibody or antibody-fragment) able to bind specifically to the TNFα binding site of infliximab or of adalimumab that is added to the well in patient's sample (or calibrator). After washing, the second reagent is a peroxidase labelled antibody able to bind the Fc region of the anti-TNFα (Figure 4). Thus fewer steps and a single reagent are used to detect well-bound anti-TNFα drug. Table 2 summarises the information describing the principle of these assays.

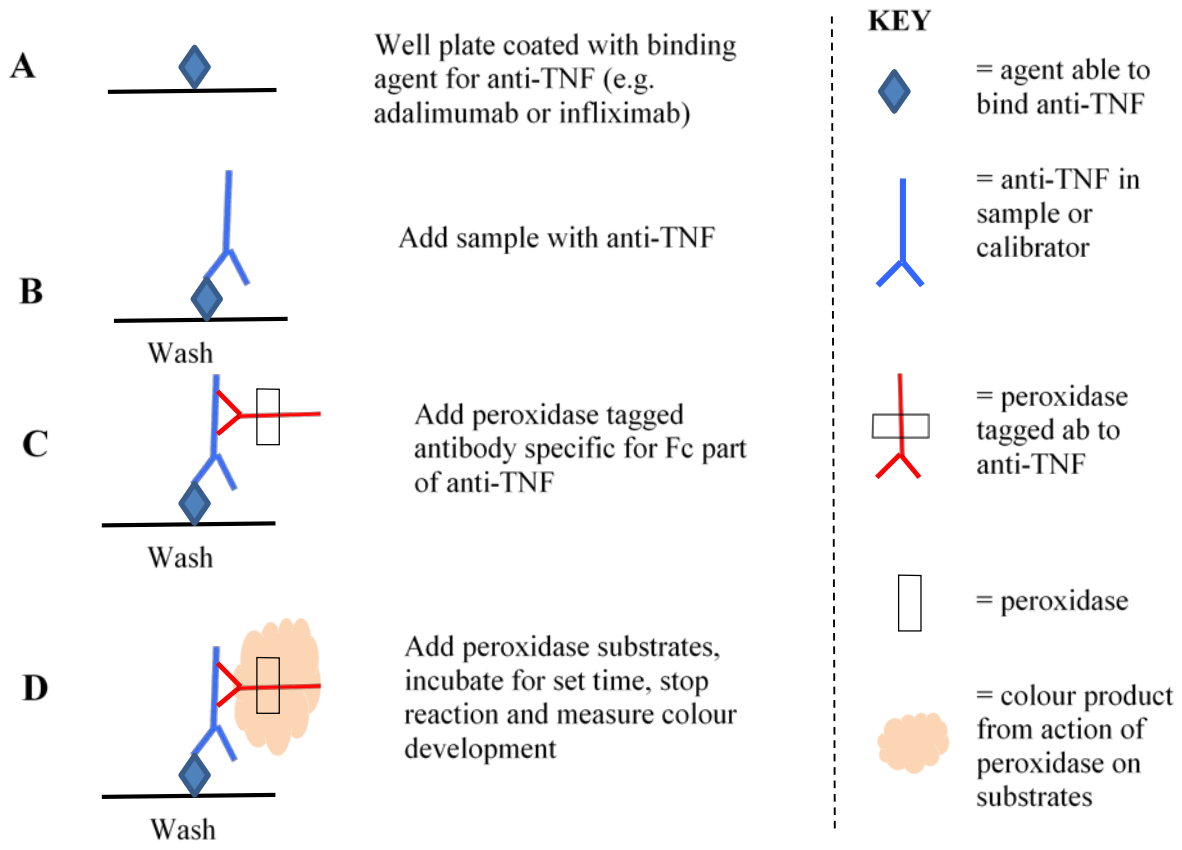


Figure 4 Diagrammatic representation of TNFα-Blocker and Promonitor assays for infliximab and adalimumab

Table 2 Summary of ELISAs to be considered in this review for detection of infliximab and adalimumab

Manufacturer (Kit)	Microplate pre-coat	Detection reagent(s)	
LISA-TRACKER Indirect ELISA	Recombinant human TNFα	Biotinylated mouse monoclonal IgG antibody directed to IgG Fc fragment	Avidin-tagged peroxidase
TNFα-Blocker ELISA ¥	Monoclonal anti-TNFα antibody *	Peroxidase labelled antibody **	
Promonitor ELISA ¥	Monoclonal anti-TNFα antibody***	Peroxidase labelled monoclonal anti-TNFα antibody****	

¥ further details supplied were labelled CIC

[redacted]

1.4.1.2 ELISAs for anti-drug antibodies (anti-drug antibodies)

The LISA-TRACKER assays for antibodies to infliximab and to adalimumab are illustrated in Figure 5.

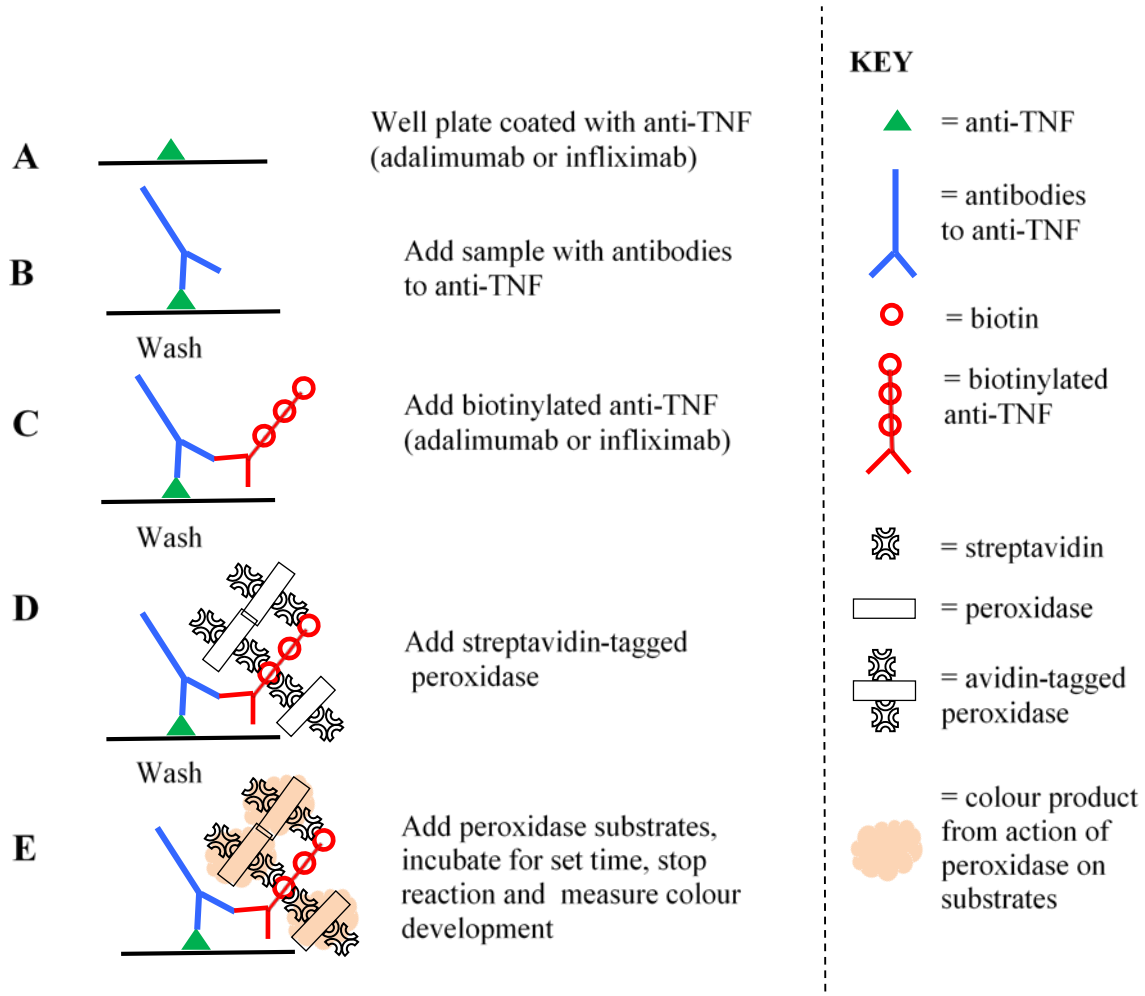


Figure 5 Diagrammatic representation of the LISA-TRACKER assay for antibodies to infliximab or to adalimumab
This is a bridging assay in which antibodies to anti-TNFα in the patient sample bridge the immobilised anti-TNFα to the biotinylated anti-TNFα. Procedural steps C and D are detection steps that function to detect the sample antibodies, ensuring a quantitative relationship between anti-TNFα antibodies and peroxidase. Step E quantifies the amount of peroxidase (and therefore anti-TNFα antibodies. Streptavidin has four very high affinity binding sites for biotin)

This assay only quantitatively estimates free antibodies to anti-TNFα. Thus anti-drug antibodies bound to the drug are lost at the first wash. The amount of bound anti-drug antibody is likely to vary between patients and is unknown. Whether anti-drug antibodies directed at non-idiotypic regions of the drugs (e.g., glycoprotein moieties, variable non-idiotypic mouse regions of infliximab etc.) are detectable or present in samples appears to be insufficiently investigated to date and is therefore uncertain. However, in-vitro tests indicate that about 90% of anti-drug antibodies bind to the TNF-

binding region of anti-TNFα drugs.[50] These, and the other anti-drug antibodies, may hasten clearance of drug from the circulation as well as neutralising its binding capacity.

TNFα-Blocker and Promonitor differ from LISA-TRACKER in employing a single reagent for detecting well-bound anti-drug antibodies rather than two (biotinylated infliximab or biotinylated adalimumab, plus avidin conjugated peroxidase). Table 3 summarises the information describing the principle of these assays.[15]

Table 3 Summary of ELISAs to be considered in this review for detection of antibodies to infliximab and adalimumab

Manufacturer (Kit)	Microplate pre-coat	Detection reagent(s)	
LISA-TRACKER (bridge ELISA)	Anti-TNFα (i.e. infliximab or adalimumab)	Biotinylated anti-TNFα that binds to the paratope of the anti-drug antibodies in the sample	Avidin-tagged peroxidase
TNFα-Blocker ELISA infliximab ¥	[redacted]	[redacted]	
TNFα-Blocker ELISA adalimumab ¥	[redacted]	[redacted]	
Promonitor ELISA ¥	[redacted]	[redacted]	
¥further details supplied were labelled CIC. [redacted]			

1.4.1.3 Brief overview of identified non-ELISA assay methods

There are no "gold standard" assays for anti-TNFα agents or for antibodies to anti-TNFα agents which might provide a robust basis for comparisons between the performances of different assays. According to the US Medical Insurance assessments "candidate" gold standards have been insufficiently investigated to establish any as a gold standard, and according to Steenholdt et al. (2013)[59] the evidence is incomplete on how these different assays may compare in practice.[14, 60-63]

There appear to be four types of assay which differ fundamentally from each other. These are:
(a) ELISAs - solid phase assays. These are available as commercial kits and several "in-house" methods are mentioned in the literature. The ELISAs generally only quantitatively measure "free" anti-TNFα and "free" anti-drug antibodies and it is acknowledged that the level of the unmeasured "bound" anti-TNFα and of "bound" anti-drug antibody may vary considerably between patients. Thus for some patient samples there is an unknown and unmeasured amount of anti-TNFα and of anti-drug antibody present, in addition to the measured "free" levels. In theory this represents a potential deficiency in the ELISA assays, although whether in practice this is serious is difficult to gauge especially in the absence of an established gold standard. This deficiency appears to have been one stimulus for the development of methods based on alternative principles. It is possible however that

the relative convenience and cheapness of ELISAs means that this inability to measure total anti-TNFα and total anti-drug antibody is supportable in practice.

(b) Radioimmunoassays (RIA) – liquid phase. These appear to be provided as a total service rather than as purchasable kits. They appear to measure total anti-TNFα and total anti-drug antibody (probably as long as the anti-drug antibody light chain is lambda class). These RIAs use 125 iodine-labelled human TNFα and 125 iodine-labelled anti-TNFαs. These are commercially available or may be relatively easily constructed from commercially available materials, however in the absence of purchasable assay kits, it is unlikely that any hospital laboratory would set up such assays for routine use. In these assays the patient's sample is mixed with a solution containing a fixed amount of 125 iodine-labelled TNFα or 125 iodine-labelled anti-TNFα further antibody (e.g., rabbit anti-human immunoglobulin λ-chain) which promotes the formation of immune complexes which are pelleted by centrifugation. Radio-iodine in the pellet is quantified in a gamma-counter. Potential disadvantages include the following i) radio-labelled reagents do not store indefinitely (125 iodine decays with a half-life of 59 days), ii) the laboratory needs to be equipped for handling hazardous (radioactive) material, iii) some staff training may be necessary, and iv) the laboratory requires a gamma counter (preferably automated for high throughput). These factors obviously have cost implications for setting up RIAs.

(c) Cell Reporter Assays

These assays utilise genetically engineered cells that respond to the presence of anti-TNFα agents by synthesising light generating enzymes. The enzymes are allowed to accumulate during an incubation period and are then supplied with appropriate substrates resulting in light emission measured with a luminometer. Samples with anti-TNFα will lead to light emission, and samples with anti-bodies to anti-TNFα will quench light emission (for further information see Appendix 2).

(d) Mobility Shift Assays

The mobility shift assay depends on detecting the shift in mobility of fluorescent probes when bound to either anti-TNFα or to anti-drug antibodies (for further information see Appendix 2).

1.4.1.4 Timing and use of assays

The anti-TNFα and anti-drug antibody assays are most frequently administered just before the next administration of the anti-TNFα agent. This is said to allow measurement of a "trough" level of anti-TNFα and has been adopted so as to minimise effects from the presence of anti-TNFα-anti-drug antibody immune-complexes in samples. For patients whose response to therapy has waned, the results of the tests are frequently dichotomised using a cut off assay result. Thus, on the basis of anti-TNFα assays patients are classified as having therapeutic levels of anti-TNFα or sub-therapeutic

levels, and on the basis of anti-drug antibody assay results they are classified as having clinically significant levels of anti-drug antibodies or insignificant levels. Such classifications yield four categories of patient for whom different explanations of failed response are possible. Algorithms have been developed prescribing treatment pathways and / or further diagnostic tests (e.g., colonoscopy) based on such classification.[15]

1.4.2 Current usage of assays in the NHS

Current practice for monitoring TNFα inhibitor antibody and drug levels in the UK is patchy due to the lack of agreed consensus and evidence for its cost effectiveness. In-house tests are performed in a few laboratories in England. However demand is low, analyses are often undertaken in batches, and it can be weeks (in some cases) before a clinician receives a result on which to act.

Whilst some centres have local monitoring protocols in conjunction with their link laboratory there is as yet no agreed algorithm for clinicians to refer to which allows for the translation of the results of the tests into coherent plans for patient management according to test outcome.

However recent emerging evidence to support anecdotal practice that such monitoring could be useful in managing patients with TNFα inhibitors, has encouraged a cautious increase in uptake.

It is anticipated that therapeutic monitoring of TNFα inhibitors, might be useful in a number of clinical scenarios in treatment of Crohn's disease in the NHS including primary and loss of response to anti-TNFα therapy and in optimisation of dosages for those who are already responding.

2 DEFINITION OF DECISION PROBLEM

This report undertaken for the NICE Diagnostics Assessment Programme examines the clinical and cost effectiveness of ELISA tests (LISA-TRACKER EISA kits, TNFα-Blocker ELISA kits, and Promonitor ELISA kits) for measurement of patient blood levels of anti-TNFα agents (Infliximab and Adalimumab; also known as TNFα inhibitors) and of antibodies to these agents (i.e., anti-drug antibody levels, anti-drug antibodies) in people with CD whose disease responds to treatment with TNFα inhibitor or who experience loss of response during TNFα inhibitor therapy. The report will help NICE to make recommendations about how well the assays work and whether the benefits are worth the cost of the tests for use in the NHS in England. The assessment will consider both clinical improvement in patients' symptoms and the cost of the tests used to measure the amount of anti-TNFα and anti-drug antibodies in patients' sera using evidence identified through systematic reviews and information submitted to NICE during the evaluation process by the companies offering the ELISA tests.

The decision questions for this project are shown in the box below:

1. Does concurrent testing of TNFα inhibitor levels and antibodies to TNFα inhibitors represent a clinically and cost-effective use of NHS resources in people with Crohn's disease whose disease responds to treatment with TNFα inhibitor?

Testing will be carried out:

a) 3 to 4 months after start of treatment or

b) 3 to 4 months and every 12 months from start of treatment

2. Does concurrent testing of TNFα inhibitor levels and antibodies to TNFα inhibitors represent a clinically and cost-effective use of NHS resources in people with Crohn's disease who experience loss of response during maintenance treatment with TNFα inhibitor?

3. Does testing of TNFα inhibitor levels followed by reflex testing of antibodies to TNFα inhibitors if drug level is undetectable represent a clinically and cost-effective use of NHS resources in people with Crohn's disease whose disease responds to treatment with TNFα inhibitor?

Testing will be carried out:

a) 3 to 4 months after start of treatment or

b) 3 to 4 months and every 12 months from start of treatment

4. Does testing of TNFα inhibitor levels followed by reflex testing of antibodies to TNFα inhibitors if drug level is undetectable represent a clinically and cost-effective use of NHS resources in people with Crohn's disease who experience loss of response during maintenance treatment with TNFα inhibitor?

2.1 Overall aim of the assessment

The overall aim of this report was to present the evidence on the clinical- and cost-effectiveness of monitoring infliximab and adalimumab and their antibodies in responders and patients with loss of response when ELISA test results are used in combination with an algorithm that prescribes treatment pathways for the management of patients with specific drug and anti-drug antibody levels.

2.2 Objectives

In the current report we addressed the following objectives:

Objective A - review of comparative performance of tests

To review and critique studies of a comparison (including relative test performance) of two or more intervention tests, or studies which compare an intervention test with a test method which can be used to perform a linked evidence assessment and to supplement these with data submitted by the relevant companies if of sufficient detail and quality.

To compare and contrast studies that reported a threshold analysis to determine the optimal drug level cut-off to predict or diagnose response.

The following objective from the protocol was moved into the introduction as it does not address the decision questions:

To provide a technical description, and evaluation, of the listed intervention tests used for Crohn's disease in therapeutic monitoring of TNFα inhibitors (infliximab and adalimumab) and their respective antibodies including what the assays measure and the mechanisms of the assays.

Objective B – description of algorithms

To describe algorithms used in studies which include data on one or more intervention test or on a test which allows a linked evidence approach to be performed (i.e., algorithms used in studies identified in Objective C). The studies are required to provide an algorithm and report clinical outcomes for the management of patients with Crohn's disease following measurement of serum levels of anti-TNFα drug and anti-drug antibodies.

To compare the algorithms used following therapeutic drug monitoring to the algorithms specified in the TAXIT study for responders,[64] and in the reporting of loss of response (algorithm adapted from the study by Scott and Lichtenstein, 2014[65]).

Objective C1 – review of clinical effectiveness of test with algorithm combinations

To systematically review the literature comparing the clinical effectiveness of [a] the intervention assays for anti-TNFα agents and/ or for anti-drug antibodies used in conjunction with a treatment algorithm in Crohn's patients treated with infliximab or adalimumab; with [b] standard care (no tests performed or test-informed algorithm used) in Crohn's disease patients treated with infliximab or adalimumab.

To assess and critique available evidence on the comparison of standard care with other test assays used in conjunction with an algorithm, and on test performance compared with the study interventions (LISA-TRACKER ELISA kits, TNFα-Blocker ELISA kits, and Promonitor ELISA kits) (see Objective A).

Objective C2 – analysis of correlation between test results and clinical state

To analyse correlation studies which investigate the relationship between tests measuring anti-TNFα and anti-drug antibody levels and clinical outcome in terms of response in patients with Crohn's disease. This objective was added because of the paucity of management studies which address the decision questions to generate information for economic modelling.

Objective D – review of cost effectiveness of test with algorithm combinations

To assess the cost-effectiveness of employing anti-TNFα and anti-TNFα antibody monitoring with LISA-TRACKER ELISA kits, TNFα-Blocker ELISA kits, and Promonitor ELISA kits in patients with Crohn's disease compared with standard care (no anti-TNFα monitoring).

To use a linked evidence approach where necessary (see Objective C above) in which evidence of clinical effectiveness is taken from studies using alternative tests and an assessment is made of the relative performance of these tests relative to the intervention assays.

3 CLINICAL EFFECTIVENESS REVIEW

3.1 Clinical Effectiveness Methods

3.1.1 Identification and selection of studies

3.1.1.1 Search strategies for clinical effectiveness

An iterative procedure was used to develop the initial Medline search, with reference to our own scoping searches and those undertaken by information specialists at NICE. Known articles were consulted and checked for relevant terms. Additional phrases were added to find relevant articles that did not include terms for the test name or type of test (e.g., Baert et al., 2003[42]) or population (e.g., Vande Casteele et al., 2012[66]) in title, abstract or indexing. This search developed for Medline was adapted as appropriate for other databases and sources. The searches for each source are provided in Appendix 3. Searches for studies for cost and quality of life were developed separately.

The search strategy comprised the following main elements:

- Searching of electronic bibliographic databases
- Contact with experts in the field
- Scrutiny of references of included studies
- Screening of manufacturers' and other relevant organisations' websites for relevant publications

Bibliographic databases:
MEDLINE; MEDLINE In-Process & Other Non-Indexed Citations; EMBASE; Cochrane Library (including Cochrane Systematic Reviews, DARE, CENTRAL, NHS EED, and HTA databases); Science Citation Index and Conference Proceedings (Web of Science); Index to Theses; DART-Europe; Dissertations & Theses; NIHR Health Technology Assessment Programme; PROSPERO (International Prospective Register of Systematic Reviews).

The following trial and patent databases were also searched: Current Controlled Trials; ClinicalTrials.gov; UKCRN Portfolio Database; WHO International Clinical Trials Registry Platform; Espacenet (European Patent Office).

Specific conference proceedings, selected with input from clinical experts and Specialist Committee Members, were also checked for the last five years:

- European Crohn's and Colitis Organisation (ECCO)
- Digestive Diseases Week (DDW) (meeting of the American Gastroenterology Association (AGA))
- British Society of Gastroenterology (BSG)

- United European Gastroenterology Week (UEGW)
- American College of Gastroenterology

The following online resources of various health services research agencies, regulatory bodies, professional societies and manufacturers were consulted via the Internet:

- International Network of Agencies for Health Technology Assessment (INAHTA) Publication http://www.inahta.org/
- FDA medical devices: http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/Databases/default.htm
- European Commission medical devices http://ec.europa.eu/health/medical-devices/
- Theradiag http://www.theradiag.com/en/
- Immundiagnostik http://www.immundiagnostik.com/en
- Proteomika http://www.proteomika.com/
- American College of Gastroenterology http://gi.org/
- European Crohn's and Colitis Organisation (ECCO)http://www.ecco-ibd.eu
- British Society of Gastroenterology (BSG) http://www.bsg.org.uk
- United European Gastroenterology (UEG) https://www.ueg.eu/
- The American Gastroenterology Association (AGA) http://www.gastro.org

The reference lists of included studies and relevant review articles were checked. Citation searches of selected included studies were undertaken using Scopus. Identified references were downloaded in Endnote X7 software. Included papers were checked for errata using PubMed. Database searches were undertaken in October and November 2014.

3.1.1.2 Inclusion and exclusion of relevant studies

During the initial inclusion / exclusion process we identified three different categories of studies which were of interest for the review. 1) Studies comparing the performance of different types of assays (assay type comparison studies addressing Objective A) 2) Studies reporting an algorithm for the management of patients with drug and /or anti-drug antibody level test results (management studies addressing Objectives B and C1). 3) Studies reporting the correlation of drug and/ or anti-drug antibody levels with patient's clinical state (response) prospectively or retrospectively (correlation studies). The third category was included in addition to the original protocol objectives. The reason behind this was the fact that we anticipated finding a limited number of management studies to answer the decision questions. The correlation studies were included for two purposes:

1) To provide an overview of the variation in drug level thresholds used to predict clinical state (for Objective A - review of comparative performance of tests)

2) To pool test outcome data for responders and non-responders as an alternative to single study data to inform the economic model (Objective C2 – analysis of correlation between test results and clinical outcomes)

Assay type comparison studies were considered in two phases because the work involved in Objective A (review of comparative performance of tests) was dependent on the available evidence in Objective C1.

In the first phase, studies were included if they compared assay performance of two or more different test assays (Objective A section 3.2.2).

Once the management studies to be included were known, the second phase included comparison studies if they compared two or more types of intervention assay or any of the intervention assays with the assays used in the management studies in order to inform a linked evidence approach.

See below for detailed inclusion and exclusion criteria for the different objectives.

3.1.1.2.1 Inclusion criteria

Objective A

Studies comparing the test performance of two or more tests for infliximab or adalimumab levels and/or for anti-drug antibodies were identified. Studies were included either if they compared two or more intervention tests, or compared an intervention test with a test method which could be used to perform a linked evidence assessment. All study designs were considered for inclusion.

Objectives B and C1

Studies which satisfied the following criteria were included:

Population Crohn's disease (CD) patients (adults and children) receiving infliximab or adalimumab. Evidence on mixed patient groups containing CD and ulcerative colitis patients was included if CD patients made up more than 50% of the study population.

Intervention Use of LISA-TRACKER ELISA kits, TNFα-Blocker ELISA kits, and Promonitor ELISA kits to estimate plasma or sera levels of anti-TNFα agents and / or of anti-drug antibodies in which test results are employed in conjunction with a treatment algorithm (Table 4). Other assay methods were considered for a linked evidence approach (Table 4).

Comparator Standard care (Treatment decisions made on clinical judgement without measuring levels of anti-TNFα and anti-drug antibodies).

Outcome Any patient outcome (e.g., CDAI score based response rate, any measure of change in severity of CD including physicians global assessment; duration of response, relapse and remission; rates of hospitalisation; rates of surgical intervention; time to surgical intervention; adverse effects of treatment; health related quality of life; and secondary if two strategies compared are found clinically equivalent: time to result; number of inconclusive results; frequency of dose adjustment; frequency of treatment switch).

Study design All study designs were considered for inclusion.

Healthcare setting Secondary and tertiary care.

Objective C2

Correlation studies were included if they provided at least one of the following:

1) A ROC threshold analysis to determine an optimal drug level threshold for predicting response (section 3.2.2.5)
2) Sufficient data to complete a 2x2 table of diagnostic accuracy of drug / anti-drug antibody level for prediction of response / LOR. (Objective C2 – analysis of correlation between test results and clinical outcomes section 3.2.5)

Population Crohn's disease (CD) patients (adults and children) receiving infliximab or adalimumab. Evidence on mixed patient groups containing CD and ulcerative colitis patients was included if CD patients made up more than 50% of the study population.

Intervention Any assay to measure anti-TNFα and / or anti-drug antibody levels.

Outcome 2x2 data of diagnostic performance of test to predict patient response / non-response and / or ROC analysis reporting optimal drug level thresholds to predict response / non-response.

Study design All study designs were considered for inclusion.

3.1.1.2.2 Exclusion criteria

Population	Studies with mixed patient groups containing <50% CD patients
Intervention	Studies reporting an algorithm where patient management was not dependent on a prescriptive algorithm
Study design	Narrative reviews
	Systematic reviews of correlation studies without meta-analysis
	Editorials / letters without original data
	Non-English language papers

Table 4 Assay methods included as interventions in the review

LISA-TRACKER assay kits (Theradiag/Alpha Laboratories)

- LISA-TRACKER Adalimumab (LTA002)
- LISA-TRACKER Infliximab (LTI002)
- LISA-TRACKER anti-Adalimumab (LTA003)
- LISA-TRACKER anti-Infliximab (LTI003)
- LISA-TRACKER Duo Adalimumab (LTA005)
- LISA-TRACKER Duo Infliximab (LTI005)

Immundiagnostik TNFα-Blocker ELISA kits (Immundiagnostik/BioHit Healthcare):

- Immundiagnostik TNFα-Blocker ADA, antibodies against infliximab (e.g. Remicade®) ELISA (K9650)
- Immundiagnostik TNFα-Blocker ADA, antibodies against adalimumab (e.g. Humira®) ELISA (K9652)
- Immundiagnostik TNFα-Blocker ADA, TOTAL antibodies against infliximab (e.g. Remicade®) ELISA (K9654)
- Immundiagnostik TNFα-Blocker ADA, TOTAL antibodies against adalimumab (e.g. Humira®) ELISA (K9651)
- Immundiagnostik TNFα-Blocker monitoring, infliximab drug level (e.g. Remicade®) ELISA (K9655)
- Immundiagnostik TNFα-Blocker monitoring, adalimumab drug level (e.g. Humira®) ELISA (K9657)

Promonitor ELISA kits (Proteomika):

- Promonitor-ADL ELISA (5080230000)
- Promonitor-IFX ELISA (5060230000)
- Promonitor-ANTI-ADL ELISA (5090230000)
- Promonitor-ANTI-IFX ELISA (5070230000)

For Objective C1 (review of clinical effectiveness of test with algorithm combinations): Tests that

were not included under the heading of intervention but for which evidence was available on comparative diagnostic performance when compared to an intervention test and where clinical outcomes were also reported, were included for the purpose of performing linked evidence modelling (these included: radioimmunoassays, cell reporter assays, liquid-phase mobility shift assays and in-house ELISAs).

3.1.2 Using the information provided by Theradiag / Alpha Laboratories, Immundiagnostik and Proteomika

The information provided by Theradiag / Alpha Laboratories, Immundiagnostik and Proteomika (see Appendix 4 for an itemised list of documents received) was screened for three purposes:

1) Additional studies not identified by our searches
2) Information for the technical description of the three intervention assays
3) Information an assay comparisons

Additionally, we sought detailed information from Theradiag / Alpha Laboratories, Immundiagnostik and Proteomika by e-mail regarding mechanisms and reactants (in particular specificities and properties of antibodies and other reagents) employed in the three intervention ELISA tests.

3.1.3 Review strategy

The general principles recommended in the PRISMA statement were used.[67] Records rejected at full text stage and reasons for exclusion were documented. Two reviewers independently screened the titles and abstracts of all records identified by the searches and discrepancies were resolved through discussion. Disagreement was resolved by retrieval of the full publication and consensus agreement. Full copies of all studies deemed potentially relevant were obtained and two reviewers independently assessed these for inclusion; any disagreements were resolved by consensus or discussion with a third reviewer.

3.1.4 Data extraction strategy

Data were extracted by one reviewer, using a piloted, data extraction form. Completed data extraction forms are available in Appendix 5. A second reviewer checked the extracted data and any disagreements were resolved by consensus or discussion with a third reviewer.

3.1.5 Quality assessment strategy

3.1.5.1 Objective A

For Objective A quality appraisal was completed using a modified QUADAS-2 checklist.[68] In the patient selection domain three questions were included, using the standard version of the tool:

- Was a consecutive or random sample of patients enrolled?
- Was a case-control design avoided?
- Did the study avoid inappropriate exclusions?

An additional question asking for the range of drug and antibody concentrations was added before a judgement was made about applicability. The applicability question was adapted to:

- Is there concern that the included patients or range of drug / antibody concentrations do not match the review question?

Regarding the index test the two standard questions were included to assess risk of bias:

- Was the threshold pre-specified?
- Were index tests interpreted without knowledge of the reference standard?

One additional question was added:

- Were the number of failed results and measurement repeats reported?

For the reference standard the two standard questions were used:

- Were the reference standard results interpreted without knowledge of the results of the index test?
- Is the comparison test likely to correctly classify the target condition?

The best reference standard to use would be use of standardised spiked samples. Use of spiked samples as a reference standard would allow the accuracy of tests to be compared with reference to the true drug and antibody levels, and would avoid the biases associated with imperfect reference standards.

However, where spiked samples were unavailable, one of the comparator tests used in the management studies and considered for a linked evidence approach was used. If the reference standard was one of the four comparator tests then it was classified as unlikely to correctly classify the target condition. This is because due to the lack of evidence they constitute an imperfect reference standard. For both comparator and index tests, judgements regarding applicability considered both the test used and the threshold applied.

For flow and timing the following standard questions were included:

- Was there an appropriate interval between intervention test and comparison test(s)?
- Did patients receive the same reference standard?

- Were all patients included in the analysis?

An additional question was included:

- Were both intervention test and reference standard conducted on all samples?

This is to measure whether some patients or samples did not receive any of the index tests. This is of particular concern if the reason for being omitted may be related to the probability of a positive or negative result.

3.1.5.2 Objective C1 review of clinical effectiveness of test with algorithm combinations

RCTs meeting the inclusion criteria were assessed using the Cochrane risk of bias tool.[69] The Downs and Black (1998) checklist[70] was used to assess the quality of non-RCTs meeting the inclusion criteria. The results of the quality assessment provide an overall description of the quality of the included studies and a transparent method of recommendation for design of future studies. Quality assessment was undertaken by one reviewer and checked by a second reviewer, any disagreements were resolved by a third reviewer through discussion.

3.1.6 Methods of analysis/synthesis

3.1.6.1 Objective A - review of comparative performance of tests

We mapped included studies according to the comparisons they undertook. A narrative was produced to summarise the studies which compared the performance of the intervention assays and assays suitable for a linked evidence approach and considering the concordance between the tests. This was assessed using the following outcomes:

1. Concordance between tests (split by reference standard results positive and reference standard results negative or clinical outcomes where available) for therapeutic drug and detectable anti-drug for all index tests and comparators
2. Characteristics of cases where there was disagreement and agreement between tests
3. Bland-Altman plots to show patterns of correlation

The specific measures of concordance used were percentage agreement between the tests (split between reference standard results positive and reference standard results negative samples where available) and Cohen's Kappa. Two main secondary outcomes were also collected. Firstly characteristics of cases where there was disagreement and agreement between tests, which may provide information about the reason for and implications of the discordant results. Secondly the shape of the Bland-Altman plots as the plot shows whether the difference between the two tests is dependent on absolute drug and antidrug levels. Mean bias and the upper and lower limits of

agreement were not particularly informative here as we are only interested in one cut-point not the whole range of concentrations. Pearson's correlation coefficient was not considered in detail as it can have high values even when clinically meaningful differences are present.[71] Where there were sufficient studies, meta-analysis of Cohen's Kappa was considered.[71]

3.1.6.2 Objective B Description of algorithms prescribing patient management following test outcomes for drug and / or anti-drug antibody levels

Algorithms used in management studies were described narratively and compared to the algorithm adapted from Scott and Lichtenstein (2014) (for loss of response)[65] and to the algorithm adapted from Vande Casteele et al. (2015) (for responders).[72] Patients or decisions non-compliant with the stated algorithm were quantified.

Time of testing, sequence of testing (drug and antibodies) and sequence of analysis were also considered.

3.1.6.3 Objective C1 Clinical studies evaluating drug monitoring for the management of Crohn's disease patients (management studies)

Depending on the available evidence, analyses were stratified according to the type of ELISA or other assay, type of drug (infliximab or adalimumab) and patient group (patients with loss of response or responders).

Study, treatment, population, and outcome characteristics were summarised and compared qualitatively and where possible quantitatively in text and graphically and in evidence tables. Pooling study results by meta-analysis was considered however meta-analysis was unsuitable for the data identified and we employed a narrative synthesis using text and tables. A detailed commentary on the major methodological problems and biases affecting the studies was also included, together with a description of how this may have influenced individual study results.

We used a linked-evidence approach[2]. Evidence on outcomes reported by studies using other test methods (radioimmunoassay, liquid-phase mobility shift assay, in-house ELISAs) for patient management was linked to evidence on comparative test performance between our intervention tests and these other methods to allow for estimates of anticipated outcome for our intervention assays.

Time of testing, sequence of testing (drug and antibodies) and sequence of analysis were also considered.

Where relevant Kaplan-Meier plots were available, individual patient data (IPD) was reconstructed using the method of Guyot 2013.[73] Parametric models were fitted to reconstructed IPD using STATA version 11.

Objective C-2

For Objective C-2 we aimed to:

a) Provide an overview of meta-analyses of studies addressing the relationship between drug and / or anti-drug antibody levels and clinical state of patients with CD by producing a narrative of identified systematic reviews with meta-analyses, presenting the reported meta-analyses results, and undertaking hierarchical meta-analysis of the data presented in the systematic reviews.
b) Pool test accuracy data for prediction of patients' clinical state (response or lost response). This was done as a potentially useful supplement to management studies for informing the economic model.

Studies which provided dichotomised test results and related these to dichotomised clinical status were identified. In particular studies were sought that reported on both drug and anti-drug antibody test results for individual patients. Two by two data for tests were extracted, together with the type of test employed (e.g. ELISA, RIA, HMSA), the anti-TNFα administered, dose regimen, patient inclusion and exclusion criteria, timing of testing, method for establishing clinical status, test cut-off used, and study design were noted where these were reported. The populations of interest were: a] responders and responders who lost response; b] patients with loss of response who continued with loss of response or who regained a response.

Meta-analyses of single test studies were undertaken: a] to provide a pooled estimate for the probability of returning a specified test result after trough anti-TNFα testing (useful for estimating reflex strategy test result probabilities); b] to provide pooled estimates for the probability of returning a specified test results by single test (i.e. anti-TNFα or anti-drug antibodies) that can be compared for consistency with the corresponding probabilities from the few identified patient level studies.

Review Manager (RevMan) 5.3 (The Cochrane Collaboration, The Nordic Cochrane Centre, Copenhagen, Denmark) was used for analysis of sensitivities and specificities. Meta-analysis was undertaken in STATA version 11 using the metandi package.[74, 75] The prevalence of clinical status was meta-analysed using random effects model with "MetaAnalyst" software.

Given the prevalence (P) of the condition tested for, and the joint sensitivity (Sens) and specificity (Spec) values from meta-analysis, the probability of returning a positive test result is:

Positive test = [P * Sens] + ([1-P]*[1-Spec])

And the probability of returning a negative test result is:
Negative test = ([1-P] * Spec) + (P*[1-Sens])

3.2 Clinical effectiveness Results

3.2.1 Search results

Figure 6 provides the PRISMA flow diagram for Objectives A, B and C. A total of 2,428 records were identified through electronic searches. Six additional records were identified from other sources. The removal of duplicates left 1,616 records to be screened, of which 1,359 were excluded at title/abstract level as these were irrelevant to the decision questions. The remaining 257 records were examined for inclusion at full-text, of which 70 (reported in 68 studies) were included in the clinical effectiveness review. Table 5 summarises the 68 included studies and refers the reader to the relevant section where they are covered. Details on the reasons for excluding studies at full text can be found in Appendix 6.

The search of on-going trials in Clinical Trials.gov, Current Controlled Trials, UKCRN Portfolio, and WHOICTRP databases (carried out between 4th and 11th November 2014) retrieved 7 relevant ongoing trials (see Appendix 7).

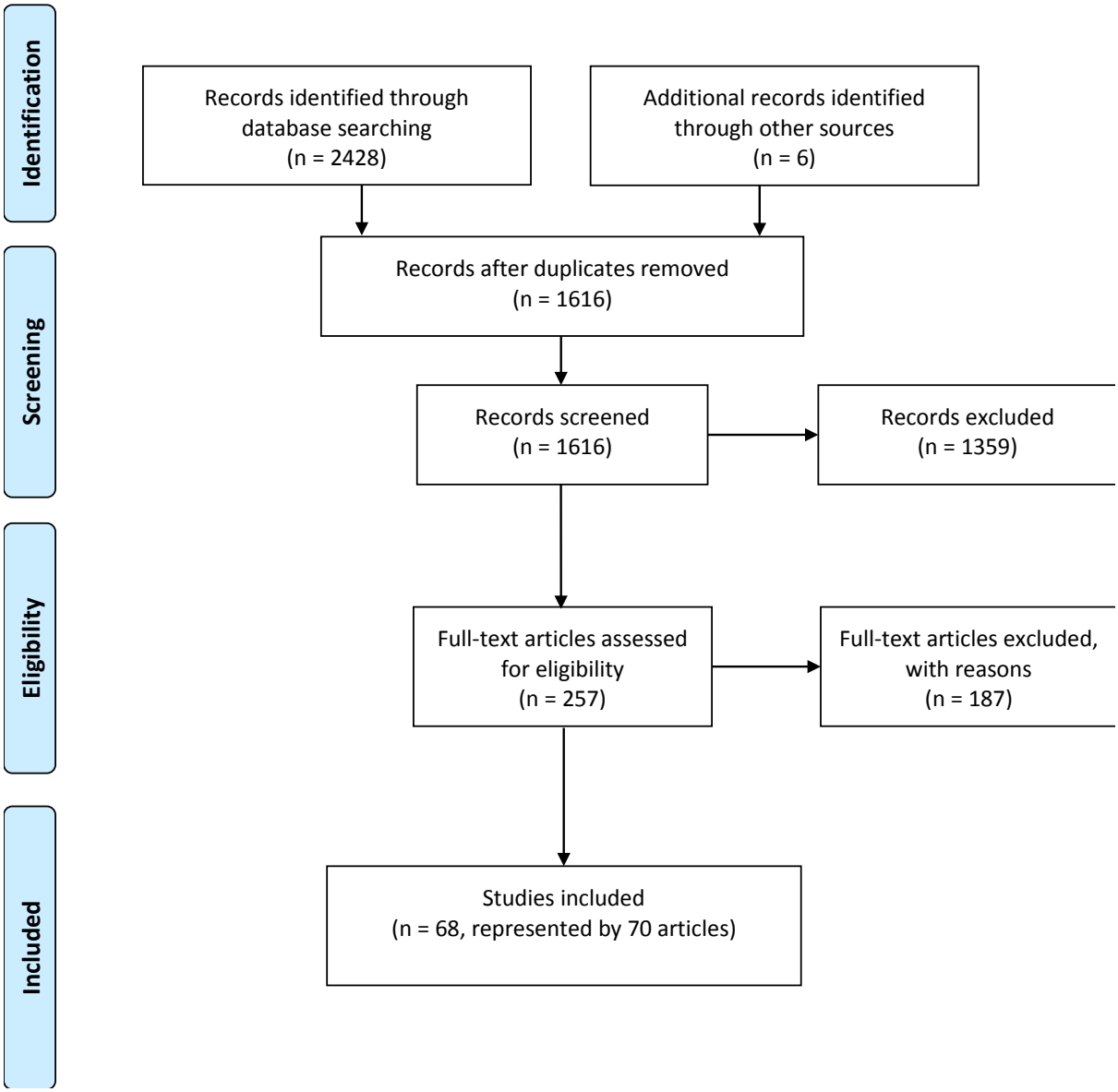


Figure 6 PRISMA diagram describing the selection of included studies for the clinical effectiveness review

Of the included studies summarised in Table 5, studies comparing assay types address the performance of the different assays for a linked evidence approach (Objective A – review of comparative performance of tests). The management studies address some aspects of the decision questions on the effectiveness of drug and anti-drug antibody monitoring (Objective B – description of algorithms and Objective C1 – review of clinical effectiveness of test with algorithm combinations). Correlation studies with a ROC threshold analysis are summarised in section 3.2.2.5 and correlation studies with sufficient 2x2 test accuracy data to contribute to the meta-analysis are reported in section (Objective C2 – analysis of correlation between test results and clinical outcomes). Columns are not mutually exclusive; studies feature in more than one section.

Table 5 Overview of utility of included studies

Reference(s) of included study	**Utility of included study and reference to relevant section in report**			
	Assay comparison Section 3.2.2	**Management stipulated by algorithm and test outcome Section 3.2.4**	**Correlation of drug / anti-drug antibodies and response – ROC Section 3.2.2.5**	**Correlation of drug / anti-drug antibodies and response – 2x2 Section 3.2.5**
Ainsworth 2008[46]	x	x	x	✓
Baert 2014[76]	x	x	✓	✓
Ben Bassat 2013[77] abstract	x	x	x	✓
Ben-Horin 2011[78]	x	x	x	✓
Ben-Horin 2012[79]	x	x	x	✓
Bodini 2014[80] abstract	✓ (mapping only)*	x	x	✓
Bortlik 2013[81]	x	x	✓	✓
Candon 2006[82]	x	x	x	✓
Chiu 2013[83]	x	x	✓	✓
Cornillie 2014[84]	x	x	✓	✓
Corstjens 2013[85]	✓ (mapping only)	x	x	x
Daperno 2013[86] abstract	✓	x	x	x
Dauer 2013[87] abstract	x	x	x	✓
Egea-Pujol 2013[88] abstract	✓	x	x	x
Eser 2013[89] abstract	✓	x	x	x
Eser 2013[90] abstract	✓	x	x	x

Reference(s) of included study	Utility of included study and reference to relevant section in report			
	Assay comparison Section 3.2.2	**Management stipulated by algorithm and test outcome Section 3.2.4**	**Correlation of drug / anti-drug antibodies and response – ROC Section 3.2.2.5**	**Correlation of drug / anti-drug antibodies and response – 2x2 Section 3.2.5**
Farrell 2003[91]	x	x	x	✓
Feagan 2012[92] abstract	x	x	✓	x
Frederiksen 2014[93]	x	x	✓	✓
Goldberg 2014[94] abstract	x	x	✓	x
Greathead 2014[95] abstract	✓ (mapping only)	x	x	x
Hanauer 2004[39]	x	x	x	✓
Hauenstein 2012[96] abstract	✓	x	x	x
Hibi 2014[97]	x	x	x	✓
Imaeda 2012[98]	✓ (mapping only)	x	x	✓
Imaeda 2014[99]	✓ (mapping only)	x	✓	✓
Imaeda 2014[100]	x	x	✓	x
Karmiris 2009[47]	x	x	✓	x
Kong 2011[101] abstract	x	x	x	✓
Kopylov 2012[102]	✓ (mapping only)	x	x	✓
Lee 2012[62]	x	x	x	✓ SR
Levesque 2014[103]	x	x	✓	x
Marits 2014[104]	x	x	✓	x
Marzo 2014[105] abstract	x	x	x	✓
Maser 2006[37]	x	x	x	✓
Mazor 2013[106] abstract	x	x	✓	x
Mazor 2014[107]	x	x	✓	✓
McTigue 2013[108] abstract	✓ (mapping only)	x	x	x
Nagore 2015 unpublished abstract provided by Proteomika	✓	x	x	x
Nagore 2015[109]	x	x	✓	✓

Reference(s) of included study	Utility of included study and reference to relevant section in report			
	Assay comparison Section 3.2.2	**Management stipulated by algorithm and test outcome Section 3.2.4**	**Correlation of drug / anti-drug antibodies and response – ROC Section 3.2.2.5**	**Correlation of drug / anti-drug antibodies and response – 2x2 Section 3.2.5**
abstract				
Nanda 2013[110]	x	x	x	✓ SR
Pallagi-Kunstar 2014[111]	x	x	✓	x
Pariente [58]	x	x	x	✓
Paul 2012[112] abstract	x	x	✓	x
Paul 2013[57]	x	x	✓	x
Paul 2014[113]	x	x	x	✓ SR
Roblin 2014[114]	x	x	✓	✓
Ruiz-Arguello 2013[115]	✓	x	x	x
Schatz 2013[116]	✓	x	x	x
Semmler 2013[117] abstract	✓ (mapping only)	x	x	x
Singh 2014[118]	x	x	✓	x
Steenholdt 2011[119]	x	x	✓	✓
Steenholdt 2013[120]	✓ (mapping only)	x	x	x
Steenholdt 2013[51]	x	x	x	✓
Steenholdt 2014[121]	✓	✓	x	x
Steenholdt 2014[122]	✓	✓	x	✓
Steenholdt 2015[123]	x	✓	x	x
Ungar 2014[124] abstract	✓ (mapping only)	x	x	x
Vande Casteele 2013[125]	✓	x	✓	✓
Vande Casteele 2012[66]	✓	x	x	x
Vande Casteele 2014[126] abstract	✓ (mapping only)	x	x	x
Vande Casteele 2015[72]	x	✓	x	x
Vaughn 2014[127]	x	✓	x	x
Wang 2010[128]	✓	x	x	x

Reference(s) of included study	Utility of included study and reference to relevant section in report			
	Assay comparison Section 3.2.2	**Management stipulated by algorithm and test outcome Section 3.2.4**	**Correlation of drug / anti-drug antibodies and response – ROC Section 3.2.2.5**	**Correlation of drug / anti-drug antibodies and response – 2x2 Section 3.2.5**
abstract				
Wang 2011[129] abstract	✓	x	x	x
Wang 2012[130]	✓	x	x	x
Ward 2013[131] abstract	x	x	✓	x
West 2008[132]	x	x	x	✓
Yanai 2012[133] abstract	x	x	x	✓
Yarur 2013[134] abstract	x	x	✓	x
Total of included references in each section	**26 (11 mapping only)**	**5**	**24**	**31 and 3 SRs**

*Studies comparing assays other than the intervention assays or as comparator for the linked evidence approach were mapped for their assay type comparison but not further considered in the assessment
SR – Systematic Review; ROC – Receiver Operating Characteristic

3.2.2 Objective A - Review of comparative performance of test assays measuring anti-TNFα and / or anti-drug antibody levels

3.2.2.1 Aim

To compare the performance of the different index tests (three specified ELISA kits) to one another, and to comparator tests which can be used to perform a linked evidence approach in order to answer the question:

Do the index tests agree with each other and with the comparator tests with regard to whether therapeutic levels of drug and detectable levels of antidrug antibodies are present, and therefore will using the tests lead to the same clinical decisions? Comparator tests here are tests with known links to improving patient outcomes from prospective studies with pre-specified algorithms (management studies).

3.2.2.2 Rationale

In a typical linked evidence approach, test accuracy studies detect cases of disease, and are linked to studies which show evidence of treatment effectiveness in cases detected. The test accuracy studies in this review produce four results:

i) drug + | anti-drug antibody +,
ii) drug + | anti-drug antibody –,
iii) drug – | anti-drug antibody +,
iv) drug – | anti-drug antibody –.

However no trials were found which used the specified index tests to direct treatments for these four patient groups.

The linked evidence approach we therefore used was to evaluate the evidence showing whether any comparator tests (drug and anti-drug antibody tests) used in CD patients improve outcomes (typically a test-treat type of trial), and to assess the accuracy of the index tests versus these comparator tests. These comparator tests form an imperfect reference standard and a simple calculation of sensitivity and specificity of the ELISA index tests against these tests (e.g. using HMSA or RIA) as reference standards might result in either over- or under-estimation due to verification bias.

We took studies which investigated if testing for drug and anti-drug antibodies can improve patient outcomes through choosing the treatment prescribed by an algorithm (i.e. in a full RCT), and linked this to the index tests. This method may work satisfactorily if the RCT is of good quality, and if there is good evidence for high concordance between the index and (imperfect) reference or comparator tests. Where there are discordant results between tests we do not know which is correct. Alternatively for spiked samples we know which test is correct, but not whether this would have any impact on clinical outcomes. Therefore the main outcome for Objective A was evaluation of the concordance between the tests. This approach is appropriate for interpreting and synthesising data when the reference standard is imperfect.[71]

3.2.2.3 Results of assay type comparison studies

The search identified 25 relevant studies (reported in 26 references) which compared two or more assays to measure anti-TNFα and / or anti-drug antibody levels in CD patients (Figure 7). Of these 10 were full texts (reported in 11 references),[66, 85, 98, 99, 102, 115, 120-122, 125, 130] and the remainder were conference abstracts including one unpublished abstract provided by Proteomika (Nagore et al., 2015).[80, 86, 88-90, 95, 96, 108, 116, 117, 124, 126, 128, 129] Of the 25 studies 11 were not further considered as they compared assays other than the intervention assays or as comparators for the linked evidence approach (Appendix 8).[80, 85, 95, 98, 99, 102, 108, 117, 120, 124, 126] Of the remaining 14 studies (15 references) which undertook relevant comparisons (including one unpublished abstract provided by Proteomika

(Nagore et al., 2015)), [66, 86, 88-90, 96, 115, 116, 121, 122, 125, 128-130] (Figure 8) only 5 (6 references) (including one unpublished abstract provided by Proteomika (Nagore et al., 2015))[66, 86, 116, 121, 122] reported concordance as numerical data or as Cohen's Kappa (Figure 14). In addition Proteomika provided information in the form of a benchmark analysis which is commercial in confidence (and therefore redacted from the text).

Four comparator tests were identified from the literature linking use of the test to clinical outcomes, these were the Radioimmunoassay,[122] the Leuven in-house ELISA,[72] the Prometheus ELISA[127] and the Prometheus HMSA.[127] All of the test comparisons identified in the search are detailed in Figure 7. The index tests are shown in green, and those tests with some literature linking use of the test to clinical outcomes are marked in blue.

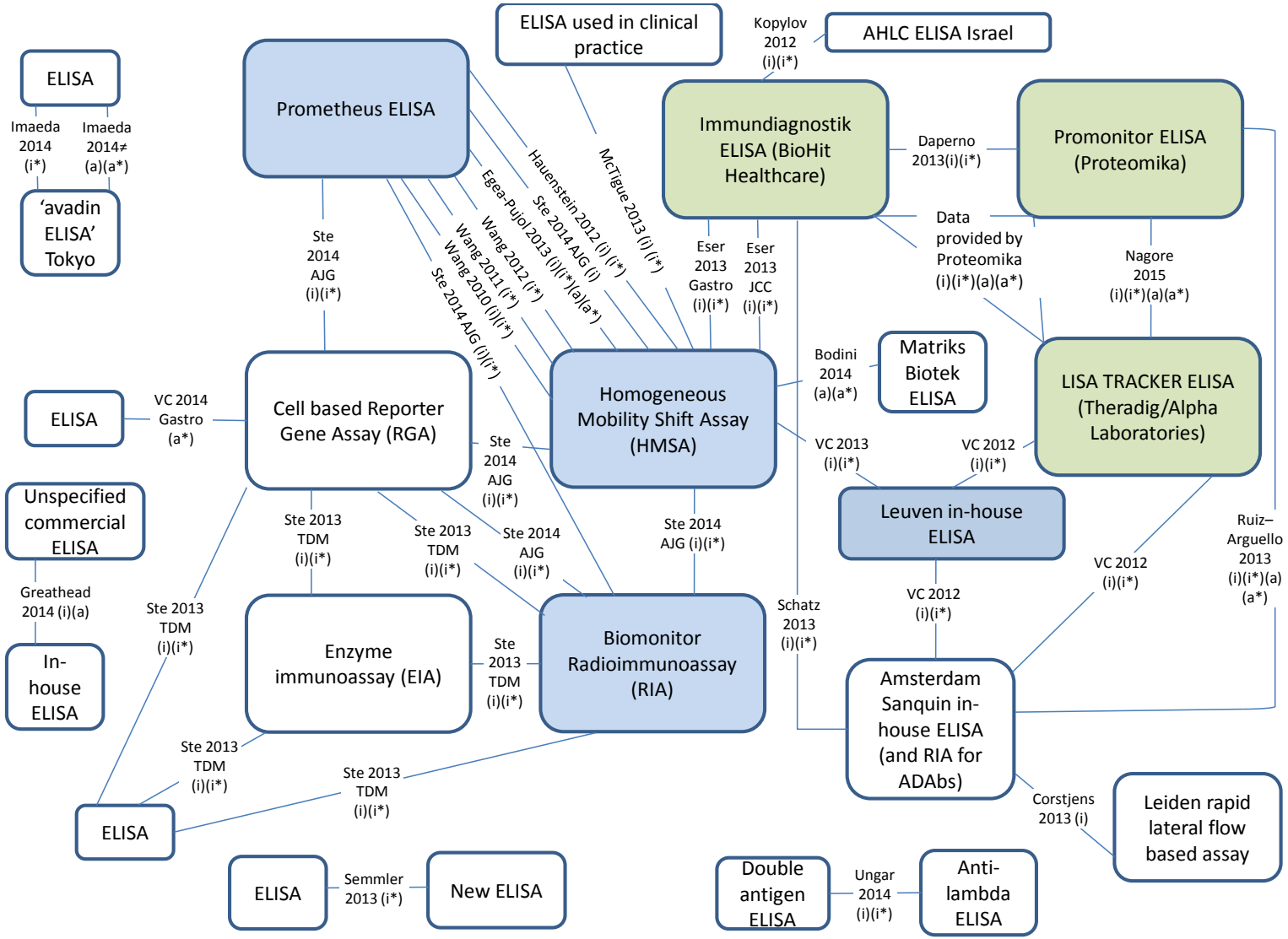


Figure 7 Summary of all of the tests for which there were comparisons in the identified literature

Key for Figure 7

The index tests are shaded green and comparator tests shaded blue. The comparisons are denoted (i) for infliximab (a) for adalimumab (i) for anti-drug antibodies to infliximab and (a*) for anti-drug antibodies to adalimumab. Bodini 2014,[80] Corstjens 2013,[85] Daperno 2013,[86] Egea-Pujol 2013,[88] Eser 2013 Gastro,[89] Eser 2013 JCC,[90] Greathead 2014,[95] Hauenstein 2012,[96] Imaeda 2014≠,[99] Imaeda 2014,[98] Kopylov 2012,[102] McTigue 2013,[108] Nagore 2015, unpublished abstract provided by Proteomika, Ruiz–Arguello 2013,[115] Schatz 2013,[116] Semmler 2013,[117] Ste 2014 AJG,[121] and [122], Ste 2013 TDM,[120] Ungar 2014,[124] VC 2012,[66] VC 2013,[125] VC 2014 Gastro,[126] Wang 2010,[128] Wang 2011,[129] Wang 2012[130]*

Only those studies which compared performance between the different index tests, or between the index and comparator tests were considered further, as shown in Figure 8. Four comparators were identified, with evidence linking use of the test to clinical outcomes, as described in section 3.2.4. Briefly, use of the radioimmunoassay was linked to outcomes in a test-treat trial[122] and use of the Prometheus ELISA and HMSA were linked to outcomes in a retrospective observational cohort.[127] The latter design is not randomised and is also subject to biases to the extent that we considered a linked evidence approach may be inappropriate. Nonetheless we included these two tests in this section for comparative purposes. An in-house ELISA from Leuven was linked to outcomes in an RCT.[72]

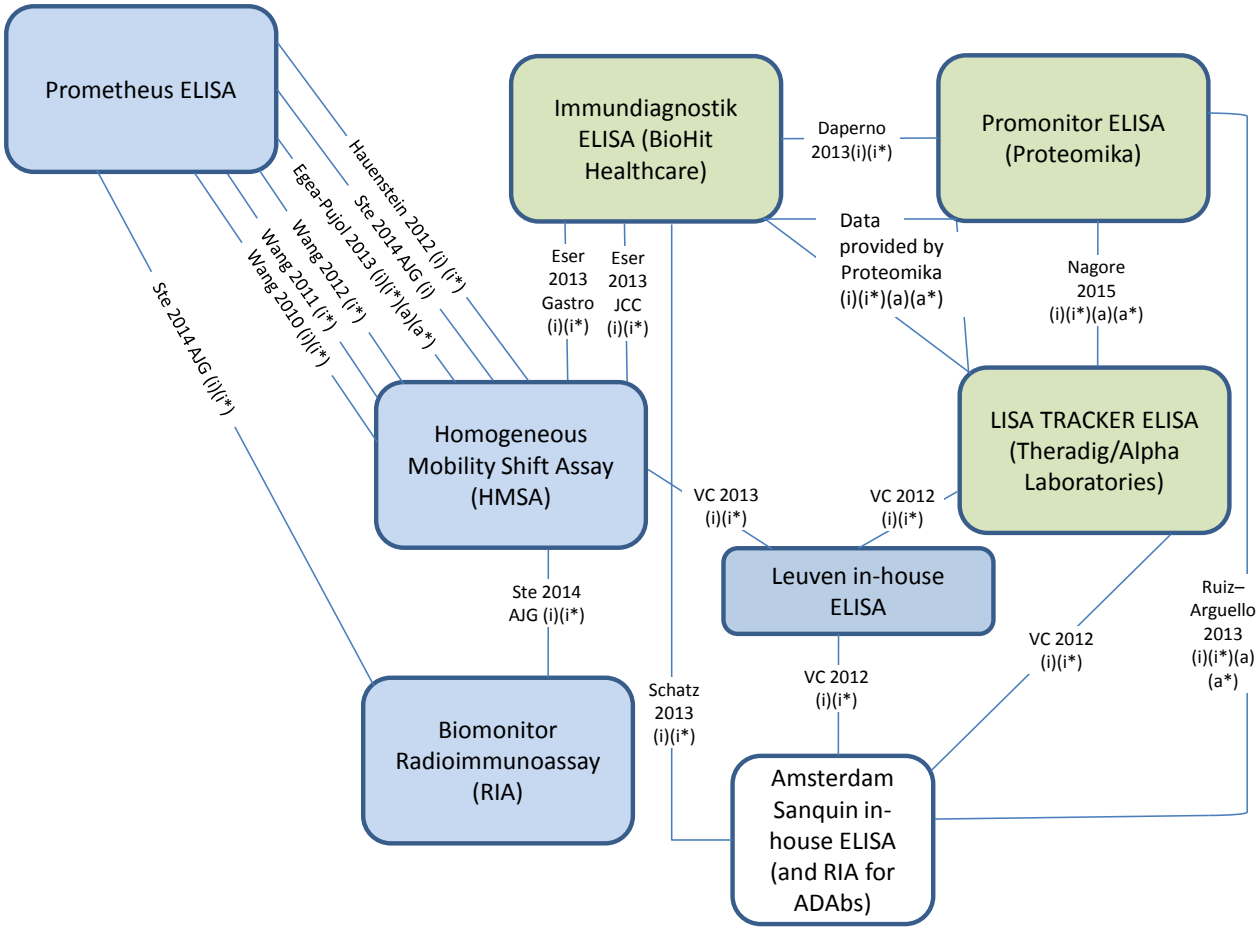


Figure 8 Comparisons that linked the index tests and comparator tests to each other
The index tests are shaded green and comparator tests shaded blue. The comparisons are denoted (i) for infliximab (a) for adalimumab (i) for anti-drug antibodies to infliximab and (a*) for anti-drug antibodies to adalimumab; All studies that compare the tests of interest are included here Daperno 2013,[86] Egea-Pujol 2013,[88] Eser 2013 Gastro,[89] Eser 2013 JCC,[90] Hauenstein 2012,[96] Nagore 2015, unpublished abstract provided by Proteomika, Ruiz–Arguello 2013,[115] Schatz 2013,[116] Ste 2014 AJG,[121] and [122] VC 2012,[66] VC 2013,[125] Wang 2010,[128] Wang 2011,[129] Wang 2012[130]*

3.2.2.3.1 Quality Appraisal

Only studies available as full texts[66, 121, 122, 125, 130] included in Figure 8 were quality assessed. The results of the quality appraisal using a tailored QUADAS-2 tool are summarised in Table 6 and Appendix 10. There was high concern regarding patient selection across all papers, with lack of clarity about the source of patients, and whether a consecutive series of patients was used. Some patients were selected from a biobank on the basis of index test results,[125] introducing a form of selection bias, and some were patients from a test-treat trial although the trial included three more patients than the comparative accuracy study and the reason for their exclusion is unclear.[121] There was also concern regarding the applicability of the patients included in these studies to our research question, in particular one study included patients who had ulcerative colitis as well as CD,[125] and

another included patients in unspecified numbers with unspecified conditions from departments of rheumatology and gastroenterology.[66]

There was low concern overall about the implementation of the index tests, but high concern about applicability where the index test measured was not one of the three ELISAs specified in our research question (i.e. they were part of a longer chain linking index tests to comparators indirectly). Risk of bias in the reference standard was high for all studies. This is because all reference standards used were imperfect, and it is difficult to determine their actual sensitivity and specificity for detecting therapeutic drug levels and levels of anti-drug antibodies. One study also interpreted the reference standard with knowledge of the index test introducing information bias.[130]

All studies had high risk of bias for flow and timing because either one test was performed at the time and the others on biobanked samples[121, 125] or the reference test were only conducted dependent on results of the index test introducing incorporation bias[130] and one study did not include all patients in the analysis.[66]

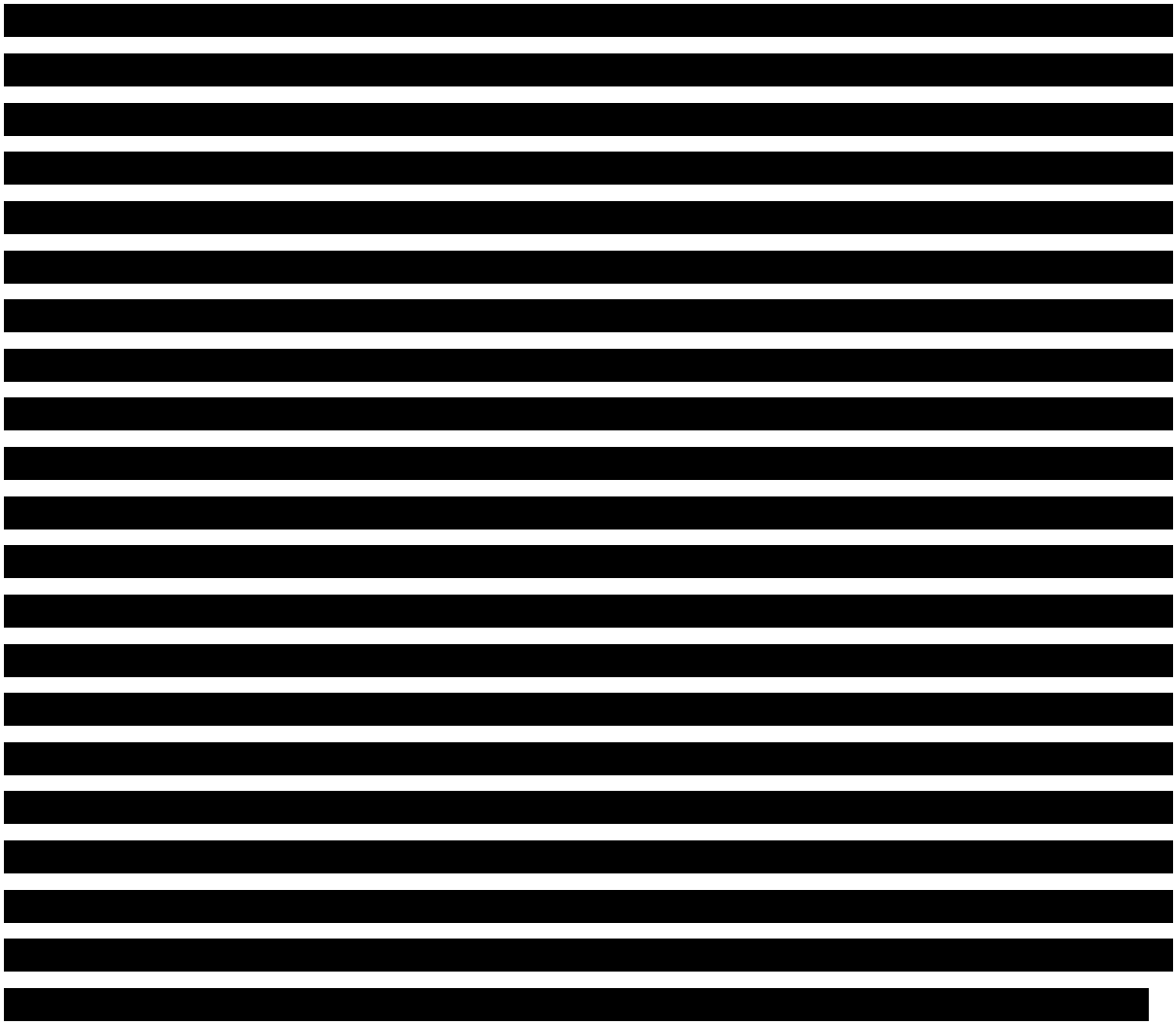
Table 6 Results of QUADAS-2[68] quality appraisal of included papers for Objective A

Study	Concerns Regarding Risk of Bias				Concerns regarding applicability		
	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
Vande Casteele 2012[66]	High	Low	High	High	High	Low	Low
Vande Casteele 2013[125]	High	Low	High	High	High	Low	Low
Steenholdt 2014[121] and [122]	High	Low	High	High	Low	High	Low
Wang 2012[130]	High	High	Unclear	High	High	High	Low

3.2.2.3.2 Comparisons between the index tests

Results are presented here for all included studies, as outlined in Figure 8. This includes 4 full papers (5 references)[66, 121, 122, 125, 130] as outlined in the quality assessment in the previous section, and 10 abstracts (including one unpublished abstract provided by Proteomika (Nagore et al., 2015)).[86, 88-90, 96, 115, 116, 128, 129]

Adalimumab



There was one unpublished abstract provided by Proteomika (unpublished abstract provided by Proteomika (Nagore et al., 2015)) comparing Promonitor to LISA-TRACKER for adalimumab. In this abstract 40 samples were used from an unspecified number of patients with IBD and an unspecified number of spiked samples. The spiked samples may be the same as described in data provided to us from the manufacturer (above). For adalimumab, drug levels were different between the different assays, 6.0 (Standard Error of Mean [SEM] 0.55) for Promonitor and 4.9 (SEM 0.39) for LISA-TRACKER. Pearson R^2 was 0.83 and the authors concluded from the spiked samples that LISA-TRACKER underestimated adalimumab levels. Additionally ten percent of samples were above the upper limit of quantification for LISA-TRACKER, and not for the Promonitor ELISA.

In summary LISA-TRACKER may underestimate adalimumab drug levels, and this underestimation will be greatest at higher absolute drug levels. The impact this would have on performance at a set threshold is unclear.

Antibodies to adalimumab

The same study that reported relationships between the index tests for adalimumab also reported some information on antibodies to adalimumab (unpublished abstract provided by Proteomika (Nagore et al., 2015)). They reported a Cohen's Kappa of 0.8 between Promonitor and LISA-TRACKER for antibodies to adalimumab, but it is unclear how many samples were included in this comparison. [redacted]

In summary we have one abstract giving a Cohen's Kappa of 0.8 between LISA-TRACKER and Promonitor, in tests for antibodies to adalimumab but it is not known how many samples and of which type were included in this analysis.

Infliximab

[redacted]

[redacted]

There was one abstract comparing Promonitor to LISA-TRACKER for infliximab (unpublished abstract provided by Proteomika (Nagore et al., 2015)). In this abstract 69 samples from an unspecified number of patients with IBD and an unspecified number of spiked samples were used. Infliximab drug levels were different between the different assays, 2.2 (SEM 0.24) for Promonitor and 3.4 (SEM 0.36) for LISA-TRACKER, Pearson R^2 was 0.98 and the authors concluded from the spiked samples that LISA-TRACKER overestimated infliximab levels. Additionally 23% of samples were above the upper limit of quantification for LISA-TRACKER, and not for the Promonitor ELISA.

Daperno et al. (2013)[86] compared in one abstract Immundiagnostik to Promonitor for infliximab drug levels. In this study Daperno et al. (2013)[86] enrolled a consecutive series of 66 patients (39 CD and 27 UC) undergoing regular infliximab dosing by IV. It is unclear if additional samples were included. Bland-Altman plots of infliximab drug levels showed mean bias of -1.8µg/mL indicating that Immundiagnostik estimates were on average lower than those of Promonitor by 1.8µg/mL and upper and lower limits of agreement of -10.8 and 7.1 respectively, indicating that 95% of infliximab drug levels measured by Promonitor were between 10.8 µg/mL lower and 7.1 UNITS higher than the same sample scores using the Immundiagnostik test. (Figure 9)

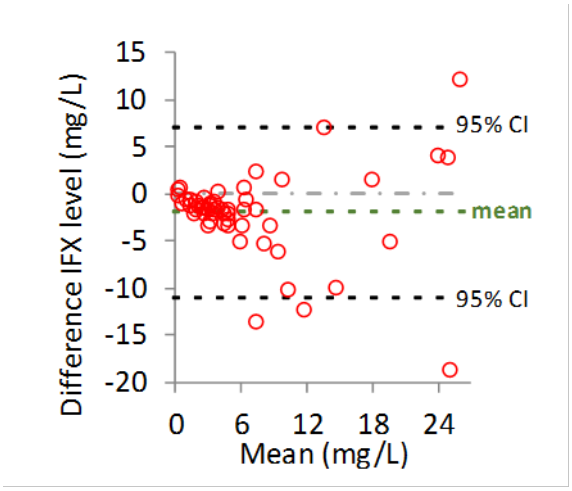


Figure 9 Reconstructed Bland-Altman plot comparing Promonitor infliximab kits (Proteomika) and Immundiagnostik TNFα-Blocker ELISA kits (Immundiagnostik). *Based on data from Daperno et al. (2013)*[86]

In summary LISA-TRACKER showed most variation in results in comparison to spiked samples, with levels of bias dependent on absolute drug levels so performance at a set threshold cannot be inferred.

[redacted]

Antibodies to Infliximab

[redacted]

One abstract described a comparison between Promonitor to LISA-TRACKER for antibodies to Infliximab (unpublished abstract provided by Proteomika (Nagore et al., 2015)). Cohen's Kappa was 1.0, indicating complete agreement between the two tests, but it is unclear how many samples were included in this comparison. For anti-drug levels 75% of samples required re-testing with LISA-TRACKER as they were above the upper limits of the measurement range.

There was one abstract by Daperno et al. (2013) comparing Immundiagnostik to Promonitor for antibodies to infliximab.[86] Daperno et al. (2013)[86] enrolled a consecutive series of 66 patients (39 CD and 27 UC) undergoing regular infliximab dosing by IV. The two tests showed identical results in just 6 out of 63 cases included in the analysis. It is unclear what is meant in the abstract by 'identical' results, but the low proportion which were identical may provide some indication that the two test for antibodies to infliximab should not be considered equivalent.

In summary one study found perfect agreement between Promonitor and LISA-TRACKER for antibodies to infliximab, another study found few (6/63, less than 10%) 'identical' results between Immundiagnostik and Promonitor.

3.2.2.3.3 Comparisons between index tests and comparator tests

Here we outline the studies linking the index tests to comparator tests which have associated evidence linking use of the test to changes in clinical outcomes

LISA-TRACKER

There are no studies linking LISA-TRACKER to any of the comparator tests for detecting adalimumab or antibodies to adalimumab.

There is one study by Vande Casteele et al. (2012) linking LISA-TRACKER to the Leuven in-house ELISA for infliximab and antibodies to infliximab.[66] This same study also compares LISA-TRACKER to the Amsterdam Sanquin in-house ELISA and RIA. These tests from the Amsterdam group are not included as comparators, but form part of the linkage pathway between the other index tests and the comparator Leuven in-house ELISA and so their relationship to LISA-TRACKER is also included here for interest.

Vande Casteele et al. (2012)[66] used 62 plasma samples from departments of gastroenterology and rheumatology. Of these 36 were clinical samples from patients, and 26 were spiked samples described as including 10 spiked with infliximab, 10 spiked with antibodies to infliximab, one spiked with adalimumab and three spiked with antibodies to adalimumab and two healthy controls. The results for these different types of sample are not fully reported separately, but parts are included. Four samples were removed as they were above the upper limit of quantification for the LISA-TRACKER assay. In detecting infliximab LISA-TRACKER gave positive results for 11 samples which were negative using either Amsterdam or Sanquin in-house ELISAs. Five of these were false positive spiked samples which did not contain infliximab, but did contain antibodies to infliximab (2 samples) and antibodies to adalimumab (3 samples). The remaining 6 samples were clinical so the true result is not known, but the authors report high levels of antibodies in these samples. The one sample spiked with adalimumab was a true negative for infliximab for both LISA-TRACKER and Leuven assays, but a false positive for the Amsterdam in-house ELISA. The Bland-Altman plots show that whilst the relationship between the Leuven and Amsterdam ELISA appears to be independent of absolute drug levels, as drug levels increase measurements using the LISA-TRACKER assay appear to increase more slowly than those using either Leuven or Amsterdam. (Figure 10) This means that the levels of concordance between the LISA-TRACKER and the other assays will be dependent upon the particular threshold used. The Bland-Altman plot showed no pattern between the Leuven and Amsterdam infliximab tests, and from visual inspection the bias was near zero, with upper and lower limits of agreement between -10mg/mL and 10mg/mL. The performance in detecting antibodies to infliximab is less clear, with discordant results reported but not the type of sample in which these results are found. The Amsterdam in-house RIA detected antibodies to infliximab in 5 samples where they were

not detected by either of the other two assays, and both the Amsterdam and LISA-TRACKER tests detected antibodies in 3 samples which tested negative using the Leuven in-house ELISA. The thresholds used for drug and antidrug levels respectively were 0.1mg/L and 10µg/L for LISA-TRACKER, 0.3mg/L and 1mg/L for Leuven, and 0.002mg/L and 12 AU/mL (1 AU/mL equals approximately 10µg/L) for Amsterdam. This higher threshold for the Leuven antibody ELISA may explain the fewer cases detected.

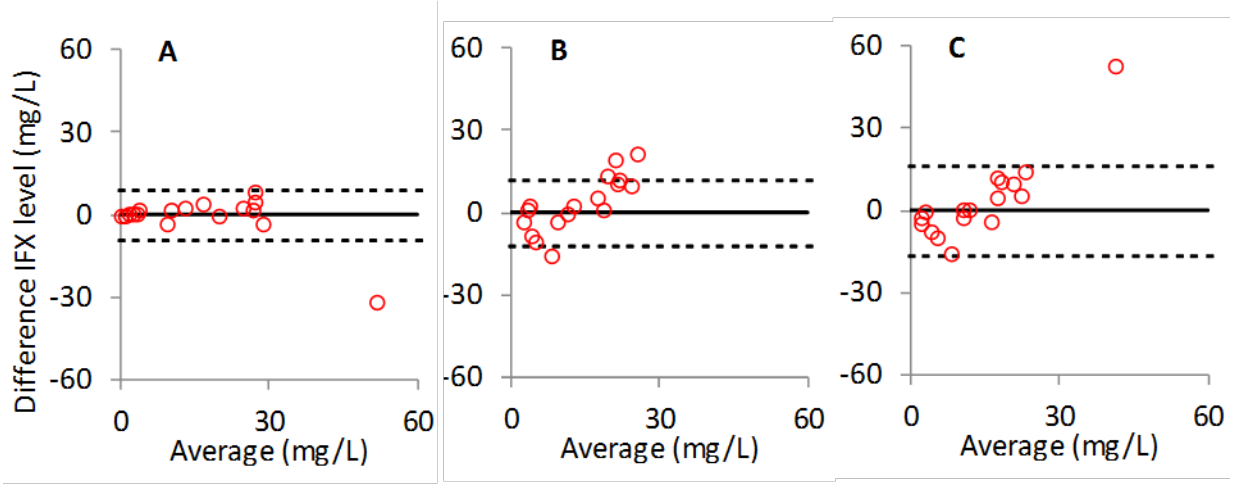


Figure 10 Reconstructed Bland-Altman plots of infliximab levels (mg/L) comparing A Amsterdam in-house infliximab ELISA and Leuven in-house infliximab ELISA, B Amsterdam in-house infliximab ELISA and LISA TRACKER Premium Infliximab kit, C Leuven in-house infliximab ELISA and LISA TRACKER Premium infliximab kit. *Based on data from Vande Casteele et al. (2012)*[66]

In summary in one study using a range of clinical and spiked samples there is some evidence that LISA-TRACKER may give false positive results for infliximab in the presence of antibodies to infliximab or adalimumab, where the Leuven and Amsterdam in-house ELISAs do not. There is also some evidence that the Amsterdam RIA is most likely to detect antibodies, followed by LISA-TRACKER then the Leuven in-house ELISA assay, but whether these antibodies detected are true positives or false positives is unclear.

Promonitor

One letter by Ruiz–Arguello (2013)[115] compared the Promonitor ELISA to the Amsterdam Sanquin ELISA for infliximab and adalimumab, and the Amsterdam Sanquin RIA for both anti-drug antibodies. In addition, Vande Casteele et al. (2012)[66] then linked the Amsterdam Sanquin ELISA and RIA to the Leuven in-house ELISA, which is one of the comparator tests.

The study comparing the Promonitor assay to the Amsterdam Sanquin tests used 120 spiked samples in total, designed to cover concentrations in the clinically meaningful range, 30 samples over the range 0.001–8, µg/mL for infliximab, 30 samples over the range 0.001–5 µg/mL for adalimumab, and

30 samples over the range 1–5000 AU/mL for both anti-drug antibodies to infliximab and adalimumab.[115] The study defines set cut points, and describes both assays as having no false positive results for drug levels. However there are no details given of the sensitivity of each assay at those set cut points, or concordance between them. The analytical sensitivity, meaning the lowest level at which the drug/antibodies are detectable was given. Analytical sensitivity of the Amsterdam Sanquin assay was slightly higher than that of Promonitor, 10ng/mL and 30ng/mL for Infliximab, respectively, 2ng/mL and 20ng/mL for adalimumab, respectively. Bland-Altman plots of each assay in comparison to the known spiked concentrations gave a mean bias of -0.467 ± 1.027 and 0.066 ± 0.196 for the Amsterdam Sanquin and Promonitor tests for infliximab respectively, and -1.140 ± 2.713 and -0.159 ± 0.488 for the Amsterdam Sanquin and Promonitor tests for adalimumab respectively. The plots are not provided, but the authors describe a systematic overestimation of the drug levels by the Amsterdam Sanquin ELISA which increases with increased drug levels, which would explain the greater confidence intervals for the mean bias estimate for the Sanquin ELISA. The authors describe this overestimation as occurring at drug levels of greater than 2µg/mL. For antibodies to infliximab and adalimumab only correlation coefficients and analytical sensitivity were reported. Analytical sensitivity of the Promonitor assay was higher than that of Amsterdam Sanquin RIA, 4ng/mL and 20AU/mL for Infliximab respectively, 2ng/mL and 30AU/mL for adalimumab respectively. Therefore, there is some evidence using spiked assays that the Amsterdam Sanquin assay may overestimate drug levels at higher concentrations, where the Promonitor assay may not. In a letter of response Rispens and van der Kleij (2013)[135] however describe an update to their testing procedure which may have corrected the overestimation.

The second link between the Amsterdam and Leuven tests has been described in detail in the previous section examining the LISA-TRACKER linkages. However to recap in brief, for drug levels the Bland-Altman plot showed no pattern so the relationship between the two tests is not dependent on threshold, and from visual inspection the bias was near zero, with upper and lower limits of agreements between -10mg/mL and 10 mg/mL. For anti-drug antibodies to infliximab the Amsterdam RIA detected more cases than the Leuven assay, with the actual veracity of these unclear.

In summary whilst we have some information linking Promonitor to the Amsterdam Sanquin tests, and further information linking these to the Leuven ELISA it is not in a format from which we can calculate the concordance between the tests at clinically relevant thresholds.

Immundiagnostik

There are no studies linking Immundiagnostik to any of the comparator tests for adalimumab or antibodies to adalimumab. Eser et al. (2013) compared in two abstracts[89, 90] the Immundiagnostik ELISA to the Prometheus HMSA, and one abstract by Schatz et al. (2013)[116] compared the

Immundiagnostik ELISA to the Amsterdam Sanquin in-house tests (for infliximab and antibodies to infliximab), which are in turn compared to the Leuven in-house ELISAs by Vande Casteele et al. (2012).[66]

The two abstracts by Eser et al. (2013)[89, 90] comparing Immundiagnostik ELISA to the Prometheus HMSA method used samples from 90 patients (66 CD and 24 UC). The authors report that HMSA was able to detect anti-drug antibodies to infliximab at the mid-infusion point, while ELISA returned inconclusive results due to interference from infliximab. However no numerical data were presented comparing the two methods so few if any conclusions can be drawn from the study.

The study by Schatz et al. (2013) linking Immundiagnostik ELISA to the Amsterdam in house tests (ELISA for drug levels, RIA for antibody levels)[116] compared performance of the two tests for infliximab and anti-drug antibodies to infliximab. They used serum samples from 202 paediatric patients of which 125 had been exposed to infliximab and 77 were infliximab naïve. Samples were considered positive for infliximab if they were above the limit of detectability, which was <0.8 µg/ml for the Immundiagnostik ELISA and <0.002µg/ml for Amsterdam in-house ELISA. Overall agreement using Cohen's Kappa was 0.792. Considering only the infliximab exposed patients, 25 were below the lower limit of detectability for both tests, leaving 87 who tested positive for both, 11 who only tested positive using the Amsterdam ELISA, (measurements ranged from 0.1-2.3 µg/ml, so some of these will be below the lower limit of detectability using the Immundiagnostik test) and 2 whose results are not reported. For anti-drug antibodies to infliximab 88 samples were concordant positive, 27 samples were concordant negative, and 10 were detected only by the Amsterdam RIA and not the Immundiagnostik ELISA.

The second link between the Amsterdam and Leuven tests has been described in detail in the previous two sections. For drug levels the Bland-Altman plot showed no pattern so the relationship between the two tests is not dependent on threshold, and from visual inspection the bias was near zero, with upper and lower limits of agreements between -10mg/mL and 10 mg/mL and for anti-drug antibodies to infliximab the Amsterdam RIA detected more cases than the Leuven assay, with actual veracity of these unclear.

In summary whilst there are good data linking Immundiagnostik to the Amsterdam in-house ELISA, with agreement for 114 out of 125 samples for infliximab and 115/125 samples for anti-drug antibodies to infliximab, the link to the Leuven ELISA is not known in terms of agreement of the two tests at a clinically relevant threshold.

3.2.2.3.4 Relationship between different comparator tests

There was one study by Steenholdt et al. (2014) comparing the performance of the Biomonitor RIA used in the test-treat trial to Prometheus HMSA and Prometheus ELISA for infliximab and antibodies to infliximab.[121] Vande Casteele et al. (2014) compared the Leuven in-house ELISA to Prometheus HMSA.[125] One full study by Wang et al. (2012)[130] and four further abstracts[88, 96, 128, 129] compared the performance of HMSA to Prometheus ELISA for infliximab and antibodies to infliximab, Egea-Pujol et al. (2013) also made the same comparison for adalimumab and antibodies to adalimumab.[88] However, three of these abstracts (Hauenstein et al., 2012,[96] Egea-Pujol et al., 2013,[88] Wang et al., 2011[129]) did not provide data on concordance, Cohen's Kappa or numbers of false positives, true positives, false negatives or true negatives test results and will not be described further here.

The studies by Wang et al. (2010)[128] and (2012)[130] compared the performance of HMSA to Prometheus ELISA. Wang et al. (2010)[128] described 20 patients with IBD who had relapsed from treatment with infliximab. ELISA detected infliximab in 15/20 and anti-drug antibodies to infliximab in 15/20. HMSA detected infliximab in 15/20 and anti-drug antibodies to infliximab in 18/20. It is not clear whether it was the samples from the same 15 patients which tested positive for both tests. The focus of Wang et al. (2012)[130] was to validate the performance of HMSA, rather than comparing it to ELISA. Out of 100 healthy controls 3 were false positive for antibodies to infliximab for HMSA. This was to be expected as the cut point was determined from the same samples as mean +2SD. Repeat measurements of these three resulted in them being below the cut-point, presumably regression to the mean. ELISA results for the 100 healthy controls were not reported. Out of 100 inflammatory bowel patients selected as positive for antibodies for infliximab on ELISA, 5 did not test positive on HMSA. The authors attribute this to elevated levels of non-specific binding in the ELISA. As we do not have the equivalent data for ELISA results on samples which tested positive using HMSA it is difficult to draw any conclusions at all. The only comparative data given constitutes a plot of correlation which does not appear to show high correlation. The studies therefore did not provide useful concordance data for evaluation.

Vande Casteele et al. (2013)[125] compared the Leuven in-house ELISA to Prometheus HMSA. Whilst the paper does describe some discordant results, the focus of the paper is on outcomes in patients with differing results rather than on comparisons between the two tests. HMSA appears to perform better at detecting antibodies to infliximab in the presence of infliximab, quantifying this is difficult due to reporting focussed on other research questions. It is described in the discussion as HMSA detecting median 9 weeks earlier. However in the absence of infliximab and with the HMSA cut-off for antibodies to infliximab set at 7.95U/ml, the Leuven in-house ELISA detected four more cases with, antibodies to infliximab. The authors report that Prometheus have since lowered the threshold to 3.13 U/mL.

Steenholdt et al. (2014)[121] took 66 frozen patient samples (CD patients with loss of response to infliximab) from the test-treat trial[122] and re-analysed them using Prometheus HMSA and Prometheus ELISA. All of these patients had RIA results using the samples before freezing, which determined treatment pathway at the time of the study. Threshold for positivity for infliximab was unclear, lower limit of quantification was clearly defined as ≥0.15μg/mL, ≥1μg/mL and ≥1.4μg/mL for RIA, HMSA and ELISA respectively.[121] However in the original RCT paper[122] whilst the same lower limit of quantification thresholds were given, additional cut points of ≥0.5μg/mL and ≥3μg/mL were given for defining therapeutic drug levels for RIA and HMSA respectively. However no additional cut-point was given for ELISA. Thresholds for anti-drug antibodies to infliximab were ≥10AU/mL, ≥3.13AU/mL and ≥1.69μg/mL for RIA, HMSA and ELISA respectively. Using RIA 54/66 (82%) tested positive for infliximab, in comparison to 58/66 (88%) for HMSA and 50/66 (76%) using ELISA. The concordance between the three tests is shown in Figure 11. In eight patients infliximab was undetectable using all three tests, in 50 patients infliximab was detectable using all three tests. Four patients had infliximab detected by RIA and HMSA but not ELISA, and a further four patients had infliximab detectable by only HMSA.

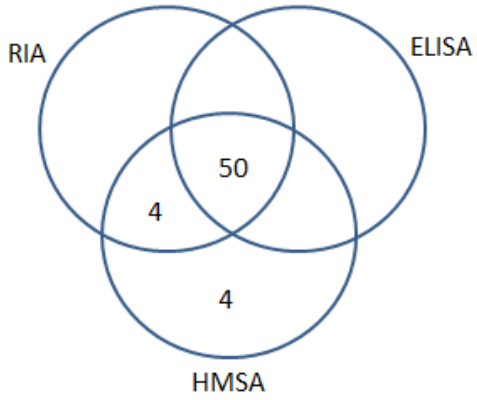


Figure 11 Concordance between RIA, HMSA and ELISA for detecting infliximab in 66 CD patients with loss of response to infliximab. *In eight patients infliximab was undetectable using all three tests*

Using RIA 18/66 (27%) tested positive for anti-drug antibodies to infliximab, in comparison to 22/66 (33%) for HMSA and 6/66 (9%) using ELISA. The concordance between the three tests is shown in Figure 12. In 43 patients anti-drug antibodies to infliximab were undetectable using all three tests, in 6 patients anti-drug antibodies to infliximab were detectable using all three tests. Eleven patients had anti-drug antibodies to infliximab detected by RIA and HMSA but not ELISA, and a further five patients had anti-drug antibodies to infliximab detectable by only HMSA, and one patient had anti-drug antibodies to infliximab detectable only by RIA.

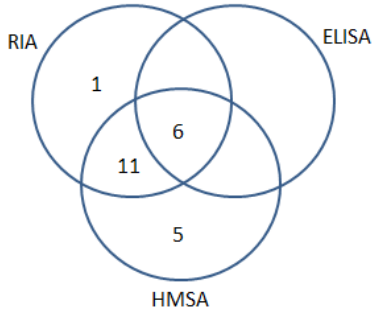


Figure 12 Concordance between RIA, HMSA and ELISA for detecting anti-drug antibodies to infliximab in 66 CD patients with loss of response to infliximab. *In 43 patients anti-drug antibodies were not detectable in any of the tests.*

In summary for infliximab: RIA and HMSA agree for 62/66 patients, with the remaining 4 testing positive on HMSA and not RIA, HMSA and ELISA agree on 58/66 patients, with 8 testing positive on HMSA and not ELISA, and RIA and ELISA agree on 62/66 patients with the remaining 4 testing positive on RIA and not ELISA. The Bland-Altman plots comparing HMSA to RIA and ELISA to RIA showed a pattern with increasing drug concentration, meaning that these two comparisons are dependent on the absolute values for thresholds chosen. (Figure 13) The relationship between HMSA and ELISA appears independent of absolute drug concentrations.

For anti-drug antibodies to infliximab: RIA and HMSA agree for 60/66 patients, with 5 testing positive on HMSA and not RIA, and 1 testing positive on RIA and not HMSA. HMSA and ELISA agree on 50/66 patients, with 16 testing positive on HMSA and not ELISA, and RIA and ELISA agree on 54/66 patients with the remaining 12 testing positive on RIA and not ELISA.

Therefore there is an indication that RIA and HMSA detect a greater number of patients with infliximab in comparison to the Prometheus ELISA, and this effect is more pronounced with anti-drug antibodies to infliximab. We do not know the true measurements for these patients.

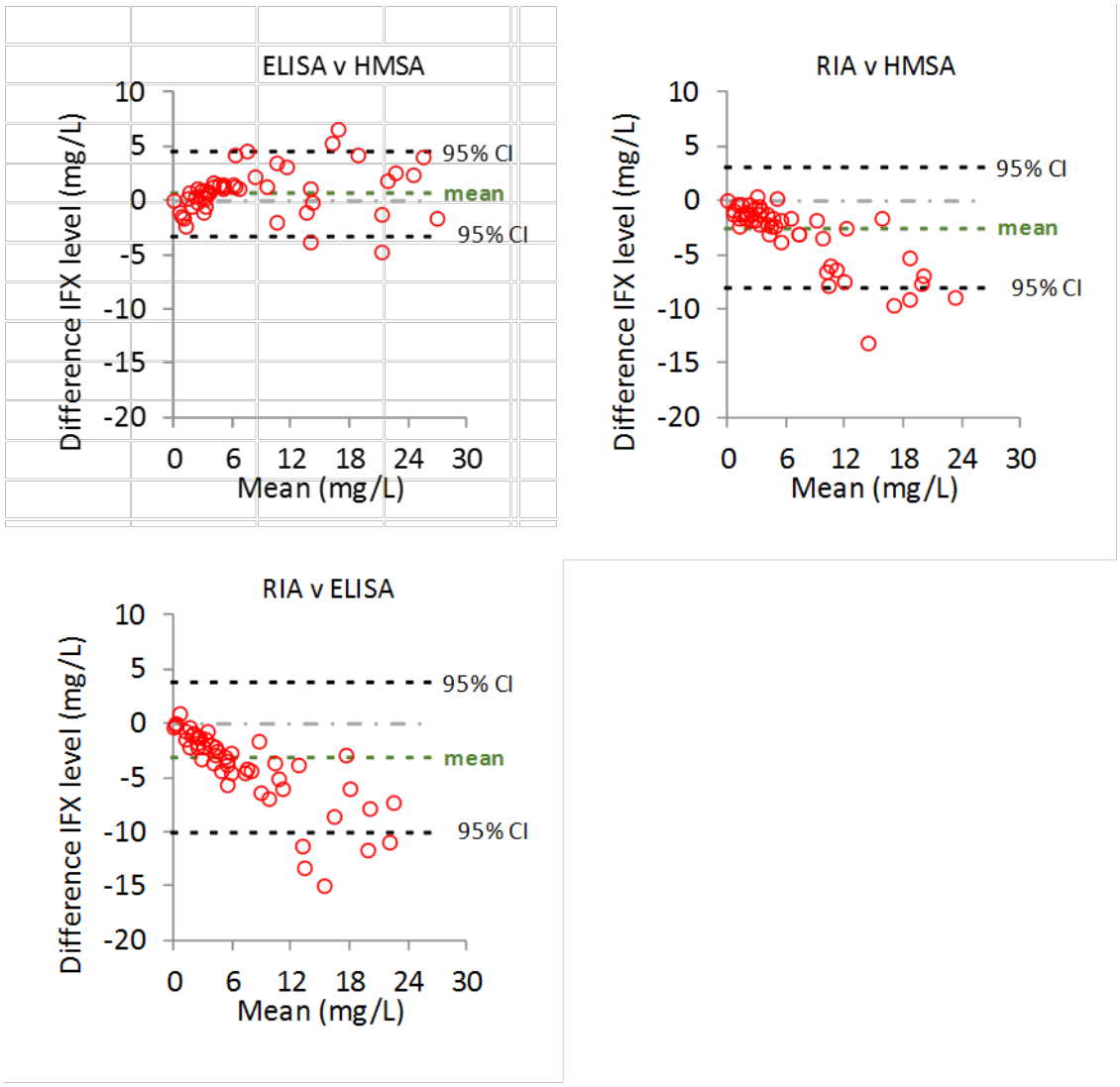


Figure 13 Reconstructed Bland-Altman plots comparing Prometheus ELISA and HMSA and the radioimmunoassay. *Based on data from Steenholdt et al. (2014)*[121]

3.2.2.4 Summary

Figure 14 summarises the studies which quantify the link between the index tests and the comparator assays using concordance data. In comparing the three index tests to one another we have data from two (including one unpublished abstract provided by Proteomika (Nagore et al., 2015)) abstracts.[86]

For anti-drug antibodies one abstract (unpublished abstract provided by Proteomika (Nagore et al., 2015)) describes complete agreement between Promonitor and LISA-TRACKER for anti-drug antibodies to infliximab, and a Cohen's Kappa of 0.8 for anti-drug antibodies to adalimumab, but does not describe how many samples were included. They also describe the upper limits of the measurement range for LISA-TRACKER as low. Daperno et al. (2013)[86] compared Immundiagnostik to Promonitor for anti-drug antibodies to infliximab and found that the two tests showed identical results in just 6 out of 63 cases, in a consecutive series of 66 patients (39 CD and 27 UC) but the

definition of 'identical' was not given. It is not possible from these data to link the three index tests as part of a linked evidence approach.

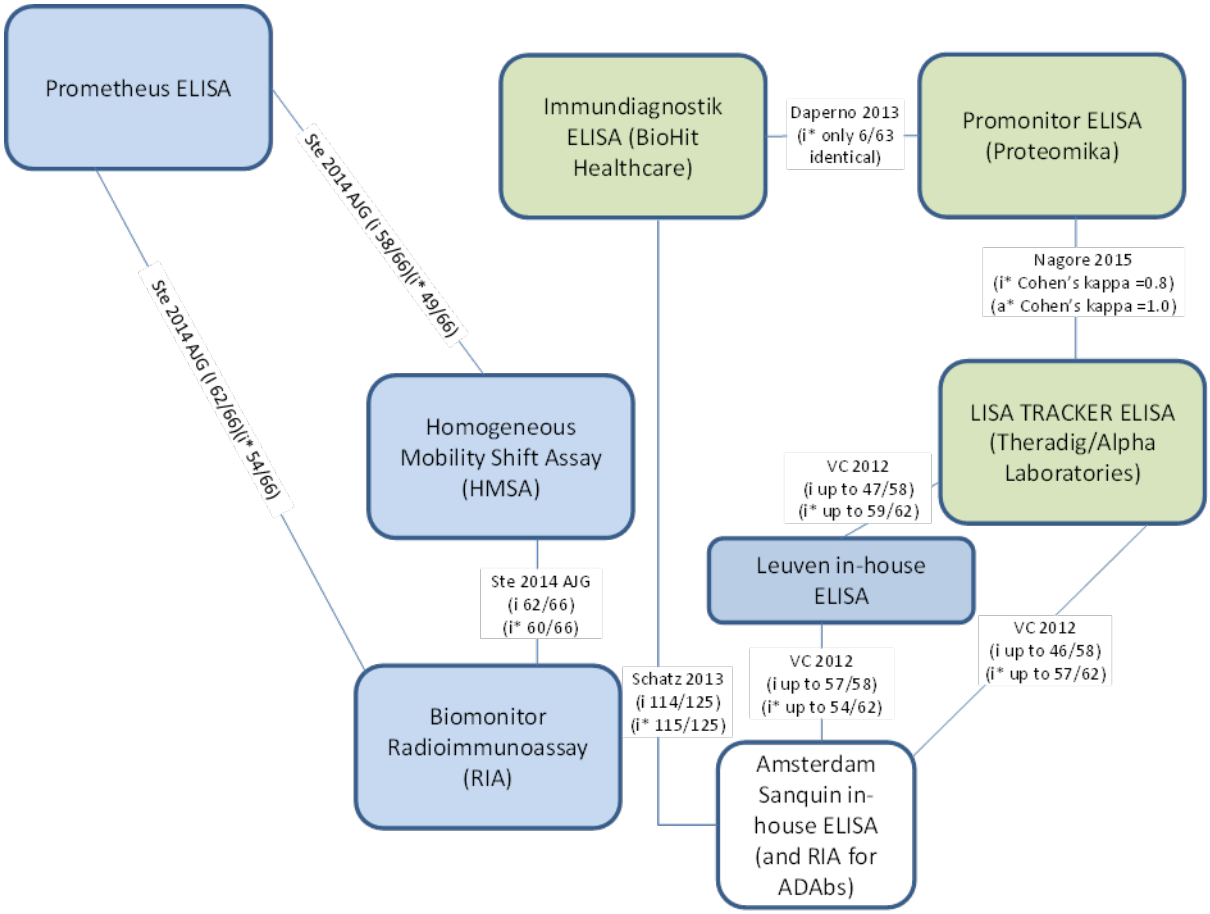


Figure 14 Results of comparisons which linked the index tests and comparator tests to each other of studies reporting concordance data
The index tests are shaded green and comparator tests shaded blue. Results are listed as either Cohen's Kappa or concordance levels displayed as a fraction for (i) for infliximab (a) for adalimumab (i) for anti-drug antibodies to infliximab and (a*) for anti-drug antibodies to adalimumab. Only studies which provide concordance at set threshold or Cohen's Kappa for the comparisons between tests are included here. Daperno 2013,[86] Nagore 2015, unpublished abstract provided by Proteomika, Schatz 2013,[116] Ste 2014 AJG,[121] and [122] VC 2012[66]*

For each index test we investigated all links to the comparator tests. There was one link to the Amsterdam Sanquin tests (ELISA for infliximab, RIA for anti-drug antibodies to infliximab)[116] which showed agreement for 114/125 samples for infliximab and 115/125 samples for anti-drug antibodies to infliximab. However there was only one study then linking the Amsterdam tests to the Leuven in-house ELISA,[66] which reported disagreement on at least 8/62 samples, with lack of clarity regarding the remainder of results. Similarly there was only one study linking LISA-TRACKER to the Leuven in-house ELISA,[66] and whilst it reported disagreement for infliximab of at least 11/58 samples, and for anti-drug antibodies to infliximab for at least 3/62 samples, the results for the remainder were

unclear. However we found no concordance data linking any of the index tests to any of the comparator tests at a clinically meaningful threshold.

In comparing the comparator tests to each other there was one study[121] which described re-analysing the same samples previously used in a test and treat trial.[122] There was agreement between Prometheus ELISA and Biomonitor RIA for infliximab for 62/66 samples and for anti-drug antibodies to infliximab for 54/66 samples, agreement between Prometheus ELISA and HMSA for 58/66 samples for infliximab and 49/66 samples for anti-drug antibodies to infliximab, and finally agreement between RIA and HMSA for 62/66 samples for infliximab and 60/66 samples for anti-drug antibodies to infliximab. We found no ongoing link to the index tests.

Overall there was insufficient evidence linking any of the index tests (LISA-TRACKER, Immundiagnostik or Promonitor) to any of the comparators with links to clinical outcomes (HMSA, RIA, Prometheus ELISA, or Leuven in-house ELISA).

3.2.2.5 Results of threshold analysis studies

The search identified 24 studies[47, 57, 76, 81, 83, 84, 92-94, 99, 100, 103, 104, 106, 107, 109, 111, 112, 114, 118, 119, 125, 131, 134] which reported a ROC threshold analysis to determine optimal cut-off levels predictive of clinical response for infliximab,[57, 76, 81, 84, 92, 100, 103, 104, 109, 111, 112, 118, 119, 125] adalimumab[47, 83, 93, 99, 106, 107, 114, 131, 134] or both.[94] Table 7 summarises the studies in terms of the threshold reported, the diagnostic performance of using the threshold, the clinical marker used for assessment of response, the assay used and the value for the area under the curve (AUC). As the area under a ROC curve in this case quantifies the overall ability of the test to discriminate between those individuals who respond and those who do not respond, an AUC value greater than 0.5 indicates an informative test while an AUC of 0.5 represents an uninformative test without discriminatory power (sensitivity + specificity =1). When considering the reported thresholds we need to bear in mind that response and loss of response are poorly defined and studies using different definitions will measure different outcomes, which will have implications on the reported thresholds.

3.2.2.5.1 Infliximab

Studies measuring infliximab used a range of assays including commercial ELISA kits, in-house or academically developed ELISAs, HMSA and RIA. Studies generally tried to optimise diagnostic performance by finding a cut-off with maximum sensitivity and specificity and maximising the area under the curve (AUC). However, according to the reported performance measures the diagnostic performance of the tests overall was only moderate. One study aimed for high sensitivity (0.90) at a trade-off of specificity (0.37),[94] while another favoured high specificity (1.00) at the expense of good sensitivity (0.33).[118] The reported infliximab cut-off ranged between 0.6[100] and 4.1µg/ml[104] and was

reported to be as high as 7µg/ml for the test with 100% specificity requiring a minimum drug trough level of 7µg/ml at treatment week 14 to be predictive of good response at week 54.[118] One study reported a trough difference before and after dose optimisation as predictive of clinical outcome[57] and another reported the trough level that predicts response after re-initiation of infliximab treatment.[76] While the trough levels for the HMSA and RIA were not too dissimilar, one reported trough levels for predicting anti-drug antibodies to infliximab using HMSA which were exceptionally high (13µg/ml).[125] There was great variation in the clinical marker used to assess clinical response which ranged from mainly subjective physician's assessment and disease activity scores to laboratory markers such as C-reactive protein (CRP) and faecal calprotectin (FC) and objective assessments of mucosal healing. Six studies used a combination of different markers for the assessment.[76, 81, 94, 100, 103, 104]

3.2.2.5.2 Adalimumab

Studies of adalimumab used mainly ELISAs with only one study each using HMSA and RIA.[93, 134] One study in the form of an abstract did not report the test type used.[106] The reported thresholds for clinical markers such response and clinical remission ranged from 3µg/ml[94] to 6.85µg/ml.[93] However, sustained clinical benefit as reported by patients and defined as 'lasting control of disease with possible dose escalation' was predicted with a high sensitivity of 95% by adalimumab levels in one study of only 0.33µg/ml or above.[47] One study reported the different threshold values for a test with maximum sensitivity 14.5µg/ml, maximum specificity 0.35µg/ml and sensitivity=specificity 6.85µg/ml.[93]

All but one study reported AUC values considerably higher than 0.5, classing them as fair to good tests. However, one study reported AUC values for three different time points of just over 0.5 (0.5, 0.57 and 0.58) and was unable to identify an adalimumab concentration associated with clinical remission (CDAI <150). This study therefore questioned the clinical utility of measuring adalimumab concentrations.[83]

Table 7 Cut-offs for drug levels from ROC analyses to predict clinical response

Reference	Cut-off in µg/ml	Performance measures				AUC (95% CI)	Clinical marker	Drug	Assay
		Sens	Spec	PPV	NPV				
Bortlik 2013[81]	3	0.70	0.62	0.41	0.84	0.70 (0.57-0.83)	Sustained response (no treatment failure or drug intolerance, no surgery, IS introduction, steroids or infliximab increase)	IFX	c ELISA
Cornillie 2014[84]	3.5	0.64	0.78	0.56	0.83	0.75	Sustained response (CDAI score change)	IFX	nc ELISA
Goldberg 2014[94] Abstract	3	0.90	0.37	NR	NR	0.75	Disease activity (physicians global assessment and CRP levels)	IFX	c ELISA

Reference	Cut-off in µg/ml	Performance measures				AUC (95% CI)	Clinical marker	Drug	Assay
		Sens	Spec	PPV	NPV				
Imaeda 2014[100]	0.6	0.73	0.62	NR	NR	0.67 (0.60-0.81)	CRP ≤0.3mg/dL Serum albumin (≥ 4.0mg/dL) FC (≤ 300µg/g) MH (Rutgeerts scoring system 0 or 1)	IFX	nc ELISA
	1.0	0.67	0.71	NR	NR	0.72 (0.50-0.73)			
	1.1	0.72	0.56	NR	NR	0.63 (0.55-0.65)			
	4.0	0.71	0.70	NR	NR	0.63 (0.56-0.70)			
Marits 2014[104]	4.1	0.87	0.44	NR	NR	0.74 (SE 0.037)	Remission (HBI <5 and CRP < 3 mg/l)	IFX	nc ELISA
Nagore 2015[109]	0.8	0.86	0.75	NR	NR	0.86 (0.76-0.96)	Active disease	IFX	c ELISA (Promonitor)
Pallagi-Kunstar 2014[111]	3.01	NR	NR	NR	NR	NR	Detecting anti-drug antibodies	IFX	c ELISA
Paul 2012[112] abstract	2	0.76	0.82	NR	NR	0.60	Remission (CDAI score <150)	IFX	c ELISA (LISA-TRACKER)
Paul 2013[57]	0.5 (trough after optimisation minus trough before optimisation)	0.88	0.76	0.78	0.86	0.91 (0.83-1.0)	Mucosal healing (FC <250µg/g)	IFX	c ELISA (LISA-TRACKER Premium)
Singh 2014[118]	4	0.53	0.75	0.76	0.52	0.64 (0.51-0.75)	Week 14 infliximab levels as predictor of week 54 clinical remission according to CDAI	IFX	c ELISA (up to July 2012) or HMSA (from July 2012)
	7	0.33	1.00	1.00	0.50	0.67 (0.58-0.75)			
Baert 2014[76]	2 (after re-exposure to infliximab)	NR	NR	NR	NR	0.76 (0.62-0.90)	Long term response (clinical assessment [HBI] and CRP levels[<3mg/l])	IFX	HMSA
Levesque 2014[103]	3	NR	NR	NR	NR	NR	Disease activity at week 8 (≥70 point increase in CDAI and CRP >5µg/l)	IFX	HMSA
Vande Casteele 2013[125]	13 (TL week 6)	0.72	0.81	NR	NR	0.87 (SE 0.06)	anti-drug antibody formation	IFX	HMSA
Steenholdt 2011[119]	0.5	0.86	0.85	NR	NR	0.93 (0.85-1.0)	Maintained response (good response to induction therapy at 0, 2 and 6 weeks followed by good response to maintenance therapy)	IFX	RIA
	2.2 (TL week 14)	0.79	0.94			0.93 (SE 0.04)			
Feagan 2012[92] Abstract	3	NR	NR	NR	NR	0.74	Disease activity	IFX	HPLC based fluid phase assay
Chiu 2013[83]	No	NR	NR	NR	NR	Week 4: 0.51	Clinical remission	AD	nc ELISA

Reference	Cut-off in µg/ml	Performance measures				AUC (95% CI)	Clinical marker	Drug	Assay
		Sens	Spec	PPV	NPV				
	adalimumab concentration identified associated with clinical remission at any time point so clinical utility of measuring adalimumab concentrations was difficult to assess					Week 24: 0.58 Week 56: 0.57	(CDAI <150)	A	
Goldberg 2014[94] Abstract	3	0.83	0.63	NR	NR	0.8	Disease activity (physicians global assessment and CRP levels)	ADA	c ELISA
Imaeda 2014[99]	5.9	0.67	0.92	NR	NR	0.83 (0.80-0.95)	CRP ≤0.3mg/dL	ADA	nc ELISA
Karmiris 2009[47]	0.33	0.95	NR	0.81	NR	NR	Sustained clinical benefit (patient reporting lasting control of disease with possible dose escalation)	ADA	nc ELISA
Mazor 2014[107]	5.85	0.68	0.71	NR	NR	0.75 (0.66-0.84)	Remission according to 2 physicians' assessment	ADA	nc ELISA
Roblin 2014[114]	4.85 4.9	0.81 0.66	0.67 0.85	0.84 0.88	0.57 0.51	0.73 0.77	Clinical remission (CDAI <150) MH (disappearance of all ulcerations on endoscopy)	ADA	c ELISA (LISA-TRACKER Premium)
Ward 2013[131] Abstract	4.9	0.83	0.65	NR	NR	0.75	Remission	ADA	c ELISA (LISA-TRACKER Premium)
Yarur 2013[134] Abstract	5	NR	NR	NR	NR	0.71	Elevation of CRP	ADA	HMSA
Frederiksen 2014[93]	14.5 0.35 6.85	1.00 0.50 0.69	0.12 0.96 0.69	0.41 0.89 0.58	1.00 0.76 0.78	0.77 (0.62-0.93)	LOR (physician's global assessment)	ADA	RIA
Mazor 2013[106] Abstract	5	NR	NR	NR	NR	0.77 (0.67-0.86)	Clinical response and normal CRP	ADA	NR

(Abbreviations: NR not reported; ADA adalimumab; IFX infliximab; LOR loss of response; RIA radioimmunoassay; c commercial; nc non-commercial; CRP C-reactive protein, sens sensitivity; spec specificity; PPV positive predictive value; NPV negative predictive value; HBI Harvey Bradshaw Index; CDAI Crohn's Disease Activity index; HMSA Homogenous mobility shift assay; MH mucosal healing; FC faecal calprotectin; TL trough level)

3.2.2.6 Summary

The range of cut-offs illustrates that no validated threshold has been established to date. Cut-offs strongly depend on the test assay used, the drug measured and the clinical marker investigated as well as the method of determination of the clinical marker. It is uncertain how clinically meaningful the reported thresholds are, as the reported sensitivities and specificities have been optimised to a varying degree across the studies and studies use different definitions of response and loss of response. An additional variable that impacts on the threshold of anti-TNFα drug levels (which is insufficiently depicted in the table because of poor reporting in the studies) is the time of testing and the time of clinical assessment.

3.2.3 Objective B Description of algorithms prescribing patient management following test outcomes for drug and / or anti-drug antibody levels

3.2.3.1 Aim

B1] To provide a narrative description of algorithms used in studies which report clinical outcomes for patients whose treatment options were directed by a test-informed algorithm; and, *B2*] to compare these with related algorithms identified in the literature during scoping as relevant to the NHS.[65, 72]

3.2.3.2 Results

Studies and reviews reporting on test results and clinical status of patients, have frequently proposed test-based treatment algorithms, but most of these have never been tested or implemented in CD patients.[43, 55, 56, 60, 63, 136-138] Here we describe the test-based algorithms that have actually been implemented in studies with CD patients and briefly compare these with the most similar "precursor" algorithms identified during scoping as relevant to the NHS.

3.2.3.2.1 Algorithms from management studies

No management studies were found for CD or IBD patients treated with adalimumab. Three infliximab studies fulfilled inclusion criteria for this objective: an RCT in CD patients with loss of response to infliximab (Steenholdt et al., 2014)[122], the TAXIT RCT[72] of IBD patients responding to infliximab, and a retrospective observational study of IBD patients responding to infliximab (Vaughn et al., 2014).[127]

Table 8 summarises the algorithm used by Steenholdt et al. (2014).[122] Tests were done using commercially available RIAs. Both drug and anti-drug antibody tests are dichotomised so that concurrent testing classifies each patient into one of the four possible combinations of test results:

1) infliximab – & anti-drug antibody +,

2) infliximab – & anti-drug antibody –,

3) infliximab + & anti-drug antibody –,
4) infliximab + & anti-drug antibody +.

An important feature is the proposal of different causes for secondary treatment failure in the groups 1 to 3; these proposals rest on interpretations of the supposed underlying mechanisms leading to the observed test results. The treatment options are prescriptive for two of the groups (1 and 2), but less so for group 3 (patients with loss of response who have therapeutic levels of infliximab and lack detectable anti-drug antibodies). Thus, the treatment received for group 3 requires further investigation and reflection by the treating clinician and may or may not include relatively expensive biological agents. Since most patients fall into this group these less prescriptive aspects add to uncertainty about treatment cost of the algorithm-based strategy and whether discretion relating to cost might play a part in decision making for group 3 (an expensive biological may or may not be adopted because of perceived cost implications). Furthermore, treatments for this group may be difficult to replicate between different groups of clinicians who may be subject to different health pressures in relation to costs and/or to differing licensing regulations for biological therapies. Results for group 4 (positive test for both infliximab and anti-drug antibody) are reviewed with suspicion and require retesting in case of error.

Table 8 Summary of the concurrent testing-based algorithm used by Steenholdt et al. (2014)[122]

	Detectable anti-infliximab antibodies	Undetectable anti-infliximab antibodies
Sub-therapeutic infliximab < 0.5 ug/mL	Group 1	Group 2
	Insufficient infliximab bioavailability due to induced immunogenicity of infliximab	Insufficient infliximab bioavailability due to non-immune mediated pharmacokinetics of infliximab
	↓	↓
	Change to different TNFα-inhibitor: Adalimumab 80 mg sc at inclusion followed by 40 mg sc every other week: dose intensification allowed	Intensify infliximab treatment: infliximab 5 mg/kg iv every 4 weeks
Therapeutic infliximab ≥ 0.5 ug/mL	Group 4	Group 3
	Consider: (A) pharmacodynamics (B) non-functional anti-infliximab antibodies (C) false positive test	Pharmacodynamics: inhibition of TNFα is ineffective due to non-TNFα driven disease.
	↓	↓
	Repeat infliximab and anti-infliximab antibody analyses and handle accordingly. If unchanged results, then act as group 3	TNFα-inhibitors not effective is discontinued. Review of clinical condition at discretion of the investigator: - if relapse of CD, use drug(s) with other target, e.g conventional immune-suppressives, glucocorticoids, and/or other biological agents. Consider surgery if appropriate. - if no relapse, treat underlying problem

Abbreviations: CD Crohn's disease; sc subcutaneous

The TAXIT algorithm for patients responding to infliximab is based on the hypothesis that an infliximab trough level between 3 and 7µg/mL is optimum for successful maintenance of clinical response; it proposes a strategy of prospective dose adjustment to achieve this target range; this likely requires trough tests before each infusion.

Error! Reference source not found. Figure 15 summarises the TAXIT algorithm as described in the recently published paper.[72] Patients are categorised into four groups according to their trough infliximab level: a] undetectable, b] low level, c] optimum level (3 to 7ug/mL), and d] high level (> 7ug/mL). Group a] are reflex-tested for anti-infliximab antibodies and subdivided into two groups on the basis of anti-drug antibody test results: for those with high anti-drug antibody levels infliximab therapy is stopped; those with anti-drug antibody at lower level (< 8ug/mL) receive a dose increase of infliximab. Group b] (detectable low trough levels of infliximab; < 3ug/mL) first receive a dosing interval reduction and secondly, if necessary, a dose increase, in attempts to bring infliximab trough concentrations within the "optimal" range (3 to 7ug/mL). Group c] already in the optimum range do

not have dose adjustment. Group c] (high trough levels of infliximab) have their dose interval increased, and if required a subsequent dose reduction. In the trial an "optimisation phase" occurred during which the algorithm was implemented to bring patients into the optimum range, and this preceded randomisation. Only those patients already successfully optimised were randomised; thus if the hypothesis is correct we would potentially expect poor generalisability with higher rates of successful maintenance in the trial than in a broader spectrum of responders.

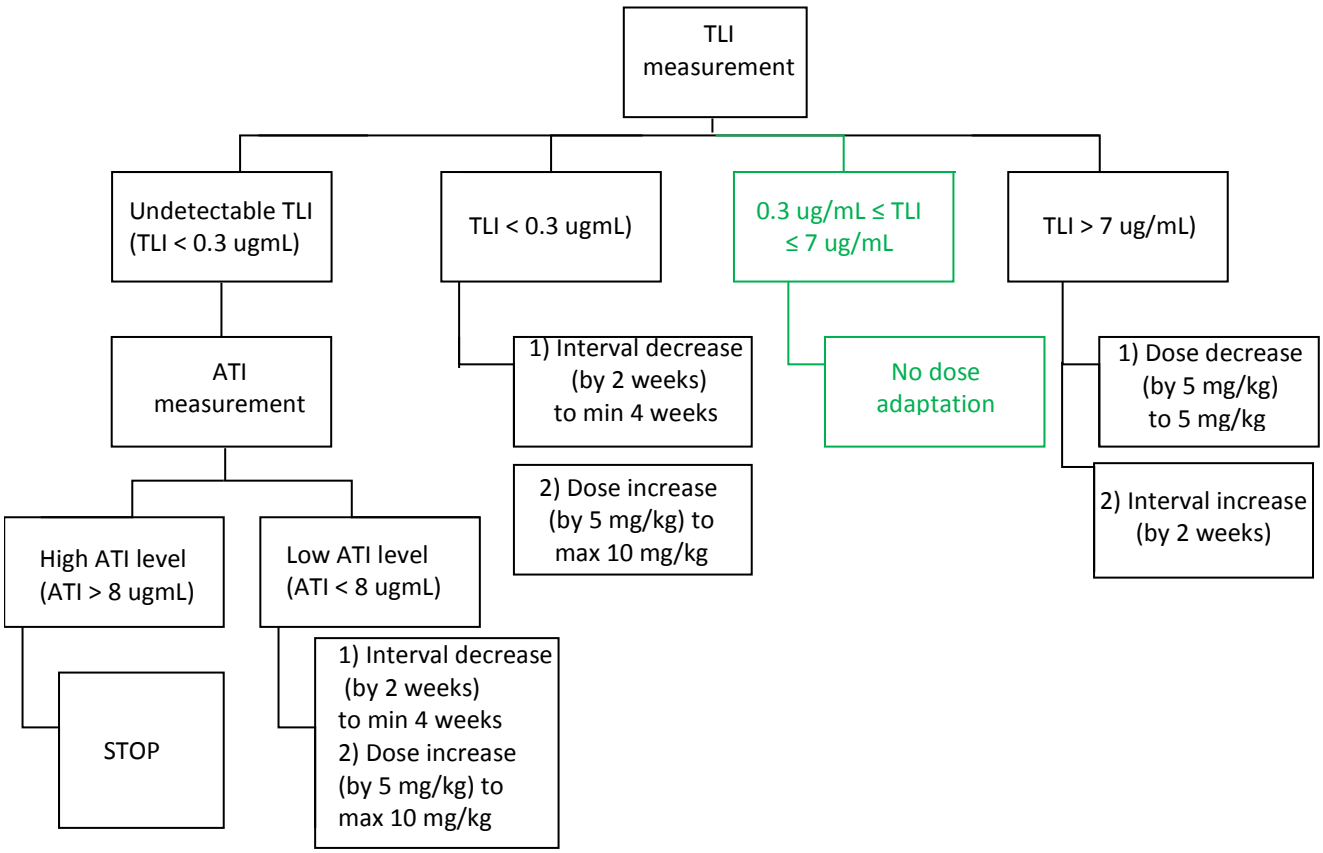


Figure 15 TAXIT algorithm presented in Vande Casteele et al. (2015) [72]
TLI – trough level of infliximab; ATI – antibodies to infliximab

Vaughn et al. (2014)[127] describe trough monitoring of infliximab as a guide to dose adjustment in a group of retrospectively identified IBD patients. Tests were done with a commercial ELISA for the earliest identified patients while for those identified later the commercial HMSA method was used. Initially dose adjustments aimed to bring trough infliximab into the detectable range, but later the target range was changed to 5 to 10ug/mL. The authors quote a typical dose adjustment for those with undetectable trough infliximab to be an increase in dose to 7 mg/Kg with 6 week interval to next infusion, followed by a return to 8 week infusion intervals. The authors state that for trough levels

<5ug/mL the dose of infliximab was increased by "*50 or 100 mg*". For patients with trough infliximab > 10ug/mL on two testing occasions the dose was decreased or, if the patient was already receiving 5 mg/kg (in the full paper this is given as "*5 mg/mL*" and is an assumed typo) the infusion interval was increased. No dose adjustment was made for those in target range. This algorithm would be somewhat difficult to implement without assuming that the authors preferred policy is to use HMSA testing with a target range 5 to 10ug/mL and to manipulate dosage and dose intervals at the clinicians' discretion so as to bring the patient into the target range.

3.2.3.2.2 Comparison to algorithms which are clinically relevant to the NHS which were identified during scoping

Scoping for this report identified two algorithms likely to be relevant for the NHS; one for patients with loss of response for use with an unspecified anti-TNFα (based on Scott and Lichtenstein, 2014[65] that specifies infliximab), and one for responders to infliximab based on the public domain description of the TAXIT trial. The algorithm proposed by Scott and Lichtenstein[65] (Figure 16) requires concurrent testing.

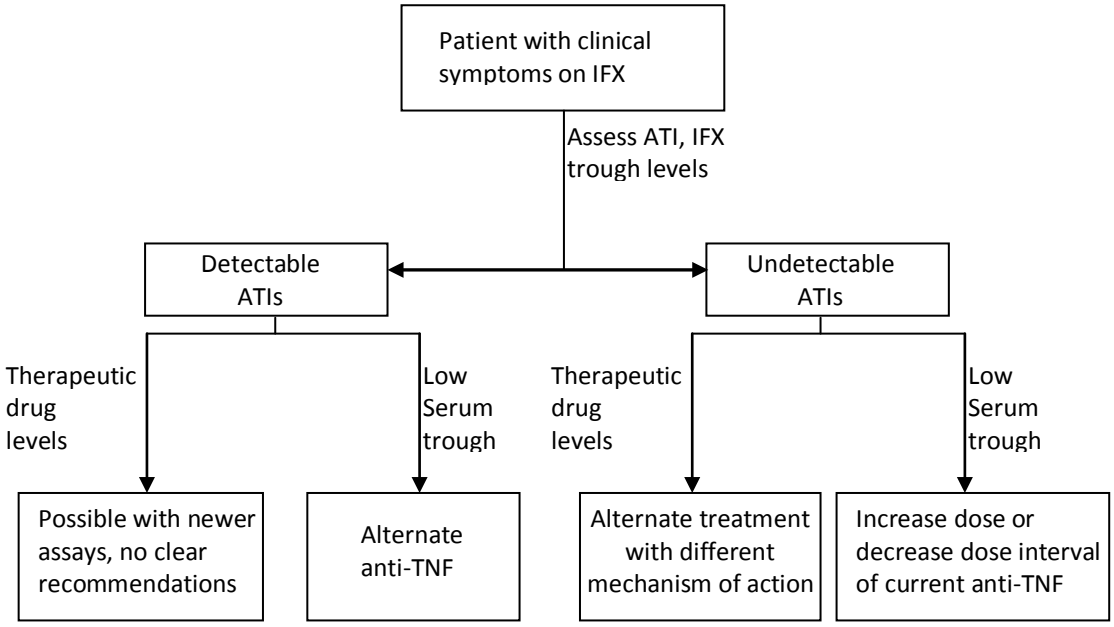


Figure 16 The Scott and Lichtenstein (2014) [65] algorithm for patients with loss of response *IFX infliximab; ATI antibodies to infliximab*

This algorithm is similar to that of Steenholdt et al. (2014)[122] (seen in Table 8) in categorising patients with loss of response into four groups on the basis of dichotomised drug and anti-drug antibody test results. Drug trough levels are classified as therapeutic or low rather than therapeutic or sub-

therapeutic as in Steenholdt et al. (2014).[122] The suggested treatments for "anti-drug antibody + | drug trough level low", "anti-drug antibody – | drug trough level therapeutic", and "anti-drug antibody – | drug trough level low" groups are the same as those by Steenholdt et al. (2014),[122] namely: use alternative anti-TNFα, use therapy with different mechanism of action, escalate drug exposure, respectively. For the "anti-drug antibody + | drug trough level therapeutic" group Scott and Lichtenstein (2014)[65] make no recommendations, but Steenholdt et al. (2014)[122] recommend redeployment to an appropriate group after repeat testing. This is a "generalised" algorithm and therefore differs in detail from that of Steenholdt et al. (2014)[122] in that cut off levels for drug trough levels are not specified and therapies are less prescriptive.

Scoping identified the algorithm presented by Scott and Lichtenstein (Figure 17) which is similar to that of Steenholdt: low drug levels are termed "suboptimal"; undetectable anti-drug antibodies are termed "low or undetectable"; for the group with "suboptimal drug" and "low or undetectable" anti-drug antibodies it is recommended that patient adherence to treatment should be checked and that an immunomodulator may be added; for the group with "therapeutic drug" level and "low or undetectable anti-drug antibodies" this version suggests confirmation of "diagnosis".

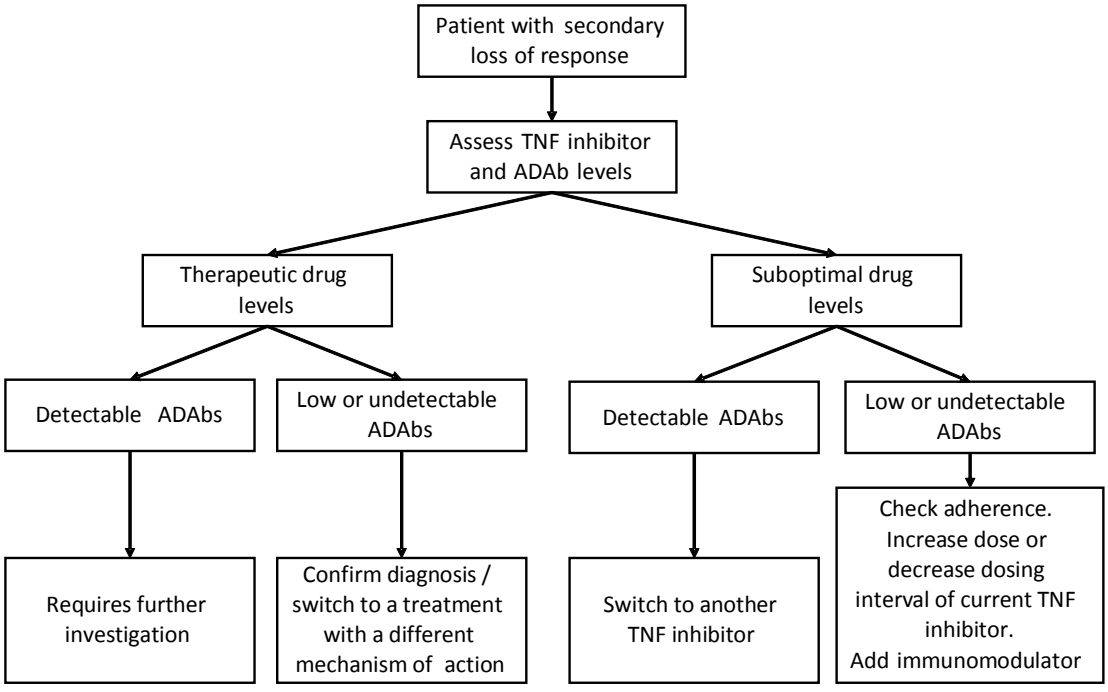


Figure 17 The precursor algorithm for loss of response identified in scoping (based on Scott and Lichtenstein 2014[65]). *TNFα tumour necrosis factor α; ADAb anti-drug antibody; ADAbs anti-drug antibodies*

The scope precursor version of the TAXIT algorithm is almost identical to the public domain version shown in Figure 18. The differences are a] no specific trough drug levels are specified; b] the addition

of an immunomodulator is recommended for the group with undetectable trough drug levels and low anti-drug antibody levels.

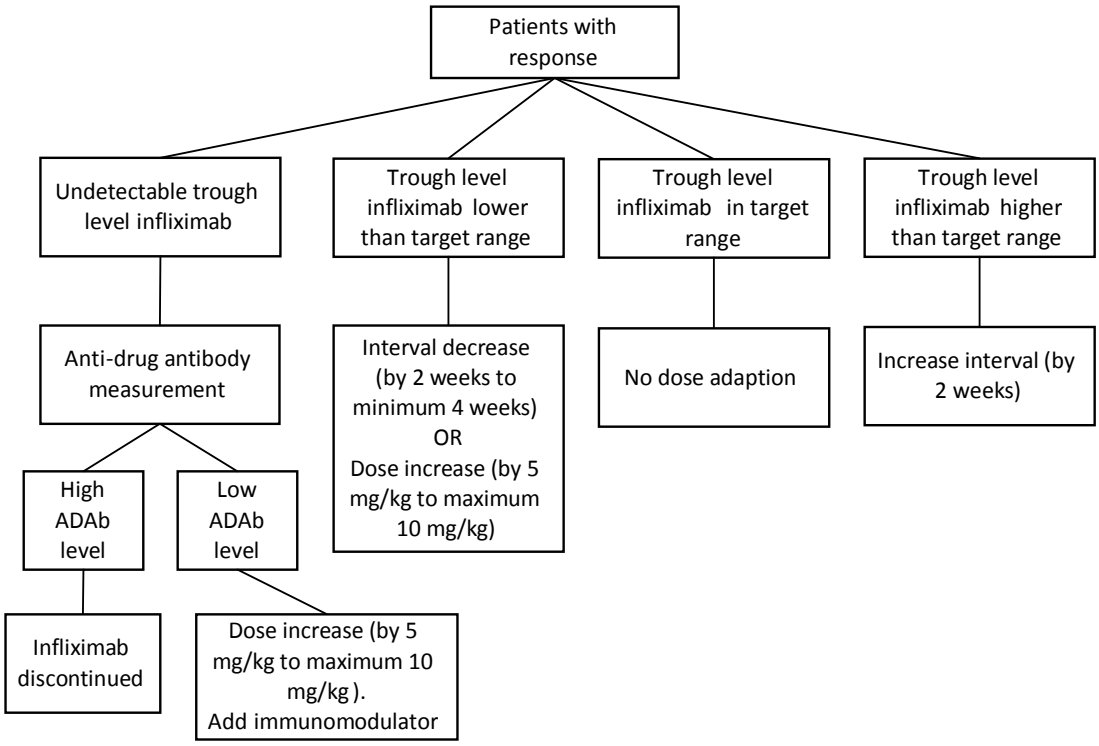


Figure 18 The precursor algorithm based on the TAXIT trial algorithm for infliximab responders *ADAb anti-drug antibody*

3.2.3.2.3 Summary

Only three management studies which used a test-informed algorithm to prescribe treatment of patients were identified.

Vaughn et al. (2014)[127] recommended trough infliximab testing for IBD patients to bring trough infliximab into the presumed therapeutic range (5-10 µg/mL). The algorithm was not adequately prescriptive to allow for easy replication.

The TAXIT trial algorithm for infliximab responders hypothesised therapeutic target range of 3 to 7ug/mL based on analyses using the HMSA method.[139] The trial used an in-house ELISA. The algorithm prescribes dose adjustments for patients with trough infliximab level <3ug/mL, and above >7ug/mL to bring the trough drug level to within target range. Patients with trough infliximab < 0.3 µg/mL were reflex tested for anti-drug antibody and dichotomised as above or below 8 µg/mL, algorithm recommended cessation of infliximab for those > 8 µg/mL. The algorithm is sufficiently detailed to be replicable.

The Steenholdt et al. (2014)[122] algorithm for patients with loss of response to infliximab employs concurrent testing for infliximab and anti-drug antibody and generates four categories of patient: 1) infliximab – & anti-drug antibody +, 2) infliximab – & anti-drug antibody –, 3) infliximab + & anti-drug antibody –, 4) infliximab + & anti-drug antibody +. The trial used RIA assays and cut-offs for dichotomising test results were based on a previous study.[119] The algorithm specifies treatments for each category of patient that are based on hypothesised mechanisms underpinning the loss of response. Treatment for patients with a + test for drug and a – test for anti-drug antibodies was not sufficiently prescriptive to be easily replicated and would likely vary between clinician(s).

The precursor / scoping algorithms represent minor differences of those proposed by Steenholdt et al. (2014)[122] and the TAXIT trial[72] investigators.

In addition to the cut-off levels used in the management studies many more have been suggested by various authors and are summarised in Table 7. This table demonstrates that cut-off levels are study specific and are not readily generalisable.

3.2.4 Objective C1 Clinical studies evaluating drug monitoring for the management of Crohn's disease patients (management studies)

3.2.4.1 Aim

The aim of this section is to assess the evidence from studies which report on the clinical impact of implementing a test-informed treatment algorithm for anti-TNFα recipients with Crohn's disease (CD).

3.2.4.2 Assessment of the risk of bias in the management studies

The risk of bias assessment using the Cochrane risk of bias tool[69] for the two included RCTs [72, 122] is summarised in Table 9 and Figure 19. Steenholdt et al. (2014)[122] described a treatment algorithm in patients with loss of response and, more recently, provided updated longer-term data (20 weeks follow-up).[123] The quality assessment of the retrospective observational pilot study by Vaughn et al. (2014)[127] was assessed using the Downs and Black checklist.[70] Further details on the quality assessment of these three studies are provided in Appendix 10 and below.

3.2.4.2.1 Randomised controlled trials (RCTs)

One RCT reported an adequate method for random sequence generation [72] and one[122, 123] was judged at unclear risk of bias since a block size of 20 for such a small study may not be ideal. Both RCTs had adequate (low risk of bias) treatment allocation concealment, attrition bias (i.e. outcome data) and reporting bias (i.e. complete reporting of outcomes, subgroups and analysis). Steenholdt et al.

(2014),[122] 2015[123]) had a high risk of performance bias as patients were blinded to randomisation group and results of serum analyses, but the physicians were not completely blinded, because they were required to use the results of analyses of serum infliximab and infliximab antibodies in the treatment of those patients who were randomised to the algorithm group. The TAXIT study[72] was considered at unclear risk of performance bias since there was insufficient information about blinding to infliximab trough and antibodies to infliximab concentrations. Both studies had an unclear risk of detection bias as no further information was provided on the blinding of outcomes assessors. Finally, although the TAXIT study [72] was considered adequate in terms of other potential bias (e.g. funding source, statistical methods used, analysis, baseline characteristics), there was concern about potential high risk of bias in the Steenholdt et al. (2014[122] 2015[123]) as 42% of patients were not treated in accordance with the algorithm resulting in patient's crossing over to the "comparator-like" treatment. Overall, Steenholdt et al. (2014,[122] 2015[123]) was rated at high risk of bias and Vande Casteele et al. (2015)[72] was rated as unclear risk of bias according to the Cochrane Handbook on summarising risk of bias.[140]

Table 9 Risk of bias by study: summary of reviewer's judgments on each risk of bias item

Bias item	First author, year, study ID	
	Steenholdt, 2014,[122] 2015[123]	Vande Casteele, 2015[72]
Selection bias Random sequence generation	?	+
Selection bias Allocation concealment	+	+
Performance bias Blinding of participants / personnel	-	?
Detection bias Blinding of outcome assessors	?	?
Attrition bias Incomplete outcome data	+	+
Reporting bias Selective reporting of the outcome, subgroups, or analysis	+	+
Other bias Funding source, adequacy of statistical methods, type of analysis [ITT/PP], baseline imbalance in important characteristics	-	+
- = high risk of bias + = low risk f bias ? = unclear risk of bias		

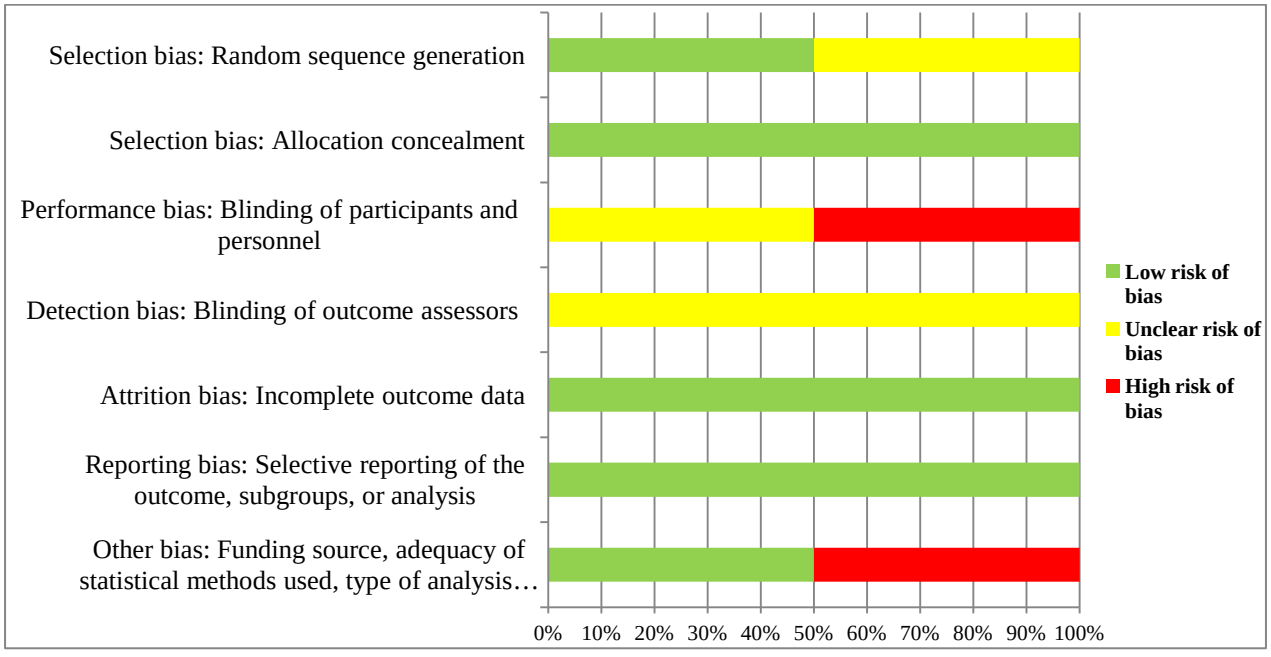


Figure 19 Risk of bias graph across two included RCTs: reviewers' judgments about each risk of bias item

3.2.4.2.2 Non-randomised study

The retrospective observational pilot study by Vaughn et al. (2014)[127] was of adequate quality. The hypothesis, main outcome, characteristics of patients, interventions and main findings were all described appropriately. The remaining reporting items were rated less favorably: a) there was no list of principal confounders in each group; b) it wasn't possible to determine if those participants selected or invited or agreeing to take part in the study were representative of a target population; c) we were unable to determine if staff, places, and facilities where the patients were treated, was representative of the treatment for the majority of patients received. The internal validity items were in general of adequate quality with no data dredging, adjustments made for different lengths of follow-up and appropriate statistical analyses, outcome measures and compliance with the interventions; however, there was no blinding of participants or assessors of the main outcomes. The internal validity selection bias items were rated less favorably than the other items, with concerns raised about whether the recruited participants in each group might be from the same population and the same time period and concerns regarding method of adjustment for confounding and power.

3.2.4.3 Results

After screening 2,428 studies we identified three which matched our inclusion criteria for management studies. All three investigated patients treated with infliximab. There were no management studies for patients treated with adalimumab. The three studies that fulfilled our inclusion criteria (Steenholdt et al., 2014;[122] Vaughn et al., 2014;[127] Vande Casteele et al., 2015[72]) address several aspects of the decision questions. Other studies were identified (Roblin et al. 2014;[56] Paul et al. 2013;[57] Pariente et al. 2012;[58] Afif et al. 2010 [55]) that investigated the clinical utility of therapeutic drug monitoring of anti-TNFα to predict response to a change in treatment mainly due to loss of response to anti-TNFα. On the basis of their results the studies retrospectively suggested a test-informed treatment algorithm for IBD patients, but did not prospectively investigate implementation of a test-algorithm strategy for patient outcomes and compare that with a strategy that might be similar to standard care. For completion these studies have been summarised and their algorithms detailed in Appendix 9. However, as the treatment change in the studies was not prescribed by a standardised algorithm (retrospective studies) and reflected only one treatment change for all patients regardless of test outcome (prospective studies), the studies did not satisfy our inclusion criteria and the outcomes reported were not useful for the health economic evaluation.

3.2.4.3.1 Overview of the included management studies

Two of the management studies included a substantial minority of UC patients along with CD patients. Steenholdt and colleagues (2014)[122] described management with a treatment algorithm in CD patients with loss of response and compared this with standard dose intensification treatment in an

RCT design. Vaughn and colleagues (2014)[127] investigated the impact of proactive drug concentration monitoring of infliximab in retrospectively identified cohorts of IBD patients who were in remission on infliximab (responders); proactive drug monitoring was compared with standard dose intensification treatment for relapse. The TAXIT study[72] was an RCT comparing clinical management versus management using dose adjustment based on trough drug levels, in IBD patients previously brought to target trough level of infliximab using a dose adjustment algorithm.

During the review process, Steenholdt and colleagues were contacted and provided further information and clarification (included below as AIC); also the authors drew our attention to an extension study of the original RCT.[123]

The three studies were heterogeneous with regard to populations, treatment algorithms, and methods of testing for infliximab and antibodies to infliximab. These difference precluded meaningful pooling of study outcomes. Table 10 summarises the major features and findings from the three studies which is followed by a detailed description of the three studies.

Table 10 Summary of the main features of the management studies

	Steenholdt 2014/15[122]	**Vaughn 2014 [127]**	**Vande Casteele 2015 [72]**
Patient population	LOR to infliximab	Responders to infliximab	Responders to infliximab
Type of IBD	CD	71.4% CD, 27% UC *	68% CD; 31% UC
Study design	RCT	Retrospective pilot study	RCT
Setting	6 Danish Centres	Tertiary health care centre	University hospital
Follow-up	12 weeks; 20 weeks	~ 4 years	One year
Aim of the study	Assess cost-effectiveness of algorithm (AL) based treatment vs. dose intensification (II) treatment	Investigate the usefulness of proactive drug monitoring to bring TLI target vs. no-proactive drug monitoring	To compare clinical and biological remission in ClinBD vs. ConBD at one year after randomisation
Comparisons	AL vs. II	proactive drug monitoring vs. no-proactive drug monitoring	ClinBD vs. ConBD
Algorithm by drug & antibody levels	Yes, both	IFX mainly (some reflex testing of antibodies)	IFX mainly (antibodies for those with undetectable IFX)
Test used	RIA	ELISA & HMSA	Leuven in-house ELISA
Time of analysis	At IFX treatment failure	Unclear	Before each infusion time
Drug & abs: Cut-off / target	RIA: drug ≥ 0.5 µg/L; antibodies detectability	Target IFX range: initially detectable, later 5-10 µg/mL	Target IFX: 0.3-0.7 µg/mL antibodies to infliximab (if IFX-): > 8 µg/mL
Limit of quantification	RIA: IFX 0.15 µg/mL; abs: 10 arbitrary units/mL	Variable during study	IFX 0.3 µg/mL antibodies to infliximab 1.0 µg/mL
Definition of clinical response	≥ 70 CDAI reduction from baseline (LuDi); 50%	Unclear / physicians' judgement	Symptom-free OR clear clinical improvement & decrease of

	Steenholdt 2014/15[122]	Vaughn 2014 [127]	Vande Casteele 2015 [72]
	reduction in active fistulas of from baseline (FisDi)		disease activity but with clinical symptoms
Definition of clinical remission	CDAI score ≤of 150 & complete closure of all fistulas despite gentle pressure	Lack of symptoms attributable to underlying IBD (by treating physicians' documentation)	HBI ≤4 for CD & PMS ≤2 with no individual subscore >1 for UC. Biological remission = CRP≤5 mg/L
Definition of clinical progression	Withdrawal for lack of effect of treatment	Not given	Not given
Definition of relapse	NA	Not given	Need of IFX dose escalation, or addition of steroids, or switch to another anti-TNFα (on physician's assessment)
Major findings			
Clinical response by subgroup	II / AL at 12 weeks	NA	NA
Group 1§ (n=14; AL n=5, II n=9)	ITT: 4 (44%)/2 (40%) PP: 4 (44%)/ 2(40%)	NA	NA
Group 2§ (n=3; AL n=1, II n=2)	ITT: 1 (50%)/ 0 (0%) PP: 1 (50%)/ 0(0%)	NA	NA
Group 3§ (n=48; AL n=26, II n=22)	ITT: 12 (55%)/ 16 (62%) PP: 12 (55%)/ 7 (54%)	NA	NA
Group 4§ (n=4; AL n=1, II n=3)	ITT: 2 (67%)/ 0 (0%) PP: 2 (67%)/ 0(0%)	NA	NA
Clinical response 20 weeks (all)	ITT: 56%/76% PP: 56%/74%	NA	NA
Clinical remission	II /AL at 20 weeks ITT: 39%/55% PP: 39%/58%	Not given	Clinical + biological (1 y) ConBD 68.8% ClinBD 65.9%; p=0.880 CD only: ConBD 63% ClinBD 55%; p=0.353
Probability of remaining on infliximab	NA	proactive drug monitoring vs. no-proactive drug monitoring: HR 0.3; 95% CI 0.1- 0.6, p=0.0006 At 5 years in treatment: proactive drug monitoring 86% vs. 52% no-proactive drug monitoring	Not reported

AL = algorithm; CD = Crohn's disease; ClinBD = clinically-based dosing; ConBD = concentration-based dosing; ELISA = enzyme linked immunosorbent assay; HMSA = homogeneous mobility shift assay; IFX = infliximab; II = IFX intensification; ITT = intention to treat analysis; LOR = loss of response; NA = not applicable; PP = per protocol analysis; RIA = radio-immunoassay; TLI = trough level IFX; UC = ulcerative colitis. ***§*** *Group identities of patients with lost response to infliximab were based on concurrent test results as follows: Group 1 infliximab negative anti-drug antibodies positive; Group 2 infliximab negative anti-drug antibodies negative; Group 3 infliximab positive anti-drug antibodies negative; Group 4 infliximab positive anti-drug antibodies positive*

3.2.4.3.2 Steenholdt et al. (2014)[122]

Study design

This was a single-blind randomised controlled trial of 69 adults with Crohn's disease who were previously responsive to maintenance therapy with "regular" infusions of infliximab (at 5 mg/kg) (i.e. patients on maintenance infliximab with loss of response). Participants were randomised to either an infliximab intensified arm (n=36) or to an algorithm arm (n=33). In the former, the dose frequency of 5 mg/kg infliximab was increased to every 4 weeks. In the latter, participants received treatment according to a defined algorithm based on serum concentrations of infliximab and of antibodies to infliximab. It is unclear if the randomisation method was the most appropriate since a block size of 20 for a small study with 69 patients may potentially threaten efficiency of allocation concealment. Follow up was 12 weeks. The study objective was to compare the cost of treatment and the level of disease control of dose intensification (standard) treatment (intensification in infliximab exposure) with that using an algorithm directed treatment strategy informed by concurrent test results. The trial was powered for non-inferiority in disease control and was undertaken in six Danish Centres.

Timing and frequency of testing

Serum samples were collected at the time when infliximab failure was reported and were analysed by radioimmunoassay (samples were stored for retrospective analysis by alternative assays methods including ELISA and HMSA). The RIA cut-offs used were: therapeutic infliximab ≥0.5 μg/L; sub-therapeutic infliximab <0.5 μg/L; the cut off for anti-drug antibody was the limit of quantification (10 arbitrary units/mL).These samples were taken immediately before infliximab infusion. No further tests were undertaken during the 12 weeks follow up.

Treatment algorithm

According to the treatment algorithm, a patient could be categorised into one of the four groups each to receive a defined treatment as shown in Table 8.

The authors suggests the following underlying mechanisms for loss of response:
Group 1 insufficient bioavailability of infliximab due to immunogenicity of infliximab
Group 2 insufficient bioavailability of infliximab due to non-immune mediated pharmacokinetics
Group 3 inhibition of TNFα ineffective due to non-TNFα driven disease
Group 4 patients are classified with group 3 if test results are replicated.

Patient characteristics and concurrent treatments

Participants could receive concomitant therapy of thiopurines, methotrexate or antibiotics or stable doses of topical agents, loperamide, oral hydrocortisone or budesonide. Participants were followed up

every 4 weeks. Mean age was 37 years (range 19 to 81 years) and the majority were female (61%). The mean duration of disease was slightly greater in those in the infliximab intensified arm than in the algorithm arm (10 years, range 1 to 35 years vs. 7 years, range 1 to 27 years). Around one quarter (26%) of the participants gave a history of smoking and 30% had undergone previous surgery while about 20% of patients had received anti-TNFα therapy previously. Mean treatment duration at anti-TNFα failure was about 2 years (657 days range 97 to 3313 days). Mean CRP level was 9 mg/mL (range 2 to 22 mg/mL).

Primary and other outcomes

The dual primary outcome consisted of mean cost of treatment over 12 weeks and the proportion of patients with "clinical response" at 12 weeks. Clinical response was defined as: ">70 point reduction in CDAI score from baseline in luminal disease and a reduction in active fistulas of>50% from baseline in fistulising disease". The study objective of estimating "disease control" in each arm was done using the proportion of patients with the primary outcome of clinical response. Other secondary outcomes included: (i) the proportion with remission, defined as an absolute CDAI score < 150 and complete closure of all fistulas despite gentle pressure (at baseline mean CDAI was 296, range 221 – 526, and301, range 230 – 487, respectively in algorithm and control arms; 3 and 4 patients in each arm had fistulising disease); (ii) the proportion with a CDAI 100 response (a reduction of CDAI score of ≥ 100 from baseline). Mean decrease in CDAI and PDAI scores, and mean increase in IBDQ scores were also reported, together with changes in laboratory measures (WBC, Hb and albumin).

Intention to treat (ITT) and per protocol (PP) populations and handling of withdrawals

Outcome analyses were reported for ITT and for PP populations. All 36 patients in the dose intensification arm received allocated treatment; there were 8 withdrawals for lack of effect or severe infusion reaction.

In the intervention arm patients in group 1 received adalimumab therapy during the study according to local guidelines at the participating centres. Some used adalimumab in the dosing registered by EMA (80 mg at week 0 followed by 40 mg every other week) while others used a more intensive regimen (160 mg at week 0, 80 mg at week 2, 40 mg at week 4 and then 40 mg every other week). Dose optimization of adalimumab was allowed (Personal communication Dr C. Steenholdt, Herlev University Hospitale, Denmark, 25/01/2015).

In the algorithm arm 14 of 33 patients did not receive treatment allocated according to the algorithm, leaving a PP population of 19. Most of these 14 non PP patients continued to receive IFX. Of these, 12 patients continued infliximab (9 patients were in group 3, and 1 patient was in group 4). The applied infliximab regimen was (all received 5 mg/kg): infliximab q8 regimen (2 infusions during the

trial, i.e. week 0 and 8): n=5 ; infliximab q4 regimen (4 infusions during the trial, i.e. week 0,4,8,12): n=2; infliximab q4 regimen but not throughout the entire trial (3 infusions during the trial): n=1; infliximab q4 regimen but not throughout the entire trial (2 infusions during the trial): n=2; infliximab q4 regimen but not throughout the entire trial (1 infusions during the trial): n=2.The remaining 2 patients had been switched to adalimumab due to misinterpretation of test results (Figure 2). Both patients were in group 3. The applied adalimumab regimen was: adalimumab induction (160-80-40) and followed by 40 mg every other week. Adalimumab induction (80-40) and followed by 40 mg every other week (Personal communication Dr Casper Steenholdt, Herlev University Hospitale, Denmark, 25/01/2015).

There were 2 and 8 withdrawals in the algorithm and dose intensified arms respectively. Patients who dropped out were also included in the statistical analyses at subsequent study visits using the last observations carried forward for efficacy (response and remission), CDAI, PDAI, biochemical variables and safety and by using the actual direct medical costs related to CD. There remains some ambiguity since it is unclear if the 8 patients who withdrew from dose intensification contributed medical costs carried forward (but which they did not receive) or if post-withdrawal drug costs were zero.

Test results according to ITT and PP populations

In the algorithm arm concurrent testing categorised patients to the four groups as follows: 26/33 to group 3, 5/33 to group 1, and 1 each to groups 2 and 4. Similar results (not known to the treating physicians) were found for the dose intensified arm. The test results are summarised in Table 11.

Table 11 Proportion of patients according to concurrent testing (ITT population)

	Algorithm arm (n=33)	**Infliximab intensified arm (n=36)**	**All (n=69)**
Grouping in algorithm, n (%)			
Group 1: sub-therapeutic infliximab and anti-drug antibody +	5 (15)	9 (25)	14 (20)
Group 2: sub-therapeutic infliximab and anti-drug antibody undetectable	1 (3)	2 (6)	3 (4)
Group 3: therapeutic infliximab and anti-drug antibody undetectable	26 (79)	22 (61)	48 (70)
Group 4: therapeutic infliximab and anti-drug antibody +	1 (3)	3 (8)	4 (6)

The 14 patients not treated PP in the algorithm arm were all in group 3 (13 patients) or group 4 (1 patient). This left the distribution of groups in the PP population as shown in Table 12.

Table 12 Proportion of patients in each algorithm group (PP population)

	Algorithm arm (n=19)	**Infliximab intensified arm (n=36)**	**All (n=55)**
Grouping in algorithm, n (%)			
Group 1: sub-therapeutic infliximab and anti-drug antibody +	5 (26)	9 (25)	14 (26)
Group 2: sub-therapeutic infliximab and anti-drug antibody undetectable	1 (5)	2 (6)	3 (5)
Group 3: therapeutic infliximab and anti-drug antibody undetectable	13 (68)	22 (61)	35 (64)
Group 4: therapeutic infliximab and anti-drug antibody +	0 (0)	3 (8)	3 (5)

These test results imply that loss of response is most commonly associated with therapeutic drug levels in the absence of detectable anti-drug antibodies (Group 3 represents 70% of 69 patients). The authors' mechanistic interpretation is that "*inhibition of TNFα-alpha is ineffective due to non-TNFα driven disease. TNFα inhibitors not effective and is discontinued*". The algorithm treatment for this group is subject to discretion and requires further investigation and reflection by clinicians.

Primary Outcome results

For the ITT population, the rate of clinical response was similar in the algorithm arm (18/33; 58%) and the dose-intensification arm (19/36; 53%): RR = 1.09; 95% CI: 0.713 to 1.673 (p=0.810). For the PP population the rates were again similar (19/36; 53%) in the dose-intensification arm and (9/19; 47%) in the algorithm arm: RR = 0.898; 95% CI: 0.510 to 1.580, p=0.781).

Table 13 summarises the rates of clinical response in ITT and PP populations according to test defined subgroups. Group 3 (i.e. therapeutic infliximab levels with undetectable anti-infliximab antibodies) contributed the majority of the patients (ITT: 66.7%; PP: 63.6%) and also most of the clinical responses (75.6% ITT, 67.9% PP) and thereby greatly influences the overall comparison between arms.

Table 13 Clinical response according to test defined subgroups

		Response: n/N (%)		RR (95% CI) p
Subgroup	**Pop**	**infliximab intensified arm**	**Algorithm arm**	**algorithm v. infliximab intensified arm**
Group 1: sub-therapeutic infliximab & anti-drug antibody + *Insufficient infliximab bioavailability due to induced immunogenicity of infliximab*	ITT	4/9 (44)	2/5 (40)	RR 0.900 (0.246 to 3.297 p 1.00
	PP	4/9 (44)	2/5 (40)	RR 0.900 (0.246 to 3.297) p = 1.00
Group 2: sub-therapeutic infliximab & anti-drug antibody undetectable *Insufficient infliximab bioavailability due to non-immune mediated pharmacokinetics*	ITT	1/2 (50)	0/1 (0)	NC
	PP	1/2 (50)	0/1 (0)	NC
Group 3: therapeutic infliximab and anti-drug antibody undetectable *Inhibition of TNFα-alpha ineffective due to non-TNFα driven disease*	ITT	12/22 (55)	16/26 (62)	RR 1.128 (0.693 to 1.837) p = 0.770
	PP	12/22 (55)	7/13 (54)	RR 0.987 (0.525 to 1.856) p = 1.00
Group 4: therapeutic infliximab & anti-drug antibody + *Pharmacodynamics or non-functional anti-infliximab antibodies or false positive test*	ITT	2/3 (67)	0/1 (0)	NC
	PP	2/3 (67)	0/0 (0)	NC
All 4 subgroups	ITT	19/36 (53)	18/33 (58)	RR 1.09 (0.713 to 1.673) p= 0.810
	PP	19/36 (53)	9/19 (47)	RR 0.898 (0.510 to 1.580) p =0.781
ITT = intention to treat; PP = per protocol; Pop = population; RR =relative risk; NC = not calculated				

More than half (55%) of group 3 patients in the dose intensification arm had regained response at 12 weeks. This appears surprising if symptoms are driven by a non-TNFα mechanism; however other explanations than intensified infliximab may explain regain of response including changes in or improved effectiveness of concomitant therapies and the natural relapse remission cycling characteristic of CD in these relatively small patient groups.

Quite high response at 12 weeks was found for group 3 algorithm patients (16/26, 62%); of these 26 group 3 patients about half received infliximab; again various infliximab-independent explanations for regain of response include changes in or improved effectiveness of concomitant or introduction of alternative therapies and the natural relapse remission cycling characteristic of CD.

For the ITT population the co-primary outcome measure of mean cost was less in the algorithm arm €6,038 (SD €4146) than the dose-intensification arm €9,178 (SD €2,058) (mean difference: – €3,141 95% CI: – €4,617 to – €1,373; P <0.001). For the PP population mean costs were €4062 (SD €2,763) in the algorithm arm versus €9,178 (SD €2,058) in the dose intensification arm (mean difference: – €5,116, 95% CI: – €6,482 to – €3,561; P <0.001). Table 14 summarises the mean cost in ITT and PP populations according to test-defined subgroups.

Table 14 Mean cost according to test defined subgroups

		Mean (SD) €		**Mean difference (95% CI), p (Algorithm – infliximab intensification)**
Subgroup		**Infliximab intensified arm**	**Algorithm arm**	
Group 1: sub-therapeutic infliximab & anti-drug antibody + *Insufficient infliximab bioavailability due to induced immunogenicity of infliximab*	ITT	8,299 (1,796)	6,837 (990)	– 1,462(– 2,819 to 712) p = 0.090
	PP	8,299 (1,796)	6,837 (990)	– 1,462(– 2,819 to 712) p = 0.090
Group 2: sub-therapeutic infliximab & anti-drug antibody undetectable *Insufficient infliximab bioavailability due to non-immune mediated pharmacokinetics*	ITT	8,666 (1,111)	9,814 (NA)	1,148(NA) p NA
	PP	8,666 (1,111)	9,814 (NA)	1,148(NA) p NA
Group 3: therapeutic infliximab and anti-drug antibody undetectable *Inhibition of TNFα-alpha ineffective due to non-TNFα driven disease*	ITT	9,898 (1,901)	5,728 (4,606)	–4,169(– 5,968 to – 1,788) p = 0.001
	PP	9,898 (1,901)	2,552 (1,639)	–7,349(– 8,557 to – 6,032) p < 0.001
Group 4: therapeutic infliximab & anti-drug antibody +	ITT	6,883 (2,309)	6,003 (NA)	–880(NA) p NA
	PP	6,883	NA	NA

		Mean (SD) €		Mean difference (95% CI), p (Algorithm – infliximab intensification)
Subgroup		**Infliximab intensified arm**	**Algorithm arm**	
Pharmacodynamics or non-functional anti-infliximab antibodies or false positive test		(2,309)	(NA)	p NA
All 4 subgroups	ITT	9,178 (2,058)	6,038 (4,146)	–3,141 (– 4,617 to – 1,373) p < 0.001
	PP	9,178 (2,058)	4,062 (2,763)	– 5,116 (– 6,482 to – 3,561) p < 0.001
ITT = intention to treat; PP = per protocol; pop = population; NA = not applicable; SD = standard deviation				

As for response, the co-primary outcome of mean cost was predominantly contributed from group 3 patients. The total cost of 12 weeks of treatment for the 36 dose intensified patients was € 330,408, of which group 3 patients contributed 65.9% (€ 217,756). The corresponding total cost for 33 algorithm patients was € 199,252 of which 74.7% (€ 148,928) was contributed by group 3 patients. The total 12 week cost for 19 PP algorithm group 3 patients was € 48,488 so that the non PP algorithm group 3 patients (n=13) cost € 100,440 at mean cost of € 7,726.

Remission

According to ITT and PP analysis, more patients in the infliximab intensified arm achieved clinical remission than in the algorithm arm (ITT: 14/36 (39%) vs. 10/33 (30%); PP: 14/36 (39%) vs. 4/19 (21%)); the difference did not reach statistical significance (ITT: RR 0.779, 95% CI 0.403 to 1.507, p=0.613; PP: RR 0.541, 95% CI 0.207 to 1.417, p=0.234. RR is for algorithm versus dose intensification).

Extension study

Clinical outcome findings to 20 weeks and mean cost to 52 weeks were published recently.[141] Of 69 patients included in the trial, 45 (17 in the algorithm arm; 28 in the infliximab intensified arm) completed 12 weeks of per protocol treatment and 29 patients (16 in the algorithm arm; 13 in the infliximab intensified arm) completed 20 weeks of per protocol treatment. Results at 20 weeks were reported for the following populations: ITT (N=69), PP at 12 weeks (N=55), completed PP at 12 weeks (N=45; i.e. 55 minus 10 withdrawals), and completed PP at 20 weeks (N=29). Table 15 summarises the results.

Table 15 Clinical response and remission at 20 weeks (Steenholdt et al. 2015[141])

	Response: n/N (%)		RR (95% CI) p
Population	**Infliximab intensified arm**	**Algorithm arm**	**Algorithm v. infliximab intensified arm**
ITT (N=69)	20/36 (56)	25/33 (76)	RR 1.4 (1.0 to 1.9) p= 0.128
PP (N=55)	20/36 (56)	14/19 (74)	RR 1.3 (0.9 to 2.0) p= 0.248
Completed PP to 12 weeks (N=45)	15/28 (54)	12/17 (71)	RR 1.3 (0.08 to 2.1) p= 0.351
Completed PP to 20 weeks (N=29)	10/13 (77)	11/16 (69)	RR 0.9 (0.6 to 1.4) p =0.697
	Remission: n/N (%)		
	Infliximab intensified arm	**Algorithm arm**	
ITT (N=69)	14/36 (39)	18/33 (55)	RR 1.4 (0.8 to 2.4) p =0.232
PP (N=55)	14/36 (39)	11/19 (58)	RR 1.5 (0.9 to 2.6) p =0.256
Completed PP to 12 weeks (N=45)	10/28 (36)	10/17 (59)	RR 1.7 (0.9 to 3.1) p =0.216
Completed PP to 20 weeks (N=29)	7/13 (54)	9/16 (54)	RR 1.1 (0.5 to 2.0) p = 1.000
ITT = intention to treat; PP = per protocol; pop = population; RR = relative risk			

None of the differences between dose intensified and algorithm groups reached statistical significance. According to ITT analyses in this and the original study, the 18/33 with response at 12 weeks in the algorithm arm improved to 25/33 by week 20, whereas in the dose intensified arm the 19/36 responders at week 12 improved to 20/36 at week 20. For remission there were 14/36 dose intensified patients in remission at both 12 weeks and 20 weeks, whereas in the algorithm arm the proportion increased from 10/33 to 18/33 by week 20. These results imply quite large clinical improvement between weeks 12 and 20 in the algorithm arm and relatively stable clinical status in the dose intensified arm.

In this extension study cost results were reported in US dollars rather than Euros as in the earlier report. This made it problematical to compare cost over the first 12 weeks with those subsequently accumulated to weeks 20 or 52. According to ITT analysis mean cost related to CD at 20 weeks were $11,940 versus $17,236for algorithm and dose intensified arms respectively (mean difference –

$5,296, 95%CI –8,453 to –1,566; p = 0.005). At 52 weeks (ITT analysis) corresponding values were $22,066 versus $29,072 (mean difference –$7,006, 95%CI –12,848 to –874; p = 0.022).

Summary and conclusions

One published study described the implementation of an algorithm in CD patients with loss of response. The authors concluded that the treatment of LOR to infliximab using an algorithm based on concurrent infliximab plus anti-drug antibody measurements significantly reduces average treatment costs per patient compared with routine infliximab dose escalation and without any apparent negative effect on clinical control of disease. These conclusions are supported by the available data, however a number of weaknesses in the study should be borne in mind: the population was small; withdrawals accounted for >20% of patients in the infliximab intensification arm; follow up was short; a large proportion of patients in the algorithm arm did not receive the algorithm-recommended treatment (42%) suggesting a question about whether the efficacy of the algorithm has in fact been tested.

In addition little information was provided on the components contributing to the co-primary outcome of mean cost. Test cost was not reported and it was unclear if or how these were incorporated into the cost analysis; nearly all patients fell into a single algorithm group which was unfortunately the one for whom treatments were least well described and where treatments largely depended on clinicians' judgment and reflection which is unlikely to be replicable between clinicians (n.b. further details of treatments were provided in the extension study).

3.2.4.3.3 Vaughn et al. (2014)[127]

Study design and conduct

The aim was to investigate the usefulness of proactive therapeutic concentration monitoring and titration of infliximab to a target concentration.[127] This was a retrospective observational pilot study of patients with IBD in clinical remission receiving infliximab at tertiary health care centres; patients were identified from records and classified into those who received proactive drug monitoring and those who did not (control group); patients who did not achieve remission were excluded. For both proactive drug monitoring and control group, clinical remission was defined as '*lack of symptoms attributable to underlying IBD based on the treating gastroenterologist's documentation'*.

The infliximab and antibodies to infliximab concentrations were measured initially using solid phase ELISA (Prometheus laboratories) and later with the HMSA assay (Prometheus laboratories). The latter test could detect infliximab as low as 1ug/ml compared to 1.4ug/mL for ELISA. In the proactive drug monitoring group serum trough infliximab levels were used to guide dose modifications to achieve target drug levels. Initially the target was detectable infliximab, later the target was changed

to an infliximab concentration between 5 to 10μg/mL. Typical changes in dose administration in the proactive drug monitoring group were as follows:

- For patients with undetectable trough drug levels, the dose of infliximab infusion was increased to 7.5 mg/kg after which the next infusion was given after 6 weeks, and after this infliximab was given every 8 weeks.
- For patients with detectable trough drug levels that were <5μg/mL, the dose of infliximab was increased by 50 or 100 mg.
- In patients with trough drug levels of >10μg/mL on at least two occasions, the dose was reduced. However, those patients who were already receiving 5 mg/kg of infliximab, instead of dose modification, the treatment interval was increased.
- In patients who had trough levels in the range of 5 to 10μg/mL, no changes were made.

(A trough concentration was defined as '*any infliximab concentration performed within 7 days of the next infusion'*).

Reactive testing was done in both groups either for loss of response (LOR) or if there was a concern for side effects due to antibody formation. For patients in the control group with LOR the dose of infliximab was increased at the treating physician's discretion according to a standard of care guideline (typically to 10mg/kg, but the dose did not reach > 10 mg/kg every four weeks).

Patient populations

There were 48 and 78 IBD patients in the proactive drug monitoring and control groups respectively. They were followed from the start of maintenance therapy until August 2013 or until their last documented clinical encounter. Proactive drug monitoring was initiated at some time during patients' maintenance and was adopted as a strategy "starting in 2009". The determining difference between groups was that testing was only performed reactively in the control group but both reactively and also proactively in the proactive drug monitoring group so as to determine any dose changes judged necessary to reach target trough concentration; furthermore when dose was escalated in the control group infliximab exposure was likely doubled (e.g. to10 mg/kg), but dose escalations in the proactive drug monitoring group were of much smaller magnitude (e.g. by 50 to 100 mg; for a 70kg individual this raises the dose from 5 mg/kg to between 5.7 and 6.4 mg/kg). Dose de-escalation (to <5 mg/kg) only occurred for the proactive drug monitoring group.

Two patients were "IBD unclassified", and 90 (69%) and 34 (29%) were diagnosed as CD and UC respectively. Almost 70% were male. Median age at infliximab initiation was 34.9 and 35years in proactive drug monitoring and control arms (IQR range 26.2 to 49.7). The median age at diagnosis was 23.5 and 25 in the proactive drug monitoring and control arms; 30% of patients had undergone

IBD surgery previously (40% of the proactive drug monitoring group but only 25% of the control group); 10% of patients were current users of tobacco, 25% were former users and 56% had never used tobacco; 52 patients (41%) received combination therapy (44% and 40% of the proactive drug monitoring and control group respectively). The median duration of infliximab before proactive drug monitoring was 43 weeks (IQR 32 to 72 weeks).

The main reported outcomes comparing proactive drug monitoring and control groups were: time remaining on treatment (Kaplan-Meier analysis) and reasons for stopping infliximab. Further details about dose changes and trough levels in the proactive drug monitoring group were also provided.

Outcomes: Time remaining on infliximab.

Patients identified as belonging to the proactive drug monitoring group remained on infliximab treatment longer than those identified as belonging to the control group. At 5 years (260 weeks) the probabilities of remaining on treatment were 86% and 52% respectively. Figure 20 shows the reconstructed Kaplan-Meier comparison between groups. Beyond 5 years there are very few patients at risk; the median duration of infliximab before proactive drug monitoring implementation was reported to be 43 weeks (interquartile range, 32-72 weeks) .In multiple Cox regression analysis, the probability of patients remaining on infliximab therapy was found to be significantly related only to proactive drug monitoring of infliximab.

Figure 20 Kaplan-Meier analysis of time to stopping infliximab treatment IPD reconstructed using the method of Guyot et al. (2012)[73] *The vertical line represents 5 years. Inset shows proportional hazards test plot*

The authors also reported on the subgroup of patients that started maintenance infliximab after 1/1/2009. Since the implementation of proactive drug monitoring was reported to be from 2009 this subgroup would appear to be the more relevant population. Figure 21 shows the reconstructed Kaplan-Meier comparison between the proactive drug monitoring and control subgroups. The reported HR was 0.3 (95% CI: 0.1 to 0.7; p =0.003). The reconstructed HR was similar 0.24, (95%CI: 0.12 to 0.51); it is possible the authors stratified their analysis by baseline variables (e.g. mono or combination therapy, previous surgery, etc.). Parametric modelling based on reconstructed IPD is provided in Appendix 11.

Figure 21 Kaplan-Meier analysis of time to stopping infliximab treatment, patients starting maintenance January 2009 IPD reconstructed using the method of Guyot et al. (2012)[73]

Outcomes: reasons for stopping infliximab

The reasons for stopping infliximab are summarised in Table 16 Reasons for stopping infliximab therapy; the most frequent causes for the control group were recurrence of IBD symptoms and acute infusions reactions. Adverse events and high antibody levels were the main causes for the proactive drug monitoring group.

Table 16 Reasons for stopping infliximab therapy

	Proactive drug monitoring	No proactive drug monitoring
Recurrent IBD symptoms	0	15
Adverse events		
Pneumonia	0	1
Drug-induced lupus	1	0
Psoriasis	1	0
High antibody concentration	1	0
Infusion reactions		
Acute	0	6
Delayed	1	0
Other unrelated to infliximab *	1	2

**Includes: unable to afford co-payment, surgery for adhesive small bowl obstruction, and proactive drug monitoring group trough levels and dose changes colectomy for flat low grade dysplasia.*

Outcomes: Proactive drug monitoring group trough levels and dose changes

For the proactive drug monitoring group the authors reported data concerning proactive tests undertaken; in these data all reactive tests were omitted. Of initial proactive tests (n=48) 37 used ELISA and 11 HMSA methodology. Of subsequent proactive tests (n=40) 7 used ELISA and 33 HMSA.

Median trough infliximab at initial testing was 5ug/mL (IQR 2.8 to 9.9) while median subsequent trough level was 7.6ug/mL (IQR 4.3 to 12.3).

Dosing adjustment after the initial proactive test involved 17 patients (35%); the adjustments were described as: 12 (71%) dose escalation, 3 (18%) dose decrease, 2 (12%) stopped therapy. Dosing adjustment in subsequent proactive tests involved 10 patients (25%); these adjustments were described as: 8 (80%) dose escalation, 2 (20%) dose decrease.

Following proactive drug monitoring, the median dose increment was 100 mg (range 50 to 250 mg) and the median duration of infliximab therapy was 144 weeks (range 36 to 685 weeks).

In 75% of patients (36/48) were reported to have reached a trough level of infliximab 5ug/mL or higher (lower end of the later target range). Of those that reached this level of infliximab, none developed antibodies to infliximab or an IR. One patient stopped infliximab after colectomy which was undertaken for flat low-grade dysplasia

Authors' conclusions

The authors concluded that proactive trough concentration monitoring of infliximab frequently identified patients with low or undetectable trough concentrations and resulted in a greater probability of remaining on infliximab. In this study the treatment algorithm was ill defined and test methods were adjusted during the study. The retrospective observational design means the selection of patient groups was at risk of bias.

3.2.4.3.4 Vande Casteele et al. (2015) - TAXIT (Trough level Adapted infliXImab Treatment) study[72]

Study design

This RCT[72] included 251 IBD patients (CD = 173; UC = 78) with stable response to infliximab therapy who were randomised (1:1) to two different treatment strategies: a] clinically based dosing (n=123); or b] infliximab trough concentration based dosing (n=128) that targeted infliximab trough level to 3 to 7ug/mL. Prior to randomisation a consecutive cohort of 275 IBD patients (186 CD, 89 UC) were screened and subjected to an optimisation phase using an algorithm for dose adjustment so as to identify patients whose trough infliximab levels could be successfully brought to the target range. All randomised patients entered with trough infliximab levels within target range. For patients randomised to the clinically based dosing arm subsequent infliximab dosing was according to clinical symptoms and CRP levels (recorded at each infusion) and followed standard clinical criteria. For those randomised to the trough concentration-based dosing arm infliximab dosing continued according to the algorithm. Patients were followed for 52 weeks post randomisation.

Of the 275 consecutive patients, 12 were excluded due to loss to follow up, or ineligibility, or because their trough infliximab levels was undetectable and antibodies to infliximab were detected at > 8ug/mL (n=6). The remaining 263 proceeded to optimisation. For 12 patients the optimisation algorithm failed to bring trough infliximab levels into target range; the remaining 251 were randomised to continued dosing based on trough infliximab levels or to the clinically based dosing strategy.

Specified primary and secondary outcomes

The primary outcome was the rate of clinical plus biological remission one year after randomisation; clinical remission required a Harvard-Bradshaw index of≤4 for CD and partial MAYO score of ≤2 for UC; biological remission required a CRP of ≤5 mg/L). Early terminations were considered failures for the primary endpoint (criteria for termination included: safety and failure of infliximab therapy defined as persisting clinical symptoms (HBI >4 or PMS >2) on two consecutive visits (including

unscheduled visits) AND active inflammation based on increased CRP concentration OR endoscopic activity).

Secondary outcomes included: durable remission, relapse, trough infliximab levels in target range, anti-drug antibody positivity, EQ-5d QOL, and total cost of treatment. In the recently accepted paper, an objective was also to compare cost-effectiveness and safety of trough level based dosing to clinically-based dosing of infliximab.

Optimisation phase

During optimisation patients were first categorised into 4 groups on the basis of trough infliximab levels: 1] >7ug/mL trough infliximab levels; 2] in target range trough infliximab levels (3 to 7ug/mL); 3] <3ug/mL trough infliximab levels; and 4] undetectable trough infliximab levels and anti-drug antibodies < 8ug/mL (patients with undetectable trough infliximab levels but anti-drug antibodies > 8ug/mL were excluded at screening). Each category was administered dose adjustments according to the algorithm shown in **Error! Reference source not found.**.

Of 72 patients with trough infliximab levels < 3ug/ml (categories 3] or 4]) dose escalation brought 69 to target trough level of infliximab; 115 category 2] patients were in range and were randomised; 67 of 72 group 1] patients were dose de-escalated to target range. A total of 251 patients were randomised, 128 to trough infliximab levels monitored dosing and 123 to clinically based dosing.

Infliximab and antibody measurement

The trough infliximab and anti-drug antibody levels were measured using an in-house developed ELISA (Leuven in-house ELISA). The trough infliximab level was measured using direct ELISA and anti-drug antibody levels using bridging ELISA. The lowest quantification value for infliximab and anti-drug antibody limit for the test were 0.3ug/ml and 1.0ug/ml respectively.

Patient characteristics

The authors reported baseline characteristics and results according to two phases: optimisation phase and maintenance phase (i.e. post-randomisation).

Optimisation phase (n=263): Mean age of patients was 41.0 years (range 30 to 48.5 years). 77.2% of patients were in remission, mean CRP level was 1.7 mg/L and the mean infliximab trough level 4.6ug/mL (2.5 to 7.7ug/ml). Around 5% of patients were receiving immunomodulators.

Maintenance phase (n=251): About 55% of patients were female; most patients (69%) were diagnosed with CD; approximately 30% had previously undergone surgery; median duration of disease was 12.5 years (6.3 to 19.9 years); median duration of disease at first infliximab exposure was 5.8 years (range 1.7 to 13.5 years); median time since first infliximab was 4.6 years (2.1 to 7.5 years); 82.5% of patients were in remission (79.8% CD; 88.5% UC); mean CRP and mean infliximab trough concentration were 1.4 (range 0.6 to 4.2) mg/L and 4.9 (range 3.9 to 8.5) µg/ml respectively.

Results: optimisation phase

The results for CD patients are summarised in Table 17.

Table 17 Remission rates for CD patients; comparison of post-optimisation versus pre-optimisation

Patient group	After optimisation Clinical remission	Before optimisation Clinical remission	Statistic (after v. before)
All CD patients §	138/173 (79.8%)	131/178 (73.6%)	RR 1.053, 95%CI: 0.936 to 1.186
CD patients dose escalated §§	38/43 (88.4%)	28/43 (65.1%)	OR 4.071, 95%CI: 1.324 to 12.524 RR 1.297, 95%CI: 1.008 to 1.669
CD patients dose reduced §§	35/51 (69.4%)	41/51 (80.4%)	OR 0.534, 95%CI: 0.215 to 1.325 RR 0.854, 95%CI: 0.678 to 1.074
§ Intention to treat analysis. §§ per protocol analysis, numbers of patients estimated from reported percentages. RR = relative risk; OR = odds ratio.			

Of 178 CD patients entering optimisation 131 were in clinical remission (HBI ≤ 4). After optimisation 138/173 were in remission (intention to treat analysis RR = 1.053, 95%CI: 0.936 to 1.186).

Of 44 CD patients in the dose escalation group entering optimisation 43 achieved target trough infliximab levels. Of these 43, 28 were in clinical remission at entry and this rose to 38/43 after optimisation (reported per protocol OR=4.1; 95% CI: 1.3 to 12.5; p = 0.020. RR=1.297, 95%CI: 1.008 to 1.669). These patients also showed a significant decrease in mean CRP at the end of optimisation (from 4.3 mg/L to 3.2 mg/L (p <0.001). Corresponding results for UC patients did not reach statistical significance (p =1.0 and p = 0.16 respectively). For CD patients who had dose reduction during optimisation (per protocol N = 51) the proportion in remission decreased from 80.4% to 69.4% (per protocol RR = 0.854, 95%CI: 0.678 to 1.074).

No statistically significant changes in clinical remission or in mean CRP concentration by the end of optimisation were observed for CD or UC patients who achieved target trough infliximab levels (p= 0.3 and p=1.0 for clinical remission and p=0.56 a p =0.86 for CRP levels respectively).

For the dose escalation group an average of 2.1 optimisations were required to reach target trough infliximab levels, and at the end of optimisation the median infusion interval was 6 weeks (range 4 to

8 weeks). For the dose reduction group a mean of 1.4 optimisations was required and the median infusion interval was 8 weeks (range 6 to 12).

Results: maintenance phase 52 week primary outcome

Almost 90% of patients completed the maintenance phase. The reasons for not completing in the clinically-based and concentration-based dosing arms respectively were: discontinuation due to active disease (4 and 4), serious adverse event (1 and 1), lost to follow-up (2 and 1), pregnancy (3 and 1), inability to maintain the target trough level (none and 1) and other reasons (2 and 1).

Similar primary outcome rates were observed for both randomised arms (88/128 (68.8%) with the concentration-based dosing and 81/123 (66%) for clinically based dosing (p=0.686)). Corresponding results for CD and UC patients separately were: 63% vs. 55% (p=0.353), and 88% vs. 84% (p=0.748) respectively.

Results did not change when analysis was restricted to only those in remission at the start of maintenance.

Results: maintenance phase secondary outcomes

There was little difference between groups in probability of maintaining durable remission (26% and 27% in concentration based and clinically based dosing arms respectively; p=0.88)

More patients in the concentration-based dosing arm than in the clinically based dosing arm (74% vs. 57%) had the infliximab trough concentration between 3 and 7ug/mL (p<0.001) whereas the risk of patients in the clinically based arm having undetectable trough levels of infliximab was significantly greater (RR 3.7; 95% CI 1.7 to 8.0; p<0.001). None of the patients in the concentration-based dosing arm were positive for anti-drug antibodies but, three patients in the clinically based arm were (p=0.116).

No deaths occurred in any group but, two patients in the clinically based dosing arm required hospital admission, one due to acute appendicitis and another due to ileostomy complications. There were 12 and 13 discontinuations in the clinically based dosing and concentration-based dosing arms respectively.

More patients in the clinically based arm (n=21, 17%) than in the concentration-based dosing group (n=9, 7%) relapsed and needed rescue therapy (RR of 2.4, 95% CI 1.2 to 5.1; p=0.018). Relapse defined as *'the need for infliximab dose escalation (interval decrease and/or dose increase), the*

addition of steroids or switch to another anti-TNFα and was based on the physician's global assessment'. In those relapsing and requiring rescue therapy, comparatively greater numbers of patients (9/21) in the clinically based arm than in the concentration-based arm (2/9) had trough infliximab levels < 3ug/ml.

Relapse free-survival time was superior in the concentration based dosing arm than in the clinically based dosing arm (Figure 22).

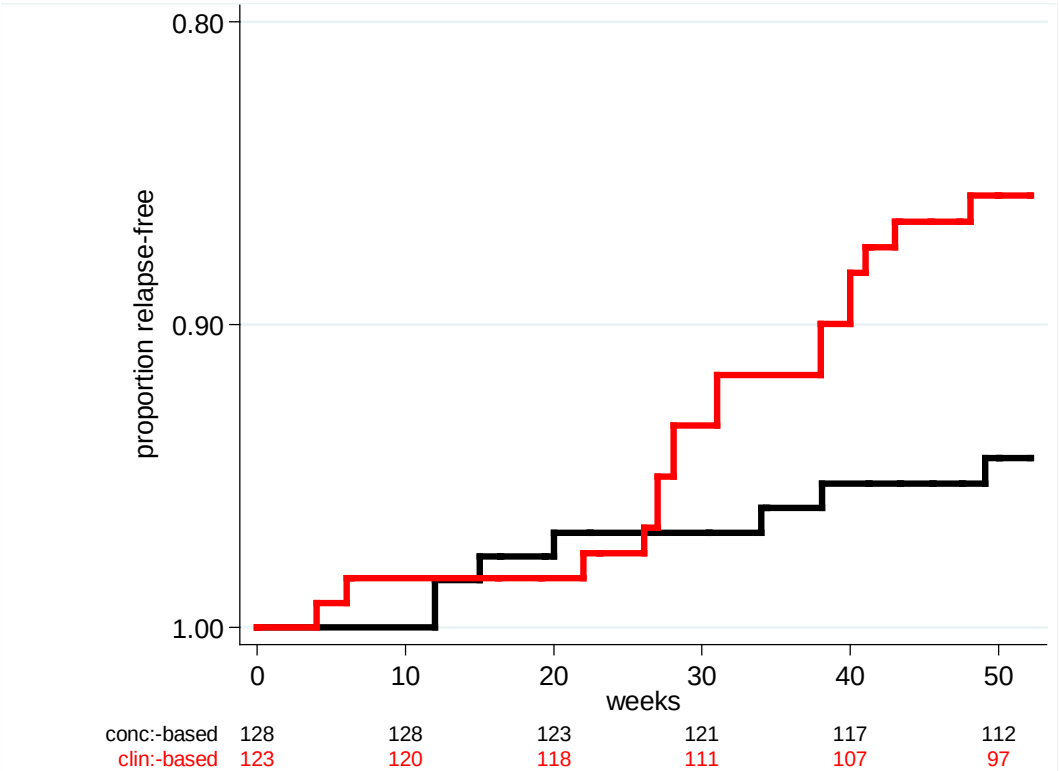


Figure 22 Kaplan-Meier analysis of time to relapse during maintenance phase. *IPD reconstructed using the method of Guyot et al. (2012)*[73]

Authors' conclusions

The authors concluded that optimisation of infliximab dose to achieve the target trough levels of 3 to 7ug/ml is more efficient and is cost-effective relative to clinically based adjustment. Therefore, the authors recommended using dose-to-target optimisation of infliximab to achieve the target trough infliximab levels and to re-evaluate the level after 6 months. It should be borne in mind that both arms received target dose optimisation prior to randomisation and therefore even the comparator group which received a clinically guided dosing regimen had already received a phase of trough level monitoring and dose adjustment.

3.2.4.4 Summary of major findings from three management studies

Three disparate studies were identified that implemented a test-informed algorithm and reported clinical outcomes. One examined CD patients only, the other two both CD and UC patients. Two were RCTs and the other a retrospective observational study. None employed designated intervention tests (LISA-TRACKER, TNFα-Blocker, or Promonitor ELISAs).

The Steenholdt et al. (2014) RCT[122] used concurrent RIA testing prior to implementation of a treatment algorithm for CD patients with lost response to infliximab; the comparator group received infliximab intensification. At 12 weeks after randomisation there was no clinical benefit from the test-algorithm strategy relative to dose intensification. For 64% of patients in the algorithm arm (those with therapeutic infliximab but no anti-drug antibodies) the algorithm recommended cessation of infliximab therapy. That cessation of infliximab therapy was not associated with reduced disease control suggests that infliximab may not be useful for most CD patients with LOR; however the criteria for loss of response may have been imprecise so that patients appeared to regain response 12 weeks later. Weaknesses of the study include: short duration, small size, substantial withdrawals, that many participants did not receive algorithm-prescribed treatments, and unclear or high risk of bias in several risk of bias domains. The authors reported cost savings for the test-algorithm group relative to the dose escalation group that are probably attributable to less use of infliximab in the algorithm arm. Further studies are required to test the reliability of findings.

The TAXIT RCT (Vande Casteele et al., 2015[72]) used a test-algorithm for infliximab responders so as to optimise the trough infliximab level to a set target range. Tests employed "in-house" ELISAs. Trough optimisation with dose adjustments did not change the proportion of CD patients in clinical remission (RR for post- versus pre-optimisation = 1.053, 95%CI: 0.936 to 1.186). After trough-level optimisation patients were randomised to continued trough-test monitoring or to clinical monitoring. For the primary outcome (rate of clinical plus biological remission at 52 weeks) there was no difference between groups for CD patients (54.9% vs. 62.6%; P = 0.353). Time to relapse for CD plus UC patients was superior with test monitoring versus clinical monitoring (P = 0.018). Total cost post-randomisation (CD plus UC patients) was slightly lower with test-monitoring (€20,723 vs. €21,023). The risk of bias was low for most risk of bias domains.

The retrospective observational study of Vaughn et al. (2014)[127] compared proactive trough concentration monitoring with dosing based on clinical judgment (test results did not influence treatment). The clinically managed patients' dose escalations were likely to involve a doubling of infliximab exposure whereas in the trough-monitoring group some dose changes were dose reductions and dose escalations were considerably more moderate than in the clinically-managed group. The

authors' major finding was that relative to clinical monitoring trough monitoring was associated with far superior retention on infliximab treatment (HR 0.3, 95% CI 0.1 to 0.7; p = 0.003). The observational design of the study and the retrospective identification of participants based on medical records mean that the study was at considerable risk of bias.

3.2.4.5 Limitations of the review of management studies

The most important limitation of this review was that the relevant management studies did not directly address our research questions; a further difficulty was the very limited supply of studies. Three management studies were found in which patients treated with infliximab were investigated but no corresponding studies were found for adalimumab. The timing of testing specified in the research questions did not correspond with that used in any of the studies. Furthermore, two of the three available studies investigated a mixture of CD and UC patients and only one study (Steenholdt et al., 2014[122]) reported the impact of treatment algorithm in clinical outcomes for patients exclusively diagnosed with CD. In this study there were few patients and follow up was short so that the power to detect differences in clinical outcome between randomised groups over a clinically meaningful period was limited. However, this study provided some evidence that, at least in the short term, a dose escalation strategy for loss of response to infliximab may be more costly than the alternative strategy proposed in the author's treatment algorithm. Further investigation is required to establish that cost savings are not associated with deterioration in disease control. The TAXIT RCT investigated IBD patients with stable response to infliximab.[72] Patients in both randomised groups were optimised to a target trough level of infliximab so that a comparator group in which from outset "*Treatment decisions made on clinical judgment without measuring levels of TNFα inhibitor and antibodies to TNFα inhibitors*" did not exist. The pilot study of Vaughn et al. (2014)[127] was a retrospective observational study and therefore findings should be viewed with considerable caution. This review of management studies clearly highlights gaps in the evidence and indicates that further studies are needed.

3.2.4.6 Evidence taken forward to the economic evaluation

Data from the three management studies[72, 122, 127] have been taken forward for economic evaluation. The two RCTs, one for responders and the other for infliximab recipients with loss of response, have informed model structure and provided information for the base case economic analysis. The study by Vaughn,[127] which reports substantial clinical advantage for a test-algorithm strategy in terms of retention in infliximab treatment, has been used in economic evaluation sensitivity analysis.

3.2.5 Objective C2 Studies relating test results to clinical state of patients (correlation studies)

3.2.5.1 Search results

The search identified three systematic reviews with meta-analytic pooling of results from multiple studies[62, 110, 113] and 31 primary studies[37, 39, 46, 51, 58, 76-84, 87, 91, 93, 97-99, 101, 102, 105, 107, 109, 114, 119, 122, 125, 132, 133] that reported the relationship between test outcomes and clinical status of patients in sufficient detail allowing 2x2 data being extracted of diagnostic performance when using a drug and / or anti-drug antibody test to diagnose / predict response or loss of response. The systematic reviews are summarised in section 3.2.5.2 and the primary studies are analysed in section 3.2.5.3.

3.2.5.2 Published meta-analyses of studies relating test results to clinical state of patients

3.2.5.2.1 Aim

To present an overview of meta-analyses of studies addressing the relationship between drug and / or anti-drug antibody levels and clinical state of patients with CD.

3.2.5.2.2 Rationale

In order to use anti-TNFα drug and anti-drug antibody levels as tests to aid the management of CD patients on anti-TNFα drugs, the test results are used to predict response or loss of response which will prompt appropriate action. How good the tests are will therefore not only depend on the choice of treatment (change) following the test results (prescribed by the algorithms discussed in section 3.2.3) but also on the diagnostic performance of the test to predict response or lack of response correctly. We therefore reviewed systematic reviews with meta-analyses of studies addressing the relationship between drug and / or anti-drug antibody levels and clinical state of patients with CD to assess the diagnostic performance of the various assays in predicting response and LOR. It should be kept in mind that the definitions of response and remission are not standardised, and that the standard that the tests are measured against is clinical assessment which is far from perfect.

3.2.5.2.3 Results

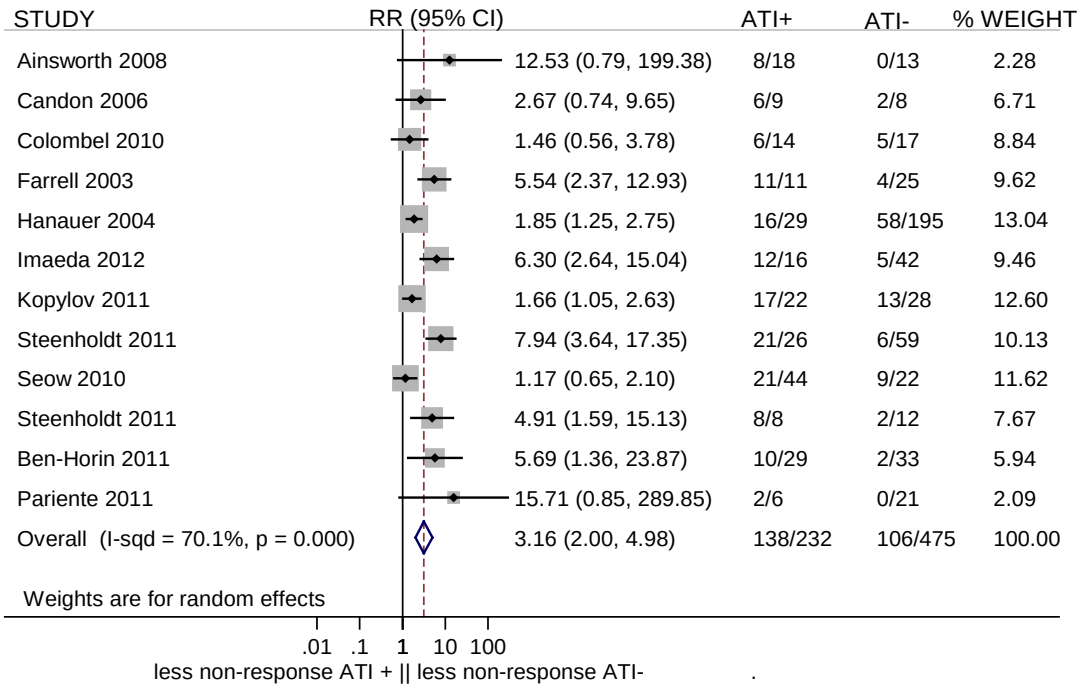
The literature search yielded several reviews which addressed the relationship between test results and the clinical state of patients with IBD.[35, 43, 48, 60-62, 110, 113, 138] Of these reviews, four were systematic with meta-analytic pooling of results from multiple studies.[48, 62, 110, 113] One meta-analysis (MA) encompassed several inflammatory conditions in addition to IBD and is not considered further here.[48] The three remaining meta-analyses considered anti-drug antibodies and one also examined drug trough-level tests.[113] Although many of the primary studies included in the MAs presented data in terms of diagnostic or predictive tests (e.g. sensitivity, specificity and other test accuracy measures), the meta-analyses addressed the risk of a particular test result (e.g. negative) in patients with a

particular clinical state e.g. LOR and calculated a relative risk of a negative test result in LOR relative to state no LOR, or conversely relative risk of LOR in patients with negative test relative to those with a positive test. Viewing the tests as diagnostic/ predictive, permits hierarchical (bivariate) MA that incorporates covariance between sensitivity and specificity estimates. The relative risk statistic does not formally allow for covariance between estimated associations. Below each of the MAs is considered in turn.

Nanda et al. (2013)[110]

The authors estimated the pooled relative risk of LOR to infliximab in patients with a positive test for anti-drug antibodies relative to those with a negative test for anti-drug antibodies (a greater risk of LOR in antibody positive patients compared to antibody negative patients generates a RR > 1.0). Eleven studies were included, one with only UC patients, three studies with mixed IBD populations (one of which reported results separately by UC and CD) and seven studies of CD patients. The comparative numbers of events and patients were reported. The pooled estimate (RR = 3.16; 95%CI 2.00 to 4.98. I^2 = 70.1% Figure 23 upper panel) indicated about a three-fold greater risk of LOR in those with a positive anti-drug antibodies test than in those with a negative test.

When viewed as a predictive / diagnostic test (Pepe et al., 2003[142]) the same data can be analysed to estimate the sensitivity and specificity, and meta-analysed to generate a pooled joint sensitivity-specificity value (and other test accuracy parameters).[143] In this a positive test for anti-drug antibodies is viewed as predictive / diagnostic of LOR. Figure 23 middle panel indicates marked heterogeneity amongst the studies and the trade-off between sensitivity and specificity in the different studies. Meta-analysis of sensitivity and specificity is not recommended; estimates of statistical heterogeneity unexplained by chance (I^2) were 83.7 % and 87.9 % for sensitivity and specificity respectively. The lower panel summarises summary receiver operating characteristic (sROC) MA results. The MA test accuracy results are summarised in Table 18. The large RCT-based study by Hanauer et al. (2004) [39] was identified as both influential and as an outlier; including or excluding this study made little difference to the summary test accuracy estimates but substantially decreased the 95% CI around the prediction region in sROC space (lower panel Figure 23). This study differed from the others in having the lowest ratio of positive to negative test results probably resulting from the number of tests classified as inconclusive.



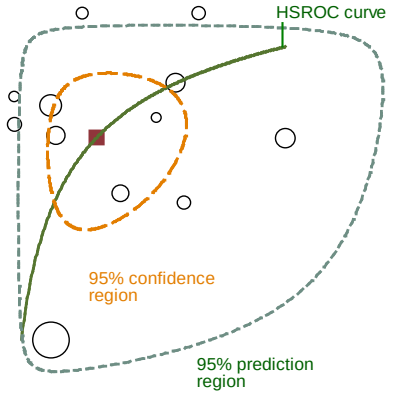


Figure 23 Meta-analysis of data based on that from Nanda et al. (2013)[110]
Upper panel: random effects MA of the relative risk of LOR to infliximab (anti-drug antibodies + versus anti-drug antibodies –). Middle panel: sensitivity and specificity forest plot of the included studies. Lower panel: sROC bivariate MAs of sensitivity specificity pairs (hollow symbols); pooled point estimate square solid symbol (the sROC plot on the right excludes an influential outlier study). HSROC = hierarchical summary receiver operating characteristic; ATI – antibodies to infliximab

Table 18 Test accuracy parameters generated by hierarchical meta-analysis[143]

		Sensitivity	Specificity	Diagnostic OR	Likelihood ratio +	Likelihood ratio -
Excludes outlier	Point estimate *95% CI*	0.72 *0.64-0.78*	0.79 *0.64-0.89*	9.87 *4.07-23.92*	3.49 *1.85-6.60*	0.35 *0.26-0.48*
All studies	Point estimate *95% CI*	0.70 *0.55-0.82*	0.81 *0.67-0.89*	9.81 *4.09-23.54*	3.63 *2.04-6.45*	0.37 *0.24-0.58*

The implication of these test accuracy results was explored in terms of predictive values as suggested in the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy.[140] Since predictive values are influenced by prevalence of the target condition, we determined a pooled random effects estimate of prevalence (LOR) amongst the studies (34.7%; 95% CI 25.1% to 44.4%). The point estimates for positive (PPV) and negative (NPV) predictive values at this prevalence were 65% and 84% respectively. The influence of prevalence on these values is illustrated in Figure 24 across the range of prevalence of the included studies and the 95% CI around the pooled prevalence.

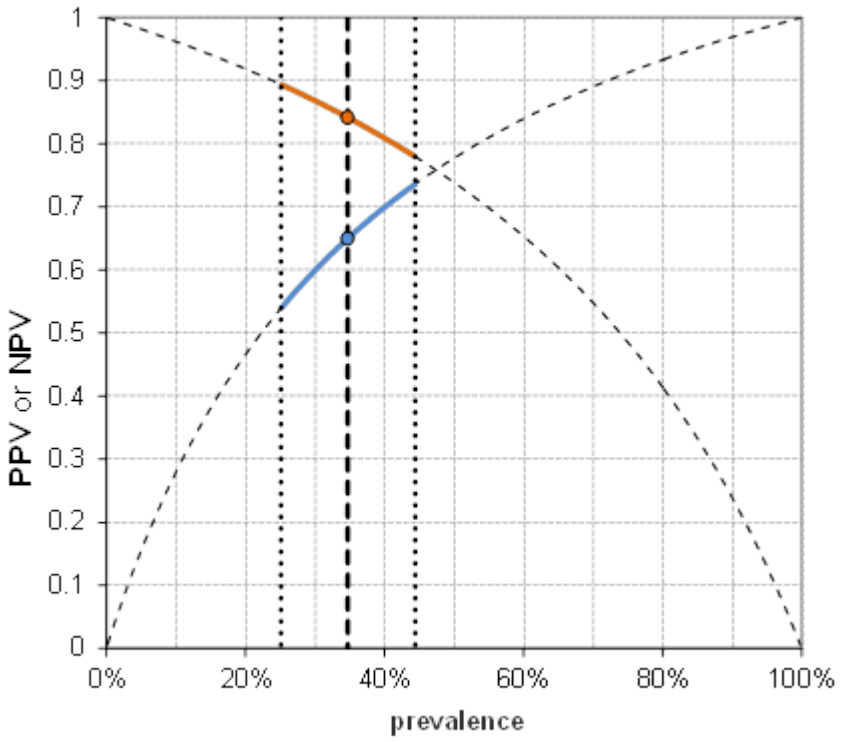


Figure 24 PPV and NPV according to prevalence of LOR at the sROC model estimate of sensitivity and specificity, as prevalence increases PPV increases and NPV decreases
Data points = PPV and NPV at sROC sensitivity and specificity and pooled prevalence. Dashed vertical lines = pooled prevalence and 95% CI. Thick curves = PPV and NPV for hierarchical model sensitivity and specificity at the pooled prevalence and 95% CI

The MA results indicate that the anti-drug antibody test has only moderate accuracy performance in predicting / detecting LOR to infliximab.

Lee et al. (2012)[62]

The authors estimated the pooled relative risk of remission in patients with a positive test for anti-drug antibodies to infliximab relative to those with a negative test for anti-drug antibodies (a RR < 1.0 indicates anti-drug antibodies are associated with lower risk of remission, consistent with the hypothesis that anti-drug antibodies reduce response to infliximab therapy). Comparative numbers of events and patients were reported. The fixed effects and random effects RR are 0.90 (95% CI: 0.79 to 1.02) and 0.96 (95% CI: 0.77 to 1.19) respectively. Statistical heterogeneity unexplained by chance was 37% (I^2 statistic). When the presence of antibodies to infliximab are considered as predictor of, or diagnostic of, a lack of remission then MA yielded low joint sensitivity specificity values of 0.42 and 0.69 respectively (Figure 25).

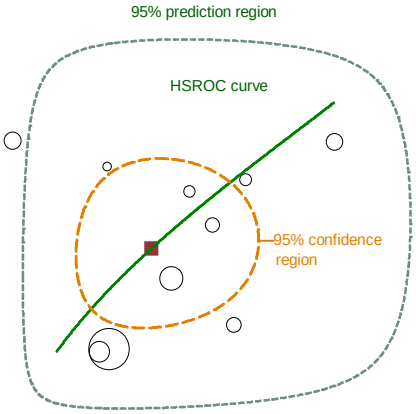


Figure 25 Meta-analysis of data based on that from Lee et al. (2012)[62]
Left panel: Fixed effects MA of the relative risk of remission (presence of antibodies to infliximab versus absence of antibodies to infliximab. Right panel: sROC bivariate MA of sensitivity specificity pairs (hollow symbols); pooled point estimate square solid symbol HSROC = hierarchical summary receiver operating characteristic.

The results indicate that presence of anti-drug antibodies does not strongly increase the risk of lack of remission and that a positive test for the presence of anti-drug antibodies has poor discriminatory power for predicting / diagnosing a lack of remission.

Lee and colleagues[62] also reported a meta-analysis examining the association between the development of anti-drug antibodies and the use of immunosuppressant therapies. Eleven studies were included, they generated a fixed effects relative risk (antibodies present with suppressants versus antibodies present with no suppressants) of 0.50 (95% CI: 0.42 to 0.59; I^2 = 43.4%) indicating a 50% reduction in risk of developing anti-drug antibodies when suppressants are administered. Fixed effects and random effects meta-analyses are illustrated in Figure 26.

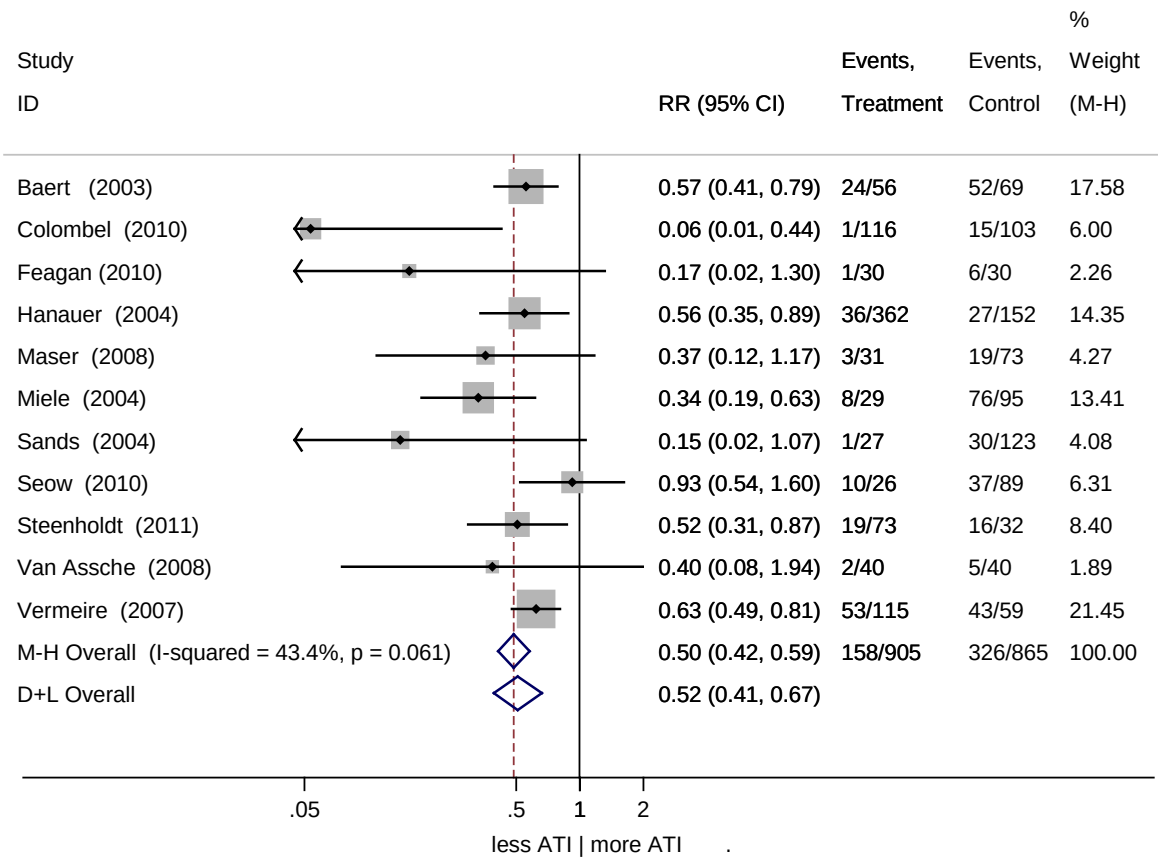


Figure 26 Relative risk of anti-drug antibodies with immuno-suppressants versus without suppressants. *Data based on Lee et al. (2012)[62] M-H Mantel-Haenszel; D+L = DerSimonian-Laird; ATI antibodies to infliximab*

Paul et al. (2014)[113]

The authors estimated the pooled ratio for the odds of lack of response in association with a negative test for adalimumab (i.e. sub-therapeutic) versus the odds of lack of response in those with a positive test for adalimumab (an OR > 1.0 indicates that sub-therapeutic adalimumab levels are associated with lack of clinical response). The comparative numbers of events and patients were not reported and it was difficult to verify the included data from the references provided. The pooled OR differed between 3 studies of adults with CD (7.5; 95% CI: 3.58 to 13.90) and 2 studies of children with CD (1.59; 95% CI: 1.00 to 2.54). The overall pooled OR was (2.60; 95% CI: 1.79 to 3.77).

The reported ORs are equivalent to the diagnostic odds ratio (DOR) of a diagnostic test in which sub-therapeutic drug levels are the test for lack of clinical response. As such they are modest relative to the DOR value of 9.6 (odds pooled sensitivity/odds pooled specificity) for studies included in the Nanda review[110] using the test for anti-bodies to infliximab as predictor of lack of response.

The authors also estimated the pooled ratio for the odds of lack of response in association with a negative test for anti-drug antibodies to adalimumab versus the odds of lack of response in those with

a positive test for anti-drug antibodies (an OR > 1.0 indicates that presence of anti-drug antibodies are associated with lack of clinical response). The reported OR of 10.15 (95% CI: 3.90 to 26.40) is equivalent to the DOR for presence of antibodies (to adalimumab) as predictor for lack of clinical response. This value is similar to the DOR for antibodies to infliximab as predictor for lack of response (using the pooled sensitivity specificity pair (0.72 and 0.79) derived from the studies in the Nanda review[110] which provides a DOR of (0.72/0.28) / (0.21/0.79) = 9.67).

3.2.5.2.4 Summary

When viewed as predictive tests for lack of response and or lack of remission, the published MAs indicate modest accuracy of tests for trough drug levels or for presence of anti-drug antibodies. Typically the predictive values indicate substantial proportions of false positive and false negative test results.

3.2.5.3 Analysis of correlation studies of anti-TNFα / anti-drug antibodies level and response

3.2.5.3.1 Aim

To pool test outcome data from correlation studies for responders and patients with loss of response as an alternative to single study data to inform the economic model.

3.2.5.3.2 Rationale

No published and tested multivariable prognostic models were found that incorporated test results with other variables to predict clinical status. The bulk of the identified studies about tests for anti-TNFα and anti-drug antibody levels were classified as correlation studies (see Table 5). These reported correlations or associations between test results and other patient dependent variables. Only one published management study (Steenholdt et al., 2014[122]) used test results to guide treatment options according to an algorithm, thus most of the evidence about tests does not directly address the clinical effectiveness decision questions. Some of the studies dichotomised test and related test results to clinical status; they can provide probabilities that a patient will return a particular type of test result and the probability that the test outcome is associated with response or lack of response; information that may be useful for economic modelling.

The decision questions identify two testing strategies, concurrent and reflex:

i) concurrent testing for both anti-TNFα levels and anti-drug antibodies levels; when tests are dichotomised using cut offs, these generate four patient categories:
   Anti-drug antibodies+/Anti-TNFα –, anti-drug antibodies +/ Anti-TNFα+, anti-drug antibodies – /Anti-TNFα –, anti-drug antibodies – /Anti-TNFα+.

ii) reflex-testing in which tests for anti-TNFα levels precede subsequent testing for anti-drug antibodies, and anti-drug antibodies tests are only done for those with sub-therapeutic levels of anti-TNFα; when tests are dichotomised using cut offs, these generate three patient categories: Anti-TNFα+, Anti-TNFα – /anti-drug antibodies+, and Anti-TNFα – /anti-drug antibodies –.

Test result probabilities by patient category can be obtained from studies that reported both drug and anti-drug antibody test results for each patient. Very few such studies were found. For reflex testing, test result probabilities for the first test may additionally be obtained from studies where test results are reported by group rather than individual; where several such studies are available the option of meta-analysis of multiple studies is available so as to gain greater power. Studies that undertook both anti-TNFα and anti-drug antibody tests but did not provide test results for each patient (but only by group) were not useful for getting estimates of concurrent testing probabilities because contingency probabilities could not be calculated (e.g. the probability that an individual with a negative drug test was either negative or positive for the anti-drug antibody test). However these may be meta-analysed so as to provide a comparison (by single test result) with the few available patient-level studies in order to gauge consistency of test results from patient-level studies with those across multiple studies.

3.2.5.3.3 Results

The studies identified as correlation (N=136) adopted several perspectives in reporting test results. Most commonly the association of test results with another variable, usually correlation of drug levels and anti-drug antibodies levels, was assessed and correlation reported as Pearson's correlation coefficient or Spearman's rank correlation coefficient. Other associations investigated were between anti-drug antibodies and or anti-TNFα levels and measures of serum CRP, or of faecal calprotectin, or estimates of clinical status. Those that dichotomised test results and related these to dichotomised clinical state (e.g. response or lack of response) were considered potentially useful for the decision questions. Data from this type of study can be represented in a 2x2 diagram like that shown below in Table 19.

Table 19 Illustration of 2x2 table data from correlation studies

	Clinical state A	Clinical state B	Total
Test positive	TP a	FP b	a + b
Test negative	FN d	TN c	d + c
Total	a + d	b + c	a + b + c + d

Those studies from which the values for a, b, c and d, could be obtained were taken further (N=31) and those (N=105) that had insufficient data were excluded (see Appendix 6). Some viewed test

results as diagnostic or predictive of the clinical state of interest and test accuracy parameters were reported (e.g. sensitivity, specificity, or ROC plots). Other studies considered the risk of a particular test result (e.g. positive) in patients with a particular clinical state (e.g. state A) and calculated a relative risk of a positive test result in state A relative to state B: ([a/a+d] / [b/b+c]), or conversely relative risk of state A in patients with positive test relative to those with a negative test: ([a/a+b] / [d/d+c]).

The 31 studies[37, 39, 46, 51, 58, 76-84, 87, 91, 93, 97-99, 101, 102, 105, 107, 109, 114, 119, 122, 125, 132, 133] taken forward for meta-analysis (Table 47; Appendix 12) were heterogeneous in terms of populations, treatments, tests used, completeness of reporting, test cut-offs used for dichotomising test results, definitions of clinical response, and the time from treatment initiation to that at which clinical status was assessed. Most studies were retrospective and used convenience sample populations where data from medical records about clinical state were available and serum samples had been collected and stored for future assay. The commonest threats to validity of study findings in this collection of studies is selection bias, lack of power and the use of subjective measures to establish clinical status.

Concurrent testing: test result probabilities

Three studies reported both drug and antibody test results for the same individuals in relation to clinical status.[98, 99, 122] These allowed calculation of the number of patients in each of the two dichotomised clinical states distributed to each of the four possible combinations of test result (i.e. [drug + | antibody +], [drug + | antibody –], [drug – antibody +], [drug – | antibody –]).[98, 99, 122] The results summarised in Table 20 to Table 22 indicate the probability of loss of response according each possible test result category for the three studies.

Table 20 Concurrent testing for responders receiving adalimumab

Imaeda 2014[99]	ADAbs +	ADAbs –	TOTAL	Population & anti-TNFα therapy; Tests
Anti-TNFα –	LOR = 8 RESP = 0	LOR = 2 RESP = 2	LOR = 10 RESP = 2	Responders on adalimumab maintenance. ELISA. Prevalence of LOR = 37.5%
Anti-TNFα +	LOR = 2 RESP = 4	LOR = 3 RESP = 19	LOR = 5 RESP = 23	
TOTAL	LOR = 10 RESP = 4	LOR = 5 RESP = 21	LOR = 15 RESP = 25	
The probability of a patient returning each of the four possible test result combinations was: ADAbs +/ Anti-TNFα – = 0.200; ADAbs +/Anti-TNFα + = 0.150; ADAbs –/Anti-TNFα – = 0.10; ADAbs –/Anti-TNFα + = 0.550. The probabilities of losing response according to category of test result were: 1.00, 0.333, 0.500 and 0.136 respectively. ADAbs – anti-drug antibodies; RESP – responders; LOR – loss of response				

Table 21 Concurrent testing for responders receiving infliximab

Imaeda 2012[98]	ADAbs +	ADAbs –	TOTAL	Population & anti-TNFα therapy; Tests
Anti-TNFα –	LOR = 9 RESP = 1	LOR = 0 RESP = 7	LOR = 9 RESP = 8	Responders on infliximab maintenance. ELISA. Prevalence of LOR = 29.3%
Anti-TNFα +	LOR = 3 RESP = 3	LOR = 5 RESP = 30	LOR = 8 RESP = 33	
TOTAL	LOR = 12 RESP = 4	LOR = 5 RESP = 37	LOR = 17 RESP = 41	
The probability of a patient returning each of the four possible test result combinations was: ADAbs +/ Anti-TNFα – = 0.172; ADAbs +/ Anti-TNFα + = 0.103; ADAbs – /Anti-TNFα – = 0.121; ADAbs – /Anti-TNFα + = 0.603. The probabilities of losing response according to category of test result were: 0.900, 0.500, 0.000 and 0.143 respectively. ADAbs – anti-drug antibodies; RESP – responders; LOR – loss of response				

Table 22 Concurrent testing for people with loss of response receiving infliximab

Steenholdt 2014[122]	ADAbs +	ADAbs –	TOTAL	Population & anti-TNFα therapy; Tests
Anti-TNFα –	NOR = 8 RESP = 6	NOR = 2 RESP = 1	NOR = 10 RESP = 7	Failure on infliximab, continued failure or gain of response at 12 weeks. RIA. Prevalence of NOR = 44.9%
Anti-TNFα +	NOR = 1 RESP = 3	NOR = 20 RESP = 28	NOR = 21 RESP = 31	
TOTAL	NOR = 9 RESP = 9	NOR = 22 RESP = 29	NOR = 31 RESP = 38	
The probability of a patient returning each of the four possible test result combinations was: ADAbs +/ Anti-TNFα – = 0.203; ADAbs +/ Anti-TNFα + = 0.058; ADAbs – /Anti-TNFα – = 0.0.043; ADAbs – /Anti-TNFα + = 0.696. The probabilities of failing to gain a response according to category of test result were: 0.571, 0.250, 0.667 and 0.417 respectively. ADAbs – anti-drug antibodies; RESP – responders; LOR – loss of response; NOR – no regain of response				

Reflex testing: test result probabilities

The test results for studies which reported both drug and antibody test results for the same individuals in relation to clinical status can be condensed to provide test results for three groups of patients: Anti-TNFα+, Anti-TNFα – /anti-drug antibodies+, and Anti-TNFα – /anti-drug antibodies –. The results summarised in Table 23 to Table 25 indicate the probability of loss of response according to each possible test result category.

Table 23 Reflex testing for responders receiving adalimumab

Imaeda 2014[99]	ADAbs +	ADAbs –	TOTAL	Population & anti-TNFα therapy; Tests
Anti-TNFα –	LOR = 8 RESP = 0	LOR = 2 RESP = 2	LOR = 10 RESP = 2	Responders on adalimumab maintenance. ELISA. Prevalence of LOR = 37.5%
Anti-TNFα +	LOR = 5 RESP = 23		LOR = 5 RESP = 23	
TOTAL			LOR = 15 RESP = 25	
The probability of a patient returning each of the three possible test result combinations was: Anti-TNFα+, 0.700; Anti-TNFα – /ADAbs+, 0.200; Anti-TNFα – /ADAbs –, 0.100. The probabilities of losing response according to category of test result were: 0.179, 1.00 and 0.500 respectively. ADAbs – anti-drug antibodies; RESP – responders; LOR – loss of response				

Table 24 Reflex testing for responders receiving infliximab

Imaeda 2012[98]	ADAbs +	ADAbs –	TOTAL	Population & anti-TNFα therapy; Tests
Anti-TNFα –	LOR = 9 RESP = 1	LOR = 0 RESP = 7	LOR = 9 RESP = 8	Responders on infliximab maintenance. ELISA. Prevalence of LOR = 29.3%
Anti-TNFα +	LOR = 8 RESP = 33		LOR = 8 RESP = 33	
TOTAL			LOR = 17 RESP = 41	
The probability of a patient returning each of the three possible test result combinations was: Anti-TNFα+, 0.707; Anti-TNFα – /ADAbs+, 0.172; Anti-TNFα – /ADAbs –, 0121. The probabilities of losing response according to category of test result were: 0.195, 0.00 and 0.900 respectively. ADAbs – anti-drug antibodies; RESP – responders; LOR – loss of response				

Table 25 Reflex testing for people with loss of response receiving infliximab

Steenholdt 2014[122]	ADAbs +	ADAbs –	TOTAL	Population & anti-TNFα therapy; Tests
Anti-TNFα –	NOR = 8 RESP = 6	NOR = 2 RESP = 1	NOR = 10 RESP = 7	Failure on infliximab, continued failure or gain of response at 12 weeks. RIA. Prevalence of NOR = 44.9%
Anti-TNFα +	RESP = 31 NOR = 21		NOR = 21 RESP = 31	
TOTAL			NOR = 31 RESP = 38	
The probability of a patient returning each of the three possible test result combinations was: Anti-TNFα+, 0.754; Anti-TNFα – /ADAbs +, 0.203; Anti-TNFα – /ADAbs –, 0.044. The probabilities of not gaining response according to category of test result were: 0.404, 0.667 and 0.571 respectively. ADAbs – anti-drug antibodies; RESP – responders; NOR – no regain of response				

Meta-analytic test result probabilities: trough infliximab levels

Meta-analysis results for single test studies using trough infliximab levels as a test for LOR in responders and failure to regain response in patients with LOR are summarised in Appendix 12.1. For responders the probability of returning a positive test result (i.e. infliximab undetectable) was 0.367 at the pooled prevalence (the range based on 95% CI for prevalence was 0.340 to 0.385; this does not take into account uncertainty in the summary point estimate); for a negative test result the probability was 0.632 (the range based on 95% CI for prevalence was 0.615 to 0.659; this does not take into account uncertainty in the summary point estimate). The probability of a positive test reduced to 0.271

when prevalence was set to that of the single available patient level study of Imaeda et al. (2014)[99] which returned a similar positive test probability of 0.293 (95% CI: 0.181 to 0.427).

Only two studies were available for patients with loss of response so that a meaningful pooled estimate could not be undertaken

Meta-analytic test result probabilities: antibodies to infliximab

Meta-analysis results for single test studies using antibodies to infliximab as a test for LOR in responders and failure to regain response in patients with LOR are summarised in Appendix 12.2. The probability of returning a positive test result (i.e. anti-IFX antibodies undetectable) was 0.345 at the pooled prevalence (the range based 95% CI for prevalence was 0.324 to 0.365); for a negative test result the probability was 0.655 (the range based on 95% CI for prevalence was 0.635 to 0.686; this does not take into account uncertainty in the summary point estimate). The probability of a positive test reduced to 0.274 when prevalence was set to that of the single available patient level study of Imaeda et al. (2014)[99] which returned a similar positive test probability of 0.276 (95% CI: 0.167 to 0.409).

Seven heterogeneous studies[46, 58, 76, 78, 82, 122, 125] were available for patients with loss of response (Appendix 12 Table 47). The probability of returning a positive test result (i.e. anti-infliximab antibodies present) was 0.387 at the pooled prevalence (the range based on 95% CI for prevalence was 0.331 to 0.442; this does not take into account uncertainty in the summary point estimate); for a negative test result the probability was 0.613 (the range based on 95% CI for prevalence was 0.558 to 0.669; this does not take into account uncertainty in the summary point estimate). The probability of a positive test increased to 0.425 when prevalence was set to that of the single available patient level study of Steenholdt et al. (2014),[122] which returned a much lower positive test probability of 0.261 (95% CI: 0.163 to 0.381).

Meta-analytic test result probabilities: trough adalimumab levels

Meta-analysis results for single test studies using trough adalimumab levels as a test for LOR in responders and failure to regain response in patients with LOR are summarised in Appendix 12.3. The probability of returning a positive test result (i.e. adalimumab undetectable) was 0.444 at the pooled prevalence (the range based on 95% CI for prevalence was 0.389 to 0.499; this does not take into account uncertainty in the summary point estimate). The probability of a positive test reduced to 0.390 when prevalence was set to that of the single available patient level study of Imaeda et al. (2012)[98] which returned a lower positive test probability of 0.300 (95% CI: 0.166 to 0.465).

A single study related trough adalimumab levels to clinical outcome for patients with loss of response. No patient level dual test studies were available for a comparison of test probabilities.

Meta-analytic test result probabilities: anti-adalimumab antibody levels

Meta-analysis results for single test studies using trough anti-adalimumab antibody levels as a test for LOR in responders are summarised in Appendix 12.4.

The probability of returning a positive test result (i.e. anti-adalimumab antibodies present) was 0.253 at the pooled prevalence. The probability of a positive test reduced to 0.230 when prevalence was set to that of the single available patient level study of Imaeda et al. (2012)[98] which returned a higher positive test probability of 0.350 (95% CI: 0.206 to 0.517).

3.2.5.3.4 Summary

Available evidence

Only three studies were found that reported the results of both drug and anti-drug antibody tests for individual patients (one for infliximab treated responders, one for infliximab treated patients with loss of response, and one for adalimumab treated responders). These studies allowed estimation of the proportion of patients that would enter each of the treatment categories following from concurrent or reflex testing strategies.

Representativeness of available evidence

Since only a single patient level study was available for each of the different CD patient populations the test results from these studies were compared with test results from the meta-analysis of multiple single test studies. In view of the considerable uncertainties, due in part to the small number of studies and their small size, the meta-analysis test results were sufficiently similar to those of the three patient level studies to conclude that the latter were reasonably representative for the patient populations of interest.

Accuracy of tests as predictors of clinical condition

The test accuracy of drug level tests and anti-drug antibodies tests as predictors of clinical status was moderate (Appendix 12). Positive and negative predictive values across clinical prevalence ranges indicated that 20% to 30% of positive and negative test results were incorrect at plausible prevalence settings for clinical status (Appendix 12.5).

3.2.5.3.5 Evidence taken forward to the economic evaluation

The only correlation studies that provided input for economic evaluation were the concurrent testing study by Imaeda et al. (2012)[98] of patients treated with infliximab and that of Steenholdt et al. (2014)[122] of patients with loss of response to maintenance infliximab. Because the Steenholdt et al. (2014)[122] study coupled testing results with prospective implementation of a treatment algorithm it was used in the base case economic analysis. Data from the Imaeda study[98] was used in a sensitivity analysis in the cost effectiveness comparison of testing strategies versus standard care. The reason for the lack of usefulness of most of the correlation studies was that very few reported extractable data for concurrent or reflex testing.

3.3 Summary of clinical effectiveness findings

Assays based on different principles have been developed to measure anti-TNFα agents and antibodies to anti-TNFs in blood samples. There is little consensus about the most appropriate assay to use in clinical practice and no gold standard is established against which assay performance can be assessed. Studies have examined the predictive ability of tests to discriminate clinical condition of IBD patients; meta-analysis of such studies has indicated that the tests have only moderate predictive utility. Irrespective of imperfect test accuracy, when tests are used in tandem with an appropriate treatment algorithm they may deliver equal or better patient outcomes than a standard care strategy undertaken without testing. No RCT was found that tested this possibility for CD patients responding to anti-TNFα agents. The TAXIT trial described outcomes when a test-algorithm strategy based on trough infliximab levels was implemented for IBD patients responding to infliximab, but a standard care comparator population was not available because all randomised TAXIT patients received test-directed optimisation of infliximab dosing. A single retrospective case series of IBD patients responding to infliximab reported better retention in infliximab treatment for those whose dose changes were based on prospective testing compared to those whose dose was not based on prospective testing. However this study design was at appreciable risk of bias particularly with respect to selection bias. One randomised study compared a test-algorithm strategy versus an intensified dose strategy in CD patients who had lost response to infliximab. No difference in clinical outcome was observed but cost savings were reported for the test-algorithm strategy. The study was of short duration (data at 12 and 20 weeks only), was small (69 patients), and about half of the intervention patients received treatment that did not conform to the algorithm and a substantial proportion received unspecified therapy decided according to clinical judgement. The generalisability of findings and the longer term implications of the study are difficult to gauge.

The available evidence provides a limited platform for deciding if testing for anti-TNFα agents and or antibodies to anti-TNFα drugs provides a clinical advantage over standard anti-TNFα strategies used

for responders or for patients with loss of response. Ongoing trials may deliver more relevant data to inform a decision.

The main points of the clinical effectiveness can be summarised:

- ELISA assays are susceptible to interference to a greater extent than other assays such as RIA and HMSA
- There is uncertainty about which assay is optimal for drug monitoring as well as when and how often assessments should take place and whether levels of drug, anti-drug antibodies or both should be determined
- The clinical significance of measuring accurate and very low levels of drug / anti-drug antibodies is not known
- Transient anti-drug antibodies might be the result of drug masking anti-drug antibodies from detection by forming complexes particularly after dose intensification
- The evidence on concordance between the three intervention assays is contradictory. Overall there is insufficient evidence to make claims about the comparative performance between the three intervention assays or in relationship to other assays for a linked evidence approach
- The available evidence, although scarce, showed varying degree of disagreement between assays
- Studies determined their own cut-off values which vary greatly between studies. This reflects the fact that cut-offs are study specific and not readily generalisable
- Two RCTs with evidence on the clinical utility of testing and test informed algorithm that are sufficiently prescriptive were identified – one for patients with LOR and one for responders
- The algorithms in the RCTs are slightly different to the ones presented to us in the NICE scope for this work, reflecting the influence of the variation in clinical judgment
- The RCTs recruited different patients groups (LOR / responders), used different tests and different testing strategies addressing different aspects of the decision questions (concurrent testing for patients with LOR and reflex testing (dose optimisation) for responder)).
- Drug monitoring might be cost saving without loss of effectiveness mainly due to reduced administration of infliximab in patients who do not require infliximab (drug positive and anti-drug antibodies negative) according to one RCT
- Drug optimisation during induction phase in responders might lead to an increase in clinical remission and savings in drug costs according to one RCT
- Trough level based dosing during maintenance may increase the probability of remaining on infliximab treatment according to one observational study
- Problems with the RCTs included:
    - Mixed patient populations

    - Short follow up
    - Small patient numbers
    - No evidence on adalimumab
    - Timing of testing did not correspond with decision questions
- Meta-analyses of correlation studies showed that the diagnostic performance of the assays is only moderate when measured against clinical assessment
- Single patient level study outcomes in correlation studies were sufficiently similar to meta-analyses of multiple single test studies to use outcomes as estimates of proportions of people entering each treatment category for concurrent and reflex testing

The clinical effectiveness review provided information that was useful for the modelling in the following ways: The three management studies[72, 122, 127] informed both the structure of the economic model and provided some of the required data to populate it. The model structure was also informed by clinical expert advice about the relevant patient treatment pathways that addressed the decision problem. This extended the model to a time horizon well beyond the data from the two RCT management studies and necessitated considerable data input from studies not included in the clinical effectiveness review. A single correlation study delivered some input for the economic evaluation; however the usefulness of the correlation studies for economic analysis was limited because concurrent or reflex testing results were rarely reported (most studies only correlated clinical status with either test results for anti-TNFα or results for antibodies to anti-TNF). Although the correlation studies provide some indication of the test accuracy of currently used tests this is irrelevant for the economic decision because any deficiency in test accuracy is subsumed within the combined test + algorithm intervention.

4 COST-EFFECTIVENESS REVIEW AND HEALTH ECONOMIC MODELLING

4.1 Systematic review of existing cost-effectiveness evidence

This chapter will explore and review all published studies on the cost-effectiveness of LISA-TRACKER ELISA kits, TNFα-Blocker ELISA kits, and Promonitor ELISA kits for measuring levels of TNFα inhibitors and of anti-drug antibodies in detail.

4.1.1 Aim

To review all cost-effectiveness studies including any existing models and to identify any suitable data such as resource use, costs, utilities and transition probabilities to help inform our economic model for the evaluation of the cost-effectiveness of LISA-TRACKER ELISA kits, TNFα-Blocker ELISA kits, and Promonitor ELISA kits for measuring levels of TNFα inhibitors and of anti-drug antibodies in detail.

4.1.2 Methods

4.1.2.1 Search strategy

A comprehensive search of the literature for published economic evaluations (including any existing models), cost studies and quality of life (utility) studies was performed. The systematic search included searching the following electronic databases during December 2014 (from the 12th to 17th December):

- MEDLINE (Ovid) - 1946 to Week 3 November 2014
- MEDLINE In-Process Citations and Daily Update (Ovid) - December 11, 2014
- EMBASE (Ovid) - 1947 to 15 December 2014
- NHS Economic Evaluation Database (NHS EED) (Cochrane Library)
- Science Citation Index (Web of Knowledge) - 1970 – present
- Cost-effectiveness analysis (CEA) registry
- EconPapers (RePEc)
- School of Health and Related Research Health Utilities Database (ScHARRHUD)

The search included terms for CD, anti-TNFα drugs and the different assay kits, combined with economic and quality of life (QoL) terms. The search was limited to studies published in the English language. The search strategy developed was based on the clinical effectiveness review with input from a health economist. Details of the full search strategies are provided in Appendix 3.

4.1.2.2 Inclusion criteria

Only studies meeting the following inclusion criteria were included in the review:

- Study type: Fully published economic evaluations (including economic models)
- Population: People with Crohn's disease
- Intervention: Anti-TNFα drugs (adalimumab and infliximab) and antibody drug testing (LISA-TRACKER ELISA kits, TNFα-Blocker ELISA kits, and Promonitor ELISA kits) for any dosage or treatment regimen
- Comparator: Standard care treatment - anti-TNFα drugs (adalimumab and infliximab) for any dosage or treatment regimen
- Outcomes: Cost-effectiveness or cost-utility studies reporting outcomes as clinical effectiveness measures or utility measures (utility, EQ-5D or SF-6D score or QALYs).

4.1.2.3 Exclusion criteria

Studies meeting the following exclusion criteria were excluded from the review:

- Non-English-language publications
- Studies in the health areas where these anti-TNFα drugs have also been used such as ulcerative colitis, rheumatoid *arthritis, psoriasis, and* tuberculosis

4.1.2.4 Assessment of eligibility and data extraction

All retrieved records (citations and abstracts) were collected in a specialist database (Endnote) and duplicate records were identified and removed. Two reviewers independently reviewed titles and abstracts to identify potentially relevant papers for inclusion. Any discrepancies were resolved by discussion. See Appendix 13 for the table of full text studies excluded with reason.

Data extraction was carried out in two stages by one reviewer using standardised data extraction sheets (see Appendix 14) and was then checked by a second reviewer. Stage one considered all eligible studies (fully published economic evaluations including any economic models) and stage two considered studies assessed for usefulness for populating the economic model. Data extracted during stage one included the following:

- study details: author names, source of publication, language and publication type
- baseline characteristics: population, intervention, comparators, outcomes, and type of economic evaluation
- methods: target population and subgroups, setting and location, study perspective, time horizon, discount rate, measurement of effectiveness, measurement and valuation preference based outcomes, resource use and costs, currency, price date and conversion, model type, assumptions and analytical methods

- results: study parameters, incremental costs and outcomes and characterising uncertainty
- discussion: study findings, limitations, generalisability and conclusions
- other: sources of funding, conflicts of interest and comments

4.1.2.5 Quality assessment

The quality of full economic evaluation studies that were identified were assessed using the Consolidated health economic evaluation reporting standards (CHEERS) checklist (see Appendix 15) by one reviewer and cross-checked by a second reviewer. The CHEERS checklist comprises six dimensions which include title and abstract, introduction, methods, results, discussion and other. Under these dimensions, a series of questions check whether the criteria have been clearly reported. Any studies containing an economic model were further assessed using the framework for the quality assessment of decision analytic modelling by Philips et al (2004)[144] (see Appendix 15). The Philips' checklist contains two main dimensions, structure of the model and data used to parameterize the model. Under these dimensions several questions assess whether the criteria has been clearly reported.

4.1.2.6 Data synthesis

Information extracted from the included studies were summarised and tabulated. Findings from individual studies were compared narratively.

4.1.3 Results

4.1.3.1 Search results for Objective D

The literature search identified 2,466 records through electronic database searches and other sources. After removing duplicates, 1,527 records were screened for inclusion. On the basis of title and abstract sift only, 1,518 records were excluded. The remaining nine records were included for full-text screening. A further five articles[5, 25, 145-147] were excluded at the full-text stage, as these studies did not contain any assay kits for measuring levels of TNFα inhibitors and anti-drug antibodies. The literature search identified four studies [72, 122, 123, 148] which included cost-effectiveness of different assay kits for measuring levels of TNFα inhibitors and of anti-drug antibodies (see Figure 27 for more detail)

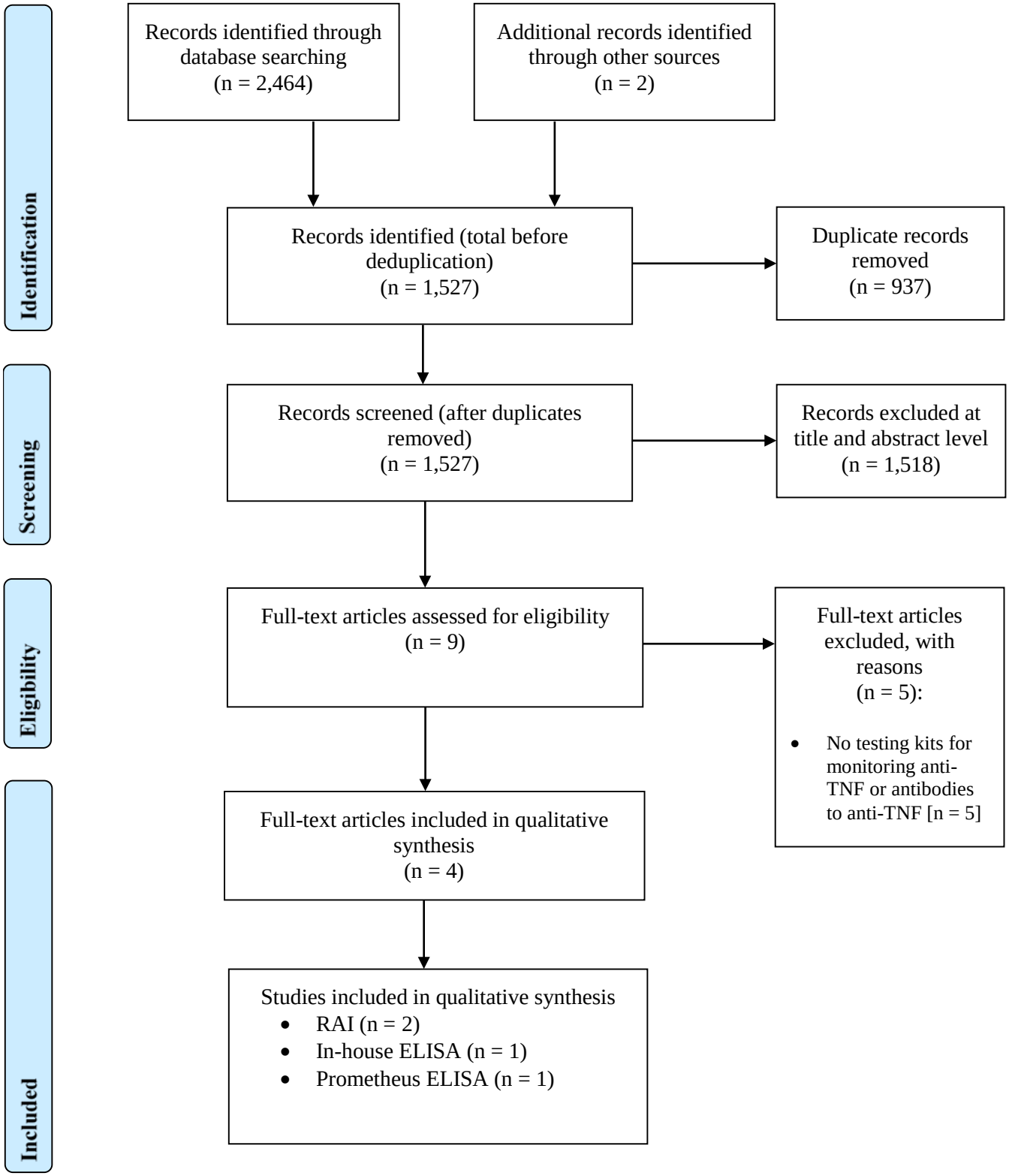


Figure 27 PRISMA diagram of cost-effectiveness studies

In developing the economic model, we have consulted the previous Technology Appraisal Guideline and Health Technology Assessment (HTA) report by Dretzke et al. (2011).[5] Even though this former work did not include any assay kits for measuring levels of TNFα inhibitors and of anti-drug antibodies. The aim of this Diagnostic Assessment Review as specified by NICE was to build upon this previous work. The next section contains a summary of this previous HTA report and then the results of the cost-effectiveness review including quality assessment will be outlined.

4.1.3.2 Summary of the Health Technology Assessment report by Dretzke et al. (2011)[5]

The main aim of this HTA report was to assess the cost-effectiveness of anti-TNFs in the management of adult patients with moderate-to-severe Crohn's disease in the UK National Health Service (NHS). The authors described induction therapy as the use of anti-TNFα therapy with the aim to achieve remission (a repeated re-induction treatment was considered, instead of a one-off induction therapy) and maintenance therapy as the use of anti-TNFα therapy to maintain remission in patients who have responded (and continue to respond) to anti-TNFα therapy when in relapse. Response by the authors was defined as remission within eight weeks.

The authors developed a Markov model from an NHS and Personal Social Services (PSS) perspective to estimate the incremental cost per quality-adjusted life year (QALY) gained for both adalimumab and infliximab (anti-TNFα therapy) compared with standard care. Mortality was not included in the model as the authors found no differential in the mortality rates which were reported in the clinical trials reviewed and therefore felt that a lifetime horizon would not improve the precision of the cost-effectiveness estimate. Instead, the time horizon for the model was 1 year and the cycle duration was 4 weeks. The model for both induction and maintenance therapy started with a cohort of patients in the standard care refractory relapse health state. The model had four main health states and at any time and on any given treatment, a patient was in remission, in relapse, undergoing surgery or in post-surgery remission.

Transition probabilities for the standard care health states were based on the Silverstein et al. (1999).[149] Transition probabilities for both the induction and maintenance model were assigned a treatment effect by using relapse to remission probabilities from RCT evidence; however, for the maintenance model there was a lower remission to relapse rate.

The majority of utility values for the model were based on study which used the time-trade off measure to estimate the health-related quality of life in Crohn's disease by Gregor et al. (1997)[19]. For surgery, the utility value was not available in published literature, therefore it was assumed that the

average utility value for surgery would be equivalent to EQ-5D health state 22222 with a utility weight of 0.516.

The direct costs to the NHS were sum of the anti-TNFα costs and type-specific health state costs. For anti-TNFα therapy for both induction and maintenance, the costs were derived from the BNF (BNF 2007; 2008) and administration costs were also included for adalimumab. Type specific health state costs included: costs for surgery which were modelled as the cost of inpatient IBD interventions and post-surgery remission costs were based on outpatient surgical gastrointestinal follow-up. Moderate and severe relapse costs were modelled as the cost of IBD outpatient major and intermediate interventions. Relapse costs were based on a gastrointestinal admission to hospital. Remission costs were modelled using literature. The majority of health state unit costs were obtained from the NHS reference cost database (NHS reference costs 2005-06).

Incremental cost-effectiveness ratios and cost-effectiveness acceptability curves were presented. One-way sensitivity analyses and probabilistic sensitivity analyses using 10,000 simulations were conducted to characterise uncertainty in the model.

For induction therapy for severe Crohn's disease, both adalimumab and infliximab dominated standard care (i.e. cheaper and more effective). For maintenance therapy for severe Crohn's disease, neither drug was cost-effective (well above NICE thresholds). For moderate Crohn's disease, for maintenance therapy for both drugs and induction therapy for infliximab, these were not cost-effective (well above NICE thresholds); however, for induction therapy for adalimumab dominated standard care.

Sensitivity analysis found that patients who had severe disease, infliximab induction treatment was found to be cost-effective relative to maintenance treatment and standard care in over 99% of cases at all points up to £100,000 per QALY. Likewise, adalimumab induction treatment was found to be cost-effective relative to maintenance treatment and standard care for thresholds up to £100,000 per QALY.

The key limitations of this model was a short time frame (one-year time horizon); the exclusion of death from the model; no randomised controlled data available for maintenance therapy; and the use of Silverstein et al. (1999)[149] data for transition probabilities which inherently had its own problem, that is, surgery rates were higher and relapse rates much lower than in routine practice.

4.1.3.3 Overview of included studies

The literature search identified four studies[72, 122, 123, 148] which met our inclusion criteria (studies looking at the cost-effectiveness of different assay kits for measuring levels of TNFα inhibitors and of anti-drug antibodies) and were reviewed. Below we present an overview of the included studies by population (responders and loss of response) of interest.

4.1.3.3.1 Vande Casteele et al. (2015)[72]

Vande Casteele and colleagues aimed to determine whether concentration-based infliximab dosing was more cost-effective than clinically-based infliximab dosing. These authors conducted a randomised-controlled trial and assigned people with moderate-to-severe Crohn's disease or ulcerative colitis to receive concentration-based or clinically-based infliximab dosing. Included patients where those that were treated with maintenance infliximab therapy for at least 14 weeks and who had a stable clinical response. These authors defined clinical response as being 'symptom-free (full responder) or having clinical improvement with an obvious decrease of disease activity but with clinical symptoms still present (partial responder)'[72] pg6. People eligible for the study were dose optimised until infliximab trough concentrations between 3-7μg/mL were reached. At the assessment of each trough concentration using an in-house developed ELISA, the dosing regimen was changed to reflect the proposed treatment algorithm, until people had trough concentration between 3-7μg/ml. Briefly, according to infliximab trough concentration, people received an increase dose of infliximab treatment, no dose adaptation or a decrease in infliximab treatment. The study was prospective, and was undertaken at a tertiary referral centre in Belgium. The study was conducted from the perspective of the third-party payer and the time horizon was one year. The EuroQol five-dimensions (EQ-5D) was used to calculate quality-adjusted life years (QALYs), and any differences in baseline utility scores were adjusted for by the use of a multiple regression approach. Resource use and costs were not reported in detail, apart from the drug costs per patient per year. All costs were expressed in Euros in 2012 prices. The base case results were expressed as an incremental cost-effectiveness ratio (ICER) based on the outcome of cost per quality-adjusted life-years (cost per QALY) gained. Uncertainty in incremental QALYs and costs were determined by non-parametric bootstrapping consisting of 1,000 iterations and plotted onto a cost-effectiveness plane. The base-case results demonstrated that concentration-based dosing was slightly less effective (0.8227 versus 0.8421) and less costly (€20,700 versus €21,000) than clinically-based dosing, but overall differences were small.

4.1.3.3.2 Steenholdt et al. (2014)[122]

Steenholdt and colleagues assessed the cost-effectiveness of receiving treatment based on serum concentrations of infliximab and infliximab antibodies at the time of infliximab treatment failure in accordance with the algorithm (for further details on the algorithm, see section 3.2.3) compared with

receiving infliximab at an increased dose frequency of 5mg/kg every four weeks. The study included patients with failure to infliximab treatment while on maintenance treatment. Failure to infliximab treatment was defined in the study as recurrence of active disease with a CDAI ≥220 and/or a minimum of one draining fistula. Serum infliximab and infliximab antibodies were analysed using radioimmunoassay. Samples were stored and further analysed using ELISA and homogenous mobility shift assay (HMSA) after study completion. The study was a randomised controlled single blind trial set in six Danish hospitals. Study perspective was not clearly stated. Cost-effectiveness was assessed at 12 weeks with visits scheduled at 0, 4, 8 and 12 weeks. Effectiveness was based on clinical response rates – that is, regaining response or continuing to lose of response to infliximab therapy. Resource use and costs were based on infliximab doses and all inpatient and outpatient contacts in hospitals which also included diagnostic and treatment procedures which were recorded in the National Patient Registry database. Costs were reported in Danish Kroner and converted to Euros in 2012 prices. The base-case results were expressed as cost per intention-to-treat and cost per-protocol population. Costs were compared using arithmetic means and were assessed by non-parametric bootstrapping. One-way sensitivity analyses of key primary and secondary endpoints were conducted. The base-case results showed that costs were significantly lower in the algorithm group than in the infliximab intensification group in both the intention-to-treat population and the per-protocol population.

4.1.3.3.3 Steenholdt et al. (2015)[123]

In follow-up to their study published in 2014,[122] Steenholdt and colleagues extended the time horizon to one year to assess the long-term costs and clinical outcomes of treatment of Crohn's disease in people with loss of response to infliximab maintenance therapy using with a proposed algorithm compared with intensified infliximab treatment. Serum infliximab and infliximab antibodies were analysed using radioimmunoassay, and were further analysed using ELISA and homogenous mobility shift assay (HMSA) after study completion. Infliximab levels were classified as therapeutic and sub-therapeutic measured at ≥(0.5µg/ml) and (<0.5 µg/mL), respectively. Infliximab antibodies were classified as detectable or undetectable. Costs were assessed at the 20-week scheduled trial visit and again at one year. Clinical outcomes were assessed after the 20 weeks. Costs were reported in Danish Kroner and converted in to US dollars in 2012 prices. The base-case results were expressed as cost per intention-to-treat, cost per-protocol population, cost per-protocol completion at end of trial week 12 and cost per-protocol completion at end of follow-up week 20. Sensitivity analyses on inclusion of estimated costs for administering biologic agents, use of actual infliximab dosing and price reduction in 3.5 and 7% on biologic agents were conducted to determine the robustness of the base-case results. The study found that the algorithm group had significantly lower costs than the infliximab intensification group at the 20 week follow-up and this was maintained throughout the one year. Base-

case results in terms of intention-to-treat for people randomised to the algorithm group was approximately US$11,900 versus US$22,100 at the 20-week and one-year follow-up, respectively. Results at one-year follow-up for people randomised to the infliximab intensification group was US$17,200 versus US$29,100, respectively. Results in terms of per protocol, those randomised to the algorithm and the infliximab groups, at the 20-week follow-up was approximately US$8,700 and US$17,200, respectively. Whilst at the one-year follow-up, was approximately US$15,700 and US$29,100 in the algorithm and intensification groups, respectively. Results from the sensitivity analyses showed similar findings to the base-case results.

4.1.3.3.4 Velayos et al. (2013)[148]

Velayos and colleagues used a decision analytical model to assess the cost-effectiveness of a testing-based strategy with an empiric dose escalation strategy for patients with moderate-to-severe Crohn's disease who become unresponsive to therapy with infliximab. These authors used the algorithm proposed by Afif et al (2010)[55] to form the basis of the testing-based strategy, whilst the empiric dose escalation strategy was informed by the consensus statement from the World Congress of Gastroenterology (Velayos et al., 2013).[148] The study was conducted from the perspective of the third party payer and a time horizon of one year with a four-week cycle length. Outcomes were reported as quality-adjusted life years. Quality-adjusted life years gained were derived based on utility values obtained from the study undertaken by Gregor et al. (1997).[19] Briefly, utility scores were obtained using various elicitation methods (standard gamble, time trade-off, visual analogue scale) on 180 individuals with Crohn's disease. Gregor and colleagues suggested that the standard gamble technique reflected the true value for health states related to people with Crohn's disease. Resource use and costs included the cost of interventions – infliximab, adalimumab, certolizumab, natalizumab, and surgery; and the cost of diagnostics: anti-infliximab antibody/serum infliximab measurement, CT enterography and colonoscopy. Costs were expressed in US dollars but the price year was not reported. The base-case results were expressed as an incremental cost-effectiveness ratio (ICER) based on the outcome of cost per QALY gained. Extensive one-way sensitivity analyses were conducted and populated with data to run the model probabilistically to represent the uncertainty in key model input parameters. The base-case results demonstrated that that the testing strategy was cheaper and marginally more effective, thus dominating the empiric strategy. Results from the sensitivity analyses showed that empiric strategy was less expensive when the cost of surgery was 5-fold more than in the base-case. Additionally, reducing the utility value for the health state of the 'mild/minimal inflammation with symptoms' from 0.80 to 0.70 resulted in marginally greater QALYs in the empiric group compared with the testing-based group. Furthermore, increasing the cost for testing to 25-fold, resulted in the testing-based strategy to be more expensive than empiric strategy. Results from the probabilistic sensitivity analysis showed that testing-based strategy is approximately

69% probability of being cost-effective compared to empiric dose-escalation at willingness-to-pay of US$50,000 per QALY.

4.1.3.4 Comparison of the included studies

All four studies included in this review have been summarised in Table 26. Three studies were based on randomised clinical trials[72, 122, 123] and only one study[148] presented an economic model. From the clinical trials, two studies[122, 123] were conducted in Denmark and one study[72] was conducted in Belgium. All four studies[72, 122, 123, 148] conducted cost-effectiveness analyses: Vande Casteele et al. (2015)[72] compared concentration-based with clinician-based dosing; Steenholdt et al. (2014)[122] and Steenholdt et al. (2015)[123] compared infliximab treatment failure using a treatment algorithm compared with infliximab dose increasing; whereas Velayos et al. (2013)[148] compared a testing-based strategy with an empiric dose escalation strategy. All studies[72, 122, 123, 148] clearly stated the type of assay used to analyse serum levels and antibodies to anti-TNFs. Two studies[122, 123] used radioimmunoassay in the base-case, one study[72] used an in-house developed assay, and the remaining study[148] used a Prometheus ELISA.

The patient populations for three studies[72, 122, 123] included eligible moderate-to-severe Crohn's patients, whereas the study by Vande Casteele et al. (2015)[72] included ulcerative colitis patients. The study perspective was not reported in two studies;[122, 123] whereas the other two studies[72, 148] conducted the analysis from a third-party payer perspective. The time horizon varied from 12-weeks to one year between studies. Steenholdt et al. (2014)[122] have undertaken their analysis based on a 12-week horizon, whilst the other three studies[72, 123, 148] used a one-year time horizon to estimate the cost-effectiveness of the different strategies.

For two studies,[72, 148] outcomes were reported as cost per quality-adjusted life-years gained. Vande Casteele and colleagues used the EQ-5D measure to estimate QALYs; whereas Velayos and colleagues did not explicitly report how the QALYs were estimated, except that they were obtained from a secondary source (Gregor et al., 1997).[19] The two studies by Steenholdt and colleagues reported outcomes in terms of cost per intention-to-treat and cost per-protocol.

Three studies[122, 123, 148] provided quite a comprehensive breakdown of resource use and costs; whereas the study by Vande Casteele and colleagues have not elaborated on resource use, apart from the drug costs. Three studies[72, 122, 123] reported costs in 2012 prices, whereas Velayos et al. (2013)[148] did not report the price year explicitly, but we assumed that costs are most likely in 2012 prices as the study was published in 2013.

No studies conducted discounting for either the costs or benefits as the time horizon for these studies were one year or less.

The results and conclusions reported differed between studies, Vande Casteele and colleagues demonstrated that concentration-based dosing was slightly less effective and less costly than clinically-based dosing, but overall differences were small; whereas Steenholdt and colleagues showed that the intervention based on the algorithm achieved similar clinical and life quality outcomes to dose intensification, but at a lower cost at 12 weeks. These results were maintained at both 20 weeks and at one-year[123] Velayos and colleagues showed that the testing strategy was cheaper and more effective than the empiric strategy.

All four studies[72, 122, 123, 148] conducted sensitivity analyses to deal with uncertainty around key parameters. The sensitivity analyses ranged from the most simplistic one-way sensitivity analyses[122, 123] to the more sophisticated probabilistic analyses.[148]

Table 26 Summary characteristics of the economic studies comparing ELISA kits

Study ID (First author, year, and country)	Aim of the study	Study characteristics (study design, perspective, setting	Intervention	Kits used to analyse serum levels and antibodies to anti-TNFs	Outcome measure(s)	Model type	Health states	Utility values	Results (base case and sensitivity analysis)
Responder									
Vande Casteele, 2015, Belgium[72]	To determine whether continued concentration-based dosing is superior to clinically-based dosing of infliximab for maintaining remission in patients with moderate to severe CD and UC	Randomised-controlled trial with a cost-effectiveness analysis, third-party payer, tertiary referral centre	Concentration-based dosing	In-house developed ELISA	Cost per QALY	Not applicable	Not applicable		Clinically-based dosing was the more cost-effective strategy with an ICER of €15,525 per QALY. Results from the PSA showed that 58.4% of simulations were in quadrant three where concentration-based dosing was less costly and less effective
Loss of response									
Steenholdt, 2014, Denmark[122]	To determine whether individualised therapy is more cost-effective than dose intensification in patients	Randomised controlled trial with a cost-effectiveness analysis, perspective not reported, six Danish hospitals	Individualised therapy based on an algorithm based on results of concurrent testing for serum levels	Radioimmunoassay	Cost per intention-to-treat and cost per-protocol population	Not applicable	Not applicable	Not applicable	Costs lower in the algorithm group compared to infliximab intensification group in both the ITT

Study ID (First author, year, and country)	Aim of the study	Study characteristics (study design, perspective, setting	Intervention	Kits used to analyse serum levels and antibodies to anti-TNFs	Outcome measure(s)	Model type	Health states	Utility values	Results (base case and sensitivity analysis)
	with CD who lose response to anti-TNFα treatment		and antibodies to anti-TNFs						population (mean difference per patient €-3,141 and the per-protocol population €-5,116. ICERs not reported Analysing serum samples using ELISA and HMSAs resulted in similar classification for the proposed algorithm in 72-78% of people
Steenholdt, 2015, Denmark[123]	To assess the cost-effectiveness of individualised therapy is a long term method compare to	Randomised controlled trial with a cost-effectiveness analysis, perspective not reported, six Danish hospitals	Receive treatment based on serum concentrations of infliximab and antibodies to infliximab at the time of	Radioimmunoassay	Cost per intention-to-treat and cost per-protocol population	Not applicable	Not applicable	Not applicable	Incremental costs in favour of the algorithm group. Results from the sensitivity analyses showed

Study ID (First author, year, and country)	Aim of the study	Study characteristics (study design, perspective, setting	Intervention	Kits used to analyse serum levels and antibodies to anti-TNFs	Outcome measure(s)	Model type	Health states	Utility values	Results (base case and sensitivity analysis)
	dose intensification in CD patients failing infliximab		infliximab treatment failure in accordance with the algorithm						similar findings to the base-case results
Velayos, 2013, USA[148]	To determine whether a testing-based strategy is more cost-effective than an empiric dose-escalation strategy	Cost-effectiveness analysis, third party payer, setting not reported	Testing-based strategy	Prometheus ELISA	Cost per QALY	Decision tree structure	Remission, response, dead	Medical remission: 0.89 Surgical remission: 0.86 Mild/minimal inflammation with symptoms: 0.80 Response: 0.77 Active disease: 0.62 Dead: 0	Testing strategy yielded similar QALYs compared with the empiric strategy (0.801 vs 0.800, respectively) but was less expensive (US$31,870 vs US$37,266, respectively). Testing strategy dominated the empiric strategy

CD; Crohn's disease; ICER, incremental cost-effectiveness ratio; ITT, intention-to-treat; NHS, National health service; PSS, personal social services; QALY, quality-adjusted life-year; UC, ulcerative colitis

4.1.3.5 Quality assessment

We present in Appendix 15 a summary of the reporting quality of the studies included in the current review against the CHEERS checklist.[150] Using a 25-point CHEERS checklist, one article[72] did not identify the study as an economic evaluation in the title. All studies provided background information to the study and clearly outlined the objectives of the study. Two studies[72, 148] reported the viewpoint of the economic analysis. All studies described the comparators fully and reported the time horizon. However, due to the short time horizon no studies conducted discounting of costs and benefits. In addition, the choice of health outcomes were well reported by all four studies;[72, 122, 123, 148] however only one study[72] reported how these health states were valued. Resource use and costs were well reported by three studies[122, 123, 148]apart from Vande Casteele et al. (2015),[72] who only described the drug costs. Majority of the studies[72, 122, 123] conducted an economic analysis alongside a randomised controlled trial, whilst one study[148] developed an economic model. In terms of analytical methods, study parameters, incremental costs and outcomes and uncertainty were well reported by all four studies. Limitations were provided by all four studies and generalisability was only partially reported by three studies.[122, 123, 148]

From the studies identified, one study[148] conducted a model-based economic analysis to determine whether a testing-based strategy was more cost-effective than an empiric dose-escalation strategy. We present in Appendix 15 a summary of the reporting quality of this study against the Philips' checklist.[144] In general, Velayos and colleagues conformed to best practice for reporting model-based economic evaluations in terms of clearly stating their decision problem, adequately outlining the objectives, clearly stating the viewpoint of the analysis, and the model structure, which represented the clinical pathway people with Crohn's disease may follow. Time horizon and cycle length were stated and justified. In terms of the data required to populate the model, Velayos and colleagues have adequately provided references, but it was unclear on the choices made between data sources and the quality of information used in the model. Additionally, it was unclear whether any expert opinion had been used when choosing baseline information for the model. The other limitation identified was the lack of explanation of pre-model analysis (e.g. calculation of transition probabilities and methods and assumptions used to extrapolate short-term results into final outcomes), and the omission of half-cycle correction.

4.1.4 Discussion and conclusion

The evidence available on the cost-effectiveness of LISA-TRACKER ELISA kits, TNFα-Blocker ELISA kits, and Promonitor ELISA kits for measuring levels of TNFα inhibitors and of anti-drug antibodies appears to be limited. We identified four cost-effectiveness analyses[72, 122, 123, 148] which

comprised three economic analyses conducted alongside clinical trials, and one model-based economic analysis.

The majority of the populations included in these studies had moderate-severe Crohn's disease and were considered responders to infliximab maintenance treatment. Studies (N=2) mainly used radioimmunoassay kits to analyse serum levels and antibodies to anti-TNFs. We appraised these analyses against frameworks for best practice for reporting economic evaluation and economic modelling. In general, all studies provided background information on the decision problem, clearly outlined the objectives of the study, adequately described and justified the choice of comparators, and reported the time horizon. In addition, Velayos and colleagues[148] clearly stated the viewpoint of their model-based economic analysis, and outlined the model structure. These studies all provide useful information in this developing area, but are subject to limitations. First, the definition for responder was not clear and varied between studies. Additionally, the definition for moderate to severe Crohn's disease patients varied across studies. Second, due to the small sample sizes this may not be reflective. Third, the short time horizon may not capture the longer-term costs and benefits of the use of testing to monitor serum anti-TNFα levels and antibodies to anti-TNFs. Fourth, it was unclear on how the authors made choices between data sources and the quality of information used in the model. From the two studies that reported their outcomes in terms of cost per QALY, only one study reported the generic preference-based measure used to estimate QALYs. This highlights a lack of transparency of the information used in the model. Other concerns relate to the lack of justification on the 4-week cycle length used in the modelling study by Velayos and colleagues.[148] Further concerns include the lack of detail on the resource use and costs in the study conducted by Vande Casteele and colleagues,[72] and the transparency on how transit probabilities were obtained and derived in the model-based evaluation.

In summary, all of these studies indicated that a testing strategy might be less costly than alternatives with variable small effects on effectiveness – some indicating small reduced benefits and some small increased benefits. Use of standard checklists suggested that all the studies are subject to some limitations.

In section 4.3.2, we outline the development of economic models to determine the cost-effectiveness of various assays to inform on the treatment algorithm for people who are considered responders and people with LOR.

4.2 Considerations of using the former HTA model by Dretzke et al. (2011) [5] to inform the current model structure

The previous HTA model[5] used Natural History data which is now outdated. The current model for the standard care arm is restricted to starting with infliximab (through lack of data for adalimumab) but otherwise adopts the general approach used in the HTA model, however using updated natural history data (for surgery, for maintenance of response, for dose escalation, and for other minor parameters, together with more recent clinical expert advice). Clearly the HTA model structure is not easily transferable to the current intervention arm since the latter requires considerable added complexity since it is based on drug and anti-drug antibody testing; however this arm conforms to the HTA approach and is designed for comparison with standard care on infliximab.

4.3 Health economic methods

4.3.1 Objective

To assess the cost-effectiveness of employing anti-TNFα and anti-TNFα antibody monitoring with LISA-TRACKER ELISA kits, TNFα-Blocker ELISA kits, and Promonitor ELISA kits in patients with Crohn's disease compared with standard care.

Standard care for people during maintenance of disease (responders) is shown in Figure 28.

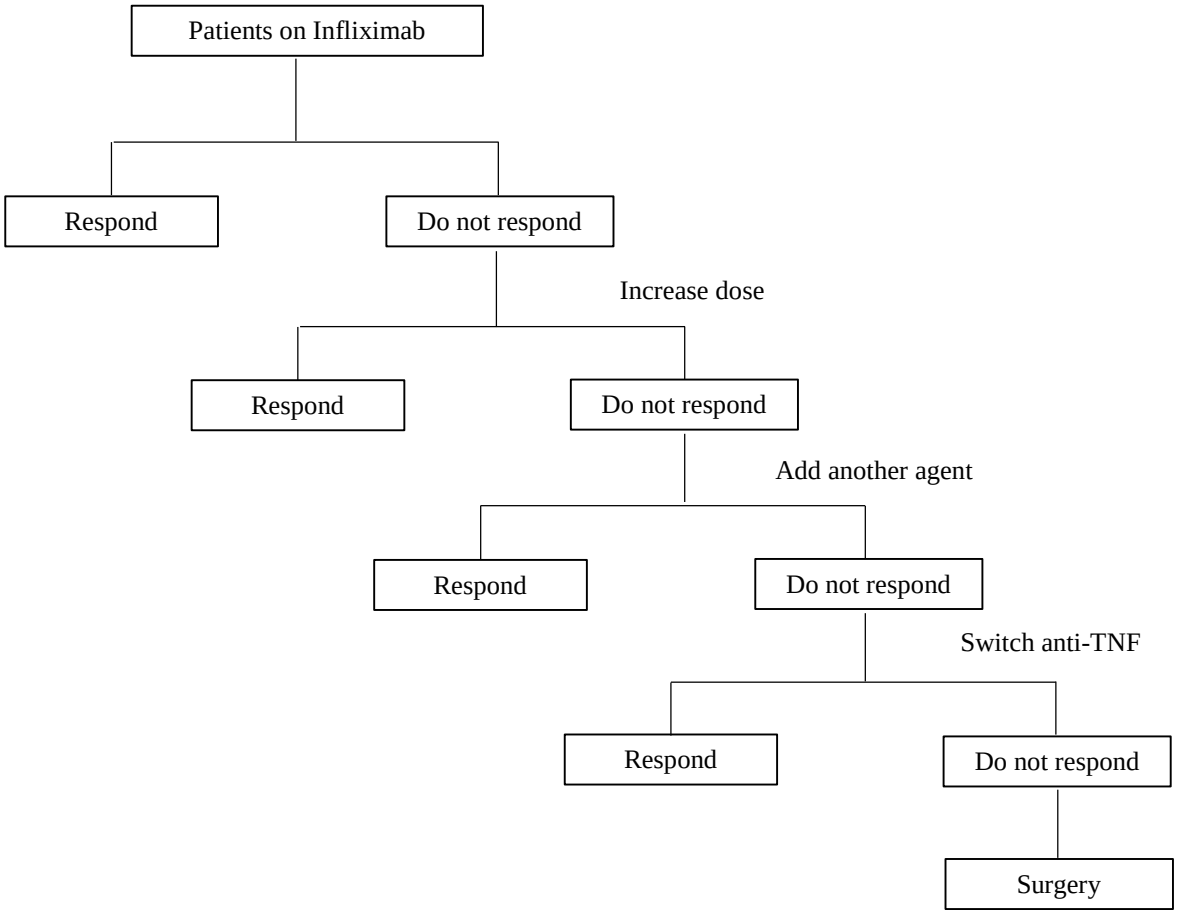


Figure 28 Standard care pathway for people on maintenance therapy

Standard care of people with CD may vary across hospitals in the UK. Based on expert clinical input, we assumed people categorised as responders to continue receiving infliximab maintenance therapy every eight weeks until they lose response. People who lose response will receive an increase in their dose. As a result of increasing the dose, people may respond to this increase or continue with lost response. People who continue with lost response will receive another agent in addition to their current treatment. As a result of adding another agent, people may regain response or continue with lost response. People who lose response will receive a switch to their anti-TNFα treatment. People who do not respond will be considered for surgery. We have assumed that people who have responded to treatment will remain on this treatment until they lose response. We assume that people who are in the post-surgery health state might receive various treatment options (anti-TNFα, a combination of anti-TNFα and immunosuppressant or no treatment). Patients who experience loss of response post-surgery are expected to follow the standard care treatment pathway as for responders entering the model who subsequently lose response; that is they will receive an increased dose of infliximab and follow the same treatment regime until they require recurrent surgery.

4.3.2 Developing the model structure

We developed a Markov model using TreeAge Pro 2013 software program (TreeAge Software, Willamstown, MA, USA). The model was developed with clinical input, and represents the clinical pathway people would undergo while being treated for moderate to severe Crohn's disease. The illustrative model structures for responders and for those who lose response is shown in Figure 29 and Figure 30, respectively. More detailed decision trees on the patient pathways can be found in Appendix 16. In the models, we compared concurrent and reflex testing conducted every three months compared to standard care for responders and those who experience loss of response:

- Standard care
- Concurrent testing: testing for TNFα inhibitor levels and antibodies to TNFα inhibitors
- Reflex testing: testing for TNFα inhibitor levels followed by testing of antibodies to TNFα inhibitors depending on the drug test level

NICE guidance on model-based economic analyses suggests adopting a time horizon long enough to capture the costs and effects of an intervention; normally a lifetime horizon, because chronic conditions may reduce life expectancy.[5] To our knowledge, no clinical trials exist that provide evidence of significant difference between testing and standard regimens in CD mortality.[5] Hence, we assumed a 10-year time horizon with four-week cycle lengths to be appropriate to capture all benefits of testing and treatment.

Table 27 shows the health states required for the responder and loss of response models.

Table 27 Definition of health states included in the Markov model

Health state	Definition
Responder	Maintenance treatment when the patient has supportable active symptoms of abdominal pain, diarrhoea, rectal bleeding or weight loss.
Loss of response	Recurrence of active symptoms while on treatment with maintenance regime, after having responded to treatment
Loss of response (no anti-TNF)[a]	Recurrence of active symptoms having discontinued anti-TNFα treatment with maintenance regime, but receiving best supportive care
Regain response	Maintenance treatment when the patient has no active symptoms having previously lost response
Post-surgery	Medication/no medication after inpatient surgical procedure
Dead	By definition

[a] People who have discontinued anti-TNF, but are receiving best supportive care

Below we discuss the testing strategies (concurrent and reflex testing) to be compared in both models (responders and people with LOR).

4.3.2.1 Concurrent testing

In the concurrent testing strategy, people would receive tests to analyse serum anti-TNFα levels and antibodies to anti-TNFs simultaneously, and on test result would receive the proposed algorithm. As a result of testing patients may be classified in various ways, for example: drug absent and antibodies present, drug absent and no antibodies, drug present and antibodies present and drug present and no antibodies. Alternatively categorisation may be based on multiple levels of drug with or without specified antibody levels (e.g. as in TAXIT[72]) Details of test results and proposed algorithms from Steenholdt et al. (2014)[122] for people with LOR and from Vande Casteele et al. (2015)[72] for responders are presented in section 3.2.3.

Responder

Based on the results from concurrent testing in the responder group, various treatment options may be adopted depending on the treatment algorithm used. In the model the treatment options are based on those used in the TAXIT study[72], the only clinical study of an implemented and defined algorithm for responders:

1. If drug is absent and antibodies are > 8 mg/mL, people receive a switch in TNFα inhibitor
2. If drug is absent and antibodies are < 8 mg/mL, people receive an increase dosage of current treatment (i.e. infliximab dose to 5mg/kg every 4 weeks)
3. If the drug is present (there is no need to measure antibodies), and depending on the trough levels, people would have either a decrease in the dosing interval (if trough level below the target range), no dose adaptation (if trough level is within the target range) or an increase the dosing intervals (if trough level is above the target range)

Following adoption these algorithm-treatments, people may remain responders, lose response (move to the loss of response health state) or die.

4.3.2.1.1 Loss of response

After loss of response to anti-TNF, testing and algorithm treatments are based on those used in the Steenholdt study[122] of patients who have lost response to anti-TNFα (infliximab); this is the only clinical study of an implemented algorithm for patients with lost response:

1. Drug absent and antibodies present, people would receive a switch in TNFα inhibitor
2. Drug absent and no antibodies, people would receive an increase dosage of current treatment
3. Drug present and antibodies present, we have assumed that some people will have symptoms not requiring surgery and discontinue anti-TNFα treatment or have active symptoms that

require surgery. People in the former would discontinue maintenance treatment and move to the loss of response health state (discontinuation of anti-TNF) and receive best supportive care. People who develop active symptoms that require surgery move to the post-surgery health state or could die

4. Drug present and no antibodies, the pathway for people with drug and antibodies present is identical to the pathway for people with drug present without antibodies

As a result of the treatment algorithm, people may remain with loss of response, or regain response or die.

4.3.2.1.2 Loss of response health state (discontinuation of anti-TNF)

People who occupy this health state are those who have discontinued anti-TNFα maintenance treatment, and are receiving best supportive care. As above, we have assumed that people who remain in this health state have symptoms of Crohn's disease that do not require surgery. People who develop active symptoms that require surgery move to the post-surgery health state or can die.

4.3.2.1.3 Re-gain response health state

Patients who move to the 'regain response' health state are tested for drugs and antibodies concurrently. Here we have assumed that they would follow the same treatment algorithm as a patient who was classed as a responder (see TAXIT[72] algorithm in section 3.2.3). As a result of the treatment algorithm, people can remain in the regain response health state, lose response and move to the loss of response health state, or die.

4.3.2.1.4 Post-surgery (remission) health state

For patients who move to the post-surgery health state, treatment options are to receive: an anti-TNF, immunosuppressant, a combination of anti-TNFα and immunosuppressant or no treatment. People who are receiving an anti-TNFα or a combination of anti-TNFα and immunosuppressant can regain response or lose response. For people who regain response or who lose response, we have assumed that the pathway is similar to people in the regain response health state or the loss of response health state, respectively, as noted above. People who are receiving immunomodulators or no treatment, the options are to remain in the post-surgery health state until further surgery is required or die.

4.3.2.2 Reflex testing

In the reflex testing strategy, people would receive a test to analyse serum anti-TNFα levels. As a result of testing, two test outcomes are likely, drug absent or drug present. Based on the drug result, people would undergo further testing for the presence/absence of antibodies. Below we outline the

health states and the pathways for people undergoing reflex testing for both responder and loss of response models. No study was identified that tested an algorithm for reflex testing. The algorithm followed in the model was therefore based on that of the TAXIT[72] trial for responders and the Steenholdt[122] algorithm for people with LOR using concurrent testing. Further details of test results and proposed algorithms are presented in section 3.2.3.

4.3.2.2.1 Responder

Based on the results from reflex testing in the responder group, various treatment options are available:

1. If drug is absent, test for antibodies. People with antibodies present would receive a switch in TNFα inhibitor. People with no antibodies who receive an increase dosage of current treatment (i.e. infliximab dose to 5mg/kg every 4 weeks)
2. If drug is absent and there are no antibodies, people would receive an increase dosage of current treatment (i.e. infliximab dose to 5mg/kg every 4 weeks)
3. If the drug is present and depending on the trough levels, people would have either a decrease in the dosing interval (if trough level below the target range), no dose adaptation (if trough level is within the target range) or an increase the dosing intervals (if trough level is above the target range)

As a result of the treatment algorithm, people could remain responders, lose response (move to the loss of response health state) or die.

4.3.2.2.2 Loss of response

1. Drug absent and antibodies present, people would receive a switch in TNFα inhibitor
2. Drug absent and no antibodies, people would receive an increase dosage of current treatment
3. Drug present and antibodies present, we have assumed that some people will have symptoms not requiring surgery and discontinue anti-TNFα treatment or have active symptoms that require surgery. People in the former would discontinue maintenance treatment and move to the loss of response health state (discontinuation of anti-TNF) and receive best supportive care. People who develop active symptoms that require surgery move to the post-surgery health state or could die

As a result of the treatment algorithm, people could remain in the loss of response state, or regain response or die.

4.3.2.2.3 Loss of response health state (discontinuation of anti-TNF)

People who occupy this health state are those who have discontinued anti-TNFα maintenance treatment, and who are receiving best supportive care. As above, we have assumed that people who remain in this health state have symptoms of CD that do not require surgery. People who develop active symptoms that require surgery move to the post-surgery health state or can die.

4.3.2.2.4 Regain response health state

Those patients who move to the regain response health state would receive reflex testing for drug levels, and if required, testing for antibodies to anti-TNFα. As above, we have assumed that they would follow the same treatment algorithm for people categorised as responders (see TAXIT[72] algorithm in section 3.2.3). As a result of the treatment algorithm, people can remain in the regain response health state, lose response and move to the loss of response health state, or die.

4.3.2.2.5 Post-surgery (remission) health state

For patients who move to the post-surgery health state, the treatment options are to receive an anti-TNFα, immunosuppressant, a combination of anti-TNFα and immunosuppressant or no treatment. People who are receiving an anti-TNFα or a combination of anti-TNFα and immunosuppressant can regain response or lose response and follow the same pathways as outline above. For people who are receiving immunomodulators or no treatment, the modelled options are to remain in the post-surgery health state until further surgery is required or to die.

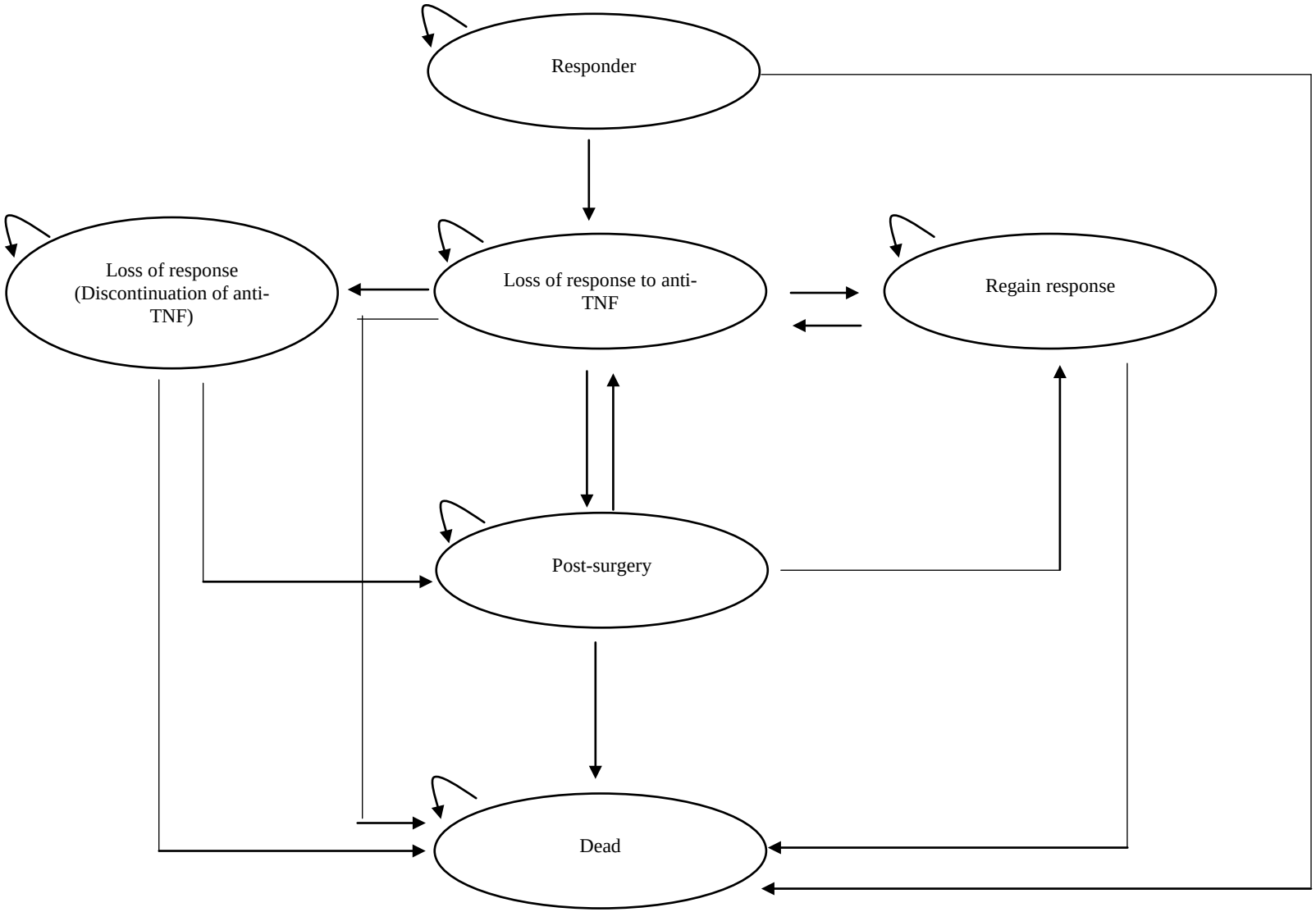


Figure 29 Illustrative structure for responders

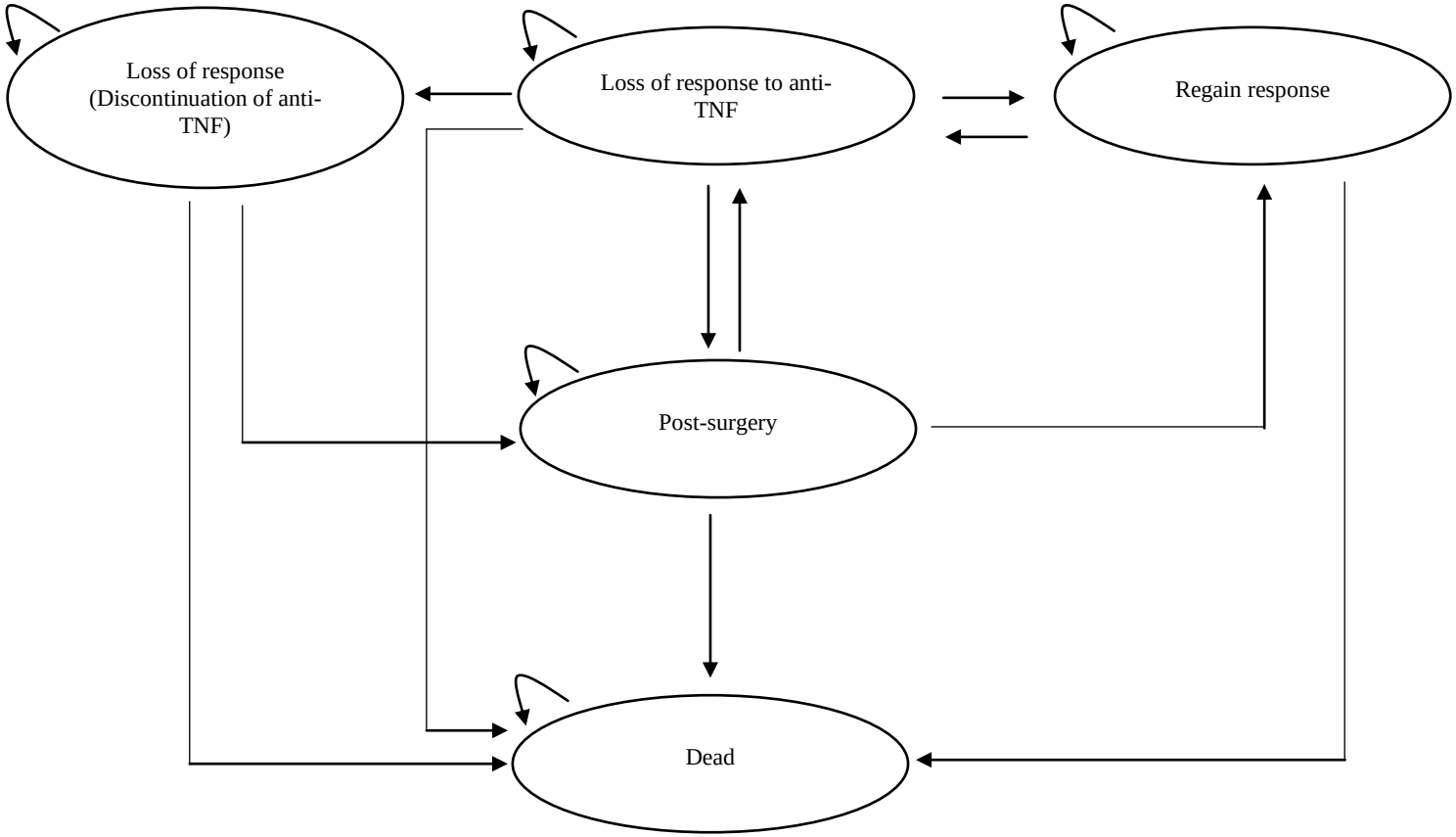


Figure 30 Illustrative structure for loss of response

4.3.3 Model assumptions

A number of assumptions were required to develop a workable model structure to enable the analyses to be undertaken. These assumptions are:

1. In our base-case, the model starts with a hypothetical cohort of 30-year olds with moderate to severe Crohn's disease
2. People were assumed to have received intravenous infusions of infliximab of 5mg/kg at week 0, two and six. Here we assumed that people were weighing >70kg
3. People who regained response have the same utility as those who are considered to be responders
4. We have assumed that people with Crohn's disease are not at increased risk of dying from the disease, and that there is no difference in mortality between testing and standard care. However, for people who have undergone surgery, we have included an increased risk of 0.0015 of dying from the procedure
5. Treatment effects for people receiving a dose escalation (from 5mg/kg to 10mg/kg of infliximab) and a decreased interval (from eight week to six week intervals) are the same
6. People who are categorised as a responder and who have trough concentration within the range that the treatment algorithm suggests receive no dose adaptation
7. In the base case we have assumed transition probabilities to be the same as standard care and used those derived from the Juillerat et al. (2015)[151]
8. People who remain in the loss of response health state (discontinuation of anti-TNF) have symptoms of Crohn's disease that in time may require surgery. People will receive best supportive care until active symptoms develop that require surgery

4.3.4 Data required for the model

The model was populated with clinical information from the current effectiveness review, and supplemented with information from secondary sources. Information required to parameterise the model included proportions, transition probabilities, resource use and costs, and utilities.

4.3.4.1 Proportions

The proportions of patients required to populate various model decision tree branches were obtained from secondary sources (e.g. management studies described in section 3.2.4) and where data were lacking from clinical input. Proportions that were estimated included: partitioning of patients by presence or absence of infliximab and of antibodies to infliximab in responders and in those with loss of response; partitioning of responders according to defined by infliximab trough levels; and partitioning by treatment options following surgery.

Table 28 summarises the partitioning of infliximab responders based on the study of Imaeda et al. (2012)[98] discussed in section 3.2.5.3 that used concurrent monitoring for the absence or presence of Infliximab and antibodies to Infliximab.

Table 28 Proportions derived based on concurrent testing of patients responding to infliximab

Result	Proportion	Source
Infliximab absent and antibodies to Infliximab present	0.17241	Imaeda et al., 2012[98]
Infliximab absent and antibodies to Infliximab absent	0.12069	
Infliximab present	0.7069	

The proportions of infliximab responders with various trough infliximab concentration levels were based on Vande Casteele et al. (2015)[72] discussed in section 3.2.4.3.4. These authors screened a cohort of inflammatory bowel disease patients who were receiving maintenance Infliximab treatment, and further categorised people by drug concentration levels based on test result. People with concentration levels < 3 µg/mL were considered below target range, people with concentration levels between 3 to 7 µg/mL were considered within range, and those > 7 µg/mL were above target range. Table 29 below shows the drug concentration levels and the proportions of responders derived from this study.

Table 29 Proportions according to infliximab trough levels of patients responding to infliximab

Trough concentration level	Threshold	Proportion	Source
Trough 1	< 3µg/mL	0.2310	Vande Casteele et al., 2015[72]
Trough 2	3 to 7µg/mL	0.4821	
Trough 3	> 7µg/mL	0.2869	

The partitioning of infliximab patients with loss of response according to concurrent test monitoring of infliximab and antibodies to infliximab was based on information obtained from Steenholdt et al. (2014)[122]. Table 30 below summarises these proportions.

Table 30 Proportions based on concurrent testing of patients with loss of response to infliximab

Result	Proportion	Source
Infliximab absent and antibodies to Infliximab present	0.1515	Steenholdt et al., 2014[122]
Infliximab absent and antibodies to Infliximab absent	0.0303	
Infliximab present and antibodies to Infliximab present	0.0303	
Infliximab present and antibodies to Infliximab absent	0.7879	

People who have undergone surgery may receive post-operative treatment to maintain remission. These options include anti-TNF, immunosuppressant, combination of anti-TNFα and immunosuppressant or no treatment. Table 31 below shows these proportions based on the study of Van der Have et al. (2014).[152]

Table 31 Treatment following surgery

Result	Proportion	Source
Anti-TNF	0.1250	Van der Have et al., 2014[152]
Immunosuppressant	0.5000	
Combination of anti-TNFα and immunosuppressant	0.1250	
No treatment	0.2500	

Table 32 summarises the proportions of infliximab responders based on the Imaeda et al. (2012)[98] that used reflex testing for the absence or presence of infliximab. People with infliximab present, we used the proportions according to infliximab trough concentration levels based on the Vande Casteele study, as shown in Table 29.

Table 32 Proportions derived based on reflex testing of patients responding to infliximab

Result	Proportion	Source
Infliximab absent	0.2931	Imaeda et al., 2012[98]
Infliximab present	0.7069	

The partitioning of infliximab patients with loss of response according to reflex test monitoring of infliximab was based on information obtained from Steenholdt et al. (2014).[122] Table 33 below summarises these proportions.

Table 33 Proportions based on reflex testing of patients with for loss of response to infliximab

Result	Proportion	Source
Infliximab absent and antibodies present	0.2029	Steenholdt et al., 2014[122]
Infliximab and antibodies absent	0.0435	
Infliximab present	0.7536	

4.3.4.2 Time to event transition probabilities

Table 34 summarises the transition probabilities for time to event outcomes used in the models.

Table 34 Summary of parametric models used for estimating transition probabilities for time to event outcomes

Standard care			
Transition	**Model**	**Source**	**Comments / Assumptions**
1. Infliximab maintenance to loss of response	Weibull λ 0 .02616269 γ 0.70142692	Juillerat et al. (2015)[151]	Observed data to 10 years, CD patients only, Weibull model provided a good fit
2. Infliximab maintenance to loss of response after infliximab escalation	Weibull λ 0.0020961 γ 0.502925	Ma et al. (2014)[153]	Time to loss of response after dose escalation. Observed data to > 6 years, Weibull model provided a good fit
3.Adalimumab after infliximab failure to loss of response	Exponential TP: C-1= 0.484; C-2 on = 0.032369 (λ 0.032904)	Sandborn (2007)[154] Karmaris (2009)[47]	RCT of adalimumab for CD (Sandborn 2007[154]) Prospective study 168 patients (Karmaris 2009[47]) exponential model provided a good fit
All			
Transition	**Model**	**Source**	**Comments / Assumptions**
4. Time to surgery	Weibull,§ λ 0.0350475 γ 0.4165309	Nguyen et al. (2011)[155]	Large study seven years of data, surgery incidence similar to small UK study; Weibull model provided a good fit. Model assumes patients diagnosed 10 years previously
5. Time to recurrent surgery	Gompertz,§ λ 0.1083159 γ -0.3677309	Nguyen et al. (2011)[155]	As above; Gompertz model provided a good fit
6. Time to post-surgical relapse on no therapy	Exponential λ 0.052305	Gordon et al. (2014)[156]	Assumed constant hazard; limited data

7. Time to post-surgical relapse on immunosuppressant	Exponential λ 0.0306871	Gordon et al. (2014)[156]	Assumed constant hazard; limited data
8. Time to post-surgical relapse on anti-TNF	Exponential λ 0.02100296	Baert et al., 2014[76]	Limited data, assumes applicability of study population. Exponential model provided a reasonable fit
9. Time to post-surgical relapse on anti TNFα and immunosuppressant	As above	Lack of data	Assumed as anti-TNFα alone
Standard care; sensitivity analysis			
Transition	**Model**	**Source**	**Comments / Assumptions**
10. Infliximab maintenance to loss of response after infliximab escalation	Weibull λ 0 .02616269 γ 0.70142692	Juillerat et al. (2015)[151]	Time to loss of response after dose escalation. Assumes 21% of 10 years spent in treatment with escalated dose
Intervention arm; test-algorithm strategy			
Transition	**Model**	**Source**	**Comments / Assumptions**
11. Infliximab maintenance to loss of response (dose escalation group)	Weibull λ 0 .02616269 γ 0.70142692	Juillerat et al. (2015)[151]	No evidence for advantage relative to standard care
12. Infliximab maintenance to loss of response (dose unchanged group)	Weibull λ 0 .02616269 γ 0.70142692	Juillerat et al. (2015)[151]	No evidence for difference according to trough group
13. Infliximab maintenance to loss of response (dose decreased group)	Weibull λ 0 .02616269 γ 0.70142692	Juillerat et al. (2015)[151]	No evidence for difference according to trough group
14. Regained response on adalimumab to loss of response (group 1, infliximab - / antibodies to infliximab +)	Exponential λ 0.032904	Karmaris et al. (2009)[47]	Prospective study 168 patients; exponential model provided a good fit
15. Regained response on intensified infliximab to loss of response (group 2, infliximab - / antibodies to infliximab -)	Weibull λ 0.0020961 γ 0.502925	Ma et al. (2014)[153]	Time to loss of response after dose escalation. Observed data to > six years, Weibull model provided a good fit
16. Regained response on un-prescribed treatment for loss of response (group 3 or 4 infliximab +/	Exponential TP: Cys-1,2 & 3= 0.16004/ cycle (λ	Steenholdt 2014[122] Rutgeerts (1999)[157]	Steenholdt: 13/27 in non-response at week 12. After week 12 assumed constant hazard for loss of regained response based on placebo group from the Rutgeerts 1999

antibodies to infliximab + or -	0.17442) C-4 on = 0.086173/cycle (λ 0.032904)		RCT
Intervention arm; test-algorithm strategy: sensitivity analyses for LOR			
Transition	**Model**	**Source**	**Comments / Assumptions**
17. Infliximab maintenance to loss of response (all trough groups)	Weibull λ 0.02228365 γ 0.70142692	Imaeda et al. (2012)[98] Juillerat et al. (2015)[151]	Assumes 3.9% advantage from test strategy based on limited data from TAXIT
18. Infliximab maintenance to loss of response of remission (all trough groups)	Exponential λ 0.0360311	Vande Casteele[72]	79.77% in remission at start, 62.6% at week 52; Assumes constant hazard for loss of remission
Intervention arm; test-algorithm strategy: sensitivity analyses for treatment after LOR			
Transition	**Model**	**Source**	**Comments / Assumptions**
19. Infliximab maintenance to loss of response of remission (all trough groups)	Exponential λ 0.0360311	Vande Castelle[72]	79.77% in remission at start, 62.6% at week 52; Assumes constant hazard for loss of remission
§ uses transition probabilities for cycles 130 to 260. C = four-week cycle; Cys = four-week cycles			

4.3.4.2.1 Transition probabilities from time to event studies

The transition probabilities provided in Table 34 are mainly derived from analyses of various time to event studies judged to provide relevant information consistent with the model structure. Further details regarding the derivation of, and justification for, these are provided in Appendix 17.

4.3.4.2.2 Resource use and costs

The resource use and costs included were those directly incurred by the NHS. Costs for reagents for monitoring trough concentration of anti-TNFs and antibody measuring kits, treatment for Crohn's disease, laparoscopic ileocolic resection were all included in the analysis. Resource use and costs associated with occupying all health states except dead were also included. Unit costs are presented in Table 35. The majority of the cost information used in the analyses were obtained from secondary sources.

Cost for infliximab and antibodies to infliximab monitoring kits were obtained from Theradiag / Alpha Laboratories (information provided by Theradiag / Alpha Laboratories) In Appendix 18, we present a breakdown of the resource use and costs associated with infliximab and antibodies to infliximab monitoring kits. In the models, we used a cost of £39.58 for concurrent testing, for

monitoring infliximab and antibodies to infliximab per person. For reflex testing, we used a cost of £43.48 for people being monitored for infliximab, and if results suggest drug absent then an antibodies test was undertaken. People whose result suggest drug present, no antibodies monitoring test was undertaken, hence we used a cost of £21.74.

Costs for maintenance treatment were obtained from the British National Formulary (BNF 2013/14)[158]. Costs of treatment associated with the induction phase (weeks 0 to six) were not included. Infliximab treatment costs comprised of its acquisition and administration costs. In the base-case, we assumed that people receiving maintenance therapy have received infusions of infliximab 5mg/kg every eight weeks, and that people weighed on average 70kg. For infliximab maintenance, we derived a cost of £1966.41 (assuming four 100mg vials at £419.62 plus administration costs of £287.93 per infusion) every eight weeks. For people switching to adalimumab, we derived a cost of £1408.28 (2 x £704.28; assuming 40mg of adalimumab is required every two weeks) per four-week cycle. We assumed people would administer adalimumab; hence no administration costs were included.

Estimated costs for management (outpatient visits to consultants and further investigations) associated with occupying all health states except the dead state were obtained from the NHS Reference cost database 2013/14[159] and in consultation with clinical expert. These health state costs include outpatient visits, colonoscopy, and magnetic resonance imaging (MRI). In Table 35, we present the unit costs per year associated with each health state.

Costs obtained from published sources were adjusted to 2013/14 prices using the Hospital and Community Health Service (HCHS) pay and price index Curtis et al. (2014)[160] and future costs were discounted at a rate of 3.5% per annum, as recommended by National Institute for Health and Care Excellence (NICE).

Table 35 Resource use and costs and utilities used in the models

Variable	Base-case value	Range for SA	Distribution	Reference(s)
Resource use and costs				
Monitoring infliximab	21.74		Fixed	
Monitoring antibodies to infliximab (reflex testing)	41.98		Fixed	NICE
Monitoring infliximab and antibodies to infliximab (concurrent testing)	38.83		Fixed	
Maintenance infliximab [a]	1966.41		Fixed	BNF 2013/14[158]

Variable	Base-case value	Range for SA	Distribution	Reference(s)
Maintenance adalimumab [b]	704.28		Fixed	BNF 2013/14[158]
Azathioprine [c]	8.40		Fixed	
Mercatopurine [d]	100.94		Fixed	BNF 2013/14[158]
Predinsolone [e]	14.25		Fixed	and expert opinion
Nutritional therapy (Modulen) [f]	15.06		Fixed	
Laparoscopic ileocolic resection [g]	6908		Fixed	
Responder [h]	725.69		Fixed	NHS reference
Loss of response [h]	1241.38		Fixed	costs 2013/14[159]
Regain response [h]	725.69		Fixed	and expert opinion
Post-surgery [h]	790.69		Fixed	
Utility values				
Responder	0.77	(0.70, 0.84)	Beta (117.04,34.96)	Velayos et al. (2013)[148]
Loss of response	0.62	(0.59, 0.66)	Beta (465,750)	Derived from Gregor et al. (1997)[19]
Regain response	0.77	(0.70, 0.84)	Beta (117.04,34.96)	Assumption
Surgery	0.60	(0.46, 0.73)	Beta (28.8,19.2)	Marchetti et al. (2014)[161]
Post-surgery	0.86	(0.82, 0.90)	Beta (301,49)	Velayos et al. (2013)[148]
Dead	0		Fixed	By definition
Other				
Mortality (age-specific death rates)	Life tables		Fixed	ONS 2014[162]
Mortality associated with surgical procedure	0.0015		Fixed	Velayos et al., 2013[148]
Discount rate per annum (costs and QALYs)	3.5%		Fixed	

[a] People receiving 5mg/kg of Infliximab during maintenance therapy every eight weeks. See appendices for details

[b] People receiving of Adalimumab during maintenance therapy every 40mg/kg every two weeks. See appendices for details

[c] Cost based on a 50mg (56 tablet pack) and recommended dosage of 2.5mg/kg per day

[d] Cost based on a 50mg (25 tablet pack) and recommended dosage of 1.25mg/kg per day

[e] Cost based on a 20mg/100ml single dose and recommended dosage of 30mg in week one then 5mg each week for the next three weeks

[f] Cost based on a 400g

[g] People undergoing a laparoscopic ileocolic procedure. Detail resources used are provided in the appendices

[h] Unit cost (per year) associated with occupying this health state. Please see appendices for further details on resource use

4.3.4.2.3 Outcomes

The quality-adjusted life-year (QALY) gained was the outcome measure used in our analyses. To calculate the estimated QALYs associated with the health states described in the model, we obtained utility weights from published literature[148] reported in our review of cost-effectiveness, and combined these utility values with the data on life expectancy Office of National Statistics. Utility values reported in Velayos et al. (2013)[148] were obtained from the study undertaken by Gregor et al. (1997)[19]. Gregor and colleagues compared various elicitation techniques (standard gamble, time trade-off and visual analogue scale) on 180 consecutive Crohn's disease patients. These authors suggested that the standard gamble technique reflected the true value for health states related to people with Crohn's disease, and these values may be the most appropriate to be used in an economic analysis. Table 35 shows the utility weights used in the model. In each cycle of the model, people will incur a utility payoff depending on the health state being occupied. In the model, we applied a utility weight of 0.77 for individuals categorised as responder or as having regained response. For those considered to have lost response we assigned a utility value of 0.62. Those who had undergone a surgical procedure and who remained in the post-surgery health state, were assigned a utility weight of 0.86.

4.3.5 Analysis

The model was constructed to assess the cost-effectiveness of concurrent testing, reflex testing and no testing for measuring patient blood levels of anti-TNFα agents and of antibodies to these agents in people with severe Crohn's disease. The model estimated the mean costs and effects associated with each testing strategy, and was simulated over a 10-year time horizon with four-weekly cycle lengths. The starting point for the responder population was a hypothetical cohort of people age 30 years whose disease responds to a maintenance course of TNFα inhibitor therapy. This age has been chosen because the onset for Crohn's disease is likely to occur in late teens to age 30 years (Saito et al., 2014). We define a maintenance course as receiving 5mg/kg of intravenous infliximab every eight weeks. The analysis was undertaken from an NHS perspective in an outpatient care setting, and outcomes reported as incremental cost effectiveness ratios (ICER), expressed in terms of cost per cost per quality-adjusted life-year (QALY) gained.

4.3.5.1 Sensitivity analysis

In addition to our base-case analysis, we have undertaken a number of sensitivity analyses. These analyses are summarised below:

1. Undertake concurrent testing and reflex testing every 12 months in the responder and loss of response models

2. Estimate the mean costs and effects associated with each strategy using a one-year time horizon with four-week cycle lengths
3. In the no testing strategy arm, using transition probabilities derived from Juillerat et al. for people who lose of response after dose escalation. Using transition probabilities derived from Weibull distribution with a changed scale parameter
4. In the responder model, using transition probabilities derived from Vande Casteele et al. (2015)[72] on infliximab maintenance to loss of response of remission (all trough groups)
5. Changing the proportion of people with infliximab and antibodies to infliximab present from 0.7878 to 0.2000

4.3.5.2 Probabilistic sensitivity analyses

Probabilistic sensitivity analyses (PSA) were undertaken to determine the joint uncertainty in key model input parameters of test results and expected QALYs. The PSA was undertaken based on the outcome of cost per QALY only. In probabilistic sensitivity analysis, each model parameter is assigned a distribution reflecting the amount and pattern of its variation, and cost-effectiveness results are calculated by simultaneously selecting random values from each distribution. The distributions used in the PSA are presented in Table 35. We have calculated probabilities that each strategy is the most cost-effective, at a willingness-to-pay of £20,000/QALY.

4.3.6 Results of base-case analyses and sensitivity analyses

Here we present the results of the base-case analyses based on the simplifying assumptions made in the model. In the base-case, using a hypothetical cohort of severe Crohn's disease adults aged 30 years, the results on concurrent testing, reflex testing and no testing (standard practice) based on the outcome of quality-adjusted life-year gained are presented in Table 36. At the 10-year time horizon, in the standard practice cohort, no testing resulted in 6.5146 QALYs with a corresponding mean cost of £137,600. The reflex testing cohort gained 6.3315 QALYs with a mean cost of £145,900. The concurrent testing cohort, gained 6.3215 QALYs with mean cost of £147,100. These results show that the no testing strategy was less costly and produced more QALYs, hence dominating the reflex testing and concurrent testing strategy.

Table 36 Base-case results for the analysis cost per QALY (2013/14 prices)

Strategy	Mean cost per strategy (£)	Difference in costs	Effectiveness (QALYs)	Incremental QALYs	Incremental cost-effectiveness ratio (£) (ICER)
No testing	137,600	-	6.5146	-	-

Reflex testing	145,900	8300	6.3315	0.1831	Dominated
Concurrent testing	147,100	9500	6.3215	0.1931	Dominated

Table 37 presents the results of the analyses based on an outcome of cost per QALY in the loss of response model with testing (concurrent and reflex) undertaken every three months. At the 10-year time horizon, the results show that the concurrent testing strategy resulted in 6.2600 QALYs with a corresponding mean cost of approximately £139,200. Reflex testing produced marginally more QALYs at an incremental cost of approximately £95,700 per QALY. The no testing strategy has a mean cost of approximately £199,900 and costs approximately £45,500 more than reflex testing with a total effectiveness of 6.5031 QALYs. This result indicates that in this loss of response model, the 'no testing' strategy is less cost effective than either reflex or concurrent testing. (Each additional QALY gained by adopting the 'no testing' strategy compared to reflex testing costs £257,340 in a cohort of people with loss of response).

Table 37 Base-case results for the analysis cost per QALY (2013/14 prices) (loss of response model)

Strategy	**Mean cost per strategy (£)**	**Difference in costs**	**Effectiveness (QALYs)**	**Incremental QALYs**	**Incremental cost-effectiveness ratio (£) (ICER)**
Concurrent testing	139,200	-	6.2600	-	-
Reflex testing	140,300	1100	6.2715	0.0115	95,700
No testing	199,900	59,600	6.5031	0.2316	257,340

Table 38 Univariate sensitivity analyses

Parameter varied	Mean cost per strategy	Difference in costs	Effectiveness (QALYs)	Incremental QALYs	Incremental cost-effectiveness ratio (£) (ICER)
Base case					
No testing	137,600	-	6.5146	-	-
Concurrent testing	145,900	8300	6.3315	0.1831	Dominated
Reflex testing	147,100	9500	6.3215	0.1931	Dominated
Annual testing in responder model					
Concurrent testing	116,300	-	6.2446	-	-
Reflex testing	116,400	100	6.2508	0.0062	16,100
No testing	137,600	21,200	6.5146	0.2638	80,400
Annual testing in loss of response model					
Concurrent testing	111,200	-	6.1877	-	-
Reflex testing	112,100	900	6.1946	0.0069	130,400
No testing	199,900	87,800	6.5031	0.3085	284,600
One-year time horizon in responder model					
No testing	15,200	-	0.8269	-	-
Concurrent testing	20,300	5100	0.8085	-0.0184	Dominated
Reflex testing	20,400	5200	0.8092	-0.0177	Dominated
One-year time horizon in loss of response model					
Concurrent testing	14,200	-	0.7531	-	-
Reflex testing	14,600	400	0.7562	0.0031	129,000
No testing	23,400	8800	0.8154	0.0592	148,600
Loss of response after dose escalation using transition probabilities derived from Juillerat Weibull with changed scale parameter in the standard care arm only					
No testing	143,800	-	6.5092	-	-
Reflex testing	145,900	2100	6.3315	-0.1777	Dominated
Concurrent testing	147,100	3300	6.3215	-0.1870	Dominated
Responder model: Infliximab maintenance to loss of response of remission (all trough groups) using transition probabilities derived from Vande Casteele et al. (2015)[72]					
No testing	137,600	-	6.5146	-	-

Parameter varied	Mean cost per strategy	Difference in costs	Effectiveness (QALYs)	Incremental QALYs	Incremental cost-effectiveness ratio (£) (ICER)
Reflex testing	144,600	7000	6.3115	-0.2031	Dominated
Concurrent testing	145,700	8100	6.3008	-0.2138	Dominated
Responder model: Reducing the proportion of people with infliximab and antibodies to infliximab present from 0.7878 to 0.2000					
No testing	137,600	-	6.5146	-	-
Concurrent testing	157,000	19,400	6.4038	0.1108	Dominated
Reflex testing	157,000	19,400	6.4246	0.0900	Dominated

4.3.7 Results of sensitivity analyses

We undertook a number of one-way sensitivity analyses to determine the impact of changing key model input parameters on the results. First, in the responder model, we changed the testing strategy from three months to annual testing. The results showed that concurrent testing was the cheapest strategy with a mean cost of approximately £116,300 generating a corresponding 6.2446 QALYs. In the reflex testing arm, this strategy was marginally more expensive and provided more QALYs with an ICER of approximately £16,100 per QALY. As expected, the mean cost of the no testing strategy remained unchanged. A no testing compared to a reflex testing strategy had a reported ICER of £80,400 per QALY.

Second, the effect of changing the three month testing to annual testing in the loss of response model resulted in both concurrent and reflex testing being cheaper than the no testing strategy. Third, on changing the model time horizon from 10-years to one-year with three month cycles, we found that the no testing strategy continued to dominate both testing strategies. In the loss of response model, the no testing strategy was the most expensive and most effective strategy with mean costs of approximately £23,400 and corresponding QALYs of 0.8154.

Finally, on changing the scale parameter in the Weibull model based on information from the Juillerat study, using transition probabilities derived from the Vande Casteele study, and reducing the proportion of people with infliximab and antibodies to infliximab from 0.7878 to 0.2000, we found that the no testing strategy continued to dominate the testing strategies.

In further sensitivity analyses, we varied key model input parameters to determine which inputs influence the ICER. Figure 31 and Figure 32 show the percentage change in the cost per QALY by varying these inputs by an increase and decrease of 10% of the basecase value. The results showed that the model is stable to most of these changes but sensitive to a 10% increase in the utility value for people who regain response in both reflex and concurrent testing.

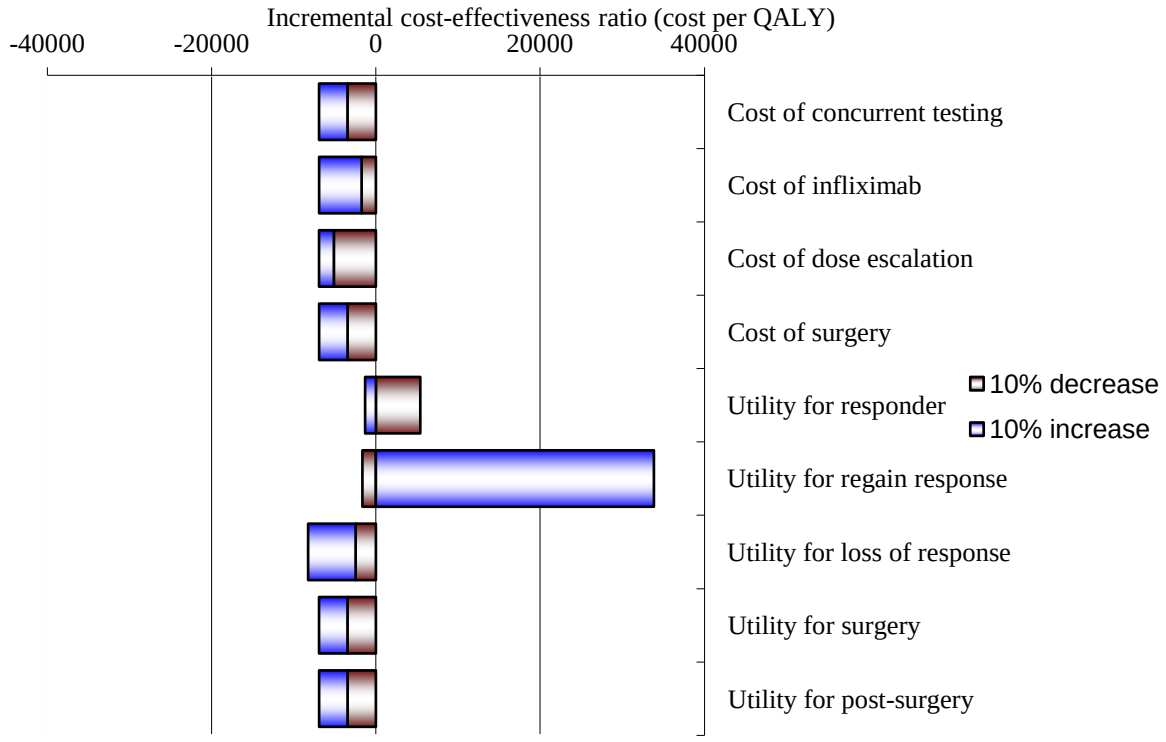


Figure 31 tornado diagram comparing no testing versus reflex testing

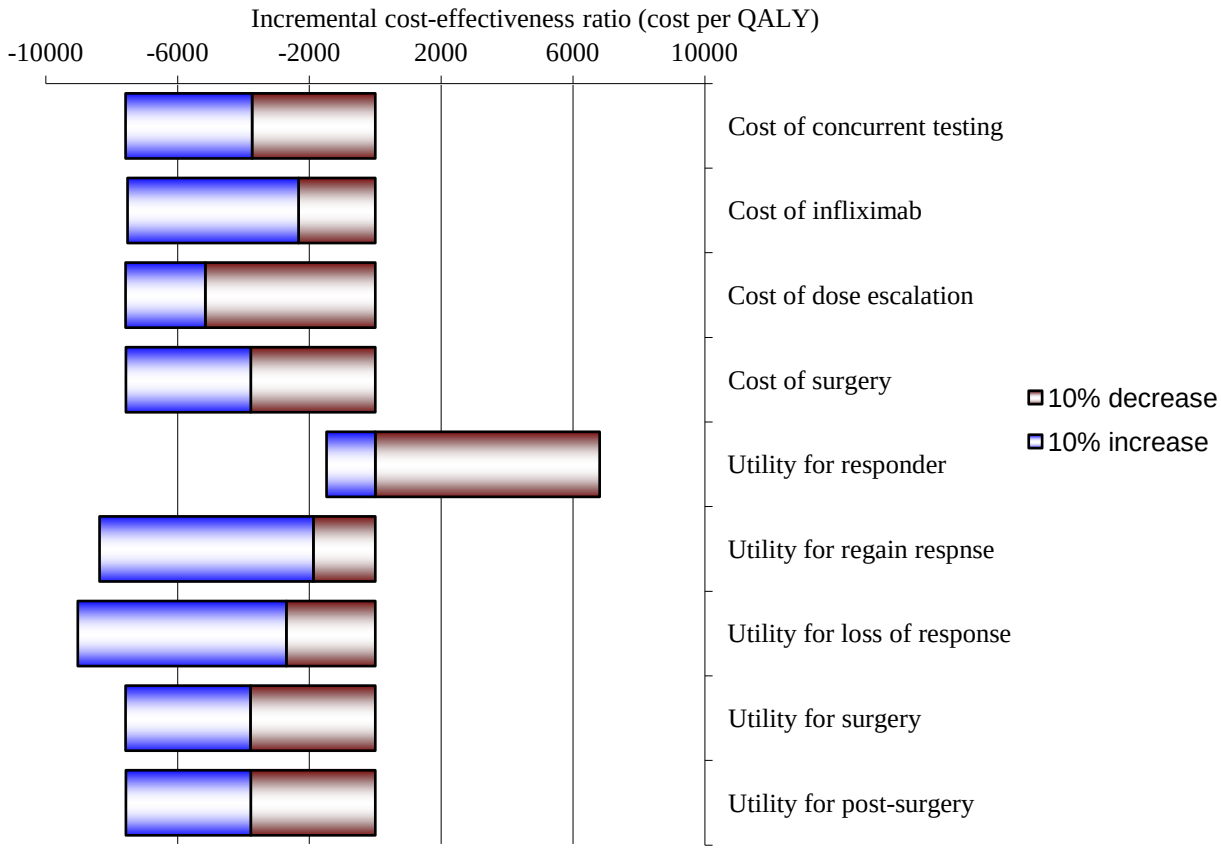


Figure 32 Tornado diagram comparing no testing versus concurrent testing

4.3.8 Results of probabilistic sensitivity analysis and CEAC

Results of the probabilistic sensitivity analysis for the base-case cost per quality-adjusted life-year outcome

Figure 33 shows the Monte Carlo simulation for the responder model. The scatterplot illustrates the uncertainty in the expected costs and QALYs based on concurrent and reflex testing compared with the no testing strategy. For the 10,000 runs of the Monte Carlo simulations, the scatterplots show considerable uncertainty around additional expected costs and QALYs.

The results for the responder model are presented in the form of cost-effectiveness acceptability curves (CEACs) in Figure 34. CEACs give the probability that a strategy is cost-effective at various willingness- to-pay values for a QALY. The willingness-to-pay threshold used by NICE is between £20-30,000 per QALY. From the information and assumptions used in the model, the results from Figure 34 show that at £20,000 per QALY the no testing strategy is 92% likely to be cost-effective compared to concurrent and reflex testing.

Figure 33 Probabilistic sensitivity analysis results for concurrent and reflex testing and no testing. Scatterplot using distributions around model input parameters

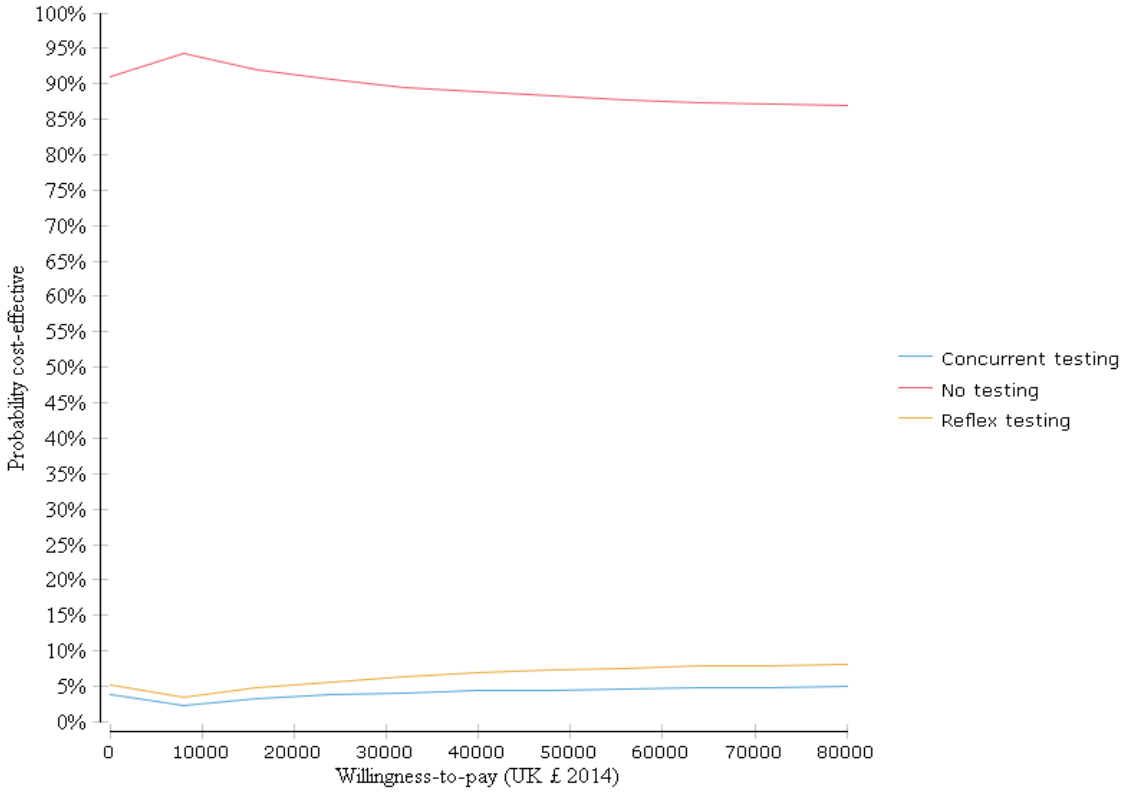


Figure 34 Cost-effectiveness acceptability curve using distributions around outcomes

4.3.9 Summary of cost effectiveness section

In summary, a de novo Markov model was built in TreeAge Pro 2013 to evaluate the cost effectiveness of test-algorithm based treatment strategies versus standard care. Two test strategies were assessed: concurrent testing of drugs and of antibodies to the drugs and sequential or reflex testing (i.e. drug test first, then anti-drug antibody test depending on indications of the drug test). The model structure was informed by studies from the clinical effectiveness review, additional published studies and analysis and expert clinical advice. The model had a 4 week cycle and a ten year time horizon and adopted an NHS and PSS perspective. Costs were adjusted to 2013/14 prices and annually discounted at 3.5%. The starting point was a hypothetical cohort of people age 30 years. Outcomes are reported as incremental cost effectiveness ratios (ICER), expressed in terms of cost per quality-adjusted life-year (QALY) gained. A linked-evidence approach was necessary. In this approach evidence from studies using tests other than the designated intervention tests was employed as a proxy for intervention test evidence. A number of sensitivity analyses were undertaken including: a shortened 1 year time horizon with four-week cycle lengths, altered transition probabilities for loss of response, altering the proportions of people in the different testing results categories and an arbitrary 10% change in the main input parameters. Probabilistic sensitivity analysis was also undertaken (10,000 model runs).

Two management studies, both RCTs of reasonable quality, have used treatment algorithms similar to those suggested in the NICE scope. The economic modelling has been built around the algorithms used in these studies. Expert opinion was sought regarding the complex patient pathways followed by CD patients and the treatment pathways dictated by the algorithms. Populating the model with information from the two management studies was problematical because of their small size, short duration, the reporting of outcomes not directly relevant to an economic model, the lack of an appropriate standard care arm for economic modelling in one study, and a lack of reporting outcomes according to testing results. Many external sources of data were required to populate the model and refining data inputs from these sources is currently still in progress.

Base case deterministic and probabilistic model results and sensitivity analysis results have been presented. The results require scrutiny using further investigations for model data inputs and sensitivity analyses, particularly with regard to frequency of testing, so as to test their robustness and to identify the main drivers of the ICER. However our conlcusions are that very similar QALY gains are likely in both arms while the cost of the testing strategy, whether undertaken concurrently or as a reflex strategy, appears to generate more than double the costs of standard care.

5 DISCUSSION

It has been proposed that measuring levels of anti-TNFα drug and antibodies raised against the drug during an immune response can aid the management of Crohn's disease patients who are on maintenance therapy. This implies that patients have responded to induction therapy of anti-TNFα with a reduction in symptoms and receive scheduled regular treatments. The main reason for drug monitoring in CD is to keep patients symptom free for as long as possible and avoid surgery by 1) optimising the dose and preventing LOR in patients who respond to drug treatment and 2) treating LOR with the most appropriate change in treatment in patients who have lost response during maintenance therapy, said to be patients with loss of response. In this assessment we investigated to what extent drug and anti-drug antibody measurements using three different type of commercially available ELISA kits can meet this aim of improved outcomes and if this approach is cost effective. The kits under assessment were: LISA-TRACKER ELISA kits, TNFα-Blocker ELISA kits, and Promonitor ELISA kits.

5.1 Decision problem and objectives

Our overall objective was to undertake a clinical and cost-effectiveness analysis of testing anti-TNFα levels and antibodies to anti-TNFα in people with Crohn's disease who are either responding to anti-TNFα treatment or have lost response to treatment during maintenance therapy. Testing strategies considered in this review were concurrent testing of drug and anti-drug antibody levels and antibody testing conditional on the absence of anti-TNFα. We aimed to systematically review the evidence on the clinical effectiveness of monitoring anti-TNFα drugs and their antibodies in responders and patients with loss of response when ELISA test results are used in combination with an algorithm that prescribes treatment pathways for the management of patients with specific drug and anti-drug antibody levels. We also aimed to identify evidence relevant to the costs of using these ELISA assays and to develop a cost-effectiveness model.

5.2 Summary of Methods and Findings

5.2.1 Clinical Effectiveness

We searched a number of databases including MEDLINE, EMBASE, the Cochrane library and the Science Citation Index. We mapped the included studies according to the studies' focus as management studies (reporting clinical outcomes following drug and anti-drug antibody testing and change in patient management according to a prescriptive algorithm for the management of CD patients), assay type comparison studies (comparing any of the three intervention ELISAs with each other or with assays used in a linked evidence approach) and correlation studies (reporting the relationship between test outcome and clinical status of tested patients). Management studies were

assessed for their clinical outcome data in relationship to assay type, test outcomes and algorithm followed. Four different test outcomes are possible when dichotomised testing for drug and anti-drug antibody levels. These are drug present/ antibodies absent | drug absent / antibodies absent | drug present / antibodies present | drug absent / antibodies present. The proportion of patients falling into these four categories according to testing and their clinical outcome data in terms of response and non-response following prescribed treatment changes were taken forward to the modelling. Another testing strategy categorises patients into groups according to several levels of anti-TNFα and prescribes appropriate treatment accordingly.

Assay type comparison studies were assessed for concordance statistics reported for relevant comparisons between assay types used in management studies and the three intervention assay kits. Our aim was to evaluate the generalisability of clinical outcome data from studies using non-intervention assays to the three intervention assays of interest in a linked evidence approach.
Correlation studies were assessed for sufficient data on the diagnostic performance of tests in predicting response / LOR in the two different patient groups and meta-analysed in order to use alternative data to the data from single management studies in the modelling.

We found 2,428 records of which 62 studies were included and an additional 6 studies were identified through other sources making up a total of 68 included studies. Of these studies three were management studies measuring levels of infliximab using RIA in patients with LOR, a commercial ELISA and HMSA in responders, and an in-house ELISA in responders, respectively. The three studies used different algorithms for the management of patients with certain test outcomes and only two of the three studies measured antibodies in addition to infliximab levels. All of the studies were small in size and none was long enough to fully assess the effect of following a treatment algorithm for the management of patients undergoing anti-TNFα therapy for CD. Furthermore, the cut-off of drug and anti-drug antibody levels used to determine therapeutic levels were not comparable in these studies. The sample collection time and analysis time were different as were the definitions used for clinical response, remission, progression and relapse. Steenholdt et al. (2014)[122] and (2015)[123] was the only RCT that compared drug monitoring and treatment change according to an algorithm with standard care (dose intensification) in patients with LOR. Their primary outcome was cost and they therefore concluded that combined measurement of drug and anti-drug antibodies reduces average treatment costs per patients compared with routine infliximab dose escalation and without any apparent negative effect on clinical efficacy. However, dose escalation was the most expensive treatment option in the standard care arm and might not be representative for UK clinical practice. Two studies investigated dose optimisation in responders: a 52 week randomised controlled trial (TAXIT)[72] indicated no benefit in clinical remission from test directed dose optimisation (RR before

versus after optimisation: 1.053, 95%CI: 0.936 to 1.186), and no difference at one year between clinically based dosing and test based dosing in clinical and biological remission (P = 0.686); the small retrospective observational study of Vaughn et al. (2014) [127] reported superior retention in infliximab treatment implying clinical benefit from test monitoring, however this study was judged to be at considerable risk of selection bias.

The links from the assays used in the management studies to the intervention assays of interest were weak and were complicated by the fact that none of the assays can be classed as gold standard; this limits the comparative data that is useful for a linked evidence approach using concordance data and/or Cohen's Kappa. The only direct link that was found was a study[66] comparing the performance of LISA-TRACKER with that of the Leuven in-house ELISA used in the TAXIT trial investigating the effectiveness of dose adjustment in responders to infliximab.[72] It reported disagreement for infliximab level measurements of at least 11/58 samples, and for anti-drug antibodies to infliximab for at least 3/62 samples, the results for the remainder were unclear.[66] Overall, there were no concordance data linking any of the index tests to any of the comparator tests at a clinically meaningful threshold. From this data it cannot be assessed which assay is more accurate or to what extent the results from the management studies are relevant to the intervention assays.

Meta-analyses of correlation studies indicated moderate test accuracy; positive and negative predictive value estimates derived from meta-analyses indicated that between 20 and 30% of positive and negative test results are likely to be inaccurate.

5.2.2 Cost-effectiveness

A comprehensive search of the literature for published economic evaluations, utility studies and cost studies was performed.

Four studies reported information on the cost-effectiveness of kits available for measuring levels of TNFα inhibitors and of anti-drug antibodies in people with severe Crohn's disease. From these, one study[148] used a decision analytic model to assess the cost-effectiveness of using test-based strategy compared to dose escalation in people who have loss responsiveness to infliximab. This review highlights that there is a paucity of economic evidence in this area.

The economic evidence was critically appraised against frameworks for best practice for reporting an economic evaluation. In terms of the quality of the reporting standards, most studies performed well against the CHEERS checklist. These studies provided useful information, but were subject to limitations. First, one study[72] has not stated in the title that an economic evaluation was conducted.

Second, resource use and costs reported in Vande Casteele et al. (2015)[72] were not comprehensive, only costs related to drug treatment were included. From the study[148] that conducted a model-based economic evaluation, these authors have adequately reported information on the decision problem, the structure of the model and its assumptions, time horizon and cycle lengths, and resource use and costs. Limitations of this study included, lack of clarity on the methods used to extrapolate short-term results into final outcomes. Second, it was unclear if the model was developed with any clinical input. Finally, these authors have not undertaken half-cycle correction nor did they justify its omission.

A de novo Markov model was built in TreeAge Pro 2013 to evaluate the cost effectiveness of test-algorithm based treatment strategies versus standard care. Two test strategies were assessed: concurrent testing of drugs and antibodies to the drugs and sequential or reflex testing (i.e. drug test first, then anti-drug antibody test depending on indications of the drug test). The model structure was informed by studies from the clinical effectiveness review, additional published studies and expert clinical advice. The model had a four-week cycle and a ten year time horizon and adopted an NHS and PSS perspective. Costs were adjusted to 2013/14 prices and annually discounted at 3.5%. The starting point was a hypothetical cohort of people age 30 years. Outcomes are reported as incremental cost effectiveness ratios (ICER), expressed in terms of cost per quality-adjusted life-year (QALY) gained. A linked-evidence approach was necessary. In this approach evidence from studies using tests other than the designated intervention tests was employed as a proxy for intervention test evidence. A number of sensitivity analyses were undertaken including: a shortened one-year time horizon with four-week cycle lengths, altered transition probabilities for loss of response, altering the proportions of people in the different testing results categories. Probabilistic sensitivity analysis was also undertaken (10,000 model runs).

In the base-case, results show that standard practice was less costly and produced more QALYs, hence dominating both the reflex testing and the concurrent testing strategy.

The results based on the outcome cost per quality adjusted life-year showed that the no testing (standard practice) dominated the reflex testing and concurrent testing strategies at the 10-year time horizon. Standard practice was least costly and produced more QALYs compared to the other strategies. In the standard practice cohort, the effectiveness of no testing resulted in 6.5146 QALYs with a corresponding mean cost of £137,600. In the reflex testing cohort, this strategy resulted in 6.3315 QALYs with a mean cost of £145,900. In the concurrent testing cohort, the total effectiveness was 6.3215 QALYs with mean cost of £147,100.

Sensitivity analyses indicated that change in testing frequency from 3 monthly to annually or reducing the time horizon to one-year changed the most cost effective option to a concurrent testing strategy. The PSA indicated a 92% likelihood that the 'no-testing' strategy was cost effective at a willingness to pay of £20,000 per QALY.

The no testing strategy continued to dominate both testing strategies when making changes to the model time horizon, statistical methods used to derive transition probabilities, and redistributing the proportion of people with the absence/presence of infliximab and antibodies to infliximab.

In most cases, the effect of varying some key model input parameters by an arbitrary 10% showed that no testing continued to dominate the testing strategies. With a 10% increase in the utility value for people who are responders, results suggested that concurrent testing was the most cost-effective testing strategy with an ICER of approximately £6,800 per QALY.

5.3 Strengths and Limitations

We undertook extensive systematic searches for relevant evidence and screened more than 30,000 titles. We used a recently developed method for analysis of published time to event data and undertook a new meta-analysis of test accuracy studies. In undertaking a linked evidence approach, and as far as evidence would allow, we rigorously examined the likely equivalence of assay methods specified as interventions compared to those used in the identified studies which investigated a test-treatment algorithm strategy. A particular strength of this work was the consideration of the additional objective (Objective C2 – analysis of correlation between test results and clinical outcomes) to include correlation studies. Correlation studies which reported both drug and anti-drug antibody test results for each patient provided test result probabilities by patient category (response and LOR). This was used in the economic model as an alternative to the probabilities reported by the management studies. Correlation studies reporting test results by group rather than individual were used to provide a pooled estimate for the probability of returning a specified test results after trough anti-TNFα testing (useful for estimating reflex strategy test result probabilities). This information was used where no evidence from management studies was available.

One of the main problems with this work is that the underlying evidence base for a 'linked evidence' approach is of concern. No test-algorithm studies employed the specified intervention tests. The only comparative evidence of monitoring drug and anti-drug antibody levels and standard care for the economic evaluation comes from studies using other assays than the three intervention assays under assessment even though a formal link between those assays and the intervention assays could not be established. All of the economic modelling depends upon the assumption that the commercial ELISA

kits with RIA and the Leuven in-house ELISA are equivalent is reasonable. However, the technical description of the assays, the differing drug thresholds used and the data from assay type comparison studies seem to suggest otherwise. Furthermore, there was insufficient evidence to link any of the index tests to any of the comparator tests with links to clinical outcomes. We looked for concordance data or Cohen's Kappa at set thresholds to determine how much different tests agreed and to use this data to undertake a sensitivity analysis. Unfortunately, the study by Steenholdt et al. (2014)[121] did not use any of the index tests and the link based on concordance data could not be established. We therefore remain uncertain to what extent the outcomes of the assessment apply to the three assay kits under evaluation and the cost effectiveness estimates which we have presented may not be reflective of the cost effectiveness of LISA-TRACKER ELISA kits, TNFα-Blocker ELISA kits, and Promonitor ELISA kits.

Furthermore, most of the available evidence about tests and algorithms prescribing treatment according to test results does not directly address the clinical effectiveness decision questions. The majority of evidence in the form of correlation studies does not generally present clinical decision making following test outcome; and studies that do present clinical decision making generally did not act on a prescriptive algorithm. While a number of algorithms have been suggested, only two have been tested in RCTs. Steenholdt et al. (2014)[122] tested a prescriptive algorithm for the clinical management of patients with LOR following combined drug and anti-drug antibody testing. This is representative for our concurrent testing strategy. However, no RCT which presented and evaluated an algorithm for reflex testing was identified. We therefore needed to assume that the Steenholdt algorithm can be adapted and the same treatment options are applicable for the three possible test outcomes following reflex testing (drug absent / antibody present | drug absent / antibody absent | drug present). As the drug positive / anti-drug antibodies positive group is treated identically to the drug positive / anti-drug antibodies negative group in the RCT by Steenholdt et al. (2014),[122] (using combined drug and anti-drug antibodies testing) the model structure of the strategies of reflex testing and concurrent testing are identical.

Whether reflex testing could be clinically viable seems unlikely since the delay in treatment due to conditional testing of antibodies on absence of drug could be up to four weeks (personal communication). The evidence for responders comes from Vande Casteele et al. (2015).[72] In this RCT reflex testing of drug and anti-drug antibodies was intended to aid drug optimisation with the aim to save drug costs and avoid adverse events in patients with high drug levels by decreasing the dose and to avoid LOR in patients with low drug levels by increasing the dose. Dose optimisation appears to be the most useful approach in responders for whom the dichotomisation of drug present / absent is not

applicable. In that respect only one of the decision questions prescribed by the NICE scope was directly addressed by the RCTs.

There are many possible test-algorithm strategies in the literature reflecting individual groups views, but the only relevant ones today are those that have been implemented prospectively with patients and compared to standard care and then only if relevant outcomes were reported. A concern for the validity of this review is the evidence on the lack of adherence to a pre-specified algorithm for the management of tested patients. A number of studies show that the main reason (generally >50%) for initiating testing is LOR, partial response or a flare.[55, 94, 163-165] Other reasons for testing included routine monitoring and adverse events. While this is reasonably constant across studies, the consideration of test results in clinical decision making in absence of a prescriptive treatment algorithm, varied widely. While one study reported that drug levels but not the presence of antibodies influenced treatment decisions,[94] another study reported that more patients with positive than negative anti-drug antibody tests received a treatment change.[166] There is also evidence that test results only impacted treatment decisions in 73% of patients tested[55] or that an appropriate treatment change of switching anti-TNFα agent in patients with positive anti-drug antibody test and dose increase in patients with sub-therapeutic infliximab levels only occurred in 57% and 21%, respectively.[164]

The rather sporadic consideration of test results in clinical decision making seems to suggest that an algorithm is needed to standardise the response to test outcomes. The study by Steenholdt et al. (2014)[122] revealed, however, that the algorithm prescribing treatment in this study was not followed in 42% of patients in the algorithm group tested for drug and anti-drug antibodies. This questions the validity of the comparative evidence, the usefulness of the algorithm and therefore the usefulness of testing anti-TNFα drug and anti-drug antibodies if no standardised treatment approach can be achieved.

In the course of the review it became apparent that the management of Crohn's disease patients varies widely between hospitals and between treating clinicians and that elements of the NICE guidance are possibly out of date. The overall aim to avoid surgery, the heterogeneity of disease symptoms, the relapsing and remitting disease pattern and possibly the personal preferences of clinicians and individual patients mean that it is difficult to establish a standardised pathway for patients with CD. This is reflected in the different algorithms identified but also in the different treatment options specified for patients with a certain test outcome and reflects the common opinion that a personalised approach to optimal anti-TNFα treatment can only be successful if multiple factors rather than just a single test result is considered in the management of patients.[29]

This presented a considerable challenge for the modelling. While the published algorithms tended to present several treatment options for patients with a certain test outcome reflecting the individual difference in disease status, previous medication and duration of disease, our model was required to prescriptive restricting treatment options to one or two possible treatments with little considertation of patient variability. It is therefore questionable to what extent the model results can predict or reflect clinical practice. A further complication was the fact that the algorithms were not developed for UK practice. The change to other non-TNFα biologics, namely vedolizumab could not be chosen as an option in our model as suggested in the Steenholdt algorithm. For this reason the proportion receiving surgery might be slightly overestimated in the model. However, clinicians advised us that patients might be referred to clinical trials of vedolizumab so still receiving this treatment option. Approval of vedolizumab in the UK for the treatment of Crohn's disease would therefore change the model outcome.

When considering using ELISAs or other assays to predict clinical outcomes (e.g. response) of patients with Crohn's disease on anti-TNFα treatment it is important not to forget that the tests are not perfect. Evaluation of the predictive performance of the assays appears to show that a considerable number of patients will have a false positive or false negative result and might receive the wrong treatment prescribed by an algorithm only considering test outcomes. Unfortunately, due to lack of outcome data we were unable to model patients as true positives/true negatives/false positives/false negatives according to test outcome. Therefore in our model, the test is not treated strictly as a diagnostic test but rather as an intervention of combined test and algorithm and the test result is not considered as a separate entity. This of course begs the question of test accuracy.

Furthermore, there is no gold standard for the assessment of response in Crohn's disease patients. This means that studies use different definitions for response, remission and relapse. As these definitions were used for patient selection and for classification of outcomes in primary studies we need to be cautious about the generalisability of outcomes. Finally, we had to deal with tests that lack a validated threshold for drug levels for the classification of response which has large implications for the generalisability of study outcomes. Appropriate test thresholds are strongly dependent on assay type and time of testing, the drug measured and whether anti-drug antibodies are measured, and the type of clinical marker used to evaluate response (CRP level, serum albumin, FC, endoscopic scoring of mucosal healing). While we are aware that the various assays do not measure the precise level of drug and anti-drug antibody due to difficulties of interference and drug / antibody complexes, the uncertainties around the drug threshold and definitions of response questions the value of knowing precise measurements of drug and anti-drug antibody levels in patients.

The included studies only recruited adult patients with Crohn's disease and the applicability of outcomes to children remains therefore unknown. However, in an abstract Turon et al. (2013)[167] reported that measurement of infliximab levels in paediatric IBD patients were informative and may improve safety and clinical symptoms in this patient group in which 47% of tests resulted in some form of modification of management.

Evidence was also lacking on the effectiveness of monitoring patients on adalimumab. While both adalimumab and infliximab are anti-TNFα agents, they are different molecules, are administered via different routes and at different doses using different schedules. It is therefore a further big assumption to treat outcomes of monitoring patients on infliximab as equivalent to outcomes for patients on adalimumab. For these reasons the economic modelling was limited to infliximab-treated patients.

The impact of immunomodulators on patient outcome was not formerly assessed in this review as it was outside the scope. However, the evidence suggests that immunomodulators generally improve patient outcome. The role that immunomodulators might play in the monitoring of drug levels and anti-drug antibody levels is an area for future research. Evidence is also emerging that faecal calprotectin can be used to monitor patients with Crohn's disease as it is a good marker for IBD activity and predicts relapse in time before symptoms return, therefore providing the clinician time to optimise therapy.[168] Future research is needed into how this marker and anti-TNFα monitoring might complement each other in the management of CD patients.

Finally, while the population of this review was restricted to patients with Crohn's disease, a substantial number of studies included in the review recruited patients with IBD. We remain uncertain about the impact the patient mix may have on the reported outcomes. Even though infliximab and adalimumab have been recently approved by NICE for the treatment of UC (positive NICE TA published in December 2014), outcomes from this assessment should not be readily transferred to UC patients; for example Bar-yoseph et al. (2013)[169] reported that in UC patients infliximab is more immunogenic and reaches lower trough levels than in patients with CD. This seems to suggest that there are may be differences in the response to infliximab between the two patient groups that may be of importance for the cost-effectiveness of monitoring anti-TNFα agents in UC patients.

Overall, due to the paucity of evidence, especially with regard to comparative studies, it is difficult to draw firm conclusions regarding the clinical effectiveness of testing strategies.

One of the strengths of the work includes the building of a de novo Markov model for the cost effectiveness assessment. However populating the economic model using outcome and test data from the management studies was problematical because of their small population size, short duration, and difficulties in allocating outcomes to categories of patients returning different defined test results. Inputs for the economic model needed to be drawn from disparate studies so that conclusions need to be tested with data from further research. The appropriateness of evidence sourced from other studies may be questioned because of differences in populations, and incomplete or ambiguous information regarding trough drug and anti-drug antibody levels. Several studies sourced for model inputs included a proportion of patients with a UC diagnosis; the impact of this on model input is difficult to gauge.

Although data is available about duration of anti-TNFα therapies, few studies report reasons for stopping these therapies; some patients may stop because of sustained remission and it was not possible to model this change in treatment satisfactorily due to lack of relevant data for the population groups explored in the model. For the same reasons it was not possible to model in the long term who are re-introduced back onto anti-TNFα treatment, for example the substantial proportion of patients who fail on infliximab and are then switched to receive a variety of non-anti-TNFα therapies that may or may not improve their clinical status, but who in the longer term are given anti-TNFα again.

All the studies used for modelling included mixed patient populations, a substantial proportion of which were already being treated with immunomodulators and who had previously been exposed to steroids. Steroids are used intermittently for flare, and information on frequency and duration of use is missing. Immunomodulators added to anti-TNF therapy are aimed at bringing back a better response to anti-TNF. The clinical effects of these agents is subsumed within the analysis of time to anti-TNF cessation. Information about the timing and frequency of addition of immune-modulators and the duration of their use is inadequate. Therefore it was difficult to model the addition of immuno-modulators and use of steroids for patients on anti-TNFα and clinical expert opinion was sought regarding the proportion of time over ten years patients treated with anti-TNFα agents would spend using steroid and immunomodulators. This data was assumed to apply for both testing and standard care arms of the model and was used for costing purposes.

We were not able to include adverse events occurring as a result of drug treatments in the models. Also, we have not included any health states or costs for people who may have complications following surgery. As a result, this may underestimate the costs.

6 CONCLUSIONS

The systematic review evidence gives some indication that the use of testing could be cost effective. But the RCTs on which this finding is based are small and lack validity. The tests themselves appear to generate substantial rates of false positives and false negatives. Base case deterministic and probabilistic model results have been presented. These indicate that very similar QALY gains are likely in both arms while the cost of the testing strategy, whether undertaken concurrently or as a reflex strategy, appears to generate substantially higher costs than standard care. No-testing appears to be the most cost-effective option, which was robust when investigated in various sensitivity analyses.

6.1 Recommendations for further research

We are aware that more comparative evidence is becoming available with the publication of the TAILORIX trial (ClinicalTrials.gov Identifier: NCT01442025) expected to be available in 2016.This trial will provide further insight into the effectiveness of sustaining therapeutic infliximab trough levels by measuring drug levels followed by dose increase if criteria are met in CD patients. Furthermore, the PANTS (UKCRN ID 14175) study is expected to report the most comprehensive data relating test outcomes to patient status in the second half of 2015.

However, there are many shortcomings in the current evidence base. Future research should address the following questions:

- How does measuring anti-TNFα drug and their antibodies by ELISA kits vary from using RIA and other methods?
- What are the clinically significant differences in the performance of ELISAs, RIA and HMSA?
- What are the best criteria for estimating response, non-response and LOR, and at what time should an assessment take place?
- What is the most widely acceptable algorithm in the UK?
- What are the barriers of following an algorithm for clinical management according to anti-TNFα and anti-drug antibody levels?
- What is the best time point of measuring drug and antibody?
- What is the validated drug threshold that predicts clinical outcome?
- What is the effectiveness and cost-effectiveness of monitoring Crohn's disease patients on adalimumab and for paediatric patients with Crohn's disease?
- What is the relevance of co-treatment with immunomodulators in the monitoring of anti-TNFα agents and their antibodies?
- Is there a benefit of measuring total drug / antibodies over free drug / antibody only?

7 REFERENCES

1. National Institute for Health and Care Excellence. Crohn's disease: Management in adults, children and young people. CG152. London, UK: National Institute for Health and Care Excellence; 2012 [cited 06/10/2014]; Available from: https://www.nice.org.uk/guidance/cg152.
2. Merlin T, Lehman S, Hiller J, Ryan P. The "linked evidence approach" to assess medical tests: A critical analysis. *Int J Technol Assess Health Care*. 2013;29(03):343-50.
3. Clark W, Raftery J, Song F, Barton P, Cummins C, Fry-Smith A, *et al.* Systematic review and economic evaluation of the effectiveness of infliximab for the treatment of Crohn's disease. *Health Technology Assessment (Winchester, England).* 2003;7(3):1-67.
4. NHS choices. Crohn's disease. 2013 [cited 06/10/2014]; Available from: http://www.nhs.uk/Conditions/Crohns-disease/Pages/Introduction.aspx.
5. Dretzke J, Edlin R, Round J, Connock M, Hulme C, Czeczot J, *et al.* A systematic review and economic evaluation of the use of tumour necrosis factor-alpha (TNF-α) inhibitors, adalimumab and infliximab, for Crohn's disease. *Health Technol Assess*. 2011;15(6):1-244.
6. National Institute for Health and Care Excellence. Infliximab (review) and adalimumab for the treatment of Crohn's disease. TA187. London, UK: National Institute for Health and Care Excellence; 2010 [cited 06/10/2014]; Available from: http://www.nice.org.uk/guidance/TA187.
7. Hanauer SB, Sandborn W. Management of Crohn's disease in adults. *Am J Gastroenterol*. 2001;96(3):635-43.
8. Jenkins HR. Inflammatory bowel disease. *Arch Dis Child*. 2001;85(5):435-7.
9. Sostegni R, Daperno M, Scaglione N, Lavagna A, Rocca R, Pera A. Review article: Crohn's disease: monitoring disease activity. *Aliment Pharmacol Ther*. 2003;17 Suppl 2:11-7.
10. Hyams J, Crandall W, Kugathasan S, Griffiths A, Olson A, Johanns J, *et al.* Induction and maintenance infliximab therapy for the treatment of moderate-to-severe Crohn's disease in children. *Gastroenterology*. 2007;132(3):863-73; quiz 1165-6.
11. Vermeire S, Van Assche G, Rutgeerts P. Laboratory markers in IBD: useful, magic, or unnecessary toys? *Gut*. 2006;55(3):426-31.
12. Sandborn WJ, Feagan BG, Hanauer SB, Lochs H, Lofberg R, Modigliani R, *et al.* A review of activity indices and efficacy endpoints for clinical trials of medical therapy in adults with Crohn's disease. *Gastroenterology*. 2002;122(2):512-30.
13. Irvine EJ. Usual therapy improves perianal Crohn's disease as measured by a new disease activity index. McMaster IBD Study Group. *J Clin Gastroenterol*. 1995;20(1):27-32.
14. Allez M, Karmiris K, Louis E, Van Assche G, Ben-Horin S, Klein A, *et al.* Report of the ECCO pathogenesis workshop on anti-TNF therapy failures in inflammatory bowel diseases: definitions, frequency and pharmacological aspects. *J Crohns Colitis*. 2010;4(4):355-66.
15. Freeman K, Connock M, Mistry H, Taylor-Phillips S, Court R, Tsertsvadze A, *et al.* Crohn's disease: Tests for therapeutic monitoring of TNF inhibitors (LISA-TRACKER ELISA kits, TNFα-Blocker ELISA kits, and Promonitor ELISA kits): Final Protocol. National Institute for Health and Care Excellence; 2014 [cited 21/04/2015]; Available from: http://www.nice.org.uk/guidance/gid-dt24/documents/crohns-disease-tests-for-therapeutic-monitoring-of-tnf-inhibitors-lisatracker-elisa-kits-tnfablocker-elisa-kits-and-promonitor-elisa-kits-final-protocol2.
16. Lichtenstein GR, Hanauer SB, Sandborn WJ. Management of Crohn's disease in adults. *Am J Gastroenterol*. 2009;104(2):465-83; quiz 4, 84.
17. Dignass A, Van Assche G, Lindsay JO, Lemann M, Soderholm J, Colombel JF, *et al.* The second European evidence-based Consensus on the diagnosis and management of Crohn's disease: Current management. *J Crohns Colitis*. 2010;4(1):28-62.
18. British Medical Assocation and Royal Pharmaceutical Society of Great Britain. British National Formulary and British National Formulary for Children. [cited 06/10/2014]; Available from: http://www.bnf.org/bnf/index.htm.
19. Gregor JC, McDonald JW, Klar N, Wall R, Atkinson K, Lamba B, *et al.* An evaluation of utility measurement in Crohn's disease. *Inflamm Bowel Dis*. 1997;3(4):265-76.
20. Cohen RD. The quality of life in patients with Crohn's disease. *Aliment Pharmacol Ther*. 2002;16(9):1603-9.

21. Standards for the Healthcare of People who have Inflammatory Bowel Disease (IBD): 2013 Update. IBD Standards; 2013 [cited 25/02/2015]; Available from: http://www.ibdstandards.org.uk/uploaded_files/IBDstandards.pdf.
22. Bassi A, Dodd S, Williamson P, Bodger K. Cost of illness of inflammatory bowel disease in the UK: a single centre retrospective study. *Gut*. 2004;53(10):1471-8.
23. Jewell DP, Satsangi J, Lobo A, Probert C, Forbes A, Ghosh S, *et al.* Infliximab use in Crohn's disease: impact on health care resources in the UK. *Eur J Gastroenterol Hepatol*. 2005;17(10):1047-52.
24. Sprakes MB, Ford AC, Suares NC, Warren L, Greer D, Donnellan CF, *et al.* Costs of care for Crohn's disease following the introduction of infliximab: a single-centre UK experience. *Aliment Pharmacol Ther*. 2010;32(11-12):1357-63.
25. Buchanan J, Wordsworth S, Ahmad T, Perrin A, Vermeire S, Sans M, *et al.* Managing the long term care of inflammatory bowel disease patients: The cost to European health care providers. *Journal of Crohns & Colitis*. 2011;5(4):301-16.
26. Hanauer SB, Feagan BG, Lichtenstein GR, Mayer LF, Schreiber S, Colombel JF, *et al.* Maintenance infliximab for Crohn's disease: the ACCENT I randomised trial. *Lancet*. 2002;359(9317):1541-9.
27. Sands BE, Blank MA, Patel K, van Deventer SJ, Study AI. Long-term treatment of rectovaginal fistulas in Crohn's disease: response to infliximab in the ACCENT II Study. *Clin Gastroenterol Hepatol*. 2004;2(10):912-20.
28. Colombel JF, Sandborn WJ, Rutgeerts P, Enns R, Hanauer SB, Panaccione R, *et al.* Adalimumab for maintenance of clinical response and remission in patients with Crohn's disease: the CHARM trial. *Gastroenterology*. 2007;132(1):52-65.
29. Ben-Horin S, Chowers Y. Tailoring anti-TNF therapy in IBD: drug levels and disease activity. *Nature Reviews Gastroenterology & Hepatology*. 2014;11(4):243-55.
30. Targan SR, Hanauer SB, van Deventer SJ, Mayer L, Present DH, Braakman T, *et al.* A short-term study of chimeric monoclonal antibody cA2 to tumor necrosis factor alpha for Crohn's disease. Crohn's Disease cA2 Study Group. *N Engl J Med*. 1997;337(15):1029-35.
31. Hanauer SB, Sandborn WJ, Rutgeerts P, Fedorak RN, Lukas M, MacIntosh D, *et al.* Human anti-tumor necrosis factor monoclonal antibody (adalimumab) in Crohn's disease: the CLASSIC-I trial. *Gastroenterology*. 2006;130(2):323-33; quiz 591.
32. de Boer N, Lowenberg M, Hoentjen F. Management of Crohn's disease in poor responders to adalimumab. *Clin Exp Gastroenterol*. 2014;7:83-92.
33. Bendtzen K. Anti-TNF-alpha biotherapies: perspectives for evidence-based personalized medicine. *Immunotherapy*. 2012;4(11):1167-79.
34. Gisbert JP, Panes J. Loss of response and requirement of infliximab dose intensification in Crohn's disease: a review. *Am J Gastroenterol*. 2009;104(3):760-7.
35. Billioud V, Sandborn WJ, Peyrin-Biroulet L. Loss of response and need for adalimumab dose intensification in Crohn's disease: a systematic review. *Am J Gastroenterol*. 2011;106(4):674-84.
36. Carrillo-Ramos MJ, Duarte-Chang C, Maldonado-Perez B, Beltran-Castano R, Castro-Laria L, Arguelles-Arias F, *et al.* Adalimumab or infliximab for the treatment of inflammatory bowel disease patients: Which is more effective? *Gastroenterology*. 2014;1):S-196.
37. Maser EA, Villela R, Silverberg MS, Greenberg GR. Association of trough serum infliximab to clinical outcome after scheduled maintenance treatment for Crohn's disease. *Clin Gastroenterol Hepatol*. 2006;4(10):1248-54.
38. Cassinotti A, Travis S. Why don't we just measure infliximab drug levels in IBD? *Pract Gastroenterol*. 2010;34(10):11-20.
39. Hanauer SB, Wagner CL, Bala M, Mayer L, Travers S, Diamond RH, *et al.* Incidence and importance of antibody responses to infliximab after maintenance or episodic treatment in Crohn's disease. *Clin Gastroenterol Hepatol*. 2004;2(7):542-53.
40. Ungar B, Kopylov U, Yavzori M, Fudim E, Picard O, Lahat A, *et al.* Predictors of formation of antibodies to infliximab (ATI) and secondary loss of response in IBD patients treated with infliximab. *J Crohns Colitis*. 2014;8:S45.

41. Barry D, Bloomfeld RS. The prevalence of human antichimeric antibodies in patients on infliximab increases with age. *Gastroenterology*. 2012;1):S387.
42. Baert F, Noman M, Vermeire S, Van Assche G, G DH, Carbonez A, *et al.* Influence of immunogenicity on the long-term efficacy of infliximab in Crohn's disease. *N Engl J Med*. 2003;348(7):601-8.
43. Khanna R, Sattin BD, Afif W, Benchimol EI, Bernard EJ, Bitton A, *et al.* Review article: a clinician's guide for therapeutic drug monitoring of infliximab in inflammatory bowel disease. *Aliment Pharmacol Ther*. 2013;38(5):447-59.
44. Vermeire S, Noman M, Van Assche G, Baert F, D'Haens G, Rutgeerts P. Effectiveness of concomitant immunosuppressive therapy in suppressing the formation of antibodies to infliximab in Crohn's disease. *Gut*. 2007;56(9):1226-31.
45. Vande Casteele N, Cuypers L, Singh S, Hauenstein S, Ohrmund L, Chuang E, *et al.* Transient versus sustained antibodies to infliximab: Possibility to overcome low titer antibody responses by dose optimisation. *J Crohns Colitis*. 2012;6:S110.
46. Ainsworth MA, Bendtzen K, Brynskov J. Tumor necrosis factor-alpha binding capacity and anti-infliximab antibodies measured by fluid-phase radioimmunoassays as predictors of clinical efficacy of infliximab in Crohn's disease. *Am J Gastroenterol*. 2008;103(4):944-8.
47. Karmiris K, Paintaud G, Noman M, Magdelaine-Beuzelin C, Ferrante M, Degenne D, *et al.* Influence of trough serum levels and immunogenicity on long-term outcome of adalimumab therapy in Crohn's disease. *Gastroenterology*. 2009;137(5):1628-40.
48. Garces S, Demengeot J, Benito-Garcia E. The immunogenicity of anti-TNF therapy in immune-mediated inflammatory diseases: a systematic review of the literature with a meta-analysis. *Ann Rheum Dis*. 2013;72(12):1947-55.
49. Gils A, Vande Casteele N, Poppe R, Van de Wouwer M, Compernolle G, Peeters M, *et al.* Development of a universal anti-adalimumab antibody standard for interlaboratory harmonization. *Ther Drug Monit*. 2014;36(5):669-73.
50. van Schie KA, Hart MH, de Groot ER, Kruithof S, Aarden LA, Wolbink GJ, *et al.* The antibody response against human and chimeric anti-TNF therapeutic antibodies primarily targets the TNF binding region. *Ann Rheum Dis*. 2015;74(1):311-4.
51. Steenholdt C, Palarasah Y, Bendtzen K, Teisner A, Brynskov J, Teisner B, *et al.* Pre-existing IgG antibodies cross-reacting with the Fab region of infliximab predict efficacy and safety of infliximab therapy in inflammatory bowel disease. *Aliment Pharmacol Ther*. 2013;37(12):1172-83.
52. Van Assche G, Magdelaine-Beuzelin C, D'Haens G, Baert F, Noman M, Vermeire S, *et al.* Withdrawal of immunosuppression in Crohn's disease treated with scheduled infliximab maintenance: a randomized trial. *Gastroenterology*. 2008;134(7):1861-8.
53. Feagan BG, McDonald JW, Panaccione R, Enns RA, Bernstein CN, Ponich TP, *et al.* Methotrexate for the prevention of antibodies to infliximab in patients with Crohn's disease. *Gastroenterology*. 2010;1):S167-S8.
54. Colombel JF, Sandborn WJ, Reinisch W, Mantzaris GJ, Kornbluth A, Rachmilewitz D, *et al.* Infliximab, azathioprine, or combination therapy for Crohn's disease. *N Engl J Med*. 2010;362(15):1383-95.
55. Afif W, Loftus EV, Jr., Faubion WA, Kane SV, Bruining DH, Hanson KA, *et al.* Clinical utility of measuring infliximab and human anti-chimeric antibody concentrations in patients with inflammatory bowel disease. *Am J Gastroenterol*. 2010;105(5):1133-9.
56. Roblin X, Rinaudo M, Del Tedesco E, Phelip JM, Genin C, Peyrin-Biroulet L, *et al.* Development of an algorithm incorporating pharmacokinetics of adalimumab in inflammatory bowel diseases. *Am J Gastroenterol*. 2014;109(8):1250-6.
57. Paul S, Del Tedesco E, Marotte H, Rinaudo-Gaujous M, Moreau A, Phelip JM, *et al.* Therapeutic drug monitoring of infliximab and mucosal healing in inflammatory bowel disease: a prospective study. *Inflamm Bowel Dis*. 2013;19(12):2568-76.
58. Pariente B, Pineton de Chambrun G, Krzysiek R, Desroches M, Louis G, De Cassan C, *et al.* Trough levels and antibodies to infliximab may not predict response to intensification of infliximab therapy in patients with inflammatory bowel disease. *Inflamm Bowel Dis*. 2012;18(7):1199-206.

59. Steenholdt C. Use of infliximab and anti-infliximab antibody measurements to evaluate and optimize efficacy and safety of infliximab maintenance therapy in Crohn's disease. *Dan Med J*. 2013;60(4):B4616.
60. Ben-Horin S, Chowers Y. Review article: loss of response to anti-TNF treatments in Crohn's disease. *Aliment Pharmacol Ther*. 2011;33(9):987-95.
61. Cassinotti A, Travis S. Incidence and clinical significance of immunogenicity to infliximab in Crohn's disease: a critical systematic review. *Inflamm Bowel Dis*. 2009;15(8):1264-75.
62. Lee LY, Sanderson JD, Irving PM. Anti-infliximab antibodies in inflammatory bowel disease: prevalence, infusion reactions, immunosuppression and response, a meta-analysis. *Eur J Gastroenterol Hepatol*. 2012;24(9):1078-85.
63. Chaparro M, Guerra I, Munoz-Linares P, Gisbert JP. Systematic review: antibodies and anti-TNF-alpha levels in inflammatory bowel disease. *Aliment Pharmacol Ther*. 2012;35(9):971-86.
64. Vande Casteele N, Gils A, Ballet V, Compernolle G, Peeters M, Van Steen K, *et al.* Randomised Controlled Trial of Drug Level Versus Clinically Based Dosing of Infliximab Maintenance Therapy in IBD: Final Results of the TAXIT Study (OP001). *United European Gastroenterology Journal*. 2013;1(1s):A1.
65. Scott FI, Lichtenstein GR. Therapeutic Drug Monitoring of Anti-TNF Therapy in Inflammatory Bowel Disease. *Current Treatment Options in Gastroenterology*. 2014;12(1):59-75.
66. Vande Casteele N, Buurman DJ, Sturkenboom MG, Kleibeuker JH, Vermeire S, Rispens T, *et al.* Detection of infliximab levels and anti-infliximab antibodies: a comparison of three different assays. *Aliment Pharmacol Ther*. 2012;36(8):765-71.
67. Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *BMJ*. 2009;339:b2535.
68. Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, *et al.* QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. *Ann Intern Med*. 2011;155(8):529-36.
69. Higgins JP, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, *et al.* The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*. 2011;343:d5928.
70. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *J Epidemiol Community Health*. 1998;52(6):377-84.
71. Methods Guide for Medical Test Reviews. AHRQ Publication No. 12-EC017. Rockville, MD: Agency for Healthcare Research and Quality; 2012 [cited 11/03/2015]; Available from: http://www.effectivehealthcare.ahrq.gov/ehc/products/246/558/Methods-Guide-for-Medical-Test-Reviews_Full-Guide_20120530.pdf.
72. Vande Casteele N, Ferrante M, Van Assche G, Ballet V, Compernolle G, Van Steen K, *et al.* Trough Concentrations of Infliximab Guide Dosing for Patients with Inflammatory Bowel Disease. *Gastroenterology*. Forthcoming 2015.
73. Guyot P, Ades A, Ouwens M, Welton N. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC Med Res Methodol*. 2012;12(1):9.
74. Harbord R, Whiting P. metandi: Meta-analysis of diagnostic accuracy using hierarchical logistic regression. *Stata Journal*. 2009;9(2):211-29.
75. Harris RJ, Bradburn MJ, Deeks JJ, Harbord R, Altman DG, Sterne JAC. metan: fixed- and random-effects meta-analysis. *Stata Journal*. 2008;8(1):3-28.
76. Baert F, Drobne D, Gils A, Vande Casteele N, Hauenstein S, Singh S, *et al.* Early trough levels and antibodies to infliximab predict safety and success of reinitiation of infliximab therapy. *Clin Gastroenterol Hepatol*. 2014;12(9):1474-81.e2; quiz e91.
77. Ben-Bassat O, Romanova A, Iacono A, Irwin SP, Greenberg GR. Association of serum infliximab and antibodies to infliximab to long-term clinical outcome and mucosal healing in crohn's disease. *Gastroenterology*. 2013;1):S775.
78. Ben-Horin S, Yavzori M, Katz L, Kopylov U, Picard O, Fudim E, *et al.* The immunogenic part of infliximab is the F(ab')2, but measuring antibodies to the intact infliximab molecule is more clinically useful. *Gut*. 2011;60(1):41-8.

79. Ben-Horin S, Mazor Y, Yanai H, Ron Y, Kopylov U, Yavzori M, *et al.* The decline of anti-drug antibody titres after discontinuation of anti-TNFs: implications for predicting re-induction outcome in IBD. *Aliment Pharmacol Ther*. 2012;35(6):714-22.
80. Bodini G, Savarino V, Dulbecco P, Baldissarro I, Savarino E. ELISA vs. HMSA: A comparison between two different methods for the evaluation of adalimumab serum concentration and anti-adalimumab antibodies Preliminary data. *J Crohns Colitis*. 2014;8:S278.
81. Bortlik M, Duricova D, Malickova K, Machkova N, Bouzkova E, Hrdlicka L, *et al.* Infliximab trough levels may predict sustained response to infliximab in patients with Crohn's disease. *J Crohns Colitis*. 2013;7(9):736-43.
82. Candon S, Mosca A, Ruemmele F, Goulet O, Chatenoud L, Cezard JP. Clinical and biological consequences of immunization to infliximab in pediatric Crohn's disease. *Clin Immunol*. 2006;118(1):11-9.
83. Chiu YL, Rubin DT, Vermeire S, Louis E, Robinson AM, Lomax KG, *et al.* Serum adalimumab concentration and clinical remission in patients with Crohn's disease. *Inflamm Bowel Dis*. 2013;19(6):1112-22.
84. Cornillie F, Hanauer SB, Diamond RH, Wang J, Tang KL, Xu Z, *et al.* Postinduction serum infliximab trough level and decrease of C-reactive protein level are associated with durable sustained response to infliximab: a retrospective analysis of the ACCENT I trial. *Gut*. 2014;63(11):1721-7.
85. Corstjens PL, Fidder HH, Wiesmeijer KC, de Dood CJ, Rispens T, Wolbink GJ, *et al.* A rapid assay for on-site monitoring of infliximab trough levels: a feasibility study. *Anal Bioanal Chem*. 2013;405(23):7367-75.
86. Daperno M, Frigerio F, Guiotto C, Germano L, Ercole E, Arico S, *et al.* Evaluation of the diagnostic performance of two commercially available tests for infliximab trough levels (IFX-TL) and antibodies to infliximab (ATI) titration in inflammatory bowel disease (IBD). *J Crohns Colitis*. 2013;7:S213-S4.
87. Dauer RM, Yarur AJ, Abreu MT. Infliximab re-induction outcomes after a failure to treatment. *Gastroenterology*. 2013;1):S430.
88. Egea-Pujol L, Reddy R, Patel S, Christie R, Salbato J, Shah S, *et al.* Homogenous mobility shift assay (HMSA) overcomes the limitations of elisa and eclia assays for monitoring infliximab (IFX), adalimumab (ADA), and associated anti-drug antibodies in serum. *Am J Gastroenterol*. 2013;108:S548.
89. Eser A, Primas C, Hauenstein S, Lockton S, Singh S, Reinisch W. Comparison of early measurement of infliximab and antibodies-to-infliximab serum levels with standard trough analysis. *Gastroenterology*. 2013;1):S779.
90. Eser A, Primas C, Haunstein S, Lockton S, Wang S, Singh S, *et al.* Detection of anti infliximab antibodies in patients with inflammatory bowel disease (IBD) in the presence of infliximab by homogeneous liquid phase anti infliximab mobility shift assay. *J Crohns Colitis*. 2013;7:S231-S2.
91. Farrell RJ, Alsahli M, Jeen YT, Falchuk KR, Peppercorn MA, Michetti P. Intravenous hydrocortisone premedication reduces antibodies to infliximab in Crohn's disease: a randomized controlled trial. *Gastroenterology*. 2003;124(4):917-24.
92. Feagan BG, Singh S, Lockton S, Hauenstein S, Ohrmund L, Croner LJ, *et al.* Novel infliximab (IFX) and antibody-to-infliximab (ATI) assays are predictive of disease activity in patients with crohn's disease (CD). *Gastroenterology*. 2012;1):S114.
93. Frederiksen MT, Ainsworth MA, Brynskov J, Thomsen OO, Bendtzen K, Steenholdt C. Antibodies Against Infliximab Are Associated with De Novo Development of Antibodies to Adalimumab and Therapeutic Failure in Infliximab-to-Adalimumab Switchers with IBD. *Inflamm Bowel Dis*. 2014;20(10):1714-21.
94. Goldberg R, Beswick L, Van Langenberg D, Sally B, Rosella O, Gibson P, *et al.* Predictors of sub-therapeutic infliximab or adalimumab trough levels and anti-drug antibodies and their influence on therapeutic decisions. *J Crohns Colitis*. 2014;8:S223.
95. Greathead L, Kelleher P, Steel A. Development and validation of ELISA to measure serum anti TNFa levels. *J Crohns Colitis*. 2014;8:S97-S8.
96. Hauenstein S, Ohrmund L, Salbato J, Reddy R, McCowen K, Monk P, *et al.* Comparison of homogeneous mobility shift assay and solid phase elisa for the measurement of drug and anti-drug

antibody (ADA) levels in serum from patients treated with anti-TNF biologics. *Gastroenterology*. 2012;1):S538.
97. Hibi T, Sakuraba A, Watanabe M, Motoya S, Ito H, Sato N, *et al.* C-reactive protein is an indicator of serum infliximab level in predicting loss of response in patients with Crohn's disease. *J Gastroenterol*. 2014;49(2):254-62.
98. Imaeda H, Andoh A, Fujiyama Y. Development of a new immunoassay for the accurate determination of anti-infliximab antibodies in inflammatory bowel disease. *J Gastroenterol*. 2012;47(2):136-43.
99. Imaeda H, Takahashi K, Fujimoto T, Bamba S, Tsujikawa T, Sasaki M, *et al.* Clinical utility of newly developed immunoassays for serum concentrations of adalimumab and anti-adalimumab antibodies in patients with Crohn's disease. *J Gastroenterol*. 2014;49(1):100-9.
100. Imaeda H, Bamba S, Takahashi K, Fujimoto T, Ban H, Tsujikawa T, *et al.* Relationship between serum infliximab trough levels and endoscopic activities in patients with Crohn's disease under scheduled maintenance treatment. *J Gastroenterol*. 2014;49(4):674-82.
101. Kong JY, Bundell CS, Pawlik J, Hollingsworth PN, Forbes GM. Trough serum infliximab level, anti-infliximab antibody status and response to infliximab maintenance treatment in inflammatory bowel disease (IBD). *J Gastroenterol Hepatol*. 2011;26:59-60.
102. Kopylov U, Mazor Y, Yavzori M, Fudim E, Katz L, Coscas D, *et al.* Clinical utility of antihuman lambda chain-based enzyme-linked immunosorbent assay (ELISA) versus double antigen ELISA for the detection of anti-infliximab antibodies. *Inflamm Bowel Dis*. 2012;18(9):1628-33.
103. Levesque BG, Greenberg GR, Zou G, Sandborn WJ, Singh S, Hauenstein S, *et al.* A prospective cohort study to determine the relationship between serum infliximab concentration and efficacy in patients with luminal Crohn's disease. *Aliment Pharmacol Ther*. 2014;39(10):1126-35.
104. Marits P, Landucci L, Sundin U, Davidsdottir L, Nilsson J, Befrits R, *et al.* Trough s-infliximab and antibodies towards infliximab in a cohort of 79 IBD patients with maintenance infliximab treatment. *J Crohns Colitis*. 2014;8(8):881-9.
105. Marzo M, Armuzzi A, Felice C, Pugliese D, Andrisani G, Tolusso B, *et al.* Role of trough levels and antibodies to infliximab in the evaluation of loss of response and infusion reactions to infliximab therapy in inflammatory bowel disease. *Dig Liver Dis*. 2014;46:S77.
106. Mazor Y, Kopylov U, Hur DB, Almog R, Waterman M, Ben-Horin S, *et al.* Evaluating adalimumab drug and antibody levels as predictors of clinical and laboratory response in crohn's disease patients. *Gastroenterology*. 2013;1):S778.
107. Mazor Y, Almog R, Kopylov U, Ben Hur D, Blatt A, Dahan A, *et al.* Adalimumab drug and antibody levels as predictors of clinical and laboratory response in patients with Crohn's disease. *Aliment Pharmacol Ther*. 2014;40(6):620-8.
108. McTigue M, Sandborn W, Levesque B, Patel D. Clinical utility of next generation infliximab and antibodies to infliximab assay. *Am J Gastroenterol*. 2013;108:S527.
109. Nagore D, Ruiz Del Agua A, Pascual J, Llinares-Tello F, Herreros B, Martinez A. Therapeutic Cut-off of Infliximab in Patients with Inflammatory Bowel Diseases. Digestive Diseases Week (DDW); Forthcoming 2015 [cited 18/03/2015]; Available from: https://ddw2015.abstractcentral.com/login.
110. Nanda KS, Cheifetz AS, Moss AC. Impact of antibodies to infliximab on clinical outcomes and serum infliximab levels in patients with inflammatory bowel disease (IBD): a meta-analysis. *Am J Gastroenterol*. 2013;108(1):40-7; quiz 8.
111. Pallagi-Kunstar E, Farkas K, Szepes Z, Nagy F, Szucs M, Kui R, *et al.* Utility of serum TNF-alpha, infliximab trough level, and antibody titers in inflammatory bowel disease. *World J Gastroenterol*. 2014;20(17):5031-5.
112. Paul S, Tedesco ED, Marotte H, Phelip JM, Genin C, Roblin X. Interest of the dosage of serum concentration of infliximab and antibodies anti infliximab in the therapeutic response under infliximab in IBD. *Gastroenterology*. 2012;1):S354.
113. Paul S, Moreau AC, Del Tedesco E, Rinaudo M, Phelip JM, Genin C, *et al.* Pharmacokinetics of adalimumab in inflammatory bowel diseases: a systematic review and meta-analysis. *Inflamm Bowel Dis*. 2014;20(7):1288-95.

114. Roblin X, Marotte H, Rinaudo M, Del Tedesco E, Moreau A, Phelip JM, *et al.* Association between pharmacokinetics of adalimumab and mucosal healing in patients with inflammatory bowel diseases. *Clin Gastroenterol Hepatol*. 2014;12(1):80-4.e2.
115. Ruiz-Arguello B, del Agua AR, Torres N, Monasterio A, Martinez A, Nagore D. Comparison study of two commercially available methods for the determination of infliximab, adalimumab, etanercept and anti-drug antibody levels. *Clin Chem Lab Med*. 2013;51(12):e287-9.
116. Schatz SB, Prell C, Freudenberg F, Hajji M, Bufler P, Koletzko S. PA-G-0035 Comparison of different tests for determination of infliximab levels and antibodies against infliximab in pediatric IBD patients. The 46th Annual Meeting of The European Society of Paediatric Gastroenterology, Hepatology and Nutrition. *J Pediatr Gastroenterol Nutr*. 2013;56 suppl 2:19.
117. Semmler J, Pilch A, Armbruster F, Dignass A, Stein J. Development of a new immunoassay for the accurate determination of anti-infliximab antibodies in inflammatory bowel disease. *Clin Chem Lab Med*. 2013;51 (10):eA27-eA8.
118. Singh N, Rosenthal CJ, Melmed GY, Mirocha J, Farrior S, Callejas S, *et al.* Early infliximab trough levels are associated with persistent remission in pediatric patients with inflammatory bowel disease. *Inflamm Bowel Dis*. 2014;20(10):1708-13.
119. Steenholdt C, Bendtzen K, Brynskov J, Thomsen OO, Ainsworth MA. Cut-off levels and diagnostic accuracy of infliximab trough levels and anti-infliximab antibodies in Crohn's disease. *Scand J Gastroenterol*. 2011;46(3):310-8.
120. Steenholdt C, Ainsworth MA, Tovey M, Klausen TW, Thomsen OO, Brynskov J, *et al.* Comparison of techniques for monitoring infliximab and antibodies against infliximab in Crohn's disease. *Ther Drug Monit*. 2013;35(4):530-8.
121. Steenholdt C, Bendtzen K, Brynskov J, Thomsen OO, Ainsworth MA. Clinical implications of measuring drug and anti-drug antibodies by different assays when optimizing infliximab treatment failure in Crohn's disease: post hoc analysis of a randomized controlled trial. *Am J Gastroenterol*. 2014;109(7):1055-64.
122. Steenholdt C, Brynskov J, Thomsen OO, Munck LK, Fallingborg J, Christensen LA, *et al.* Individualised therapy is more cost-effective than dose intensification in patients with Crohn's disease who lose response to anti-TNF treatment: a randomised, controlled trial. *Gut*. 2014;63(6):919-27.
123. Steenholdt C BJ, Thomsen OØ, Munck LK, Fallingborg J, Christensen LA, Pedersen G, Kjeldsen J, Jacobsen BA, Oxholm AS, Kjellberg J, Bendtzen K, Ainsworth MA. Individualized therapy is a Long-Term Cost-Effective Method Compared to Dose Intensification in Crohn's Disease Patients Failing Infliximab. *Dig Dis Sci*. 2015.
124. Ungar B, Anafy A, Kopylov U, Ron Y, Yanai H, Dotan I, *et al.* The clinical and immunological significance of low level of infliximab in the absence of anti-infliximab antibodies in patients with IBD. *Gastroenterology*. 2014;1):S-245.
125. Vande Casteele N, Gils A, Singh S, Ohrmund L, Hauenstein S, Rutgeerts P, *et al.* Antibody response to infliximab and its impact on pharmacokinetics can be transient. *Am J Gastroenterol*. 2013;108(6):962-71.
126. Vande Casteele N, Peeters M, Compernolle G, Ferrante M, Van Assche GA, Vermeire S, *et al.* TNF-responsive cellular based assay reveals neutralizing capacity of anti-adalimumab antibodies in crohn's disease and ulcerative colitis patients. *Gastroenterology*. 2014;1):S-242.
127. Vaughn BP, Martinez-Vazquez M, Patwardhan VR, Moss AC, Sandborn WJ, Cheifetz AS. Proactive therapeutic concentration monitoring of infliximab may improve outcomes for patients with inflammatory bowel disease: results from a pilot observational study. *Inflamm Bowel Dis*. 2014;20(11):1996-2003.
128. Wang SL, Ohrmund L, Singh S. Measurement of human anti-chimeric antibodies (Haca) and infliximab levels in patient serum using a novel homogeneous assay. *Gastroenterology*. 2010;1):S684-S5.
129. Wang SL, Ohrmund L, Hauenstein S, Salbato J, Reddy R, Monk P, *et al.* Evaluation of a novel homogeneous mobility shift assay for the measurement of human antibodies-To-Infliximab and infliximab levels in Patient serum. *Am J Gastroenterol*. 2011;106:S475-S6.

130. Wang SL, Ohrmund L, Hauenstein S, Salbato J, Reddy R, Monk P, *et al.* Development and validation of a homogeneous mobility shift assay for the measurement of infliximab and antibodies-to-infliximab levels in patient serum. *J Immunol Methods*. 2012;382(1-2):177-88.
131. Ward MG, Kariyawasam VC, Mogan SB, Blaker PA, Patel KP, Pantelidou M, *et al.* Clinical utility of measuring adalimumab trough levels and antibodies to adalimumab in patients with inflammatory bowel diseases. *J Gastroenterol Hepatol*. 2013;28:100-1.
132. West RL, Zelinkova Z, Wolbink GJ, Kuipers EJ, Stokkers PC, van der Woude CJ. Immunogenicity negatively influences the outcome of adalimumab treatment in Crohn's disease. *Aliment Pharmacol Ther*. 2008;28(9):1122-6.
133. Yanai H, Mlynarsky L, Ron Y, Ben Yehoyada M, Yeshuron D, Santo EM, *et al.* The questionable value of infliximab trough levels during prolonged maintenance therapy. *Gastroenterology*. 2012;1):S788-S9.
134. Yarur AJ, Deshpande AR, Sussman DA, Hauenstein S, Lockton S, Barkin JS, *et al.* Serum adalimumab levels and antibodies correlate with endoscopic intestinal inflammation and inflammatory markers in patients with inflammatory bowel disease. *Gastroenterology*. 2013;1):S774-S5.
135. Rispens T, van der Kleij D. Reply to Ruiz-Arguello et al.: Comparison study of two commercially available methods for the determination of infliximab, adalimumab, etanercept and anti-drug antibody levels. *Clin Chem Lab Med*. 2013;51(12):e291-2.
136. Garces S, Demengeot J, Benito-Garcia E. Clinical impact of immunogenicity of infliximab, adalimumab and etanercept: A systematic review of the literature with a meta-analysis. *Annals of the Rheumatic Disease*. 2013;71.
137. Lichtenstein GR. Comprehensive review: antitumor necrosis factor agents in inflammatory bowel disease and factors implicated in treatment response. *Therap Adv Gastroenterol*. 2013;6(4):269-93.
138. Yanai H, Hanauer SB. Assessing response and loss of response to biological therapies in IBD. *Am J Gastroenterol*. 2011;106(4):685-98.
139. Vande Casteele N, Khanna R, Levesque BG, Stitt L, Zou GY, Singh S, *et al.* The relationship between infliximab concentrations, antibodies to infliximab and disease activity in Crohn's disease. *Gut*. 2014.
140. Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]: The Cochrane Collaboration; 2011.
141. Steenholdt C, Bendtzen K, Brynskov J, Thomsen OØ, Munck LK, Christensen LA, *et al.* Changes in Serum Trough Levels of Infliximab During Treatment Intensification but not in Anti-infliximab Antibody Detection are Associated with Clinical Outcomes after Therapeutic Failure in Crohn's Disease. *J Crohns Colitis*. 2015;9(3):238-45.
142. Pepe NS. The statistical evaluation of medical tests for classification and prediction. New York: Oxford University Press; 2003.
143. Harbord RM, Deeks JJ, Egger M, Whiting P, Sterne JA. A unification of models for meta-analysis of diagnostic accuracy studies. *Biostatistics*. 2007;8(2):239-51.
144. Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R, *et al.* Review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health Technol Assess*. 2004;8(36):1-158.
145. Blackhouse G, Assasi N, Xie F, Marshall J, Irvine EJ, Gaebel K, *et al.* Canadian cost-utility analysis of initiation and maintenance treatment with anti-TNF-alpha drugs for refractory Crohn's disease. *J Crohns Colitis*. 2012;6(1):77-85.
146. Bodger K, Kikuchi T, Hughes D. Cost-effectiveness of biological therapy for Crohn's disease: Markov cohort analyses incorporating United Kingdom patient-level cost data. *Aliment Pharmacol Ther*. 2009;30(3):265-74.
147. Kaplan GG, Hur C, Korzenik J, Sands BE. Infliximab dose escalation vs. initiation of adalimumab for loss of response in Crohn's disease: a cost-effectiveness analysis. *Aliment Pharmacol Ther*. 2007;26(11-12):1509-20.

148. Velayos FS, Kahn JG, Sandborn WJ, Feagan BG. A test-based strategy is more cost effective than empiric dose escalation for patients with Crohn's disease who lose responsiveness to infliximab. *Clin Gastroenterol Hepatol*. 2013;11(6):654-66.
149. Silverstein MD, Loftus EV, Sandborn WJ, Tremaine WJ, Feagan BG, Nietert PJ, *et al*. Clinical course and costs of care for Crohn's disease: Markov model analysis of a population-based cohort. *Gastroenterology*. 1999;117(1):49-57.
150. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, *et al*. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. *Int J Technol Assess Health Care*. 2013;29(2):117-22.
151. Juillerat P, Sokol H, Froehlich F, Yajnik V, Beaugerie L, Lucci M, *et al*. Factors associated with durable response to infliximab in Crohn's disease 5 years and beyond: a multicenter international cohort. *Inflamm Bowel Dis*. 2015;21(1):60-70.
152. van der Have M, van der Aalst KS, Kaptein AA, Leenders M, Siersema PD, Oldenburg B, *et al*. Determinants of health-related quality of life in Crohn's disease: a systematic review and meta-analysis. *J Crohns Colitis*. 2014;8(2):93-106.
153. Ma C, Huang V, Fedorak DK, Kroeker KI, Dieleman LA, Halloran BP, *et al*. Crohn's disease outpatients treated with adalimumab have an earlier secondary loss of response and requirement for dose escalation compared to infliximab: a real life cohort study. *J Crohns Colitis*. 2014;8(11):1454-63.
154. Sandborn WJ, Rutgeerts P, Enns R, Hanauer SB, Colombel JF, Panaccione R, *et al*. Adalimumab induction therapy for Crohn disease previously treated with infliximab: a randomized trial. *Ann Intern Med*. 2007;146(12):829-38.
155. Nguyen GC, Nugent Z, Shaw S, Bernstein CN. Outcomes of patients with Crohn's disease improved from 1988 to 2008 and were associated with increased specialist care. *Gastroenterology*. 2011;141(1):90-7.
156. Gordon M, Taylor K, Akobeng AK, Thomas AG. Azathioprine and 6-mercaptopurine for maintenance of surgically-induced remission in Crohn's disease. *Cochrane Database Syst Rev*. 2014;8:Cd010233.
157. Rutgeerts P, D'Haens G, Targan S, Vasiliauskas E, Hanauer SB, Present DH, *et al*. Efficacy and safety of retreatment with anti-tumor necrosis factor antibody (infliximab) to maintain remission in Crohn's disease. *Gastroenterology*. 1999;117(4):761-9.
158. BNF 66: September 2013 - March 2014. London: BMJ Group, Pharmaceutical Press; 2013.
159. Department of Health. NHS reference costs 2013 to 2014. 2014 [cited 24/03/2015]; Available from: https://www.gov.uk/government/publications/nhs-reference-costs-2013-to-2014.
160. Curtis L. Unit Costs of Health and Social Care. 2014 [cited 24/03/2015]; Available from: http://www.pssru.ac.uk/project-pages/unit-costs/2014/.
161. Marchetti M, Liberato NL. Biological therapies in Crohn's disease: are they cost-effective? A critical appraisal of model-based analyses. *Expert Rev Pharmacoecon Outcomes Res*. 2014;14(6):815-24.
162. Office for National Statistics. National Life Tables, United Kingdom, 1980-82 to 2011-13. 2014 [cited 31/03/2015]; Available from: http://www.ons.gov.uk/ons/search/index.html?newquery=National+Life+Tables%2C+United+Kingdom%2C+1980-82+to+2011-13.
163. Rai T, Navaneethan U, Dalal D, Lashner B, Shen B. Clinical implications of measuring infliximab levels and human anti-chimeric antibodies in patients with inflammatory bowel disease. *Am J Gastroenterol*. 2012;107:S634.
164. Armbruster S, Ally M, Maydonovitch C, Betteridge J, Veerappan G. The use of human anti-chimeric antibody (HACA) and infliximab levels in the management of inflammatory bowel disease. *Am J Gastroenterol*. 2012;107:S641.
165. Vande Casteele N, Drake K, Hauenstein S, Levesque BG, Singh S, Sandborn W. Infliximab and antibody to infliximab concentrations in 7,613 patients shows indication for testing, association with loss of response and provides new insights into binding characteristics of anti-drug antibodies. *Gastroenterology*. 2014;1):S-242.

166. Wolf DC, Lockton S, Hauenstein S, Carroll S, Singh S, Chuang E. A multi-center observational study in community gastroenterology practices evaluating the clinical usage of testing for serum levels of infliximab and antibodies to infliximab. *Gastroenterology*. 2013;1):S423.
167. Turon J, Langseder A, Irizarry R, Ahuja K, Rosh JR. Clinical outcome of pediatric IBD patients after measurement of infliximab drug and anti-drug antibody levels. *Gastroenterology*. 2013;1):S531.
168. Turvill J. Mapping of Crohn's disease outcomes to faecal calprotectin levels in patients maintained on biologic therapy. *Frontline Gastroenterology*. 2014;5(3):167-75.
169. Bar-yoseph H, Chowers Y, Ben-Horin S, Waterman M. Infliximab is more immunogenic and reaches lower trough levels in ulcerative colitis patients compared to crohn's disease patients. *Gastroenterology*. 2013;1):S780.
170. Vester-Andersen MK, Vind I, Prosberg MV, Bengtsson BG, Blixt T, Munkholm P, *et al.* Hospitalisation, surgical and medical recurrence rates in inflammatory bowel disease 2003-2011-a Danish population-based cohort study. *J Crohns Colitis*. 2014;8(12):1675-83.
171. Ramadas AV, Gunesh S, Thomas GA, Williams GT, Hawthorne AB. Natural history of Crohn's disease in a population-based cohort from Cardiff (1986-2003): a study of changes in medical treatment and surgical resection rates. *Gut*. 2010;59(9):1200-6.
172. Armuzzi A, Felice C, Papa A, Marzo M, Pugliese D, Andrisani G, *et al.* Prevention of postoperative recurrence with azathioprine or infliximab in patients with Crohn's disease: an open-label pilot study. *J Crohns Colitis*. 2013;7(12):e623-9.
173. Latimer N. NICE DSU technical support document 14: Survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data. Decision Support Unit, ScHARR, University of Sheffield; 2013 [cited 14/04/2015]; Available from: http://www.nicedsu.org.uk/NICE%20DSU%20TSD%20Survival%20analysis.updated%20March%202013.v2.pdf.

8 APPENDIX

8.1 Appendix 1 Details of manufacturers ELISA kits

These details are taken from the NICE final scope for this diagnostic assessment and are based on information supplied to NICE by the kit manufacturers.

LISA-TRACKER ELISA kits (*Theradiag / Alpha Laboratories*)

LISA-TRACKER assay kits are enzyme linked immunosorbent assays (ELISAs) for the quantitative determination of TNFα inhibitor levels and antibodies against TNFα inhibitor. There are 6 LISA-TRACKER ELISA kits relevant to this assessment (Table 39). Two of these kits measure the levels of free anti-drug antibodies, 2 kits measure the levels of free TNFα inhibitor and 2 kits measure the levels of both free anti-drug antibodies and TNFα inhibitor.

Table 39 LISA-TRACKER ELISA kits

Name (code)	Detects	Microplate pre-coat	Secondary reagent	Incubation times
LISA-TRACKER Adalimumab (LTA002)	Free adalimumab	TNF-α	Biotinylated anti-human IgG antibody	1 hour; 1 hour; 30 mins; 15 mins
LISA-TRACKER Infliximab (LTI002)	Free infliximab	TNF-α		
LISA-TRACKER anti-Adalimumab (LTA003)	Free anti- adalimumab antibodies	Adalimumab	Biotinylated adalimumab	1 hour; 1 hour; 30 mins; 15 mins
LISA-TRACKER anti-Infliximab (LTI003)	Free anti- infliximab antibodies	Infliximab	Biotinylated infliximab	
LISA-TRACKER Duo Adalimumab (LTA005)	As above; the Duo Adalimumab kit consists of a LISATRACKER Adalimumab kit and a LISA-TRACKER antiAdalimumab kit			
LISA-TRACKER Duo Infliximab (LTI005)	As above – the Duo Infliximab kit consists of a LISATRACKER Infliximab kit and a LISA-TRACKER antiInfliximab kit			
Note: There are 2 additional LISA-TRACKER ELISA kits which are available in some European countries, but not the UK. The LISA-TRACKER Premium Adalimumab and the LISA-TRACKER Premium Infliximab both measure 3 parameters: TNFα inhibitor, TNF-α levels and anti-drug antibody levels.				

The LISA-TRACKER ELISA kits consist of pre-coated strips of microtitre plate (96 wells), reagents, wash buffer, standards and controls. The assays can be run simultaneously or individually on any manual or automated standard ELISA based processor platform. The assay procedure is similar for all the assays but the reagents used are dependent on whether the ELISA is detecting levels of TNFα inhibitor or levels of anti-drug antibody in the patient's sera.

Detecting levels of TNFα inhibitor

Patient samples, the standards and controls are added to the pre-coated microtitre plate. The TNFα inhibitor (adalimumab or infliximab) present in the patient samples, standards and controls, binds to the coated wells during the first incubation step and any unbound substances are removed in a

subsequent washing step. The secondary reagent is then added which binds to the TNFα inhibitor attached to the coated plate. Any unbound reagent is removed by a second wash step before peroxidase labelled streptavidin is added to the plate. Streptavidin binds to the biotin-labelled antibody complex and any unbound streptavidin is removed by a final wash step. Finally, a chromogenic substrate solution is added and colour develops in proportion to the amount of TNFα inhibitor present in the patient sample. The colour change reaction is stopped by the addition of an acid solution and the optical density is read by a spectrophotometer. A range of calibration is determined based on the optical density of the standards and this is used to define the quantity of drug in each sample. The limits of detection are presented in Table 40.

Detecting levels of antibodies to TNFα inhibitor

Patient samples, the standards and controls are added to the pre-coated microtitre plate. The free anti-infliximab antibodies or free anti-adalimumab antibodies present in the patient samples, standards and controls, bind to the coated wells during the first incubation step and any unbound substances are removed in a subsequent washing step. The secondary reagent is then added which binds to the anti-drug antibodies attached to the coated plate. Any unbound reagent is removed by a second wash step before peroxidase labelled streptavidin is added to the plate. Streptavidin binds to the biotin-labelled complex and any unbound streptavidin is removed by a final wash step. Finally, a chromogenic substrate solution is added and colour develops in proportion to the amount of anti-drug antibodies present in the patient sample. The colour change reaction is stopped by the addition of an acid solution and the optical density is read by a spectrophotometer. A range of calibration is determined based on the optical density of the standards and this is used to define the quantity of antibodies to TNFα inhibitor in each patient sample. The limits of detection and assay ranges are presented in Table 40.

Table 40 Interpretation of results, limits of detection and assay ranges for LISA-TRACKER assays

Name (code)	Results interpretation	Limit of detection	Assay range
LISA-TRACKER Adalimumab (LTA002)	Quantitative. Generation of standard curve and determination of drug level in µg/mL	0.1 µg/mL	0.1 to 8 µg/mL
LISA-TRACKER Infliximab (LTI002)		0.1 µg/mL	0.1 to 8 µg/mL
LISA-TRACKER anti-Adalimumab (LTA003)	Quantitative. Generation of standard curve and determination of ADAb level in ng/mL	10 ng/mL	10 to 160 ng/mL
LISA-TRACKER anti-Infliximab (LTI003)		10 ng/mL	10 to 200 ng/mL

TNFα-Blocker ELISA Kits (Immundiagnostik AG)

There are 6 Immundiagnostik ELISA kits relevant to this assessment, which are distributed in the UK by BioHit Healthcare Ltd (Table 41). Two of these kits measure the levels of free anti-drug antibodies, 2 kits measure the levels of total anti-drug antibodies (free antibodies and antibodies already bound to the drug) and 2 kits measure the levels of free TNFα inhibitor.

Table 41 Immundiagnostik ELISA kits

Name (code)	Detects	Microplate pre-coat	Secondary reagent	Incubation times
Immundiagnostik TNFα- Blocker monitoring, infliximab drug level (e.g. Remicade®) ELISA (K9655)	Free infliximab	Monoclonal anti-infliximab antibody	Peroxidase labelled antibody	1 hour; 1 hour; 10 to 20 mins
Immundiagnostik TNFα- Blocker monitoring, adalimumab drug level (e.g. Humira®) ELISA (K9657)	Free adalimumab	Monoclonal anti-adalimumab antibody	Peroxidase labelled antibody	
Immundiagnostik TNFα- Blocker ADA, antibodies against infliximab (e.g. Remicade®) ELISA (K9650)	Free anti-infliximab antibodies	Infliximab F(ab)2 fragments	Peroxidase labelled infliximab	2 x 15 mins; 16 to 20 hours; 1 hour; 10 to 20 mins
Immundiagnostik TNFα- Blocker ADA, antibodies against adalimumab (e.g. Humira®) ELISA (K9652)	Free anti-adalimumab antibodies	Adalimumab F(ab)2 fragments	Peroxidase labelled adalimumab	16 to 20 hours; 1 hour; 10 to 20 mins
Immundiagnostik TNFα- Blocker ADA, TOTAL antibodies against infliximab (e.g. Remicade®) ELISA (K9654)	Total anti-infliximab antibodies	Streptavidin	N/A	20 mins; 1 hour; 1.5 hours; 10 to 20 mins
Immundiagnostik TNFα- Blocker ADA, TOTAL antibodies against adalimumab (e.g. Humira®) ELISA (K9651)	Total anti-adalimumab antibodies	Streptavidin	N/A	

The kits consist of strips of pre-coated microtitre plate (96 wells), reagents, buffers, standards (drug level ELISAs only) and controls. The ELISAs can be performed manually or run on an automated ELISA processor. The two ELISAs that measure free infliximab or adalimumab (K9655 and K9657) follow a standard ELISA procedure for detecting levels of TNFα inhibitor as described in section 1.4.1, except that the secondary reagent is directly labelled with peroxidase and therefore there is no biotin-streptavidin binding step. The two ELISAs that measure free anti-adalimumab antibodies or free anti-infliximab antibodies (K9650 and K9652) follow a standard ELISA procedure for detecting levels of antibodies to TNFα inhibitor as described in section 1.4.1, except that the secondary reagent is directly labelled with peroxidase and therefore there is no biotin-streptavidin binding step. Further, standards are not used in the anti-drug antibody ELISAs, therefore the results are interpreted semi-quantitatively using a cut-off control interpreted semi-quantitatively using a cut-off control. Details on the interpretation of results, limits of detection and assay measurement ranges are presented in Table 42.

The TOTAL anti-drug antibody ELISA kits (K9654 and K9651) enable the measurement of anti-drug antibodies in the presence of TNFα inhibitor. During sample preparation immune complexes between anti-drug antibodies and adalimumab or infliximab are dissociated using an acidic buffer. Biotinylated and peroxidase-labelled adalimumab or infliximab are added to the sample and form complexes with the anti-drug antibodies. The complexes bind via biotin to the streptavidin coated plate. Following a wash step a chromogenic substrate is added, the colour change reaction is stopped by the addition of an acid solution and the optical density is read by a spectrophotometer.

Table 42 Interpretation of results, limits of detection and assay ranges for the Immundiagnostik ELISAs

Name (code)	Results interpretation	Limit of blank	Assay range
Immundiagnostik TNFα-Blocker monitoring, infliximab drug level (e.g. Remicade®) ELISA (K9655)	Quantitative. Generation of standard curve and determination of drug level in µg/mL	2.0 ng/mL	0.4 to 45 µg/mL
Immundiagnostik TNFα-Blocker monitoring, adalimumab drug level (e.g. Humira®) ELISA (K9657)		2.3 ng/mL	0.4 to 45 µg/mL
Immundiagnostik TNFα-Blocker ADA, antibodies against infliximab (e.g. Remicade®) ELISA (K9650)	Semi-quantitative. Evaluated by a cut- off control (10 AU/mL) to give a positive or negative result	5.787 AU/mL	N/A
Immundiagnostik TNFα-Blocker ADA, antibodies against adalimumab (e.g. Humira®) ELISA (K9652)		N/A	N/A
Immundiagnostik TNFα-Blocker ADA, TOTAL antibodies against infliximab (e.g. Remicade®) ELISA (K9654)	Semi quantitative. Evaluated by a cut- off control (10 AU/mL) to give a positive or negative result	2.653 AU/mL	N/A
Immundiagnostik TNFα-Blocker ADA, TOTAL antibodies against adalimumab (e.g. Humira®) ELISA (K9651)		2.765 AU/mL	N/A

Promonitor ELISA Kits (Proteomika)

There are 4 Promonitor ELISA kits relevant to this assessment (Table 43). Two of these kits measure the levels of free anti-drug antibodies and 2 kits measure the levels of free TNFα inhibitor.

Table 43 Promonitor ELISA kits

Name (code)	Detects	Microplate pre-coat	Secondary reagent	Incubation times
Promonitor- ADL ELISA (5080230000)	Free adalimumab	Anti-adalimumab human monoclonal antibody	Peroxidase labelled anti- adalimumab monoclonal antibody	1 hour; 1 hour; 25 to 35 mins
Promonitor- IFX ELISA (5060230000)	Free infliximab	Anti-TNF-α human monoclonal antibody bound to human recombinant TNF-α	Peroxidase labelled anti- infliximab monoclonal antibody	1 hour; 1 hour; 10 to 20 mins
Promonitor- ANTI-ADL ELISA (5090230000)	Free anti-adalimumab antibodies	Adalimumab	Peroxidase labelled adalimumab	1 hour; 1 hour; 25 to 35 mins
Promonitor- ANTI-IFX ELISA (5070230000)	Free anti-infliximab antibodies	Infliximab	Peroxidase labelled infliximab	1 hour; 1 hour; 25 to 35 mins

The kits consist of strips of pre-coated microtitre plate (96 wells), reagents, buffers, standards, controls and ELISA cover films. The IFX ELISA and ADL ELISA follow a standard ELISA procedure for detecting levels of TNFα inhibitor as described in section 1.4.1, except that the secondary reagent is directly labelled with peroxidase and therefore there is no biotin-streptavidin binding step. The ANTI-IFX ELISA and the ANTI-ADL ELISA follow a standard ELISA procedure for detecting levels of antibodies to TNFα inhibitor as described in section 1.4.1, except that the secondary reagent is directly labelled with peroxidase and therefore there is no biotin-streptavidin binding step. The ELISAs can be performed manually or run on an automated ELISA processor. Details on the interpretation of results, the assay ranges and limits of quantification are presented in Table 44.

Table 44 Limits of quantification and assay ranges for Promonitor ELISAs

Name	Results interpretation	Limit of quantification	Assay range
Promonitor- ADL ELISA	Semi-quantitative. Evaluated using a cut-off value (0.024 µg/mL for adalimumab, 0.035 µg/mL for infliximab) to give a positive or negative result	2.9 ng/mL	0.024 to 12 µg/mL
Promonitor- IFX ELISA	Quantitative. Generation of standard curve and determination of drug level in µg/mL	1.7 ng/mL	0.035 to 14.4 µg/mL
Promonitor-ANTI-ADL ELISA	Semi-quantitative. Evaluated using a cut-off value (10 AU/mL for anti- adalimumab antibodies, 5 AU/mL for anti-infliximab antibodies) to give a positive or negative result	3.7 AU/mL	3.5 to 2000 AU/mL
Promonitor-ANTI-IFX ELISA	Quantitative. Generation of standard curve and determination of anti-drug antibody level in AU/mL	2 AU/mL	2 to 1440 AU/mL

8.2 Appendix 2 cell reporter assays and mobility shift assays

Cell reporter assays

The reporter cells are genetically engineered to contain genes for two light producing enzymes "*luciferases*" (one from the firefly which can generate red light, and one from the sea pansy which can generate blue light). The firefly gene is under the control of a TNFα signalling pathway so that when the cells are incubated in the presence of TNFα they synthesise the enzyme, after a standard incubation time appropriate substrates for the enzyme are added and the emitted red light measured with a luminometer. If anti-TNFα is present the TNFα response is partially quenched and the quenching estimated. If anti-drug antibodies are present, quenching by anti-TNFα is reduced and this can be measured. The sea pansy gene is expressed during incubation after which appropriate substrates are added and the blue light emitted measured in the luminometer. The usefulness of the blue light measure is that it allows "normalisation" of the red light emission as interfering agents in patient blood samples equally affect both firefly and sea pansy systems. Requirements in addition to appropriate cell reporter cultures and reagents include requirement for a luminometer (although these are not necessarily routinely available) and equipment for culture of growth arrested genetically engineered cells under controlled conditions (oxygen, CO_2, humidity). These assays appear to be available as a service and commercial kits are not available.

Mobility shift assays

The mobility shift is exploited using size exclusion HPLC. The mobility shift assay is a liquid phase assay based on size exclusion HPLC (SE-HPLC) which separates free probe (small size) from probe in an immune-complex (large size). The anti-drug antibody assays use fluorescent-dye-labelled anti-TNFα (D*) as the probe. In the presence of antibodies to anti-TNFα some D* form immune complexes with these (D*-anti-drug antibody complexes) and will exhibit a mobility shift on the SE-HPLC column relative to the D* which remains free. The amount of D* shifted to greater mobility is proportional to the amount of anti-drug antibody present. The amount of dye (*) present in the eluent stream coming from the HPLC column at different mobilities is measured with a fluorimeter (Figure 35).

The anti-TNFα assay uses fluorescent-dye-labelled TNFα (TNFα*) as the probe; in the presence of anti-TNFα some TNFα* forms immune-complexes with the anti-TNFα and these have greater mobility on the SE-HPLC than the free TNFα*. The amount of TNFα* shifted to greater mobility is proportional to the amount of anti-TNFα present. The amount of dye (*) present in the eluent stream coming from the HPLC column at different mobilities is measured with a fluorimeter.

In measuring anti-drug antibody the patient sample is subjected to an acid step which "unbinds" bound anti-TNFα and anti-drug antibody so that all anti-TNFα and anti-drug antibody are "free"; after neutralisation the sample is incubated with fluorescent-dye-labelled anti-TNFα (D*) as described above. Some D* will form immune complexes with the sample anti-drug antibodies (D*-anti-drug antibody complexes) and these have a different mobility on SE-HPLC than D* thus the mobility of some of the D* is shifted, the proportion of D* shifted is dependent on the level of anti-drug antibody in the sample. This assay is theoretically a candidate for a gold standard. It is more likely to measure all classes of anti-drug antibodies and also total anti-drug antibody than the ELISAs and is probably less prone to interference from serum components in samples. It does not use hazardous materials. This assay appears to be only available as a service and may not be practicable for use for UK patients. Setting up mobility shift assays in a hospital laboratory and constructing requisite reagents would be a major and expensive undertaking.

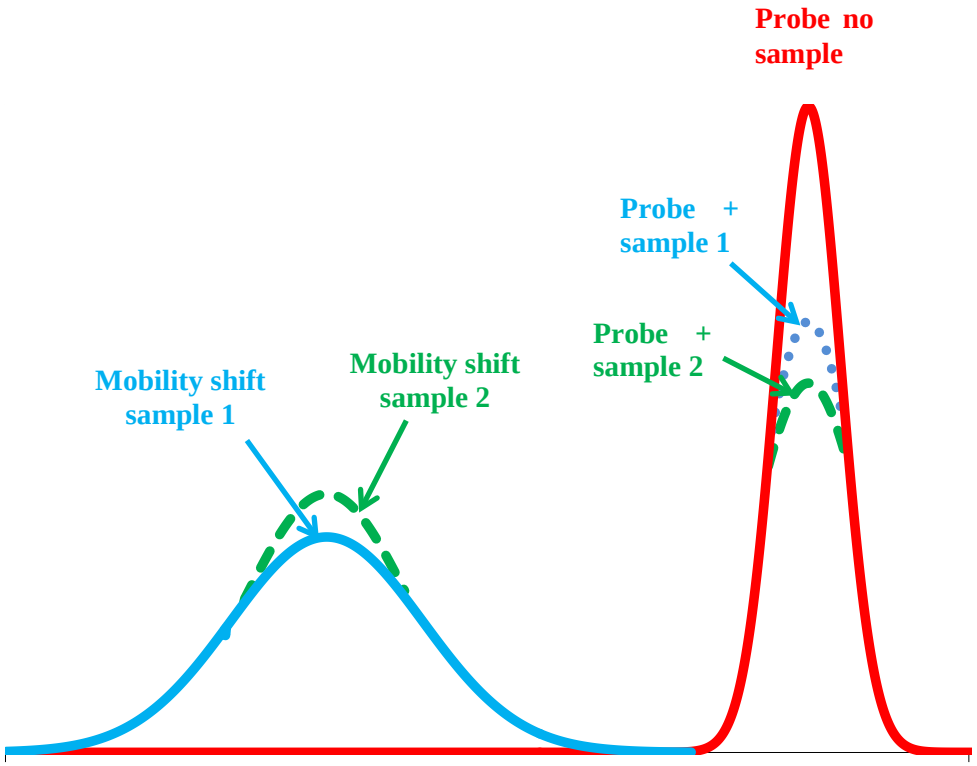


Figure 35 Illustration of chromatograms obtained after size exclusion of probe-labelled samples after size exclusion using HPLC. *The vertical axis represents the fluorescence signal*

8.3 Appendix 3 Search strategies

Clinical Effectiveness: database searches

Ovid MEDLINE(R) 1946 to October Week 2 2014, searched on 22/10/2014

1	adalimumab.mp.	3597
2	ADA.tw.	7105
3	infliximab.mp.	8842
4	IFX.tw.	326
5	((anti-TNF* or antiTNF* or TNF*) adj2 inhibitor*).mp.	2577
6	anti* tumo?r* necrosis* factor*.mp.	3007
7	Tumor Necrosis Factor-alpha/ and Antibodies, Monoclonal/	7682
8	anti* drug* antibod*.tw.	186
9	ADAb.tw.	19
10	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9	24181
11	lisa* tracker*.mp.	1
12	(immundiagnostik* or immunodiagnostik* or immunediagnostik*).mp.	159
13	(proteomika* or promonitor*).mp.	13
14	exp Enzyme-Linked Immunosorbent Assay/	129174
15	enzyme* link* immunoassay*.mp.	2873
16	enzyme* link* immuno* assay*.mp.	158537
17	ELISA*.mp.	113426
18	11 or 12 or 13 or 14 or 15 or 16 or 17	205224
19	*Radioimmunoassay/	7091
20	(radioimmuno* or radio immuno* or radio-immuno*).mp.	101819
21	RIA.tw.	17353
22	reporter* gene* assay*.mp.	3663
23	RGA.tw.	336
24	semi* fluid* phase* enzyme* immuno*.mp.	0
25	EIA.tw.	8288
26	((homogenous* or homogeneous*) adj1 mobilit* shift* assay*).mp.	4
27	HMSA.tw.	62
28	(Biomonitor* or iLite).tw.	4102
29	(Matriks* Biotek* or Shikari*).mp.	2
30	(Prometheus* or Anser IFX or Anser ADA).mp.	258
31	19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30	124775
32	((monitor* or pharmacokinetic* or measur* or level* or concentration*) adj3 (adalimumab or ADA or infliximab or IFX or Anti-TNF* or Anti-Tumour Necrosis Factor*)).mp.	1087
33	Inflammatory Bowel Diseases/	14444

34	Crohn Disease/	31596
35	crohn*.tw.	32370
36	inflammator* bowel* disease*.tw.	26840
37	IBD.tw.	11936
38	33 or 34 or 35 or 36 or 37	58401
39	(((monitor* or pharmacokinetic* or measur* or level* or concentration*) adj3 (adalimumab or infliximab or Anti-TNF* or AntiTNF* or Anti-Tumour Necrosis Factor*)) and (correlat* or associat* or test performance)).mp.	218
40	10 and 18 and 38	93
41	10 and 31 and 38	19
42	32 and 38	157
43	39 or 40 or 41 or 42	367
44	Animals/ not Humans/	3983380
45	43 not 44	349

Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations October 21, 2014, searched on 22/10/2014

1	adalimumab.mp.	469
2	ADA.tw.	426
3	infliximab.mp.	814
4	IFX.tw.	69
5	((anti-TNF* or antiTNF* or TNF*) adj2 inhibitor*).mp.	308
6	anti* tumo?r* necrosis* factor*.mp.	323
7	anti* drug* antibod*.tw.	39
8	ADAb.tw.	1
9	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8	1824
10	lisa* tracker*.mp.	0
11	(immundiagnostik* or immunodiagnostik* or immunediagnostik*).mp.	2
12	(proteomika* or promonitor*).mp.	0
13	enzyme* link* immunoassay*.mp.	133
14	enzyme* link* immuno* assay*.mp.	3996
15	ELISA*.mp.	8044
16	10 or 11 or 12 or 13 or 14 or 15	10101
17	(radioimmuno* or radio immuno* or radio-immuno*).mp.	1176
18	RIA.tw.	386
19	reporter* gene* assay*.mp.	240
20	RGA.tw.	47

21	semi* fluid* phase* enzyme* immuno*.mp.	0
22	EIA.tw.	357
23	((homogenous* or homogeneous*) adj1 mobilit* shift* assay*).mp.	0
24	HMSA.tw.	5
25	(Biomonitor* or iLite).tw.	343
26	(Matriks* Biotek* or Shikari*).mp.	1
27	(Prometheus* or Anser IFX or Anser ADA).mp.	23
28	17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27	2386
29	((monitor* or pharmacokinetic* or measur* or level* or concentration*) adj3 (adalimumab or ADA or infliximab or IFX or Anti-TNF* or Anti-Tumour Necrosis Factor*)).mp.	112
30	crohn*.tw.	2478
31	inflammator* bowel* disease*.tw.	2627
32	IBD.tw.	1480
33	30 or 31 or 32	4400
34	(((monitor* or pharmacokinetic* or measur* or level* or concentration*) adj3 (adalimumab or infliximab or Anti-TNF* or AntiTNF* or Anti-Tumour Necrosis Factor*)) and (correlat* or associat* or test performance)).mp.	30
35	9 and 16 and 33	15
36	9 and 28 and 33	0
37	29 and 33	35
38	34 or 35 or 36 or 37	57

Embase Classic+Embase 1947 to 2014 Week 42, searched on 22/10/2014

1	adalimumab.tw.	7379
2	*adalimumab/	3997
3	ADA.tw.	10848
4	infliximab.tw.	13600
5	*infliximab/	8056
6	IFX.tw.	1722
7	((anti-TNF* or antiTNF* or TNF*) adj2 inhibitor*).tw.	4663
8	anti* tumo?r* necrosis* factor*.tw.	4171
9	*tumor necrosis factor alpha inhibitor/	1283
10	anti* drug* antibod*.tw.	469
11	ADAb.tw.	44
12	*drug antibody/	1528
13	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12	35630
14	lisa* tracker*.tw.	11
15	(immundiagnostik* or immunodiagnostik* or immunediagnostik*).tw.	74

16	(proteomika* or promonitor*).tw.	27
17	*enzyme linked immunosorbent assay/	14622
18	enzyme* link* immunoassay*.tw.	3275
19	enzyme* link* immuno* assay*.tw.	71923
20	ELISA*.tw.	166866
21	14 or 15 or 16 or 17 or 18 or 19 or 20	207373
22	*radioimmunoassay/	17240
23	(radioimmuno* or radio immuno* or radio-immuno*).tw.	74895
24	RIA.tw.	20769
25	reporter* gene* assay*.tw.	4396
26	RGA.tw.	400
27	semi* fluid* phase* enzyme* immuno*.tw.	1
28	EIA.tw.	10836
29	((homogenous* or homogeneous*) adj1 mobilit* shift* assay*).tw.	39
30	HMSA.tw.	98
31	(Biomonitor* or iLite).tw.	5664
32	(Matriks* Biotek* or Shikari*).tw.	13
33	(Prometheus* or Anser IFX or Anser ADA).tw.	568
34	22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33	113752
35	((monitor* or pharmacokinetic* or measur* or level* or concentration*) adj3 (adalimumab or ADA or infliximab or IFX or Anti-TNF* or Anti-Tumour Necrosis Factor*)).tw.	2016
36	*crohn disease/	34280
37	crohn*.tw.	50039
38	inflammator* bowel* disease*.tw.	41418
39	IBD.tw.	23266
40	36 or 37 or 38 or 39	82551
41	(((monitor* or pharmacokinetic* or measur* or level* or concentration*) adj3 (adalimumab or infliximab or Anti-TNF* or AntiTNF* or Anti-Tumour Necrosis Factor*)) and (correlat* or associat* or test performance)).tw.	544
42	13 and 21 and 40	278
43	13 and 34 and 40	109
44	35 and 40	507
45	41 or 42 or 43 or 44	938
46	nonhuman/ not human/	3490973
47	45 not 46	917

Cochrane Library (Wiley), searched on 22/10/2014

#1	adalimumab:ti,ab,kw	451
#2	ADA:ti,ab	237
#3	infliximab:ti,ab,kw	767
#4	IFX:ti,ab	39
#5	((anti-TNF* or antiTNF* or TNF*) near/2 inhibitor*):ti,ab,kw	106
#6	(anti* next tumo*r* next necrosis* next factor*):ti,ab,kw	256
#7	MeSH descriptor: [Tumor Necrosis Factor-alpha] this term only	2408
#8	MeSH descriptor: [Antibodies, Monoclonal] this term only	3978
#9	#7 and #8	409
#10	(anti* next drug* next antibod*):ti,ab,kw	19
#11	(ADAb):ti,ab,kw	0
#12	#1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11	6714
#13	(lisa* next tracker*):ti,ab,kw	0
#14	(immundiagnostik* or immunodiagnostik* or immunediagnostik*):ti,ab,kw	0
#15	(proteomika* or promonitor*):ti,ab,kw	0
#16	MeSH descriptor: [Enzyme-Linked Immunosorbent Assay] explode all trees	2122
#17	(enzyme* next link* next immunoassay*):ti,ab,kw	84
#18	ELISA*:ti,ab,kw	2534
#19	#13 or #14 or #15 or #16 or #17 or #18	3958
#20	MeSH descriptor: [Radioimmunoassay] explode all trees	1176
#21	(radioimmuno* or radio next immuno* or radio-immuno*):ti,ab,kw	2761
#22	RIA:ti,ab	570
#23	(reporter* next gene* next assay*):ti,ab,kw	11
#24	RGA:ti,ab	8
#25	(semi* next fluid* next phase* next enzyme* next immuno*):ti,ab,kw	0
#26	EIA:ti,ab	339
#27	((homogenous* or homogeneous*) near/1 (mobilit* next shift* next assay*)):ti,ab,kw	1
#28	HMSA:ti,ab	1
#29	(Biomonitor* or iLite):ti,ab,kw	14
#30	(Matriks* next Biotek* or Shikari*):ti,ab,kw	0
#31	(Prometheus* or Anser next IFX or Anser next ADA):ti,ab,kw	23
#32	#20 or #21 or #22 or #23 or #24 or #25 or #26 or #27 or #28 or #29 or #30 or #31	3651
#33	((monitor* or pharmacokinetic* or measur* or level* or concentration*) near/3 (adalimumab or ADA or infliximab or IFX or Anti-TNF* or Anti-Tumour next Necrosis next Factor*)):ti,ab,kw	83
#34	MeSH descriptor: [Inflammatory Bowel Diseases] this term only	273
#35	MeSH descriptor: [Crohn Disease] this term only	997

#36	crohn*:ti,ab,kw	1512
#37	(inflammator* next bowel* next disease*):ti,ab,kw	798
#38	IBD:ti,ab	271
#39	#34 or #35 or #36 or #37 or #38	2037
#40	(((monitor* or pharmacokinetic* or measur* or level* or concentration*) near/3 (adalimumab or infliximab or Anti-TNF* or AntiTNF* or Anti-Tumour next Necrosis next Factor*)) and (correlat* or associat* or test next performance)):ti,ab,kw	33
#41	#12 and #19 and #39	8
#42	#12 and #32 and #39	1
#43	#33 and #39	18
#44	#40 or #41 or #42 or #43	49

All Results (49)

Cochrane Reviews (0)

All Review Protocol

Other Reviews (1)

Trials (47)

Methods Studies (0)

Technology Assessments (1)

Economic Evaluations (0)

Cochrane Groups (0)

Science Citation Index and Conference Proceedings – Science (Web of Science), searched on 22/10/2014

# 40	806	#39 OR #38 OR #37 OR #36 Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 39	324	#35 AND #32 Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 38	26	#35 AND #31 AND #9 Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 37	128	#35 AND #16 AND #9 Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 36	539	TS=(((monitor* or pharmacokinetic* or measur* or level* or concentration*) near/3 (adalimumab or ADA or infliximab or IFX or Anti-TNF* or ("Anti-Tumour Necrosis" near/1 Factor*))) and (correlat* or associat* or "test performance"))

		Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 35	80,743	#34 OR #33 Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 34	53,142	TS=(((inflammator* near/1 bowel*) near/1 disease*) or IBD) Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 33	50,398	TS=crohn* Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 32	1,366	TS=((monitor* or pharmacokinetic* or measur* or level* or concentration*) near/3 (adalimumab or ADA or infliximab or IFX or Anti-TNF* or ("Anti-Tumour Necrosis" near/1 Factor*))) Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 31	79,288	#30 OR #29 OR #28 OR #27 OR #26 OR #25 OR #24 OR #23 OR #22 OR #21 OR #20 OR #19 OR #18 OR #17 Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 30	713	TS=(Prometheus* or "Anser IFX" or "Anser ADA") Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 29	10	TS=((Matriks* near/1 Biotek*) or Shikari*) Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 28	8,841	TS=(Biomonitor* or iLite) Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 27	107	TS=HMSA Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 26	11	TS=((homogenous* or homogeneous*) near/1 (mobilit* near/1 (shift* near/1 assay*))) Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 25	8,832	TS=EIA Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 24	1	TS=((semi* near/1 fluid*) near/3 (enzyme* near/1 immuno*)) Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 23	0	TS=((semi* near/1 fluid*) near/2 (enzyme* near/1 immuno*)) Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 22	0	TS=(semi* near/1 fluid* near/1 phase* near/1 enzyme* near/1 immuno*) Indexes=SCI-EXPANDED, CPCI-S Timespan=All years

# 21	0	TS=(((semi* near/1 fluid*) near/1 phase*) near/1 (enzyme* near/1 immuno*)) Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 20	1,230	TS=RGA Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 19	4,518	TS=(reporter* near/1 gene* near/1 assay*) Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 18	12,773	TS=RIA Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 17	46,937	TS=(radioimmuno* or (radio near/1 immuno*) or radio-immuno*) Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 16	146,389	#15 OR #14 OR #13 OR #12 OR #11 OR #10 Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 15	113,120	TS=ELISA* Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 14	60,666	TS=((enzyme* near/1 link*) near/1 (immuno* near/1 assay)) Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 13	2,850	TS=((enzyme* near/1 link*) near/1 immunoassay*) Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 12	1	TS=(proteomika* or promonitor*) Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 11	9	TS=(immundiagnostik* or immunodiagnostik* or immunediagnostik*) Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 10	0	TS=(lisa* near/1 tracker*) Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 9	32,262	#8 OR #7 OR #6 OR #5 OR #4 OR #3 OR #2 OR #1 Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 8	35	TS=ADAb Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 7	2,534	TS=((anti* near/1 drug*) near/1 antibod*) Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 6	4,072	TS=((anti* near/1 tumo$r*) near/1 (necrosis* near/1 factor*)) Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 5	4,065	TS=((anti-TNF* or antiTNF* or TNF*) near/2 inhibitor*)

		Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 4	373	TS=IFX Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 3	13,729	TS=infliximab Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 2	8,006	TS=ADA Indexes=SCI-EXPANDED, CPCI-S Timespan=All years
# 1	4,973	TS=adalimumab Indexes=SCI-EXPANDED, CPCI-S Timespan=All years

Index to Theses, searched on 28/10/2014

((adalimumab or infliximab or AntiTNF* or Anti-TNF* or "Anti TNF" or "Anti TNFa" or "Anti TNFalpha" or (TNF* w/2 inhibitor*) or (Anti-Tum*r w/2 Necrosis) or ("anti drug" w/2 antibod*) or ADAb) AND (crohn* or "inflammatory bowel disease" or IBD))

14 document(s) retrieved

(((adalimumab or infliximab or AntiTNF* or Anti-TNF* or "Anti TNF" or "Anti TNFa" or "Anti TNFalpha" or (TNF* w/2 inhibitor*) or (Anti-Tum*r w/2 Necrosis) or "anti drug antibody" or "anti drug antibodies" or "anti-drug antibody" or "anti-drug antibodies" or ADAb) w/10 (monitor or monitoring or monitors or monitored or pharmacokinetic or pharmacokinetics or measure or measures or measurement or measuring or level or levels or concentration or concentrations)) AND ((correlate* or correlation* or associate* or association* or "test performance")))

4 document(s) retrieved

DART-Europe, searched on 28/10/2014

(adalimumab or infliximab or AntiTNF* or Anti-TNF* or "Anti TNF" or "Anti TNFa" or "Anti TNFalpha" or (TNF* and inhibitor*) or (Anti-Tum*r and Necrosis) or ("anti drug" and antibod*) or ADAb) and (crohn* or "inflammatory bowel disease" or "inflammatory bowel diseases" or IBD)

113 document(s) retrieved

Dissertations and Theses, searched on 29/10/2014

all(((adalimumab or infliximab or AntiTNF* or Anti-TNF* or "Anti TNF" or "Anti TNFa" or "Anti TNFalpha" or (TNF* n/2 inhibitor*) or (Anti-Tum*r n/2 Necrosis) or ("anti drug" n/2 antibod*) or ADAb) AND (crohn* or "inflammatory bowel disease" or "inflammatory bowel diseases" or IBD)))

21

all(((adalimumab or infliximab or AntiTNF* or Anti-TNF* or "Anti TNF" or "Anti TNFa" or "Anti TNFalpha" or (TNF* n/2 inhibitor*) or (Anti-Tum*r n/2 Necrosis) or "anti drug antibody" or "anti drug antibodies" or "anti-drug antibody" or "anti-drug antibodies" or ADAb) n/10 (monitor or monitoring or monitors or monitored or pharmacokinetic or pharmacokinetics or measure or measures or measurement or measuring or level or levels or concentration or concentrations)) and (correlate* or correlation* or associate* or association* or "test performance"))
15

NIHR HTA Programme, searched on 29/10/2014

adalimumab
16
infliximab
23
TNF
17

PROSPERO, searched on 29/10/2014

adalimumab in All fields
OR
infliximab in All fields
OR
TNF* inhibitor* in All fields
OR
AntiTNF* in All fields
OR
Anti-TNF* in All fields
29 records

ClinicalTrials.gov, searched on 04/11/2014

Search Terms (any field): adalimumab OR infliximab OR (TNF AND (anti OR inhibitor OR blocker)) OR "anti drug antibody" OR "anti drug antibodies" OR ADAb
AND
Condition: crohn OR "inflammatory bowel disease" OR "inflammatory bowel diseases"
AND
Title: monitor OR pharmacokinetic OR measure OR measuring OR level OR concentration OR assay
14 studies

Current Controlled Trials, searched on 04/11/2014

(adalimumab OR infliximab OR TNF* OR AntiTNF* OR Anti-TNF* OR anti drug antibod* OR ADAb) AND (crohn* OR inflammatory bowel disease*) AND (monitor* OR pharmacokinetic* OR measure* OR measuring OR level* OR concentration* OR assay*)

30 studies

UKCRN Portfolio Database, searched on 04/11/2014

Specialty: Gastroenterology

Research Summary: adalimumab infliximab TNF AntiTNF Anti-TNF ADAb

'Any' selected (combines terms with Boolean OR)

4 studies

WHO ICTRP, searched on 10/11/2014

Advanced Search

In Title: adalimumab OR infliximab OR AntiTNF* OR Anti-TNF* OR TNF inhibitor* OR TNFα inhibitor* OR TNF alpha inhibitor* OR TNFalpha inhibitor* OR anti drug antibody OR anti drug antibodies OR ADAb

AND

In Condition: Crohn* OR inflammatory bowel disease*

AND

In Intervention: monitor* OR pharmacokinetic* OR measure* OR measuring OR level* OR concentration* OR assay*

39 trials found

Espacenet (European Patent Office), searched on 10/11/2014

Advanced Search

Applicant(s): THERADIAG – 1 result "Methods for detecting antibodies" (relevant)

Applicant(s): Immundiagnostik – 27 results (sifted online, none relevant)

Checked how known Theradiag patent found above is classified and combined the following 2 most relevant classification numbers:

G01N2333/525 Assays involving biological materials from specific organisms or of a specific nature - Tumor necrosis factor (TNF)

G01N2800/52 Detection or diagnosis of diseases - Predicting or monitoring the response to treatment; Prognosis

Advanced Search

CPC: G01N2333/525 AND G01N2800/52 – 27 results (browsed for manufacturer's name, found relevant Proteomika patent)

Sifted online and used 'Also published as' to find English language versions

Clinical Effectiveness: conference proceedings

Searched on 22/01/2015

Specifically looked for studies with clinical outcomes and based on the use of an algorithm (i.e. 'management' studies).

European Crohn's and Colitis Organisation (ECCO)

Abstracts published in Journal of Crohn's and Colitis

2011 – 2014 Indexed in Embase. Checked and the search of Embase has picked them up.

2015 searchable via website

Sifted 2015 online. 5 potentially relevant abstracts saved.

Digestive Diseases Week (DDW) (meeting of the American Gastroenterology Association(AGA))

www.ddw.org

abstracts in Gastroenterology

2009 – 2014 Indexed in Embase. Checked and the search of Embase has picked them up.

n.b. Promonitor have sent 2 abstracts submitted to DDW May 2015.

British Society of Gastroenterology (BSG)

abstracts in Gut

Indexed in Embase (2011, 2012 and 2014). Checked and the search of Embase has picked these years up.

Checked 2010 and 2013 via organisation's website http://www.bsg.org.uk/education/meeting/index.html

Searches:

infliximab

adalimumab

TNF

Sifted online. 2 potentially relevant abstracts saved.

UEGW

2013 and 2014 in UEG journal, available via Pubmed central. Not indexed in Embase

Checked via Pubmed.

Search	Query	Items found
#4	Search ((#2 or #3)) AND #1	13 – sifted online, none with algorithms
#3	Search (inflammatory bowel disease*) OR IBD	36891
#2	Search crohn*	42212
#1	Search "United European Gastroenterol J"[Journal]	149

Previous years not available.

American College of Gastroenterology

Meeting abstracts in Am J Gastroenterol

2010 – 2013 indexed in Embase. Checked and the search of Embase has picked them up.

2014 conference website says "All abstracts submitted will be published in a supplement to the October 2014 issue of The American Journal of Gastroenterology." Check via journal website.

http://www.nature.com/ajg/journal/v109/n2s/pdf/ajg2014281a.pdf

Searches:

infliximab

adalimumab

TNF

2014 sifted online – no 'management' studies with clinical outcomes based on use of an algorithm.

Clinical Effectiveness: Websites

Searched on 02/02/2015

European Crohn's and Colitis Organisation (ECCO)

www.ecco-ibd.eu

Browsed consensus statements for Crohn's Disease, Publications and Research Projects.

No additional 'management' studies identified.

The American Gastroenterology Association (AGA)

http://www.gastro.org

Browsed: 'Technical Reviews'

Browsed: Research > Research Resource Library > Immunology, Microbiology and IBD.

No additional 'management' studies identified.

British Society of Gastroenterology (BSG)

http://www.bsg.org.uk

Browsed: 'Research' and 'Clinical' sections.

No additional ‘management’ studies identified.

United European Gastroenterology (UEG)

https://www.ueg.eu/

Browsed: ‘Research’ section.

No additional ‘management’ studies identified.

American college of gastroenterology

http://gi.org/

Browsed: ‘Research and Awards’ and ‘Clinical Guidelines’ sections.

No additional ‘management’ studies identified.

International Network of Agencies for Health Technology Assessment (INAHTA) Publication

http://www.inahta.org/

Searched within publications for:

infliximab

adalimumab

TNF

No additional ‘management’ studies identified.

FDA medical devices

http://www.fda.gov/MedicalDevices/ProductsandMedicalProcedures/

Searched for:

infliximab

adalimumab

TNF

Filtered by topic to limit to ‘Medical Devices’

No additional ‘management’ studies identified.

European Commission medical devices

http://ec.europa.eu/growth/sectors/medical-devices/

Searched for:

infliximab

adalimumab

TNF

No additional ‘management’ studies identified.

Theradiag

http://www.theradiag.com/en/

Browsed Theranostics > LISA TRACKER

Saved and sifted list of publications for LISA Tracker. No additional ‘management’ studies identified.

Immundiagnostik

http://www.immundiagnostik.com/en

Browsed website. No specific lists of publications, but manuals for relevant assays contain references. The manuals have been sent with other information from manufacturer and references already sifted.

Proteomika

http://www.proteomika.com/

Browsed website. Brochure has a list of references. Sifted. No additional ‘management’ studies identified.

Cost-Effectiveness: Searches for published cost-effectiveness studies

Ovid MEDLINE(R) 1946 to November Week 3 2014, searched 12/12/2014

1	adalimumab.mp.	3662
2	ADA.tw.	7143
3	infliximab.mp.	8957
4	IFX.tw.	335
5	((anti-TNF* or antiTNF* or TNF*) adj2 inhibitor*).mp.	2630
6	anti* tumo?r* necrosis* factor*.mp.	3048
7	Tumor Necrosis Factor-alpha/ and Antibodies, Monoclonal/	7737
8	anti* drug* antibod*.tw.	188
9	adalimumab.mp.	3662
10	ADA.tw.	7143
11	infliximab.mp.	8957
12	IFX.tw.	335
13	((anti-TNF* or antiTNF* or TNF*) adj2 inhibitor*).mp.	2630
14	anti* tumo?r* necrosis* factor*.mp.	3048
15	Tumor Necrosis Factor-alpha/ and Antibodies, Monoclonal/	7737
16	anti* drug* antibod*.tw.	188
17	ADAb.tw.	19
18	9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17	24434
19	lisa* tracker*.mp.	1
20	(immundiagnostik* or immunodiagnostik* or immunediagnostik*).mp.	159
21	(proteomika* or promonitor*).mp.	13
22	exp Enzyme-Linked Immunosorbent Assay/	129940
23	enzyme* link* immunoassay*.mp.	2879
24	enzyme* link* immuno* assay*.mp.	159574
25	ELISA*.mp.	114330
26	19 or 20 or 21 or 22 or 23 or 24 or 25	206726
27	*Radioimmunoassay/	7654
28	(radioimmuno* or radio immuno* or radio-immuno*).mp.	102645
29	RIA.tw.	17539
30	reporter* gene* assay*.mp.	3695
31	RGA.tw.	337
32	semi* fluid* phase* enzyme* immuno*.mp.	0
33	EIA.tw.	8313

34	((homogenous* or homogeneous*) adj1 mobilit* shift* assay*).mp.	4
35	HMSA.tw.	62
36	(Biomonitor* or iLite).tw.	4140
37	(Matriks* Biotek* or Shikari*).mp.	2
38	(Prometheus* or Anser IFX or Anser ADA).mp.	260
39	27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38	125716
40	Inflammatory Bowel Diseases/	14609
41	Crohn Disease/	31828
42	crohn*.tw.	32634
43	inflammator* bowel* disease*.tw.	27171
44	IBD.tw.	12128
45	40 or 41 or 42 or 43 or 44	58950
46	18 and 45	3875
47	26 and 45	1771
48	39 and 45	278
49	exp Economics/	513380
50	exp "Costs and Cost Analysis"/	190833
51	Health Status/	63445
52	exp "Quality of Life"/	126611
53	exp Quality-Adjusted Life Years/	7642
54	(pharmacoeconomic* or pharmaco-economic* or economic* or cost*).tw.	461021
55	(health state* or health status).tw.	40275
56	(qaly* or utilit* or EQ5D or EQ-5D or euroqol or euro-qol or SF-36 or SF36 or SF-6D or SF-6D or SF6D or HUI).tw.	138384
57	(markov or time trade off or TTO or standard gamble or hrql or hrqol or disabilit* or disutilit*).tw.	129972
58	(quality adj2 life).tw.	148233
59	(decision adj2 model).tw.	3980
60	(visual analog* scale* or discrete choice experiment* or health* year* equivalen* or (willing* adj2 pay)).tw.	31394
61	("resource use" or resource utili?ation).tw.	9307
62	(well-being or wellbeing).tw.	44692
63	49 or 50 or 51 or 52 or 53 or 54 or 55 or 56 or 57 or 58 or 59 or 60 or 61 or 62	1298647
64	46 and 63	458
65	47 and 63	71

66	48 and 63	9
67	64 or 65 or 66	526
68	limit 67 to english language	479

Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations December 11, 2014, searched 16/12/2014

1	adalimumab.mp.	502
2	ADA.tw.	461
3	infliximab.mp.	868
4	IFX.tw.	76
5	((anti-TNF* or antiTNF* or TNF*) adj2 inhibitor*).mp.	330
6	anti* tumo?r* necrosis* factor*.mp.	355
7	anti* drug* antibod*.tw.	45
8	ADAb.tw.	2
9	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8	1949
10	lisa* tracker*.mp.	0
11	(immundiagnostik* or immunodiagnostik* or immunediagnostik*).mp.	3
12	(proteomika* or promonitor*).mp.	2
13	enzyme* link* immunoassay*.mp.	142
14	enzyme* link* immuno* assay*.mp.	4191
15	ELISA*.mp.	8507
16	10 or 11 or 12 or 13 or 14 or 15	10654
17	(radioimmuno* or radio immuno* or radio-immuno*).mp.	1197
18	RIA.tw.	401
19	reporter* gene* assay*.mp.	250
20	RGA.tw.	49
21	semi* fluid* phase* enzyme* immuno*.mp.	0
22	EIA.tw.	379
23	((homogenous* or homogeneous*) adj1 mobilit* shift* assay*).mp.	1
24	HMSA.tw.	6
25	(Biomonitor* or iLite).tw.	390
26	(Matriks* Biotek* or Shikari*).mp.	1
27	(Prometheus* or Anser IFX or Anser ADA).mp.	23
28	17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27	2503
29	crohn*.tw.	2585

30	inflammator* bowel* disease*.tw.	2745
31	IBD.tw.	1547
32	29 or 30 or 31	4595
33	9 and 32	466
34	16 and 32	110
35	28 and 32	6
36	(pharmacoeconomic* or pharmaco-economic* or economic* or cost*).tw.	54972
37	(health state* or health status).tw.	3544
38	(qaly* or utilit* or EQ5D or EQ-5D or euroqol or euro-qol or SF-36 or SF36 or SF-6D or SF-6D or SF6D or HUI).tw.	15909
39	(markov or time trade off or TTO or standard gamble or hrql or hrqol or disabilit* or disutilit*).tw.	13731
40	(quality adj2 life).tw.	17497
41	(decision adj2 model).tw.	400
42	(visual analog* scale* or discrete choice experiment* or health* year* equivalen* or (willing* adj2 pay)).tw.	3999
43	("resource use" or resource utili?ation).tw.	992
44	(well-being or wellbeing).tw.	4897
45	36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44	101172
46	33 and 45	63
47	34 and 45	9
48	35 and 45	1
49	46 or 47 or 48	73
50	limit 49 to english language	71

OVID Embase Classic+Embase 1947 to 2014 December 15, searched 16/12/2014

1	adalimumab.tw.	7509
2	*adalimumab/	4043
3	ADA.tw.	10949
4	infliximab.tw.	13814
5	*infliximab/	8148
6	IFX.tw.	1753
7	((anti-TNF* or antiTNF* or TNF*) adj2 inhibitor*).tw.	4742
8	anti* tumo?r* necrosis* factor*.tw.	4224
9	*tumor necrosis factor alpha inhibitor/	1298

10	anti* drug* antibod*.tw.	477
11	ADAb.tw.	45
12	*drug antibody/	1542
13	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12	36094
14	lisa* tracker*.tw.	11
15	(immundiagnostik* or immunodiagnostik* or immunediagnostik*).tw.	76
16	(proteomika* or promonitor*).tw.	27
17	*enzyme linked immunosorbent assay/	14705
18	enzyme* link* immunoassay*.tw.	3301
19	enzyme* link* immuno* assay*.tw.	72608
20	ELISA*.tw.	169424
21	14 or 15 or 16 or 17 or 18 or 19 or 20	210314
22	*radioimmunoassay/	17241
23	(radioimmuno* or radio immuno* or radio-immuno*).tw.	75063
24	RIA.tw.	20852
25	reporter* gene* assay*.tw.	4446
26	RGA.tw.	401
27	semi* fluid* phase* enzyme* immuno*.tw.	1
28	EIA.tw.	10934
29	((homogenous* or homogeneous*) adj1 mobilit* shift* assay*).tw.	40
30	HMSA.tw.	99
31	(Biomonitor* or iLite).tw.	5679
32	(Matriks* Biotek* or Shikari*).tw.	14
33	(Prometheus* or Anser IFX or Anser ADA).tw.	568
34	22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33	114144
35	*crohn disease/	34603
36	crohn*.tw.	50590
37	inflammator* bowel* disease*.tw.	42049
38	35 or 36 or 37	79897
39	13 and 38	6882
40	21 and 38	2411
41	34 and 38	394
42	exp *health economics/	200481
43	exp health status/	150318
44	exp "quality of life"/	283712

45	exp quality adjusted life year/	13007
46	(pharmacoeconomic* or pharmaco-economic* or economic* or cost*).tw.	656408
47	(health state* or health status).tw.	51749
48	(qaly* or utilit* or EQ5D or EQ-5D or euroqol or euro-qol or SF-36 or SF36 or SF-6D or SF6D or SF-6D or HUI).tw.	195997
49	(markov or time trade off or TTO or standard gamble or hrql or hrqol or disabilit* or disutilit*).tw.	189075
50	(quality adj2 life).tw.	233390
51	(decision adj2 model).tw.	5912
52	(visual analog* scale* or discrete choice experiment* or health* year* equivalen*).tw.	42481
53	("resource use" or resource utili?ation).tw.	15005
54	(willing* adj2 pay).tw.	4494
55	42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50 or 51 or 52 or 53 or 54	1506135
56	39 and 55	969
57	40 and 55	143
58	41 and 55	33
59	56 or 57 or 58	1106
60	limit 59 to english language	1045

NHS Economic Evaluation Database (NHS EED) (Cochrane Library), searched 17/12/2014

ID	Search	Hits
#1	adalimumab:ti,ab,kw	522
#2	ADA:ti,ab	295
#3	infliximab:ti,ab,kw	824
#4	IFX:ti,ab	56
#5	((anti-TNF* or antiTNF* or TNF*) near/2 inhibitor*):ti,ab,kw	129
#6	(anti* next tumo*r* next necrosis* next factor*):ti,ab,kw	264
#7	MeSH descriptor: [Tumor Necrosis Factor-alpha] this term only	2420
#8	MeSH descriptor: [Antibodies, Monoclonal] this term only	3989
#9	#7 and #8	411
#10	(anti* next drug* next antibod*):ti,ab,kw	22
#11	(ADAb):ti,ab,kw	0
#12	#1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11	6872
#13	(lisa* next tracker*):ti,ab,kw	0
#14	(immundiagnostik* or immunodiagnostik* or immunediagnostik*):ti,ab,kw	0
#15	(proteomika* or promonitor*):ti,ab,kw	0

#16	MeSH descriptor: [Enzyme-Linked Immunosorbent Assay] explode all trees	2128
#17	(enzyme* next link* next immunoassay*):ti,ab,kw	88
#18	ELISA*:ti,ab,kw	2609
#19	#13 or #14 or #15 or #16 or #17 or #18	4037
#20	MeSH descriptor: [Radioimmunoassay] explode all trees	1176
#21	(radioimmuno* or radio next immuno* or radio-immuno*):ti,ab,kw	2769
#22	RIA:ti,ab	572
#23	(reporter* next gene* next assay*):ti,ab,kw	11
#24	RGA:ti,ab	8
#25	(semi* next fluid* next phase* next enzyme* next immuno*):ti,ab,kw	0
#26	EIA:ti,ab	342
#27	((homogenous* or homogeneous*) near/1 (mobilit* next shift* next assay*)):ti,ab,kw	1
#28	HMSA:ti,ab	1
#29	(Biomonitor* or iLite):ti,ab,kw	15
#30	(Matriks* next Biotek* or Shikari*):ti,ab,kw	0
#31	(Prometheus* or Anser next IFX or Anser next ADA):ti,ab,kw	24
#32	#20 or #21 or #22 or #23 or #24 or #25 or #26 or #27 or #28 or #29 or #30 or #31	3665
#33	((monitor* or pharmacokinetic* or measur* or level* or concentration*) near/3 (adalimumab or ADA or infliximab or IFX or Anti-TNF* or Anti-Tumour next Necrosis next Factor*)):ti,ab,kw	90
#34	MeSH descriptor: [Inflammatory Bowel Diseases] this term only	277
#35	MeSH descriptor: [Crohn Disease] this term only	1006
#36	crohn*:ti,ab,kw	1556
#37	(inflammator* next bowel* next disease*):ti,ab,kw	843
#38	IBD:ti,ab	304
#39	#34 or #35 or #36 or #37 or #38	2123
#40	#12 and #39	344
#41	#19 and #39	31
#42	#32 and #39	9
#43	#40 or #41 or #42	373

All Results (373)

Economic Evaluations (30)

Science Citation Index 1970 – present (via Web of Knowledge), searched 17/12/2014

n.b. reads bottom to top

# 42	784	(#41) AND LANGUAGE: (English) Indexes=SCI-EXPANDED Timespan=All years
# 41	820	#40 AND #39 Indexes=SCI-EXPANDED Timespan=All years
# 40	1,328,585	TS=("quality of life" or QoL or hrql or hrqol or ("quality adjusted life" NEAR/1 year*) or QALY* or cost* or economic* or pharmacoeconomic* or pharmaco-economic* or euro-qol or utilit* or disutilit* or euroqol or "euro qol" or EQ5D or EQ-5D or SF-36 or SF36 or SF-6D or SF6D or HUI or (time NEAR/1 trade*) or TTO or "standard gamble" or markov or (decision NEAR/2 model*) or (visual NEAR/1 analog*) or "discrete choice" or ((health* NEAR/1 year*) NEAR/1 equivalen*) or (health NEAR/1 stat*) or "willingness to pay" or "resource use" or (resource NEAR/1 utili?ation) or wellbeing or well-being) Indexes=SCI-EXPANDED Timespan=All years
# 39	8,339	#38 OR #37 OR #36 Indexes=SCI-EXPANDED Timespan=All years
# 38	246	#34 AND #31 Indexes=SCI-EXPANDED Timespan=All years
# 37	1,971	#34 AND #16 Indexes=SCI-EXPANDED Timespan=All years
# 36	6,311	#34 AND #9 Indexes=SCI-EXPANDED Timespan=All years
# 35	560	TS=(((monitor* or pharmacokinetic* or measur* or level* or concentration*) near/3 (adalimumab or ADA or infliximab or IFX or Anti-TNF* or ("Anti-Tumour Necrosis" near/1 Factor*))) and (correlat* or associat* or "test performance")) Indexes=SCI-EXPANDED Timespan=All years
# 34	80,169	#33 OR #32 Indexes=SCI-EXPANDED Timespan=All years
# 33	52,825	TS=(((inflammator* near/1 bowel*) near/1 disease*) or IBD) Indexes=SCI-EXPANDED Timespan=All years
# 32	50,019	TS=crohn*

		Indexes=SCI-EXPANDED Timespan=All years
# 31	77,531	#30 OR #29 OR #28 OR #27 OR #26 OR #25 OR #24 OR #23 OR #22 OR #21 OR #20 OR #19 OR #18 OR #17 Indexes=SCI-EXPANDED Timespan=All years
# 30	588	TS=(Prometheus* or "Anser IFX" or "Anser ADA") Indexes=SCI-EXPANDED Timespan=All years
# 29	11	TS=((Matriks* near/1 Biotek*) or Shikari*) Indexes=SCI-EXPANDED Timespan=All years
# 28	8,544	TS=(Biomonitor* or iLite) Indexes=SCI-EXPANDED Timespan=All years
# 27	102	TS=HMSA Indexes=SCI-EXPANDED Timespan=All years
# 26	13	TS=((homogenous* or homogeneous*) near/1 (mobilit* near/1 (shift* near/1 assay*))) Indexes=SCI-EXPANDED Timespan=All years
# 25	8,367	TS=EIA Indexes=SCI-EXPANDED Timespan=All years
# 24	1	TS=((semi* near/1 fluid*) near/3 (enzyme* near/1 immuno*)) Indexes=SCI-EXPANDED Timespan=All years
# 23	0	TS=((semi* near/1 fluid*) near/2 (enzyme* near/1 immuno*)) Indexes=SCI-EXPANDED Timespan=All years
# 22	0	TS=(semi* near/1 fluid* near/1 phase* near/1 enzyme* near/1 immuno*) Indexes=SCI-EXPANDED Timespan=All years
# 21	0	TS=(((semi* near/1 fluid*) near/1 phase*) near/1 (enzyme* near/1 immuno*)) Indexes=SCI-EXPANDED Timespan=All years
# 20	962	TS=RGA Indexes=SCI-EXPANDED Timespan=All years
# 19	4,550	TS=(reporter* near/1 gene* near/1 assay*) Indexes=SCI-EXPANDED Timespan=All years
# 18	12,369	TS=RIA Indexes=SCI-EXPANDED Timespan=All years
# 17	46,687	TS=(radioimmuno* or (radio near/1 immuno*) or radio-immuno*) Indexes=SCI-EXPANDED Timespan=All years

# 16	145,530	#15 OR #14 OR #13 OR #12 OR #11 OR #10 Indexes=SCI-EXPANDED Timespan=All years
# 15	112,098	TS=ELISA* Indexes=SCI-EXPANDED Timespan=All years
# 14	60,765	TS=((enzyme* near/1 link*) near/1 (immuno* near/1 assay)) Indexes=SCI-EXPANDED Timespan=All years
# 13	2,846	TS=((enzyme* near/1 link*) near/1 immunoassay*) Indexes=SCI-EXPANDED Timespan=All years
# 12	1	TS=(proteomika* or promonitor*) Indexes=SCI-EXPANDED Timespan=All years
# 11	10	TS=(immundiagnostik* or immunodiagnostik* or immunediagnostik*) Indexes=SCI-EXPANDED Timespan=All years
# 10	1	TS=(lisa* near/1 tracker*) Indexes=SCI-EXPANDED Timespan=All years
# 9	31,622	#8 OR #7 OR #6 OR #5 OR #4 OR #3 OR #2 OR #1 Indexes=SCI-EXPANDED Timespan=All years
# 8	31	TS=ADAb Indexes=SCI-EXPANDED Timespan=All years
# 7	2,570	TS=((anti* near/1 drug*) near/1 antibod*) Indexes=SCI-EXPANDED Timespan=All years
# 6	4,119	TS=((anti* near/1 tumo$r*) near/1 (necrosis* near/1 factor*)) Indexes=SCI-EXPANDED Timespan=All years
# 5	4,113	TS=((anti-TNF* or antiTNF* or TNF*) near/2 inhibitor*) Indexes=SCI-EXPANDED Timespan=All years
# 4	381	TS=IFX Indexes=SCI-EXPANDED Timespan=All years
# 3	13,827	TS=infliximab Indexes=SCI-EXPANDED Timespan=All years
# 2	7,173	TS=ADA Indexes=SCI-EXPANDED Timespan=All years
# 1	5,046	TS=adalimumab Indexes=SCI-EXPANDED Timespan=All years

CEA Registry, searched 17/12/2014

Search for: Articles

Full Search Contents: crohn

Total: 24

Search for: Articles

Full Search Contents: inflammatory bowel disease

Total: 6

Total with duplicates from search above removed: 5

Total: 29

EconPapers (RePEc), searched 17/12/2014

crohn* OR inflammatory bowel disease* among working papers and articles and books & chapters and software and authors (25)

ScHARRHUD, searched 17/12/2014

crohn* in Any field

OR

inflammatory bowel disease* in Any field

Total: 1

8.4 Appendix 4 Information provided by Theradiag / Alpha Laboratories, Proteomika and Immundiagnostik

1 Information from Theradiag / Alpha Laboratories

The submission consists of:

- Request for information
- Technologies scoping reports
    - In response to an enquiry from NHS Greater Glasgow and Cycle (No.18 October 2013)
    - Background
    - Method
    - Finding
    - Summary

Full text of abstracts:

1) Unsworth 2013- Measurement of infliximab and anti-infliximab antibodies analytical aspects and clinical implications
2) Swart 2013-Acceptance and adjustment in a districts general cohort of IBD patients: finding and implications
3) Ward 2013-Clinical utility of measuring adalimumab trough levels and antibodies to adalimumab in patients with IBD.

Full papers:

Lists of full papers included related to Alpha Labs (1) in manufacturer submission are:

1) Nanda 2013 Am J Gastro- Impact of antibodies to infliximab on clinical outcomes and TRI in IBD Meta-analysis
2) Paul 2013 Inflamm Bowel Dis-Pharmacokinetic of adalimumab SR and Meta-analysis
3) Paul 2013 Inflamm Bowel Dis- Drug monitoring if IFX
4) Steenholdt 2013 Gut-IBD economic
5) Velayos 2013 Clinical Gastro & Hepato- testing more cost-effective than empiric dose escalation
6) Ben horin 2014 Nature IBD review-Anti-TNF tailoring in IBD
7) Vande Casteele 2014 Current Gastro Rep- IBD reviews
8) Roblin 2014 AJG-Algorithm adalimumab IBD
9) Roblin 2014 CGH- Association between pharmacokinetics of Adalimumab and mucosal healing
10) Roblin 2014 IBD-pharmacokinetics of adalimumab in IBD –meta-analysis
11) Ruemmele 2014 Jour of Corhn's and Colitis-consensus paediatric CD

Presentation

The submission consists of presentation hands out on "Monitoring *anti-TNF drugs in chronic inflammatory disease-impact on tailoring therapies*"

LISA TRACKER information:

Detailed information about LISA TRACKER from the company websites in two different languages i.e. English and French

a) LISA TRACKER Duo Infliximab (Francais)
b) LISA TRACKER Duo Infliximab (English)
c) LISA TRACKER anti- Infliximab (Francais)
d) LISA TRACKER anti- Infliximab (English)
e) LISA TRACKER Infliximab (Francais)
f) LISA TRACKER Infliximab (English)
g) LISA TRACKER Duo Adalimumab (Francais)
h) LISA TRACKER Duo Adalimumab (English)
i) LISA TRACKER anti Adalimumab (Francais)
j) LISA TRACKER anti Adalimumab (English)
k) LISA TRACKER Adalimumab (Francais)
l) LISA TRACKER Adalimumab (English)

2 Information from Proteomika

This part consists of:

Annex 1:

Lists of promonitor peer review articles (indexed in Pubmed) (n=8)

1. Chen, D. Y., Y. M. Chen, W. C. Tsai, J. C. Tseng, Y. H. Chen, C. W. Hsieh, W. T. Hung, and J. L. Lan. 2014. Significant associations of antidrug antibody levels with serum drug trough levels and therapeutic response of adalimumab and etanercept treatment in rheumatoid arthritis. Ann Rheum Dis [Epub ahead of print]. PMID 24442879
2. Llinares-Tello, F., J. Rosas, l. T. de, I, L. Valor, X. Barber, and J. M. Senabre. 2014. Comparative study of both versions of an immunoassay commercialized for therapeutic drug monitoring of adalimumab in rheumatoid arthritis. Reumatol.Clin 10:105-108. PMID 24035361
3. Llinares-Tello, F., S. J. Rosas-Gomez de, J. M. Senabre-Gallego, G. Santos-Soler, C. Santos-Ramirez, E. Salas-Heredia, X. Barber-Valles, and J. Molina-Garcia. 2014. Practical application of acid dissociation in monitoring patients treated with adalimumab. Rheumatol.Int.[Epub ahead of print]. PMID 24816715
4. Llinares, F., J. Rosas-Gómez de Salazar, J. M. Senabre-Gallego, G. Santos-Soler, C. Santos-Ramírez, E. Salas-Heredia, and J. Molina-García. 2012. Analytical and clinical evaluation of a new immunoassay for therapeutic drug monitoring of infliximab and adalimumab. Clin.Chem.Lab.Med. 50:1845-1847. PMID 23089717
5. Mazilu, D., D. Opris, C. Gainaru, M. Iliuta, N. Apetrei, G. Luca, A. Borangiu, T. Gudu, A. Peltea, L. Groseanu, C. Constantinescu, I. Saulescu, V. Bojinca, A. Balanescu, D. Predeteanu, and R. Ionescu. 2014. Monitoring drug and antidrug levels: a rational approach in rheumatoid arthritis patients treated

with biologic agents who experience inadequate response while being on a stable biologic treatment. Biomed.Res.Int. 2014:702701. PMID 24982902

6. Pascual-Salcedo, D., C. Plasencia, S. Ramiro, L. Nuno, G. Bonilla, D. Nagore, A. A. Ruiz Del, A. Martinez, L. Aarden, E. Martin-Mola, and A. Balsa. 2011. Influence of immunogenicity on the efficacy of long-term treatment with infliximab in rheumatoid arthritis. Rheumatology 50:1445-1452. PMID 21427177
7. Plasencia, C., D. Pascual-Salcedo, L. Nuño, G. Bonilla, A. Villalba, D. Peiteado, J. Díez, D. Nagore, A. Ruiz del Agua, R. Moral, E. Martín-Mola, and A. Balsa. 2012. Influence of immunogenicity on the efficacy of long-term treatment of spondyloarthritis with infliximab. Ann.Rheum.Dis. 71:1955-1960. PMID 22563028
8. Ruiz-Arguello, B., A. R. del Agua, N. Torres, A. Monasterio, A. Martinez, and D. Nagore. 2013. Comparison study of two commercially available methods for the determination of infliximab, adalimumab, etanercept and anti-drug antibody levels. Clin Chem Lab Med 51:e287-e289 PMID 23917475

Annex 2:

Lists of promonitor abstracts presented at international congresses 2014 (n=34)

1. Barrios Y, Matheu V, Franco A, Delgado E, Bustabad S. Immunogenicity analysis of two anti-TNF Infliximab vs Etanercept) therapies In rheumatologic patients. The American Academy of Allergy, Asthma & Immunology ABS 5.2.0 DTD Abstracts AB185. 2014. Ref Type: Abstract
2. Daperno M, Lavagna A, Fracchia M, Guiotto C, Germano L, Rigazio C, et al. Infliximab trough levels (IFX-Tl) are higher in patients with inflammatory bowel disease (IBD) treated with immunosuppressives: clinical correlations of IFX-LT and antibodies to infliximab (ATI) in IBD. American Gastroenterological Association AGA Abstract #Tu1173. 2013. Ref Type: Abstract
3. Daperno M, Frigerio F, Guiotto C, Germano L, Ercole E, Arico S, et al. Identical diagnostic performance of two commercially available tests for infliximab trough levels (ifx-tl) and antibodies to infliximab (ati) titration in inflammatory bowel disease (ibd): promonitor and immunodiagnostik tests. American Gastroenterological Association AGA Abstract #Tu1168. 2013. Ref Type: Abstract 4.
4. Daperno M, Frigerio F, Guiotto C, Germano L, Ercole E, Arico S, et al. Evaluation of the diagnostic performance of two commercially available tests for infliximab trough levels (IFX-TL) and antibodies to infliximab (ATI) titration in inflammatory bowel disease (IBD). European Crohn´s and Colitis Organisation ECCO Abstract #P508. 2013. Ref Type: Abstract
5. Daperno M, Frigerio F, Guiotto C, Laura G, Ercole E, Lavagna A, et al. Comparison of the performance of two commercially available tests for determination of infliximab trough levels (IFX-TL) and antibodies to infliximab (ATI), Promonitor and Immundiagnostik, in inflammatory bowel disease. Digestive and Liver Disease 45S Abstract# P.03.13. 2013. Ref Type: Abstract
6. Daperno M, Lavagna A, Fracchia M, Guiotto C, Germano L, Rigazio C, et al. Clinical correlations of infliximab trough levels (IFX-TL) and antibodies to infliximab (ATI) in inflammatory bowel disease. European Crohn´s and Colitis Organisation ECCO Abstract #569. 2013. Ref Type: Abstract

7. Diana M, Iliuta M, Gainaru C, Luca G, Apetrei N, Gudu T, et al. Infliximab and adalimumab levels and antidrug antibodies detection in patients with rheumatoid arthritis (RA): an interlaboratory comparison using a commercial elisa assay. The European League Against Rheumatism EULAR Abstracts #FRI0026. 2014. Ref Type: Abstract
8. Hernández Flórez D, Valor L, Nieto JC, Martínez L, de la Torre I, del Rio T, et al. Infliximab levels and anti-infliximab antibodies comparison between two comercial elisa versions in patients with ankylosing spondylitis. Ann Rheum Dis 73(Suppl2) Abstract#SAT0340. 2014. Ref Type: Abstract
9. Hernández D, de la Torre I, Martínez L, Nieto J, Llinares F, Rosas J, et al. Establishing cut-off of infliximab and anti-infliximab antibody levels using a commercial ELISA in patients with rheumatoid arthritis. Ann Rheum Dis 72(Suppl3) Abstract #THU0215, 237. 2013. Ref Type: Abstract
10. Hernández MV, Palasti S, Inciarte J, Cabrera-Villalba S, Ruiz-Esquide V, Ramírez J, et al. Analysis of the immunogenicity induced by tumor necrosis factor antagonists in patients with chronic inflammatory arthropathies. Ann Rheum Dis 72(Suppl3) Abstract #FRI0171, 429. 2013. Ref Type: Abstract
11. Inciarte-Mundo J, Hernández MV, Cabrera S, Ruiz-Esquide V, Ramirez J, Cañete J, et al. Immunogenicity induced by tumor necrosis factor antagonists in chronic inflammatory arthropathies: retrospective study in clinical practice conditions. American College of Rheumatology ACR Abstract #1444. 2013. Ref Type: Abstract
12. Inciarte-Mundo J, Ramírez García J, Estrada P, García M, Gozález A, Saura C, et al. Drug serum levels of tnf antagonists do not correlate with subclinical synovitis by ultrasound in patients with rheumatoid arthritis and psoriatic arthritis in clinical remission or low disease activity. Ann Rheum Dis 73(Suppl2) Abstract #AB0388. 2014. Ref Type: Abstract
13. Jauregui-Amezaga A, Ordas I, Gallego M, Ramirez A, Pino S, Masamunt MC, et al. Impacto de la determinación de niveles de Anti-TNFa y títulos de anticuerpos contra el fármaco en el manejo del tratamiento con biológicos en la enfermedad inflamatoria intestinal. Asociación Española de Gastroenterología 2013. Ref Type: Abstract
14. Jauregui-Amezaga A, Ordas I, Gallego M, Ramirez A, Pino S, Masamunt MC, et al. Impact of serum drug level and human anti-drug antibody measurement on management of biologic drugs in inflammatory bowel disease. European Crohn´s and Colitis Organisation ECCO Abstract #P481. 2013. Ref Type: Abstract
15. Juan G, Alvariño A, Oltra L, Maroto N, Cano N, Ferrer I, et al. Utility of "trough levels" determination and anti-infliximab antibodies in patients with inflammatory bowel disease. Estimation of individual pharmacokinetic parameters (PK) through population pharmacokinetic model. European Crohn´s and Colitis Organisation ECCO Abstract #P302. 2014. Ref Type: Abstract
16. Llinares-Tello F, Rosas J, de la Torre I, Valor L, Senabre JM, Barber X, et al. Comparative study of both versions of an immunoassay commercialized for therapeutic drug monitoring of adalimumab. Ann Rheum Dis 72(Suppl3) Abstract #THU0207, 234. 2013. Ref Type: Abstract
17. Llinares-Tello F, Rosas J, Senabre-Gallego JM, Molina J, Salas E, Santos-Soler G, et al. Usefulness of the acid dissociation in inmunogenicity detection in patients in treatment with anti-TNF drugs. Ann Rheum Dis 73(Suppl2) Abstract #THU0166. 2014. Ref Type: Abstract

18. Martínez L, Hernández D, Valor L, Carreño L, de la Torre I. Human anti-chimeric antibodies (HACAs) in a cohort of rheumatoid arthritis (RA) patients treated with the anti-TNF-alpha agent infliximab (IFX): disease activity and IFX levels. International Congress on Autoimmunity Abstract #609. 2012. Ref Type: Abstract
19. Nuño L, Pascual-Salcedo D, Balsa A, Moral R, Lopez MT, Ruiz A, et al. Clinical significance of the presence of anti-infliximab antibodies. Ann Rheum Dis 69(Suppl3) Abstract #OP0017, 55. 2010 Ref Type: Abstract
20. Opris D, Diana M, Gainaru C, Iliuta M, Groseanu L, Saulescu I, et al. Serum drug level and anti-citrullinated peptide antibodies as biomarkers that predict eular response in rheumatoid arthritis - a new step to personalized medicine. European League Against Rheumatism EULAR Abstracts #AB0422, 946. 2014. Ref Type: Abstract
21. Pascual-Salcedo D, Plasencia C, Nuño L, Ramiro S, Bonilla G, Nagore D, et al. Immunogenicity influences the efficacy of long-term treatment with infliximab in rheumatoid arthritis. Ann Rheum Dis 70(Suppl3) Abstract #FRI0207, 412. 2011. Ref Type: Abstract
22. Pascual-Salcedo D, Bonilla MG, Nuño L, Ruiz A, Martín-Mola E, Balsa A. Influence of immunogenicity on the efficacy of long-term treatment with infliximab. American College of Rheumatology ACR Abstract #2636. 2011. Ref Type: Abstract
23. Pascual-Salcedo D, Plasencia C, Diez J, Rojo L, Bonilla G, Ramiro S, et al. The development of antibodies against a first anti-TNF influences the clinical outcome of the therapy in rheumatic patients after switching to a second TNF inhibitor. International Congress of Autoimmunity Abstract #1590. 2012. Ref Type: Abstract
24. Pascual-Salcedo D, Plasencia C, Gonzalez del Valle L, López T, Arribas F, Villalba A, et al. Therapeutic drug monitoring (TDM) in rheumatic day clinic enables to reduce pharmaceutical cost maintaining clinical efficacy. Ann Rheum Dis 2013;72:227. Abstract #THU0189
25. Plasencia C, Pascual-Salcedo D, Bonilla MG, Nuño L, Moral R, Ruiz del Agua A, et al. Influence of immunogenicity on the efficacy of long-term treatment with infliximab in spondyloarthritis. Ann Rheum Dis 70(Suppl3) Abstract #OP0045, 82. 2011. Ref Type: Abstract
26. Plasencia C, Pascual-Salcedo D, Garcia-Carazo S, Bonilla G, Lojo L, Nuño L, et al. The immunogenicity to the first anti-TNF therapy determines the outcome of switching to a second anti-TNF in spondyloarthritis patients. American College of Rheumatology ACR Abstract #546. 2012. Ref Type: Abstract
27. Rosas-Gomez de Salazar J, Llinares-Tello F, Senabre-Gallego JM, Santos-Soler G, Santos-Ramirez C, Salas-Heredia E, et al. Evaluation of anti-tumor necrosis factor levels and anti-tumor necrosis factor antibodies in rheumatic diseases treated with infliximab and adalimumab; preliminary results from a local registry. American College of Rheumatology ACR Abstract #2211. 2011. Ref Type: Abstract
28. Rosas J, Llinares F, Santos-Ramírez C, Senabre JM, Santos-Soler G, Barber X, et al. Evaluation of anti-TNF levels and anti-TNF antibodies in rheumatic diseases treated with adalimumab, etanercept and infliximab; results from a local registry. International Congress on Autoimmunity Abstract #1568. 2012. Ref Type: Abstract

29. Rosas J, Llinares F, de la Torre I, Valor L, Barber X, Santos-Ramírez C, et al. Clinical usefulness of serum level of adalimumab, in patients with rheumatoid arthritis. Ann Rheum Dis 72(Suppl3) Abstract #THU0206, 233. 2013. Ref Type: Abstract
30. Rosas J, Llinares-Tello F, Martín S, Senabre JM, Salas E, Oliver S, et al. Evaluation of serum level of golimumab and antibodies anti-golimumab in patients with rheumatic diseases: results from a local registry. Ann Rheum Dis 73(Suppl2) Abstract #AB0389. 2014. Ref Type: Abstract
31. Ruiz del Agua A, Pascual-Salcedo D, Balsa A, Ramos I, Novalbos L, Ramiro S, et al. Monitoring of anti-TNF biological treatments. Journal of Translational Medicine 8(Suppl1), P32. 2010. Ref Type: Abstract
32. Sanmartí R, Inciarte J, Estrada P, García M, González A, Narvaez J, et al. Immunogenicity of anti-TNF antagonists in patients with rheumatoid arthritis or polyarticular psoriatic arthritis in clinical remission or low disease activity: the inmunoremar study. Ann Rheum Dis 73(Suppl2) Abstract #FRI0265. 2014. Ref Type: Abstract
33. Sarmiento Guevara M, Diaz Torne C, Ortiz MA, Torres N, Nagore D, Diaz López C, et al. Association of rituximab levels to clinical response and B Cell recovery in rheumatoid arthritis patients. Ann Rheum Dis 72(Suppl3) Abstract#SAT0125, 623. 2013. Ref Type: Abstract
34. Valor L, Hernández D, de la Torre I, Llinares F, Rosas J, Yagüe J, et al. Infliximab and adalimumab levels and antidrug antibodies detection in patients with rheumatoid arthritis (RA): an interlaboratory comparison using a commercial ELISA assay. Ann Rheum Dis 73(Suppl2) Abstract #AB0396. 2014. Ref Type: Abstract

Full paper:

Rosas 2014 Clinical and Exp Rheumatology

Clinical relevance of monitoring serum levels of adalimumab in patients with rheumatoid arthritis in daily practice

Presentation:

Topic [redacted]

Report:

Progenika Biopharma reports on "method-comparison study between Promonitor®- ELISA and iLITE™ kits for the measurement of infliximab and anti-infliximab antibodies in IBD and RA patients (dated on 25th June 2012)

Technical specification:

Information on Technical specification from Proteomika:

i) Promonitor®-ADL
ii) Promonitor®-Anti-ADL
iii) Promonitor®-Anti-IFX

iv) Promonitor®-IFX

Request for information:

Responses from Proteomika SLU to request for information

Full texts of Proteomika's abstracts

ACR 2104 (n=5)

i) Ghia ACR 2014-2436- Analytical and clinical evaluation of an immunoassay for estimating immunogenicity of infliximab and etanercept in Indian population

ii) Inciarte-Mundo ACR 2014-2926- Calprotectin serum levels reflect residual inflammatory activity in patients with rheumatoid arthritis and psoriatic arthritis on clinical remission or low disease activity undergoing TNF-antagonist therapy

iii) Llinares-Tello ACR 2014-1519-Implementation of an acid dissociation procedure for immunogenicity detection in patients treated with ANTI-TNF drugs

iv) Opris ACR-2014-1539-Relation between number of previous anti TNF agents and clinical response in rheumatoid arthritis patients treated with Rituximab

v) Rosas ACR-2014 1531-Cut off level of adalimumab and prevalence of antibodies ANTI-adalimumab in patients with ankylosing spondylitis: results from local registry

Information pack and technical specification:

Information pack and technical specification about the products

i) Promonitor®-IFX

ii) Promonitor®-ANTI-IFX

iii) Promonitor®-ANTI-ADL

iv) Promonitor®-ADL

Further information from Proteomika

i) Promonitor-IFX (5060230000)

ii) Promonitor-ADL (5080230000)

iii) Promonitor-ANTI-IFX (5070230000)

iv) Promonitor-ANTI-ADL (5090230000)

3 Information from Immundiagnostik / BioHit

Evidence:

This submission contains evidence which includes full texts (n= 2), abstracts (n=4), poster (n=5) and letters to the editor (n=2)

i) Bender 2006 (Rheumatol Int)–Immunogenicity, efficacy and adverse events of adalimumab in RA patients (full text)

ii) Kopylov 2012 (inflamm bowel dis)-Clinical utility of antihuman lamda chain based enzyme linked immunosorbet assay (ELISA) versus double antigen ELISA for the detection of anti-infliximab antibodies (full text)

iii) Daperno 2013 (poster)-Identical diagnostic performance of two commercially available tests for infliximab trough levels (IFX-TL) and antibodies to infliximab (ATI) titration in inflammatory bowel disease (IBD): promonitor and immundiagnostik test

iv) Semmler 2013 (poster)-Development of a new immunoassay for the accurate determination of anti-infliximab antibodies in inflammatory bowel disease (IBD)

v) Guidi 2013 (poster)- Assessment of loss of response to infliximab therapy in inflammatory bowel disease using antibodies to infliximab and trough levels

vi) Perry 2013 (poster)-Infliximab is stable in whole blood clotted samples for 7 days at room temperature

vii) Development of a new immunoassay (2014) (poster)

viii) Eser 2012 (abstract)-Detection of anti-infliximab antibodies in patients with IBD in the presence of infliximab by homogeneous liquid phase anti infliximab mobility shift assay

ix) Jahnel 2014 (abstract)- Formation of antibodies against infliximab in paediatric crohn's disease

x) Ussia 2014 (abstract)-A prospective assessment of antidrug antibody response over time by a new ELISA in patients with IBD treated with infliximab

xi) Schatz 2012 (abstract)-Comparison of different tests for determination of infliximab levels and antibodies against infliximab in pediatric IBD patients

xii) Fritzsche 2012 (letter)-Infliximab and adalimumab use during breastfeeding

xiii) Kong 2013 (letter)-Low trough serum infliximab and antibodies to infliximab in smoker

Immundiagnostik TNF-alpha blocker ELISAs (provided via email after discussion):

Data on assays regarding the limit of blank

Manual:

There is a manual that provide information on technology in two different versions (i.e., English version and Deutsche version)

i) TNFα blocker ADA, total antibodies against adalimumab (e.g. HUMIRA®) (Deutsche and English version)

ii) TNFα blocker ADA, antibodies against adalimumab (e.g. HUMIRA®) (Deutsche and English version)

iii) TNFα blocker ADA, total antibodies against infliximab (e.g. REMICADE®) (Deutsche and English version)

iv) TNFα blocker ADA, antibodies against infliximab (e.g. REMICADE®) (Deutsche and English version)

v) TNFα blocker monitoring adalimumab drug level (e.g. HUMIRA®) (Deutsche and English version)

vi) TNFα blocker monitoring infliximab drug level (e.g. REMICADE®) (Deutsche and English version)

Request for information

8.5 Appendix 5 Data extraction sheets

Data extraction form anti-TNFα drug monitoring: Comparison of assay types

Name of first reviewer: Sian Taylor-Phillips *Name of second reviewer: Martin Connock*

Study details	
Study ID (Endnote ref)	[121, 122]
First author surname	Steenholdt
Year of publication	2014
Country	Denmark
Study design	Retrospective analysis of biobanked samples from RCT
Publication (full/abstract)	Full
Study setting	Hospital
Number of centres (by arm)	6 Danish centres
Duration of study	NR
Follow up period	12 weeks
Funding	Support for this study was provided by unrestricted grants from Aase and Ejnar Danielsen's Foundation, Beckett Foundation, Danish Biotechnology Program,Danish Colitis-Crohn Society, Danish Medical Association Research Foundation, Frode V Nyegaard and Wife's Foundation, Health Science Research Foundation of Region of Copenhagen, Herlev Hospital Research Council, Lundbeck Foundation, P Carl Petersen's Foundation, Ole Østergaard Thomsen's Research Foundation and Jørn Brynskov's Research Foundation.
Competing interests	Trial paper[122] : "CS has served as speaker for MSD and Abbvie and as a consultant for MSD and Takeda Pharmaceutical Company. JB has served as advisory board member for Abbvie. OØT has served as a speaker and consultant for UCB and Zealand Pharma, speaker for MSA, and primary investigator for Amgen, Biogen, Novo-Nordisk and Pfizer. LKM has served as a speaker for MSD and participated in a safety study with Abbvie. JF has served as primary investigator for Centocor, Abbvie, MSD and UCB and as a consultant for Abbvie and MSD. LAC has served as a speaker for Abbvie, Tillotts Pharma and Ferring, and as a consultant for MSD. JK has served as a speaker for MSD, Abbvie and Tillotts. KB has served as a speaker for Pfizer, Roche, Novo-Nordisk, Bristol-Meyers Squibb and Biomonitor and owns stocks in Novo-Nordisk and Biomonitor. BAJ has served as an advisory board member at Tillots Pharma." Comparisons paper[121]: "Casper Steenholdt has served as a speaker for MSD and Abbvie, and as a consultant for MSD and Takeda Pharmaceutical; Klaus Bendtzen has served as a speaker for Pfizer and Biomonitor, and owns stocks in Novo-Nordisk and Biomonitor; Joni Brynskov has served as advisory board member for Abbvie; Ole 0 Thomsen has served as a speaker and consultant for UCB and Zealand Pharma, speaker for MSA, and primary investigator for Amgen, Biogen, Novo-Nordisk, and Pfizer. Mark A. Ainsworth has no interests to declare."
Aim of the study	
Cost-effective guidance of therapeutic strategy in Crohn's disease patients with secondary infliximab	

(IFX) treatment failure may be achieved by serum IFX and anti-IFX antibody (Ab) measurements by radioimmunoassay (RIA). To investigate implications of using other techniques for this purpose.

Inclusion/exclusion criteria for patients

Inclusion criteria: Adult Crohn's disease patients with secondary treatment failure of infliximab, taking regular infusions of 5 mg/kg, and having previous beneficial clinical response. Loss of response defined by a Crohn's Disease Activity Index (CDAI) of greater than or equal to 220 and/or a minimum of one draining perianal fistula.

Exclusion criteria: Contraindication to continuing infliximab, short bowel syndrome, recent history of abdominal surgery or a severe medical condition, pregnancy, or drug or alcohol abuse.

Test comparison

Tests	**Name**	**Details**
Intervention test	Radioimmunoassay (RIA)	"Serum concentrations of infliximab and anti-infliximab Antibodies were measured by RIA, as previously detailed (Biomonitor A/S). In brief, infliximab was assessed as the TNFα-binding capacity of serum by incubation of patient serum with I-TNFα (PerkinElmer, Waltham, MA), followed by separation of free and IgG-bound tracer using rabbit anti-human Fc-gamma Ab (Dako,Copenhagen, Denmark), and detection of the pellet activity using ay-counter (Wallac, Alleroed, Denmark). The infliximab concentration was expressed as the equivalent activity of I-TNFα binding to a reference infliximab solution (MSD, Ballerup, Denmark; limit of quantification (LOQ) 0.15 µg/m1). The RIA for anti-infliximab Abs used antihuman X light-chain Abs to distinguish between free I-infliximab and T-infliximab in complex with any class of 2,-containing human immunoglobulin (infliximab itself is a monoclonal Ab, which consists solely of x light chains). Thus, serum was incubated with and pellet activity was determined after precipitation of immunoglobulin-bound tracer with rabbit anti-human immunoglobulin X-chain Abs (Dako). Anti-infliximab Ab concentrations were expressed as arbitrary units (U) per ml (LOQ 10 U/ml)."
Comparison test 1	ELIZA	"Serum infliximab concentrations were determined by capture ELISA and anti-infliximab Abs by bridging ELISA, as previously described (Prometheus Laboratories). In brief, infliximab sample concentrations were determined using TNFα-coated plates and by the addition of serum and, subsequently, a streptavidin horseradish peroxidase (HRP) anti-human IgG reagent. A colorimetric signal was generated with an HRP substrate (o-phenylenediamine dihydrochloride/H202), and was quantified in a microplate reader at 490 nm (LOQ 1.4 µg/m1). The bridging ELISA for anti-infliximab Abs used infliximab-coated plates, and captured anti-infliximab Abs from the serum sample were interrogated with biotinylated infliximab. A colorimetric signal was generated, similar to the infliximab assay, but by using neutravadin-HRP (LOQ 1.69 ug/m1). By definition, anti-infliximab Abs were considered inconclusive if infliximab was detectable owing to assay interference." "Enzyme immunoassay for IgG4 anti-infliximab Ab. This enzyme immunoassay measures the binding of infliximab to mouse monoclonal antihuman IgG4 Ab (Trikem, Skanderborg, Denmark) preabsorbed to microtiter plastic plates, as previously described. Variable concentrations of patient sera were then tested at a constant serum concentration of 1% (v/v), obtained by supplementing with pooled normal human serum. After

		overnight incubation, the wells were washed and biotinylated infliximab was added. The plates were then washed and HRP was added, and a colorimetric signal was generated by the addition of substrate (tetramethylbenzidine). As an IgG4 anti-infliximab Ab standard is unavailable, the results are given as relative U/ml."
Comparison test 2	HMSA	"Detection of infliximab and anti-infliximab Abs by high-pressure liquid chromatography-based HMSA (AnserIFXTM) was carried out as previously detailed (Prometheus Laboratories) (25). In brief, infliximab concentrations were determined by incubation of Alexa488-labeled TNF-oc with patient serum. After equilibration, free TNFα and TNFα-infliximab complexes were resolved by size-exclusion high-performance liquid chromatography and the peaks were quantified by fluorescence. Concentrations were determined from a standard curve of samples with known infliximab concentrations (LOQ 1 i.t.g/m1). The 1-IMSA for anti- infliximab Abs was done similarly and by the use of labelled infliximab (LOQ 3.13 arbitrary U/ml)."
Comparison test 3	RGA (functional cell based reporter gene assay)	"Functional activities of infliximab and anti-infliximab Abs at the cellular TNF-receptor level were determined by RGA as detailed previously (iLite Infliximab Bioassay and iLite Infliximab NAb Bioassay, respectively; Biomonitor A/S) (20,26). In brief, infliximab-induced TNFα-neutralizing activity in serum was assessed using a human erythroleukemic K562 cell line transfected with an NFKB-regulated firefly luciferase reporter gene construct and, in addition, a Renilla luciferase reporter gene under the control of a constitutive promoter that allows TNFα-induced firefly luciferase to be normalized relative to Renilla luciferase expression The engineered cells were exposed in duplicate to patient serum preincubated with human recombinant TNFα (R&D Systems, Minneapolis, MN), and residual TNFα activity was quantified by luminescence after the addition of Dual-Glo (Promega, Fitchburg, WI). After the addition of Dual-GloStop&Glo reagent (Promega), the Renilla luminescence was read, and the normalized infliximab-induced TNFα neutralization was expressed as infliximab equivalents in p,g/m1 (LOQ 0.65 vg/m1). Anti-infliximab Ab-induced drug neutralization was quantified using the same reporter cells as above, except that the cells were exposed in duplicates to patient serum preincubated with infliximab and, thereafter, human recombinant TNFα. After further incubation, the cells were read for firefly and Renilla luminescence as above. Anti-infliximab Abs were expressed as arbitrary U/ml (LOQ 20 U/ml)."
Details of any repeat measurements (to check reliability, performance across different laboratories)		NR

Drug type tested	
Infliximab	Yes
Anti – Infliximab	Yes
Adalimumab	No
Anti- adalimumab	No

Selection and storage of patients/plasma samples	
Description of method of selection	Samples included in RCT, recruited from 6 centres according to inclusion criteria. No further details

	given.
Description of method and duration of storage	Room temperature storage and immediate analysis by RIA. Biobanking before analysis by ELIZA, HMSA and RGA.
Number of clinical samples	66
Number of calibrator samples (spiked) for anti-TNF	0
Number of calibrator samples (spiked) for antibodies	0
Number of blank (control) samples	0
Total number of plasma samples	66

Results of comparison				
Name of test	RIA	ELIZA	HMSA	RGA
Threshold for drug	≥0.5µg/mL therapeutic [122] ≥0.15µg/mL limit of quantification [121] unclear which threshold results refer to	≥1.4µg/mL	≥3µg/mL therapeutic [122] ≥1µg/mL limit of quantification[121] unclear which threshold results refer to	≥0.65µg/mL limit of quantification
Number positive for drug	54/66 (82%)	50/66 (76%)	58/66 (88%)	49/66 (74%)
Threshold for antibodies	≥10 arbitrary units /mL limit of quantification	≥1.69µg/mL	≥3.13 arbitrary units/mL limit of quantification	≥20 arbitrary units/mL limit of quantification
Number positive for antibodies	18/66 (27%)	6/66 (9%)	22/66 (33%)	7/66 (11%)
Details of correlation/overlap between the tests	Provided for all in figures 2 and 3			
Other information:	Anti-infliximab Ab-positive patients assessed by ELISA and RGA were all found to be positive also in RIA and HMSA. RGA did not report circulating anti-infliximab activity in 15 (68%) of 22 samples testing positive for anti-infliximab Abs by HMSA (12 of which had detectable anti-infliximab Abs in the presence of detectable infliximab) and in 11 (61%) of 18 samples testing positive for anti-infliximab Ab by RIA (6 of which had detectable anti-infliximab Abs in the presence of detectable 1FX), suggesting that the Abs reported by these two binding assays were likely blocked by the drug or alternatively had no drug-neutralizing activity.			

Results of comparison of drug levels						
Name of tests to be compared:	ELISA vs RIA	HMSA vs RIA	RGA vs RIA	ELIZA vs RGA	HMSA vs RGA	HMSA vs ELIZA
Total number concordant/all tested	62/66 (50 positive 12 negative)	62/66 (54 positive 8 negative)	59 /66 (48 positive 11 negative)	61/66 (47 positive 14 negative)	55/66 (48 positive 7 negative)	58/66 (50 positive 8 negative)

Number of positive cases concordant/all positive cases	Actually positive not known as no spiked samples.					
Number of negative cases concordant/all negative cases						
Correlation of drug measurement:						
Regression method	Pearson	Pearson	Pearson	Pearson	Pearson	Pearson
Linearity test/cusum test?	Not reported	Not reported	Not reported	Not reported	Not reported	Not reported
R^2 (95%CI)	0.95	0.96	0.94	0.91	0.95	0.97
Slope (95%CI)	1.53 (1.40 – 1.65)	1.45 (1.35 – 1.55)	0.94 (0.86 – 1.03)	1.46 (1.30 – 1.62)	1.42 (1.30 – 1.53)	0.90 (0.84 – 0.96)
Intercept (95%CI)	Not reported	Not reported	Not reported	Not reported	Not reported	Not reported
From Bland-Altman plot for drug measurement:						
Percent bias (95%CI)	-3	-2.5	0	-3.5	-3	0.5
Upper limit of agreement	4	3	3.5	4	3	4.5
Lower limit of agreement	-10	-8	-3	-11	-8.5	-3
Details of outliers						
Visually is there a pattern between the mean value and the difference?	Yes	Yes	No	Yes	Yes	No
Results of comparison for antibody levels						
Name of tests to be compared:	ELISA vs RIA	HMSA vs RIA	RGA vs RIA	ELIZA vs RGA	HMSA vs RGA	HMSA vs ELIZA
Total number concordant/all tested	54/66 (6 positive 48 negative)	60/66 (17 positive 43 negative)	55/66 (7positive 48 negative)	63/66 (5positive 58negative)	51/66 (7positive 44negative)	50/66 (6positive 44negative)
Number of positive cases concordant/all positive cases	Actually positive not known as no spiked samples.					
Number of negative cases concordant/all negative cases						
Correlation of antibody measurement:						
Regression method	Pearson	Pearson	Pearson	Pearson	Pearson	Pearson
Linearity test/cusum test?	Not reported	Not reported	Not reported	Not reported	Not reported	Not reported
R^2 (95%CI)	0.82	0.77	0.80	0.96	0.78	0.81
Slope (95%CI)	0.09 (0.07 – 0.10)	0.66 (0.52-0.79)	1.16 (0.94 – 1.38)	0.07 (0.06-0.07)	0.46 (0.36 – 0.55)	6.54 (5.34-7.74)
Intercept (95%CI)	Not reported	Not reported	Not reported	Not reported	Not reported	Not reported
From Bland-Altman plot for antibody measurement:						

Percent bias (95%CI)	NR	NR	NR	NR	NR	NR
Upper limit of agreement	NR	NR	NR	NR	NR	NR
Lower limit of agreement	NR	NR	NR	NR	NR	NR
Details of outliers	NR	NR	NR	NR	NR	NR
Visually is there a pattern between the mean value and the difference?	NR	NR	NR	NR	NR	NR

Authors' conclusion

Despite variable analytical properties, common assays result in similar classifications and interventions in patients with infliximab treatment failure, and with comparable clinical outcomes. Implications are, however, profound for the minority classified differently.

Reviewer's conclusion

For drug level slope of regression line is significantly different from 1 when comparing ELIZA to RGA or RIA, and when comparing HMSA to RGA and RIA, but not when comparing HMSA to ELIZA or RIA to RGA. When the slope is not 1 then the difference between the two measurements is dependent on the drug levels. In such cases the Bland- Altman summary statistics are not applicable (as they are dependent on absolute values).

The best agreement is between RIA and RGA and HMSA and ELIZA.

Authors present Bland-Altman plots and graphically present which tests agree on drug and anti-drug presence (using unclear cut-off) which should be included in our report if possible.

Name of first reviewer: Sian Taylor-Phillips *Name of second reviewer: Martin Connock*

Study details	
Study ID (Endnote ref)	[66]
First author surname	Vande Casteele
Year of publication	2012
Country	Belgium and the Netherlands
Study design	
Publication (full/abstract)	Full
Study setting	
Number of centres (by arm)	3
Duration of study	NA
Follow up period	NA
Funding	This study was funded in part by the Fund for Scientific Research-Flanders, grant number G.0617.12.
Competing interests	Séverine Vermeire has served as a speaker, a consultant and an advisory board member for Centocor, Abbott, UCB, Shering-Plough, MSD, Ferring, Pfizer, Chiesi, Dr Falk Pharma, and has received research funding from Centocor, MSD, Abbott and UCB. Theo Rispens has served as a speaker for Pfizer and has received research funding from Genmab. Desiree van der Kleij is an employee of Sanquin Diagnostic Services. Ann Gils has served as a speaker for Pfizer and MSD. Gerard Dijkstra has received research funding from MSD BV, the Netherlands, Abbott BV, the Netherlands and Dr Falk Pharma Benelux B.V, the

		Netherlands.
Aim of the study		
To determine the correlation between three different assays for measuring infliximab and ATI. (ATI means antibodies to infliximab)		
Inclusion/exclusion criteria for patients		
All three institutions delivered serum samples derived either from the department of gastroenterology or from the department of rheumatology and Leuven and Sanquin provided quality control samples. No clinical details of patients were collected. A total of 62 samples were analysed by all three institutes. Thirty six samples were clinical samples from patients containing different concentrations of infliximab and ATI. The other 26 samples were calibrator samples in which a serum pool of healthy controls was spiked with known concentrations of infliximab (n = 10), adalimumab (n = 1), antibodies to infliximab (ATI) (n = 10) or antibodies to ADA (ATA) (n = 3) and two blank samples with only serum of healthy controls.		
Test comparison		
Tests	**Name**	**Details**
Intervention test	LISA Tracker premium infliximab kit (Infliximab and antibodies to infliximab)	"In Groningen, the commercially available LISA-TRACKER Premium Infliximab kit (BMD Biomedical Diagnostics, Marne La Vallée, France) was used to measure infliximab and antibodies to infliximab. This kit is an ELISA and has a CE-label according to Directive 97/98/CE. Lower limit of quantification for infliximab levels is 0.1 mg/L. The lower and upper limits of quantification for antibodies to infliximab are 10 and 200 lg/L. The kit was used conform to manufacturer instructions by a qualified person."
Comparison test 1	Leuven in-house ELISA	"In Leuven, an in house developed direct ELISA was used to measure infliximab based on a previously described method. Briefly, high binding 96-well plates (Costari; Corning Inc., Corning, NY, USA) were coated overnight with TNFα (PeproTech, London, UK) at 4°C. Plates were blocked with PBS/1% bovine serum albumin (BSA) (Sigma Aldrich) for 2 h at room temperature and samples were diluted in PBS/1% BSA and incubated for 2 h at 37°C. As detecting antibody horse radish peroxidase (HRP) linked monospecific rabbit polyclonal antibody (made in house) was used and plates were incubated at RT for 1 h. Plates were developed using 400 lg/mL o-Phenylenediamine (Acros Organics, Geel, Belgium) and 0.003% (v/v) H2O2 in 0.1M sodium citrate 0.2M disodium phosphate buffer pH 5. The reaction was stopped with 2M H2SO4. Absorption at 490 nm was measured using an ELX808IU reader (Bio-tek Instruments Inc.). Results were related to a titration curve of infliximab on each plate. The cut off for an infliximab positive sample was 0.3 mg/L. To measure antibodies to infliximab, an in house developed bridging ELISA was used. Briefly, high binding 96-well plates (Costar; Corning Inc.) were coated for 72 h with infliximab (Janssen Biologics, Leiden, the Netherlands) at 4°C. Plates were blocked with PBS/ 1% BSA (Sigma Aldrich) for 2 h at room temperature and samples were diluted in PBS/0.1% BSA/0.002% (v/v) Tween 80 and incubated overnight at 4°C. As detecting antibody HRP linked infliximab (made in house) was used and plates were incubated at RT for 2 h. Plates were developed using 400 lg/mL o-Phenylenediamine (Acros Organics) and 0.003% (v/v) H2O2 in 0.1M

		sodium citrate 0.2M di-sodium phosphate buffer pH 5. The reaction was stopped with 4M H2SO4. Absorption at 490 nm was measured using an ELX808IU reader (Bio-tek Instruments Inc.). Results were related to a titration curve of monospecific rabbit polyclonal antibody to infliximab on in each plate. The cut off for an ATI positive sample was 1 mg/L equivalents."
Comparison test 2	Amsterdam, Sanguin in house ELISA	"In Amsterdam, at Sanquin, an in house developed ELISA was used to measure infliximab using the same procedures as described for adalimumab. Maxisorp ELISA plates were coated overnight with 2 lg/mL monoclonal anti-TNF-7 (Sanquin) in Phosphate buffered saline (PBS) at room temperature (RT). After five times washing with PBS/0.02% Tween (PT), plates were incubated for 1 h at RT with recombinant TNFa (0.01 lg/mL) (Strathmann Biotech HmbH, Hannover, Germany) diluted in high performance ELISA buffer (HPE, Sanquin Bloodbank, Division Reagents). Next, the plates were washed and incubated for 1 h with patient serum, which was serially diluted in HPE. Subsequently, the plates were washed with PT and incubated for 1 h with biotinylated infliximab specific rabbit anti-idiotype antibody (0.25 lg/mL in HPE). After washing, streptavidin-poly-HRP (Sanquin) (1/25 000, in HPE) was added for 1 h at 37°C. After washing, the ELISA was developed with 100 lg/mL tetramethylbenzidine in 0.11M sodium acetate (pH 5.5) containing 0.003% (v/v) H2O2. The reaction was stopped with 2M H2SO4. Absorption at 450 nm was measured using an ELX808IU reader (Biotek Instruments Inc., Winooski, VT, USA). Results were related to a titration curve of infliximab on each plate. The lowest level of quantification was 0.002 mg/L. To measure antibodies to infliximab, an in house developed RIA was used. Briefly, one microlitre of serum diluted in Freeze buffer was incubated with 1 mg protein A Sepharose (GE healthcare, Chalfont St. Giles, UK) in 800 lL of total volume. After overnight incubation, samples were washed and 125 I radioactive labeled infliximab F (ab')2 fragments were added. After overnight incubation, unbound radiolabel was washed out and Sepharosebound radioactivity was measured. Results of this test are commonly expressed by Sanquin as Arbitrary Units/ mL, where 1 AE/mL equals approximately 10 lg/L. The lower limit of quantification is 12 AE/mL."
Comparison test 3	NA	NA
Details of any repeat measurements (to check reliability, performance across different laboratories)		NR
Drug type tested		
Infliximab		Yes

Anti – infliximab	Yes
Adalimumab	No
Anti- adalimumab	No

Selection and storage of patients/plasma samples	
Description of method of selection	Samples from all three institutions from the department of gastroenterology or rheumatology.
Description of method and duration of storage	Not given
Number of clinical samples	36
Number of calibrator samples (spiked) for anti-TNF	10 infliximab +1 adalimumab
Number of calibrator samples (spiked) for antibodies	10 antibodies to infliximab and 3 antibodies to adalimumab
Number of blank (control) samples	2
Total number of plasma samples	62

Results of comparison			
Name of test	LISA Tracker	Leuven in-house ELISA	Amsterdam in-house ELISA
Threshold for drug	0.1mg/L lower LOQ (unclear if this is also cut-off)	Cut off 0.3 mg/L	LLOQ 0.002 mg/L
Number positive for drug	NR	NR	NR
Threshold for antibodies	10-200µg/L (upper and lower LOQ) (unclear if this is also cut-off)	Cut-off 1mg/L	12 AE/mL (1 AE/mL equals approximately 10 µg/L)
Number positive for antibodies	NR	NR	NR
Details of correlation/overlap between the tests	Drug: "Both Leuven and LISA TRACKER infliximab assays detected infliximab in one healthy control sample spiked with adalimumab. Furthermore, LISA TRACKER infliximab assay detected infliximab in 11 out of 62 samples (18%), not detected using Leuven in-house assay. Five out of these 11 samples were calibrator samples, of which two samples only contained antibodies to infliximab and three samples only contained antibodies to adalimumab. The remaining six samples were patient samples, all containing high levels of antibodies to infliximab." Antibodies to drug: "Leuven Assay did not detect antibodies to infliximab in three patient samples with low levels of antibodies to infliximab according to the LISA TRACKER assay."		
Other information:	Four samples excluded as above the upper limit of quantification for LISA TRACKER infliximab (200µg/L), Six samples excluded as above the upper limit of quantification for LISA TRACKER antibodies to infliximab (not given) Include a version of regression and B-A graphs		

Results of comparison of drug levels			
Name of tests to be compared:	LISA-TRACKER vs Leuven	Leuven vs Amsterdam	Amsterdam vs LISA-TRACKER
Total number concordant/all tested	Up to 47/58	Up to 57/58	Up to 46/58

Number of positive cases concordant/all positive cases	Partially reported as follows: Qualitatively, infliximab assays B (Leuven) and C (LISA TRACKER) detected infliximab in one healthy control sample spiked with adalimumab, whereas in infliximab assay A (Amsterdam), this sample was negative. "Furthermore, infliximab assay C detected infliximab in 11 out of 62 samples (18%), not detected using infliximab assays A and B. Five out of these 11 samples were calibrator samples, of which two samples only contained antibodies to infliximab and three samples only contained antibodies to adalimumab. The remaining six samples were patient samples, all containing high ATI levels."		
Number of negative cases concordant/all negative cases			
Correlation of drug measurement:			
Regression method	Pearson	Pearson	Pearson
Linearity test/cusum test?	NR	NR	NR
R^2 (95%CI)	0.53 (NB reported Pearsons r=0.73)	0.83 (NB reported Pearsons r=0.91)	0.69 (NB reported Pearsons r=0.83)
Slope (95%CI)	NR	NR	NR
Intercept (95%CI)	NR	NR	NR
From Bland-Altman plot for drug measurement:			
Percent bias (95%CI)	0 (from visual inspection)	0 (from visual inspection)	0 (from visual inspection)
Upper limit of agreement	15 mg/L (from visual inspection)	8mg/L (from visual inspection)	10 mg/L (from visual inspection)
Lower limit of agreement	-15mg/L (from visual inspection)	-8mg/L (from visual inspection)	-10mg/L (from visual inspection)
Details of outliers	One with around 60mg/L difference	One with around 30mg/L difference	Five between -20 and +25mg/L difference
Visually is there a pattern between the mean value and the difference?	B-A shows no clear pattern, but regression plot shows 11 samples positive for LISA TRACKER and zero on Leuven assay. 4 samples above upper limit for LISA TRACKER excluded from analysis	B-A shows no pattern 4 samples above upper limit for LISA TRACKER excluded from analysis	B-A shows no clear pattern, 4 samples above upper limit for LISA TRACKER excluded from analysis
Results of comparison for antibody levels			
Name of tests to be compared:	LISA-TRACKER vs Leuven	Leuven vs Amsterdam	Amsterdam vs LISA-TRACKER
Total number concordant/all tested	up to 59/62	Up to 54/62	Up to 57/62
Number of positive cases concordant/all positive cases	Qualitative analysis showed that ATI assay A (Amsterdam) detected ATI in five samples with low ATI that was not detected by ATI assays B (Leuven) and C (LISA-TRACKER). ATI Assay B did not detect ATI in three patient		
Number of negative cases			

concordant/all negative cases	samples with low ATI according to ATI assays A and C.		
Correlation of antibody measurement:			
Regression method	Pearson	Pearson	Pearson
Linearity test/cusum test?	NR	NR	NR
R^2 (95%CI)	0.94 (NB reported Pearsons r=0.97)	0.90 (NB reported Pearsons r=0.95)	0.98 (NB reported Pearsons r=0.99)
Slope (95%CI)	NR	NR	NR
Intercept (95%CI)	NR	NR	NR
From Bland-Altman plot for antibody measurement:			
Percent bias (95%CI)	None as on different scales	None as on different scales	None as on different scales
Upper limit of agreement	NA	NA	NA
Lower limit of agreement	NA	NA	NA
Details of outliers	NA	NA	NA
Visually is there a pattern between the mean value and the difference?	NA	NA	NA
Authors' conclusion			
"There is a good correlation of infliximab and antibodies to infliximab measurements between these assays. Nevertheless, the Biomedical Diagnostics kit [LISA TRACKER] detected false positive infliximab levels in 18% of the samples."			
Reviewer's conclusion			
Spiked and patient samples were not reported separately which makes conclusions more difficult to draw. For measuring infliximab drug levels there is some evidence that the LISA TRACKER assay gives false positive results in the presence of adalimumab, antibodies to infliximab and antibodies to adalimumab. The Leuven assay may be less sensitive to detecting the presence of antibodies to infliximab than LISA TRACKER.			

Name of first reviewer: Sian Taylor-Phillips Name of second reviewer: Martin Connock

Study details	
Study ID (Endnote ref)	[125]
First author surname	Vande Casteele
Year of publication	2013
Country	Leuven, Belgium, and Prometheus, California
Study design	
Publication (full/abstract)	Full
Study setting	Laboratory
Number of centres (by arm)	1
Duration of study	NR
Follow up period	NR
Funding	The design and conduct of the study, data analysis, and manuscript writing was performed independently by the authors. All authors had access to the data and decided to jointly submit the manuscript. infliximab and antibody to infliximab levels were analyzed by Prometheus Laboratories, San Diego, CA, USA that also provided additional research support funding.
Competing interests	Potential competing interests: N.V.C. has no conflict of interests. A.G. has served as a speaker for Pfizer and MSD. S.S., L.O., and S.H. are employers of Prometheus Laboratories. P.R. has served as a speaker and a consultant for Centocor, Merck, and Abbott, and has received research funding from UCB, Centocor, Merck, and Abbott. S.V. has served as a speaker for UCB, Abbott, MSD, Ferring, Centocor, and Chiesi, and has received research funding from Abbott, Centocor, MSD, and UCB.

Aim of the study
Our aim was to investigate the kinetics of [antibodies to infliximab] ATI formation and drug levels in relation to inflammatory markers and the clinical evolution of the patients.
Inclusion/exclusion criteria for patients
Patients selected from inflammatory bowel disease biobank in Leuven based on testing positive to antibodies to infliximab using the Leuven in-house ELIZA.

Test comparison

Tests	**Name**	**Details**
Intervention test	Leuven in – house ELIZA	Same as Vande Castille 2012 paper.
Comparison test 1	HMSA (Promethius lanboratories)	ATI [antibodies to infliximab] were measured using a HMSA “Samples and calibrators were diluted in phosphate-buffered saline, pH 7.3 (1:25 dilutions final) and mixed with infliximab 488 containing internal control in a 0.5 ml round bottom polypropylene 96-well sample plate (Nunc°, Thermo Fisher Scientific, Waltham, MA) with a total volume of 300 [11. The reaction mixture was incubated at room temperature for 1 h on a plate shaker in the dark, and then filtered through a 0.2 pm 96-well filter plate (Millipore, Billerica, MA) to a 96-well collection plate. The filtered samples were transferred to a high-pressure liquid chromatography (HPLC) sample vial and loaded into the sample chamber with temperature maintained at 4°C during the entire run. Hundred microliters from each sample was loaded to a SEC-3000 column (Phenomenex, Torrance, CA) and

		controlled by an Agilent (Santa Clara, CA) 1200 system equipped with fluorescent detector. 1FX488 was monitored by fluorescent detector, which was optimized to the excitation and emission of 494 and 519 run. The sample was run at the flow rate of I ml/min for 20 min with phosphate-buffered saline as mobile phase. CherriSta-Lion was used to set up the run and retrieve the data. All ATI measurements were carried out after an acid dissociation step: serum samples were first incubated with 0.5M citric acid, pH 3.0 for 1 h at room temperature. Following the dissociation, infliximab 488/internal control was added and the reaction mixture was immediately neutralized with 10x phosphate-buffered saline, pH 7.3. The reaction was continued for another hour at room temperature on a plate shaker to complete the reformation of the immune complexes. The samples were filtered and analyzed with size exclusion chromatography in combination with HPLC (SEC—HPLC) (Figure la). All ATI data are expressed as Units/ml (U/ml) and a sample was considered ATI positive when ATI 7.95U/ml. Values < 7.95 U/m1 were represented as 0 U/ml." TLI [Trough levels of Infliximab] were measured using a HMSA "The procedures for the infliximab assay were similar to the ATI assay without acid dissociation. infliximab spiked in normal healthy serum was used as calibrator. The assay was performed by incubating TNF 488/internal control with serum samples or calibrator to reach equilibrium. The reaction mixture was filtered and analyzed by SEC-HPLC. All TLI data are expressed asn/m1 and a sample was considered infliximab positive when TLI ?..0.911.tg/rni. Values < 0.911.1g/m1 were represented as 0µg/ml."
Comparison test 2	NA	NA
Comparison test 3	NA	NA
Details of any repeat measurements (to check reliability, performance across different laboratories)		NR
Drug type tested		
Infliximab		Yes
Anti – Infliximab		Yes
Adalimumab		No
Anti- adalimumab		No
Selection and storage of patients/plasma samples		
Description of method of selection		We have performed a retrospective analysis of 90 IBD patients treated with infliximab between May 1999 and August 2011. Patients gave written consent to participate in the I RB-approved Vlaamse Erfelijkheidsstudie Crohn en Colitis ulcerosa (VLECC) registry (8322201213950/S53684), a biobank containing serum, DNA, and clinical characteristics of IBD patients followed at the University Hospital Leuven, Belgium. Patients were

		selected based on a retrospective screening for ATI with an in-house-developed enzyme-linked immunosorbent assay (ELISA)
Description of method and duration of storage		Within this registry, serial serum samples of anti-TNF-treated patients are prospectively collected and stored at 20°C.
Number of clinical samples		1232 from 90patients (64 Crohns, 26 ulcerative colitis)
Number of calibrator samples (spiked) for anti-TNF		0
Number of calibrator samples (spiked) for antibodies		0
Number of blank (control) samples		0
Total number of plasma samples		1232
Results of comparison		
Name of test	Leuven ELIZA	HMSA
Threshold for drug	0.3 mg/L (from VC 2012)	0.92µg/ml
Number positive for drug	Not given	D+A+=42 (3%) D+A-=701 (57%)
Threshold for antibodies	1mg/L (from VC 2012)	7.95 U/ml
Number positive for antibodies	Not given	D-A+=266 (22%) D-A-=223 18%
Details of correlation/overlap between the tests	"When measured by ELISA four more patients were classified as having "transient ATI" that were all ATI negative in the HMSA. One patient was classified as having "sustained ATI" when analysed by ELISA in comparison with having "transient ATI" when analysed by HMSA. Furthermore, one patient was ATI negative in ELISA and classified as having "transient ATI" in the HMSA." "Surprisingly, our ELISA was more sensitive in detecting low level ATI in the absence of infliximab. Based on the results obtained by ELISA, we identified four more patients with transient ATI that were all ATI negative when analyzed by HMSA (nine samples). We ruled out false positivity in those as TLI was undetectable or low in surrounding samples and also clinically all four patients suffered LOR to infliximab in the immediate vicinity of the ATI-positive sample. Furthermore, one patient was classified as having "transient ATI" in the HMSA in comparison with "sustained ATI" in our ELISA: on the last time point available of this patient, ATI levels were just below the detection limit in the HMSA but positive in ELISA. Interestingly, infliximab treatment was discontinued in this patient owing to persistent LOR to infliximab. The difference in sensitivity probably boils down to the cutoff that was used for an ATI-positive sample, based on the background or noise in each assay. The cutoff was defined by the mean of at least 20 drug-naive patient samples + 3x s.d., hence the cutoff for ATI in the ELISA was set at 1 µg/ml equivalents in comparison with 7.9511/ml for the HMSA. The different outcome measures of both assays make it difficult to compare absolute values (µg/ml equivalents vs. U/mI). Prometheus Laboratories has informed us that they have now lowered the lower limit of quantification		

	in their commercialized HMSA to 3.13 U/ml. In the presence of infliximab, the HMSA was more sensitive in detecting ATI than ELISA: one patient was classified as having transient ATI when analyzed by HMSA and as ATI negative when analyzed by ELISA. Overall, ATI were detected approximately one time point earlier by HMSA (median 16 weeks) in comparison with ELISA (median 25 weeks) after start of infliximab. This is owing to a lower susceptibility to drug when measuring ATI in the HMSA compared with ELISA and also because antibodies with a low affinity can be picked up by HMSA."
Other information:	
Results of comparison of drug levels	
Name of tests to be compared:	HMSA vs Leuven ELISA
Total number concordant/all tested	NR
Number of positive cases concordant/all positive cases	NR
Number of negative cases concordant/all negative cases	NR
Correlation of drug measurement:	
Regression method	Pearson
Linearity test/cusum test?	NR
R^2 (95%CI)	0.69 (R=0.83)
Slope (95%CI)	NR
Intercept (95%CI)	NR
From Bland-Altman plot for drug measurement:	
Percent bias (95%CI)	NR
Upper limit of agreement	NR
Lower limit of agreement	NR
Details of outliers	NR
Visually is there a pattern between the mean value and the difference?	NR
Results of comparison for antibody levels	
Name of tests to be compared:	
Total number concordant/all tested	NR
Number of positive cases concordant/all positive cases	NR
Number of negative cases concordant/all negative cases	NR
Correlation of antibody measurement:	
Regression method	Pearson
Linearity test/cusum test?	NR
R^2 (95%CI)	0.77 (R=0.88)
Slope (95%CI)	NR
Intercept (95%CI)	NR

From Bland-Altman plot for antibody measurement:	
Percent bias (95%CI)	NR
Upper limit of agreement	NR
Lower limit of agreement	NR
Details of outliers	NR
Visually is there a pattern between the mean value and the difference?	NR
Authors' conclusion	
Leuven ELIZA was more sensitive in detecting low-level ATI in the absence of infliximab. The ELISA is a suitable all round assay for analysing infliximab ad ATI in the majority of samples.	
Reviewer's conclusion	
HMSA performs better at detecting antibodies to infliximab in the presence of infliximab, quantifying this is difficult due to reporting focussed on other research questions, it is described in the discussion as HMSA detecting median 9 weeks earlier. With HMSA cut-off for anti-drug antibodies to infliximab at7.95U/ml, the Leuven in-house ELISA detected four more cases with anti-drug antibodies, and the authors report Prometheus have since lowered the threshold to 3.13 U/mL.	

Name of first reviewer: Sian Taylor-Phillips *Name of second reviewer: Martin Connock*

Study details	
Study ID (Endnote ref)	[130]
First author surname	Wang
Year of publication	2012
Country	USA
Study design	
Publication (full/abstract)	Full
Study setting	Laboratory
Number of centres (by arm)	n/a
Duration of study	n/a
Follow up period	n/a
Funding	Prometheus Laboratories
Competing interests	Authors all Prometheus employees

Aim of the study

Current methods for the assessment of anti-drug antibodies and drug levels, involving various bridging ELISA and radioimmunoassay techniques, are limited by their sensitivity, interference, and/or complexity. To overcome these limitations, we have developed a non-radiolabeled homogeneous mobility shift assay (HMSA) to measure the antibodies-to-infliximab (ATI) and infliximab levels in serum samples. Full method validation was performed on both the ATI- and infliximab-HMSA, and the clinical sample test results were also compared with those obtained from a bridging ELISA method to evaluate the difference in performance between the two assays.

Inclusion/exclusion criteria for patients

Individual serum samples from healthy controls were obtained from blood bank donors. Sera from IBD patients treated with infliximab were obtained from residual samples leftover after testing for ATI and infliximab levels in our laboratories and the patient information was de-identified.

ATI-positive sera were prepared by pooling individual patient serum samples identified as containing high concentrations of ATI and negative for infliximab

Test comparison

Tests	**Name**	**Details**
Intervention test	Prometheus ELISA	"In brief, the ATI bridging ELISA is a microplate based, double antigen formatted assay where infliximab is coated on the solid phase 96-well plate to capture the ATI from the patient serum samples. The captured ATI is detected through binding to a biotinylated infliximab. The amount of bound biotin on the microplate is determined with the addition of a neutravidin-HRP conjugate which transforms the substrate O-phenylenediamine to a chromogenic product that is measured in a microplate reader at 490 nm. In the bridging ELISA, an affinity purified polyclonal rabbit antimouse IgG F(ab')2 (Thermo Fisher Scientific, Waltham, MA) is used to generate the standard curve for calculation of the relative amount of ATI in the patient serum sample."
Comparison test 1	HMSA	"ATI homogenous mobility shift assay (ATI-HMSA, [antibodies to infliximab]) The assay was prepared in a 96-well plate format. In order to reduce interference from circulating drug, an acid dissociation step was employed. Briefly, a solution containing a 24 µL aliquot of serum sample, 5.5 µL 0.5 M citric acid (pH 3.0), and

		10.9 µL HPLC grade water were added to each well and incubated for 1h at RT to free the ATI in the patient serum samples from other bound proteins. Following the acid dissociation step, 6 µL of a 74 µg/mL infliximab-488/IC solution was added and the reaction mixture was immediately neutralized with 27.6 µL of 10x PBS (pH 7.3). The plate was incubated for another hour at RT on an orbital shaker to complete the formation of the immune complexes. The incubated serum samples were then diluted to a final serum concentration of 2% by pipetting 18.4 µL of each sample solution, 22.6 µL 10× PBS (pH 7.3), and 259 µL HPLC grade water into the wells of a new 96-well plate. In this plate, the first four wells contained, respectively: 300 µL each of HPLC buffer as a blank, aqueous SEC1 column standard (Phenomenex, Torrance, CA) to monitor the resolution of the HPLC column, acid-dissociated 2% NHS, and acid-dissociated 2% NHS with 110 ng infliximab-488/IC for calibrating the HPLC system. The next eight wells contained 300 µL each of the ATI calibration standards (0.006, 0.011, 0.023, 0.045, 0.090, 0.180, 0.360, and 0.720 µg/mL) with 110 ng infliximab-488/IC for generating the standard curve. The next nine wells contained, respectively, 300 µL each of the three QC controls (high, mid and low) in triplicate with 110 ng infliximab-488/IC to establish the precision and accuracy of the assay. The remaining wells were then filled with 300 µL of the prepared patient serum samples. After mixing on an orbital shaker for 1 min at RT, the samples were filtered through a MultiScreen-Mesh Filter plate equipped with a Durapore membrane (0.22 µm; EMD Millipore, Billerica, MA) into a 96-well receiver plate (Nunc, Thermo Fisher Scientific, Waltham, MA). The recovered solutions in the receiver plate were then transferred sequentially to the loading vials of an autosampler at 4 °C in an Agilent Technologies 1200 series HPLC system (Santa Clara, CA). A 100 µL aliquot from each vial was loaded onto a BioSep SEC3000 column (Phenomenex, Torrance, CA) and the column effluent was monitored by a fluorescent detector at excitation and emission wavelengths of 494 nm and 519 nm, respectively. The chromatography was run at the flow-rate of 1 mL/min for a total of 20 min with 1× PBS (pH 7.3) as the mobile phase. ChemStation Software (Agilent Technologies, Santa Clara, CA) was used to set up and collect data from the runs automatically and continuously. The time needed to process all the calibration standards, controls, and 35 patient serum samples was ~22 h for a single HPLC system. Infliximab homogeneous mobility shift assay (infliximab-HMSA) The procedure for the infliximab-HMSA was similar to the ATI-HMSA, except that the acid dissociation step was omitted in the preparation of the patient serum samples. infliximab spiked in pooled NHS were used as calibration standards. The assays were performed by incubating the TNF-488/IC with serum samples or calibration standards to reach equilibrium. As in the ATI-HMSA method, the reaction mixtures were then filtered and analyzed by the SE-HPLC system.”
Comparison test 2	NA	NA
Comparison test 3	NA	NA
Details of any repeat measurements (to check		Reliability of repeat measurements of HMSA detailed, but no equivalent for ELISA.

reliability, performance across different laboratories)		
Drug type tested		
Infliximab	No (no data comparing ELISA to HMSA is presented)	
Anti – Infliximab	Yes	
Adalimumab	No	
Anti- adalimumab	No	
Selection and storage of patients/plasma samples		
Description of method of selection		Individual serum samples from healthy controls were obtained from blood bank donors. Sera from IBD patients treated with infliximab were obtained from residual samples leftover after testing for ATI and infliximab levels in our laboratories and the patient information was de-identified.
Description of method and duration of storage		NR
Number of clinical samples		100 serum samples from IBD patients previously tested positive to antibodies to infliximab using bridging ELISA
Number of calibrator samples (spiked) for anti-TNF		0
Number of calibrator samples (spiked) for antibodies		0
Number of blank (control) samples		100 serum samples from infliximab drug-naïve healthy subjects
Total number of plasma samples		200
Results of comparison		
Name of test	Prometheus Bridging ELISA	HMSA
Threshold for drug	NR	0.98 µg/mL
Number positive for drug	NR	NR
Threshold for antibodies	NR	1.19µg/mL
Number positive for antibodies	NR for healthy controls 100/100 patients (used as inclusion criteria)	3/100 healthy controls (all 3 were negative upon re-testing) 95/100 patients
Details of correlation/overlap between the tests	NR	
Other information:	Regression plot for antibodies provided	
Results of comparison of drug levels		
Name of tests to be compared:	Prometheus Bridging ELISA vs HMSA	
Total number concordant/all tested	NR	
Number of positive cases concordant/all positive cases	NR	
Number of negative cases concordant/all negative cases	NR	
Correlation of drug measurement:		

Regression method	NR
Linearity test/cusum test?	NR
R^2 (95%CI)	NR
Slope (95%CI)	NR
Intercept (95%CI)	NR
From Bland-Altman plot for drug measurement:	
Percent bias (95%CI)	NR
Upper limit of agreement	NR
Lower limit of agreement	NR
Details of outliers	NR
Visually is there a pattern between the mean value and the difference?	NR
Results of comparison for antibody levels	
Name of tests to be compared:	Prometheus Bridging ELISA vs HMSA
Total number concordant/all tested	NR
Number of positive cases concordant/all positive cases	NR
Number of negative cases concordant/all negative cases	NR
Correlation of antibody measurement:	
Regression method	Spearman
Linearity test/cusum test?	NR
R^2 (95%CI)	r=0.39 (0.2-0.55) so r^2=0.15
Slope (95%CI)	NR
Intercept (95%CI)	NR
From Bland-Altman plot for antibody measurement:	
Percent bias (95%CI)	NR
Upper limit of agreement	NR
Lower limit of agreement	NR
Details of outliers	NR
Visually is there a pattern between the mean value and the difference?	NR
Authors' conclusion	
There was a high correlation between the two methods for ATI levels (pb 0.001). Significantly, the new method identified five false-positive samples from the bridging ELISA method. Validation of the mobility shift INFLIXIMAB assay also showed high assay sensitivity, precision and accuracy. The HMSA method may also be applied to other protein-based drugs to accurately detect serum drug and anti-drug antibody levels.	
Reviewer's conclusion	
The focus of this paper was validating the performance of HMSA, rather than comparing it to ELISA. Out of 100 healthy controls 3 were false positive for antibodies to infliximab for HMSA. This was to be expected as the cut point was determined from the same samples as mean +2SD. Repeat measurements of these three resulted in them being below the cut-point, presumably regression to the mean. ELISA results for the 100 healthy controls not reported. Out of 100 inflammatory bowel patients selected as positive for antibodies for infliximab on ELISA, 5 did not test positive on HMSA. The authors attribute this to elevated levels of non-specific binding in the ELISA. As we don't have the equivalent data for ELISA results on samples that tested positive using HMSA it is difficult to draw any conclusions at all. The only comparative data is a plot of correlation which does not	

appear to show high correlation.
This paper may be useful for the first part of objective A detailing how the assays work, but is not very informative in making comparisons between assays.

Data extraction form for anti-TNFα drug monitoring: Management studies

Name of the first reviewer: Deepson S Shyangdan Name of second reviewer: Martin Connock

Study details			
Study ID (Endnote ref)	[122]		
First author surname	Steenholdt		
Year of publication	2014		
Country	Denmark		
Study design	Randomised controlled, single-blind trial		
Publication (full/abstract)	Full		
Study setting	Not clear (but looking at authors affiliation, it appears that the participating centres were University hospitals)		
Number of centres (by arm)	Six Danish Centres		
Duration of study	12 weeks		
Follow up period	At week 0, 4, 8 and 12		
Funding	Aase and Ejnar Danielsen's Foundation, Beckett Foundation, Danish Biotechnology Program, Danish Colitis-Crohn Society, Danish Medical Association Research Foundation, Frode V Nyegaard and Wife's Foundation, Health Science Research Foundation of Region of Copenhagen, Herlev Hospital Research Council, Lundbeck Foundation, P Carl Petersen's Foundation, Ole Ostergaard Thomsen's Research Foundation and Jorn Brynskov's Research Foundation		
Aim of the study			
To investigate the cost-effectiveness of interventions defined by an algorithm designed to identify specific reasons for therapeutic failure.			
Inclusion/exclusion criteria for patients			
Inclusion criteria	Adult patients diagnosed with Crohn's disease and a previous beneficial clinical response to standard IFC maintenance therapy with regular infusions of 5 mg/kg. At inclusion, all patients had secondary IFX treatment failure on IFX maintenance therapy defined as recurrence of active disease with Crohn's Disease Activity Index (CDAI) score of ≥220 and/or presence of at least one draining perianal fistula.		
Exclusion criteria	Any contraindication to continued IFX, short bowel syndrome, recent history of abdominal surgery or of a severe medical condition, pregnancy, or alcohol or drug abuse		
Study design			
N of screened	95		
N of excluded (ineligible)	26		
Randomisation / blinding	Randomised to algorithm or infliximab intensification groups using block randomisation [block size = 20] using sequentially numbered opaque envelopes Patients blinded to randomisation group and results of serum analyses. Physicians were blinded to IFX and IFX Ab test results in the intensification arm only		
N randomised	36 to dose intensification, 33 to algorithm treatments		
N of non-participants	14 not treated according to algorithm protocol [n=7 continued IFX no assessment; n=5 continued IFX no inflammation; n=2 misinterpreted analyses		
Item	IFX intensified arm	Algorithm arm	All
N study sample at baseline randomised (if applicable)	NA	NA	NA
Withdrawals	8 [n=7 lack of effect; n=1 severe infusion reaction]	2 [lack of effect]	10
Lost to follow up/drop outs (sample attrition)	unclear	unclear	unclear
Study flow(consort diagram)			

Available in paper

Treatment ALGORITH for patients randomised to algorithm arm

	Detectable anti-infliximab antibodies	Undetectable anti-infliximab antibodies
Sub-therapeutic infliximab < 0.5 ug/mL	**Group 1**	**Group 2**
	Insufficient infliximab bioavailability due to induced immunogenicity of infliximab	Insufficient infliximab bioavailability due to non-immune mediated pharmacokinetics of infliximab
	⇩	⇩
	Change to different TNFα-inhibitor: Adalimumab 80 mg sc at inclusion followed by 40 mg sc every other week: dose intensification allowed	Intensify infliximab treatment: infliximab 5 mg/kg iv every 4 weeks
Therapeutic infliximab ≥ 0.5 ug/mL	**Group 4**	**Group 3**
	Consider: (A) pharmacodynamics (B) non-functional anti-infliximab antibodies (C) false positive test	Pharmacodynamics: inhibition of TNFα is ineffective due to non-TNFα driven disease.
	⇩	⇩
	Repeat infliximab and anti-infliximab antibody analyses and handle accordingly. If unchanged results, then act as group 3	TNFα-inhibitors not effective discontinued. Review of clinical condition at discretion of the investigator: if relapse of CD, use drug(s) with other target, e.g., conventional immune-suppressives, glucocorticoids, and/or other biological agents. Consider surgery if appropriate. If no relapse, treat underlying problem

Participants (characteristics and numbers)

Item	Intensification arm	Algorithm arm	All
Total number of participants at baseline (% CD) all patients CD	36	33	69
N (%) followed up	Unclear	Unclear	Unclear
N (%) included in analysis	36 ITT (100), 36 per protocol (100)	33 ITT (100), 19 per protocol (58)	69 ITT (100), 55 per protocol (80)
Patient group (responders/ secondary loss of response)	Secondary loss of response	Secondary loss of response	Secondary loss of response
Age Mean (range) years	37 (19 to 63)	36 (19 to 81)	37 (19 to 81)
Sex Women n (%)	20 (61)	22 (61)	42 (61)
Diagnostic criteria for CD	CDAI Presence of fistulas	CDAI Presence of fistulas	CDAI Presence of fistulas
Children n (%)	None	None	None
Crohn's Disease Activity Score (CDAI) Mean (range)	301 (230 to 487)	296 (221 to 526)	299 (221 to 526)
N (%) patients in remission N (%) patients with active CD	All patients at inclusion had recurrence of active disease		
CD classification (Vienna/ Montreal)	Not clear	Not clear	Not clear
Disease duration (years) mean	10 (1 to 35)	7 (1 to 27)	9 (1 to 35)
Smoking n (%)	12 (33)	6 (18)	18 (26)
Previous surgery n (%)	10 (28)	10 (30)	20 (29)
Concomitant treatment (specify) n (%) Immunomodulators: Systemic corticosteroids or budesonide	 14 (39) 1 (3)	 13 (39) 1 (3)	 27 (39) 2
Treatment duration at anti-TNF failure (days)	635 (97 to 1913)	681 (126 to 3313)	657 (97 to 3313)
Previous anti-TNF therapy n (%)	6 (17)	8 (24)	14 (20)
CRP (mg/mL)	6 (1 to 28)	9 (3 to 21)	9 (2 to 22)
Calprotectin (µg/g)	NR	NR	NR

Treatment

Item	IFX intensified arm	Algorithm arm
Anti-TNF drug (name)	Infliximab (IFX)	IFX
Anti-TNF dose	IFX at an increased dose frequency of 5 mg/kg every 4 weeks	IFX or other based on the algorithm
Duration of treatment	Not clear, planned 12 wks	Not clear planned 12 wks
Intervention test assay (please specify):		
Manufacturer	RIA (probably Biomonitor A/S, Copenhagen, Denmark) Post hoc paper ELISA and HMSA (?Prometheus Laboratories San Diego, California, USA)	
Assay type	RIA liquid phase assays; assay for antibodies detects those with lambda chains (not kappa)	
Assay name	not specified	
Time of anti-TNF / antibody measurement	'*serum samples for IFX and IFX Ab testing were collected at the time of reported IFX treatment failure. Samples were sent for immediate analysis by radioimmunoassay*'	
Frequency of anti-TNF / antibody measurement	One test time only	
Threshold of infliximab / adalimumab (therapeutic / sub-therapeutic) (in µg/mL)	RIA: therapeutic ≥0.5 µg/L; sub-therapeutic <0.5 µg/L *Post hoc ELISA: 1.4 µg/mL for IFX* *HMSA: therapeutic ≥3 µg/L; sub-therapeutic <3 µg/L*	
Limit of quantification of anti-TNF antibodies (in U/mL [arbitrary unit/mL]) for Ab detectable/ non-detectable	RIA: limit of quantification (LOQ) 10 arbitrary units/mL *Post hoc* *ELISA: 1.69 µg/mL for IFX Abs* *HMSA: LOQ 3.13 U/mL*	
Outcomes reported		
Item		
Primary outcome (s)	A] mean cost of treatment over 12 week B] proportion of patients with "clinical response" at 12 weeks. (*Clinical response was defined as: ">70 point reduction in CDAI score from baseline in luminal disease and a reduction in active fistulas of>50% from baseline in fistulising disease"*)	
Secondary study outcomes	CDAI 100 response; clinical remission; CDAI decrease; PDAI decrease; IBDQ increase; CRP change; WBC change; Hb change; Albumin change.	
Timing of assessments (including info on parallel or sequential)	Weeks 0, 4, 8 and 12	
Time to test result	Not clear; the paper states that '*serum samples for IFX and IFX Ab testing were collected at the time of reported IFX treatment failure. Samples were sent for immediate analysis by RIA.*'	
Number of inconclusive results n (%)	None (note: in group 4 of intervention arm, tests should be repeated to confirm first test result)	
Frequency of dose adjustment n (%)	NR	
Frequency of treatment switch n (%)	NR	
Measure of disease activity (e.g. CDAI, others?)	CDAI. Short Inflammatory Bowel Disease Questionnaire (IBDQ). Perianal Disease Activity Index (PDAI) Number of draining fistulas	

Item	ALGORITHM ARM	INTENSIFICATION ARM	COMPARISON
A] Rates of Response (co primary outcome) *n.b. All patients started with secondary loss of response*	ITT: 19/33 (58%) PP: 9/19 (47%)	ITT: 19/36 (53%) PP: 19/36 (53%)	ITT: RR of 1.09; 95% CI 0.713 to 1.673, p=0.810; difference= 5% (-19% to 28%) PP: RR 0.898; 95% CI 0.510 to 1.580, p=0.781; difference = -5% (-33% to 22%)

B] Rates of CDAI 100 response	ITT: 16/33 (49%) PP 8/19 (42%)	ITT: 17/36 (47%) PP 17/36 (47%)	ITT: RR 1.027; 95% CI 0.627 to 1.681, p=1.0 PP: RR 0.892 95% CI 0.475 to 1.675, p=0.781
C] Clinical Remission	ITT: 10/33 (30%) PP 4/19 (29%)	ITT: 14/36 (39%) PP: 14/36 (39%)	ITT: RR 0.779; 95% CI 0.403 to 1.507, p=0.613 PP: RR 0.541; 95% CI 0.207 to 1.417, p=0.234
Clinical response by subgroups, n (%) **Group 1 (n=14; algo arm: n=5; IFXintes arm: n=9)** [sub-therapeutic IFX + detectable anti-IFX Abs + insufficient IFX bioavailability due to induce immunogenicity of IFX]	ITT: 2/5 (40) PP: 2/5 (40)	ITT: 4/9 (44) PP: 4/9 (44)	ITT: RR 0.90; 95% CI 0.246 to 3.297, p=1.00 PP: RR 0.90; 95% CI 0.246 to 3.297, p=1.00
Group 2 (n=3; algo arm: n=1; IFX intes arm: n=2) [sub-therapeutic IFX + undetectable anti-IFX Abs + insufficient IFX bioavailability due to non-immune mediated pharmacokinetics]	ITT: 0/1 (0) PP: 0/1 (0)	ITT: 1/2 (50) PP: 1/2 (50)	ITT: Not calculable PP: Not calculable
Group 3 (n=48; algo arm: n=26; IFX intes arm: n=22) [therapeutic IFX + undetectable anti-IFX Abs + inhibition of TNF-alpha ineffective due to non-TNF drive disease]	ITT: 16/26 (62) PP: 7/13 (54)	ITT: 12/22 (55) PP: 12/22 (55)	ITT: RR 1.128; 95% CI 0.693 to 1.837, p=0.770 PP: RR 0.987; 95% CI 0.525 to 1.856, p=1.00
Group 4 in algorithm (n=4; algo arm: n=1; IFX intes arm: n=3) [therapeutic IFX + detectable anti-IFX Abs + pharmacodynamics or non-functional anti-IFX Abs or FP test]	ITT: 0/1 (0) PP: 0/0	ITT: 2/3 (67) PP: 2/3 (67)	ITT: Not calculable PP: Not calculable
Describe definition of progression: Patients who withdrew because of lack of effect of study treatment were classified as having no response and no remission at subsequent study visits			
Describe definition of remission: An absolute CDAI score of ≤ 150 and complete closure of all fistulas despite gentle pressure Definition of clinical response: ≥ 70 point reduction in CDAI from baseline in luminal disease and a reduction in active fistulas of ≥ 50% from baseline in fistulising disease			
Duration of a) Response b) Relapse c) Remission	NR	NR	NR
Rates of hospitalisation n (%)	NR	NR	NR

Rates of surgical intervention n (%)	NR	NR	NR
Time to surgical intervention y/n	NR	NR	NR
Health related quality of life y/n	Yes	Yes	Yes
Length of follow up reported y/n	Yes; 12 weeks	Yes; 12 weeks	Yes; 12 weeks
Proportion progressing to surgery n (%)	NR	NR	NR
Time to surgical intervention	NR	NR	NR
Incidence of adverse effects of treatment			
Item	Algorithm arm	IFX intensified arm	P value
	NR	NR	NR
Dose changes			
Item	Algorithm arm	IFX intensified arm	P value
Number of patients outside therapeutic range (sub-therapeutic infliximab)	Group 1: 5 Group 2: 1 Group 3: 0 Group 4: 0	Group 1: 9 Group 2: 2 Group 3: 0 Group 4: 0	Group 1: 14 Group 2: 3 Group 3: 0 Group 4: 0
Mean anti-TNF (mg/m^2/wk) (SD)	NR		
Number of patients dose increased	Unclear treatments for group3 of algorithm arm; all patient were increased in the dose intensification arm		
Number of patients dose reduced	Unclear, group 3 of algorithm arm should have stopped infliximab but many did not		
Health related quality of life			
Item	Algorithm arm Mean *SE*	IFX intensified arm Mean *SE*	MEAN DIFFERENCE
PDAI score decrease from baseline	ITT: 2.4 *0.8* PP: 1.4 *0.5*	ITT: 1.5 *0.7* PP: 1.5 *0.7*	0.9 (−1.4 to 3.2) P 0.421 −0.1(−2.1 to 1.9) P 0.911
IBDQ score increase from baseline	ITT: 8.8 *1.7* PP: 5.4 *2.0*	ITT: 8.8 *1.9* PP: 8.8 *1.9*	0 (−5.1 to 5.2) P 0.996 −3.4 (−9.6 to 2.7) P 0.264
Author's conclusion			
Treatment of secondary IFX failure using an algorithm based on combined IFX and IFX antibody measurements significantly reduces average treatment costs per patient compared with routine IFX dose escalation and without any apparent negative effect on clinical efficacy.			
Reviewer's conclusion			
The primary outcome measure of the trial concerns costs rather than a clinical outcome. However, results on clinical response rate, defined as patients with '*≥ 70 point reduction in CDAI from baseline in luminal disease and a reduction in active fistulas of≥ 50% from baseline in fistulising disease '* were reported. The trial included patients with secondary loss to response with IFX and they were randomised into two groups i.e., IFX intensification arm where IFX treatment was intensified and an algorithm arm where patients would receive interventions based on the serum IFX and IFX Ab levels using the proposed algorithm. In terms of clinical response rate, the study found no significant difference between the two groups. The clinical response rate was numerically found to very slightly favour algorithm arm using the ITT population (58% vs. 53%) whereas, the IFX intensified arm was found to be very slightly numerically superior using the PP population (53% vs. 47%), in both cases the difference was not statistically significant. The study was underpowered to detect clinical differences between groups.			

Name of the first reviewer: Martin Connock Name of second reviewer: Paul Sutcliffe

Study details			
Study ID (Endnote ref)	[72]		
First author surname	Vande Casteele		
Year of publication	2015		
Country	Belgium		
Study design	Randomised controlled trial		
Publication (full/abstract)	Full		
Study setting	Tertiary referral centre		
Number of centres (by arm)	One		
Duration of study	52 weeks from randomisation		
Follow up period	As above		
Funding	Belgian Research Foundation		
Aim of the study			
To determine whether dosing based on therapeutic drug monitoring increases rate of remission and whether continued concentration-based dosing is superior to clinically based dosing of infliximab for maintaining remission in patients with CD and UC			
Inclusion/exclusion criteria for patients			
Inclusion criteria	Moderate-to-severe CD or UC confirmed by endoscopy and histology; at least 18 years of age; on infliximab at least 14 weeks; clinically stable		
Exclusion criteria	Non-standard higher dosing regimen for secondary loss of response to infliximab therapy at time of screening; ATI >8 µg/mL equivalents		
Study flow(consort diagram)			
Available in paper			
Item	Clinical dosing arm	Concentration dosing arm	All
N of screened	Optimisation preceded randomisation		275
N of excluded (ineligible)	Unclear	Unclear	24
N included for optimisation N randomised	123	128	263 251
N of non-participants at study entry (those refused, etc…)	6 of 263 withdrew consent during "optimisation" before randomisation; 6 further either developed loss of response or could not be optimised		
N study sample at baseline randomised	123	128	251
Discontinued post randomisation	12	13	25
Lost to follow up post randomisation	2	2	4
Participants (characteristics and numbers) randomised phase			
Total number of participants at baseline (% CD)	123 (66.7% CD)	128 (71.1% CD)	251
N (%) followed up	121 (100)	126 (100)	247
N (%) included in analysis primary outcome (remission at week 52)	123 (100)	128 (100)	251 (ITT)
Patient group (responders/ secondary loss of response)	Responders	Responders	Responders
Age median IQR years	42.0 (32.0- 48.0)	41.0 (30.0-50.3)	41.0 (30.5-49.0)
Sex Women n (%)	51 (41.5)	62 (48.4)	113 (45.0)
Diagnostic criteria for CD	IBD confirmed by endoscopy and histology		
Children n (%)	None	None	None
Crohn's Disease Activity; CRP mg/L	1.3 (0.6 – 4.5)	1.5 (0.7 – 4.0)	1.4 (0.6 – 4.2)
N (%) patients in remission (for CD HBI ≤4	At randomisation (after dose optimisation): *IBD 101 (82.1) and 106 (82.8)*, for CD 63/82 (76.8) and 75/91 (82.4**)**, in clinical and concentration arms, respectively		
CD classification (Vienna/ Montreal)	Unclear	Unclear	Unclear
Disease duration (years) median IQR	12.5 (7.1-19.3)	12.0 (5.6-20.8)	12.5 (6.3 to 19.9)
Smoking n (%)	38 (30.9)	26 (20.3)	64 (25.5)
Previous surgery n (%)	70/178 CD (39.3)		All 76/263 (28.9)
Concomitant treatment (specify) n (%) Immunomodulators: Systemic corticosteroids or budesonide	7 (5.7) NR	6 (4.7) NR	13 (5.2) NR

Treatment duration at anti-TNF failure (days)	NA	NA	NA
Previous anti-TNF therapy n (%)	NR	NR	NR
CRP (mg/mL) median (IQR)	1.3 (0.6 – 4.5)	1.5 (0.7 – 4.0)	1.4 (0.6 – 4.2)
Calprotectin (µg/g)	NR	NR	NR
Treatment			
Item	Clinical dosing arm	Concentration dosing arm	
Anti-TNF drug (name)	Infliximab (IFX)	IFX	
Anti-TNF dose	Various based on clinical decisions (CRP & symptoms)	Various based on trough IFX testing	
Duration of treatment	Patients entered on infliximab		
Intervention test assay (please specify): ELISA Technical aspect of test assay:			
Manufacturer	Non-commercial ELISA " in house" (Leuven University)		
Time of anti-TNF, antibody measurement	Repeated trough IFX testing during dose optimisation phase. After randomisation testing was done before each infusion in the concentration dosing arm		
Assay type	ELISA		
Assay name	"In house" Leuven		
Type of ELISA (bridging/ capture)	Capture ELISA to measure IFX concentrations Bridging ELISA to measure IFX Abs		
Anti-TNF alpha detection: Limit of detection	In house ELISA; reference provided to previous study. Lower limit of detection 0.3 ug/mL infliximab		
Anti-body detection: Limit of detection	In house ELISA; reference provided to previous study. Lower limit of detection 1.0 ug/mL infliximab		
Outcomes reported			
Item			
Primary outcome (s)	Proportion with clinical (HB≤ 4 for CD, Mayo ≤ 2 UC) and biological (CRP ≤ 5m/L) at 52 weeks post randomisation		
Secondary study outcomes	Durable remission (as primary but throughout 52 weeks); Relapse (need for dose escalation or addition of steroids or switch treatment), EuroQol-5D, costs of treatment		
Timing of assessments (including info on parallel or sequential)	Probably at each infusion, or every 8 weeks		
Time to test result	NR		
Number of inconclusive results n (%)	Author's used defined cut offs, inconclusive results = 0%		
Frequency of dose adjustment n (%) Frequency of treatment switch n (%)	Unclear; both groups received dose adjustments, if required, during optimisation phase so as to bring drug trough levels with target range (3 to 7 ug/mL). 115 no adjustment, 76 dose escalated, 72 dose reduction (CD plus UC). Post randomisation dose adjustments unclear		
Measure of disease activity (e.g. CDAI, others?)	For CD: HBI; CRP level; relapse see below		
Rates of remission (clinical) Optimisation phase	Before 131/178 CD (73.6%) After 138/173 CD (79.8%) by ITT 138/178 (77.5%)		
After randomisation at wk 52: Clinical and biological remission	Clinical-based CD 54.9%; trough-based CD 62.6%; P 0.353 (at start of randomisation NR)		
Durable remission (clinical and biological through 52 wks)	CD+UC: clinically based 27%; concentration-based 26%; P = 0.880		
Relapse (need for dose escalation or addition of steroids or switch in treatment)	Clinically based 21 (17%); concentration-based 9 (7%); Relative risk 2.4; 95% CI: 1.2-5.1; P=0.018. Time to relapse log rank test, P = 0.017.		
Describe definition of response	Clinical response = being 'symptom-free (full responder) or having clinical improvement with an obvious decrease of disease activity but with clinical symptoms still present (partial responder)'		
Describe definition of progression: Relapse defined as the need for dose escalation or addition of steroids or switch treatment			
Describe definition of remission: Clinical remission for CD patients = Harvey-Bradshaw index score of ≤ 4 corresponds to remission. Biological remission = CRP concentration of ≤5 mg/L			
Rates of hospitalisation n (%)	2/263 hospitalized. 1 for appendectomy, 1 for ileostomy complications;		

	both patients were in the clinically based dosing group			
Rates of surgical intervention n (%)	NR / unclear			
Time to surgical intervention y/n	NR			
Health related quality of life y/n	Yes			
Length of follow up reported y/n	Yes; 52 weeks			
Proportion progressing to surgery n (%)	NR / unclear			
Time to surgical intervention	Unclear			
Incidence of adverse effects of treatment (post randomisation phase)				
Item	Clinically based dosing N=123		Concentration-based dosing N=128	
Adverse event	number	*%*	number	*%*
Pharyngitis	20	*16.3*	25	*19.5*
Upper respiratory tract infection	55	*44.7*	59	*46.1*
Pneumonia	3	*2.4*	6	*4.7*
Aphthous stomatitis	1	*0.8*	3	*2.3*
Headache	4	*3.3*	3	*2.3*
Arthralgia	37	*30.1*	33	*25.8*
Infusion reaction	6	*9.4*	3	*2.3*
Acute reaction	6	*9.4*	1	*0.8*
Delayed hypersensitivity	0	*0*	2	*1.6*
Serious adverse event	0	*0*	1	*0.8*
Dose monitoring				
Item (Please define if necessary)				
Time of anti-TNF / antibody measurement	See above			
Frequency of anti-TNF / antibody measurement	See above			
Assay type	See above			
Assay name	See above			
Threshold of infliximab / adalimumab (therapeutic / sub-therapeutic) (in µg/mL)	Trough level defined groups at start of optimisation: 1] IFX < 0.3 ug/mL ADAb < 8ug/mL; 2] IFX <3 ug/mL; 3]IFX 3 to 7ug/mL; 4] IFX >7 ug/mL			
Limit of quantification of anti-TNF antibodies	1.0 ug/mL (see above)			
Algorithm specified for management y/n (specify)	Yes			
Algorithm provided	Yes			
Number of patients outside therapeutic range	Of 263 entering optimisation phase, 12 were not optimised (withdrew, lost response or failed to get to target range)			
Mean anti-TNF (mg/m^2/wk) (SD)	NR		NR	
	Optimisation phase CD for supp; table 1			
Number of patients dose increased	IBD 76 (28.9%). CD 44/178			
Number of patients dose reduced	IBD 72 (27.4% CD 52/178			
Number of patients no change	IBD 115 (43.7%) CD 82/178			
	During randomised phase			
Number of patients dose increased	Unclear			
Number of patients dose reduced	Unclear			
Number of patients no change	Unclear			
Health related quality of life				
Item	Concentration based		Clinically based	
EQ-5D completed	Unclear		Unclear	
Author's conclusion				
Targeting patients' infliximab TCs to 3–7 µg/mL results in a more efficient use of the drug. After dose optimization, continued concentration-based dosing was not superior to clinically.				
Reviewer's conclusion				
Small gains in reduced drug costs with dose optimisation, unclear if cost of testing will offset these; no clinical				

benefit demonstrated for testing strategy other than more relapse (probably requiring dose escalation) occurred in the clinically based dosing group.

Name of the first reviewer: Deepson S Shyangdan Name of second reviewer: Martin Connock

Study details		
Study ID (Endnote ref)	[127]	
First author surname	Vaughn	
Year of publication	2014	
Country	USA	
Study design	Retrospective observational study (pilot study) with treatment algorithm	
Publication (full/abstract)	Full	
Study setting	Beth-Israel Deaconess Medical Center (Boston, MA)	
Number of centres (by arm)	One	
Duration of study	Probably start of 2009 to August 2013	
Follow up period	Variable according to analysis subgroups	
Funding	Unclear / NIHR training grant	
Aim of the study		
To describe the outcomes of proactive TROUGH CONCENTRATION MONITORING (TCM) of IFX treated patients in clinical remission on IFX using dose adjustment based on testing to bring infliximab into target range. Outcomes include: initial and subsequent IFX trough levels, dosing changes including dose escalation and de-escalation, and outcomes of patients on IFX monotherapy. The secondary aims were to assess if proactive TCM was associated with a longer duration of IFX compared with a control group (i.e. that did not receive pro-active TCM) and to assess reasons for cessation of IFX		
Inclusion/exclusion criteria for patients		
Inclusion criteria	Patients receiving IFX for IBD at the Beth-Israel Deaconess Medical Center (Boston, MA). For a patient to be considered as having had proactive TCM of IFX, the patient must have had an IFX trough concentration while in clinical remission and testing not done for a reactive purpose (i.e. for symptoms concerning for IBD or concern for and IFX-mediated side effect)	
Exclusion criteria	Patients were excluded if 1) the IFX infusions were not administered at the hospital's infusion center; 2) the IFX concentration was drawn from cord blood; 3) there was no follow-up visit after the IFX concentration was drawn; 4) the IFX concentration was not documented in a gastroenterology clinic note; or 5) patient failed to receive at least one maintenance infusion of IFX	
Study flow(consort diagram)		
Available in paper		
Item	Proactive TCM group	Control group
N of screened	88 identified from Prometheus laboratory data	84 identified from infusion centre
N of excluded (ineligible)	14 did not meet inclusion criteria; 22 did not reach clinical remission; 4 patients did not have level when in remission or level was not a trough	10 did not reach clinical remission
N of enrolled/included (eligible)	48 included as 'proactive TCM of IFX'	74 + 4 from Prometheus record=78
N of non-participants at study entry (those refused, etc…)	NA	NA
N study sample at baseline randomised (if applicable)	NA	NA
Withdrawals	NA	NA
Lost to follow up/drop outs (sample attrition)	NA	NA
Participants (characteristics and numbers)		
item	TCM group	Control (non TCM) group
Total number of participants at baseline (% CD)	48 (38/48; 79% CD)	78 (45/78; 67% CD)
N (%) followed up	48	78
N (%) included in analysis	48	78

Patient group (responders/ secondary loss of response)	Responders (in remission)	Responders (in remission)
Age, Median (range) years	35 (29 to 42.5) at start of infliximab therapy	34.9 (26.2 to 49.7) at start of infliximab therapy
Sex Women n (%)	15 (31)	33 (42)
Diagnostic criteria for CD	NR	NR
Children n (%)	None	None
Crohn's Disease Activity Score (CDAI) Mean (SD)	NR	NR
N (%) patients in remission N (%) patients with active CD	All patients in remission	All patients in remission
CD classification (Vienna/ Montreal)	NR	NR
Disease duration (years)	Not clear	Not clear
Smoking n (%) – tobacco status	Current: 5 (10) Former: 12 (25) Never: 31 (56)	Current: 7 (9) Former: 14 (18) Never: 57 (73)
Previous surgery n (%)	19 (40)	19 (25)
Concomitant treatment ("combination therapy") n (%)	21 (44)	31 (40)
Treatment duration at anti-TNF failure (weeks)	Not clear	Not clear
Line of therapy	NR	NR
Previous anti-TNF therapy n (%)	Infliximab (100%)	Infliximab (100%)
CRP (mg/mL)	NR	
Calprotectin (µg/g)	NR	
Treatment		
Item		
Anti-TNF drug (name)	Infliximab	
Anti-TNF dose	Various	
Duration of treatment	Various Time to treatment cessation = primary outcome	
Intervention test assay (please specify): ELISA and HMSA – the authors report that 'the period of the study overlapped with the use of 2 methods of IFX and ATI detection. Initially, testing was performed through solid phase ELISA and the testing was changed to a non-radiolabeled liquid phase mobility shift assay. Technical aspect of test assay:		
Manufacturer	Prometheus Laboratories (San Diego, CA) (they performed the assays)	
Time of anti-TNF, antibody measurement	Various	
Assay type / name	ELISA or HMSA	
Type of ELISA (bridging/ capture)	NR	
Anti-TNF alpha detection: ELISA HMSA details	NR	
Anti-body detection: ELISA HMSA details	NR	
Outcomes reported		
Item		
Primary outcome (s)	Primary outcome applied for the proactive TCM group only: Initial and subsequent IFX trough levels, dosing changes including dose escalation and de-escalation, and outcomes of patients on IFX monotherapy	
Secondary study outcomes	Time to infliximab treatment cessation TCM versus control group.	
Timing of assessments (including info on parallel or sequential)	Unclear	
Time to test result	NR	
Number of inconclusive results n (%)	NR	
Frequency of dose adjustment n	TCM group: first trough test: dose escalated 12/48; dose decreased 3/48;	

(%)	dose stopped 2/48; dose unchanged 31/48. Subsequent trough tests: dose escalated 8/40; dose decreased 2/40; dose unchanged 30/40	
Frequency of treatment switch n (%)	NA	
Measure of disease activity (e.g. CDAI, others?)	Physicians' judgement of remission assessed on medical notes; Cessation of treatment.	
Rates of a) Response y/n b) Relapse y/n c) Remission y/n	Time to event analysis of time to infliximab treatment cessation Others: No	
Describe definition of progression: Equivalent to infliximab treatment cessation.		
Describe definition of remission: physicians' judgement based on medical notes (Clinical remission defined as *'lack of symptoms attributable to underlying IBD based on the treating gastoenterologist's documentation'*).		
Duration of d) Response e) Relapse f) Remission	Time to event analysis of time to infliximab treatment cessation	
Rates of hospitalisation n (%)	NR	
Rates of surgical intervention n (%)	NR	
Health related quality of life y/n	No	
Length of follow up reported y/n	Various; follow up to infliximab cessation in Kaplan Meier analysis	
Incidence of adverse effects of treatment (reasons for stopping infliximab)		
Item	TCM group	No TCM group
Recurrent IBD symptoms	0	15
Adverse events Pneumonia Drug-induced lupus Psoriasis	0 1 1	1 0 0
High antibody concentration	1	0
Infusion reactions Acute Delayed	 0 1	 6 0
Other (unrelated to infliximab)	1	2
Dose monitoring		
Item (Please define if necessary)		
Time of anti-TNF / antibody measurement	Various	
Frequency of anti-TNF / antibody measurement	Unclear	
Threshold of infliximab / adalimumab (therapeutic / sub-therapeutic) (in µg/mL)	Initially undetectable infliximab was defined as sub-therapeutic, later the target range of 5 to 10 ug/mL was used as the therapeutic range, dose adjustments (in the TCM group) were made to bring patients into this range.	
Limit of quantification of anti-TNF antibodies (in U/mL [arbitrary unit/mL]) for Ab detectable/ non-detectable	HMSA for infliximab IFX 1 ug/mL ; ELISA 1.4 ug/mL	
Algorithm specified for management y/n (specify)	Yes Algorithm resulted in dose increases of typically 50 to 100 mg (for a 70 kg patient receiving 5mg/kg (total 350 mg) this represents an increase of between 14% and 28%	
Algorithm provided	Yes, provided in narrative description, but ill defined	
Number of patients outside therapeutic range	In TCM arm at first trough test 35% needed dose adjustment	
Mean anti-TNF ($mg/m^2/wk$)	Unclear	

(SD)	
Number of patients dose increased	See above
Number of patients dose reduced	See above
Health related quality of life	
Item	
	NR
Author's conclusion	
Proactive TCM of IFX frequently identified patients with low or undetectable trough concentrations and resulted in a greater probability of remaining on IFX.	
Reviewer's conclusion	
The distinction between pro-active and non-pro-active groups is that in the latter testing was done reactively for symptom worsening, this implies that identification of this group will tend to select ill patients (select patients with worsening symptoms); whereas in the pro-active group tests were not done in response to symptoms and therefore those identified are probably less likely to be ill patients than in the control group. Patients that did not reach remission were excluded, this resulted in 22 exclusions from 88 in the TCM group, but only 10 from 84 in the control group. The part of the study comparing TCM versus no TCM provided time to event outcomes for retention in infliximab treatment; In the main other outcomes referred to the TCM only. For time to event data extraction using the method of Guyot please refer to appropriate data extraction files.	

8.6 Appendix 6 Excluded studies with reason

Full text exclusions with reason

Reference	Reason for exclusion
1. Afif, W., E. V. Loftus, Jr., W. A. Faubion, S. V. Kane, D. H. Bruining, K. A. Hanson and W. J. Sandborn (2010). "Clinical utility of measuring infliximab and human anti-chimeric antibody concentrations in patients with inflammatory bowel disease." American Journal of Gastroenterology 105(5): 1133-1139.	C insufficient data M no algorithm specified / acted on
2. Baert, F., M. Noman, S. Vermeire, G. Van Assche, D. H. G, A. Carbonez and P. Rutgeerts (2003). "Influence of immunogenicity on the long-term efficacy of infliximab in Crohn's disease." N Engl J Med 348(7): 601-608.	C insufficient data
3. Balzola, F., C. Bernstein, G. T. Ho and C. Lees (2010). "Clinical utility of measuring infliximab and human antichimeric antibody concentrations in patients with inflammatory bowel disease: Commentary." Inflammatory Bowel Disease Monitor 11(2): 85-86.	Commentary no original data
4. Balzola, F., G. Cullen, G. T. Ho and R. K. Russell (2013). "Clinical utility of newly developed immunoassays for serum concentrations of adalimumab and anti-adalimumab antibodies in patients with Crohn's disease." Inflammatory Bowel Disease Monitor 14(1): 19.	Commentary no original data
5. Ben-Horin, S. and Y. Chowers (2011). "Review article: loss of response to anti-TNF treatments in Crohn's disease." Aliment Pharmacol Ther 33(9): 987-995.	Review without MA
6. Billioud, V., W. J. Sandborn and L. Peyrin-Biroulet (2011). "Loss of response and need for adalimumab dose intensification in Crohn's disease: a systematic review." American Journal of Gastroenterology 106(4): 674-684.	SR without MA
7. Cassinotti A, Travis S. Incidence and clinicalsignificance of immunogenicity to infliximab inCrohn's disease: a critical systematic review. Inflamm Bowel Dis. 2009;15(8):1264-75.	Review without MA
8. Chaparro, M., I. Guerra, P. Munoz-Linares and J. P. Gisbert (2012). "Systematic review: antibodies and anti-TNF-alpha levels in inflammatory bowel disease." Aliment Pharmacol Ther 35(9): 971-986.	SR without MA
9. Colombel JF, Feagan BG, Sandborn WJ, Van Assche G, Robinson AM. Therapeutic drugmonitoring of biologics for inflammatory bowel disease. 2012;18(2):349-58.	Review without MA
10. Ebert, E. C., K. M. Das, V. Mehta and C. Rezac (2008). "Non-response to infliximab may be due to innate neutralizing anti-tumour necrosis factor-alpha antibodies." Clinical & Experimental Immunology 154(3): 325-331.	Measurement of antibodies to TNF-alpha not anti-TNFα drugs
11. Garces, S., J. Demengeot and E. Benito-Garcia (2013). "The immunogenicity of anti-TNF therapy in immune-mediated inflammatory diseases: a systematic review of the literature with a meta-analysis." Annals of the Rheumatic Diseases 72(12): 1947-1955.	>50% RA patients
12. Hamalainen, A., T. Sipponen and K. L. Kolho (2013). "Serum infliximab concentrations in pediatric inflammatory bowel disease." Scandinavian Journal of Gastroenterology 48(1): 35-41.	C insufficient data
13. Hibi, T., A. Sakuraba, M. Watanabe, S. Motoya, H. Ito, K. Motegi, Y. Kinouchi, M. Takazoe, Y. Suzuki, T. Matsumoto, K. Kawakami, T. Matsumoto, I. Hirata, S. Tanaka, T. Ashida and T. Matsui (2012). "Retrieval of serum infliximab level by shortening the maintenance infusion interval is correlated with clinical efficacy in Crohn's disease." Inflamm Bowel Dis 18(8): 1480-1487.	C insufficient data
14. Khanna, R., B. D. Sattin, W. Afif, E. I. Benchimol, E. J. Bernard, A. Bitton, B. Bressler, R. N. Fedorak, S. Ghosh, G. R. Greenberg, J. K. Marshall, R.	SR without MA

	Panaccione, E. G. Seidman, M. S. Silverberg, A. H. Steinhart, R. Sy, G. Van Assche, T. D. Walters, W. J. Sandborn and B. G. Feagan (2013). "Review article: a clinician's guide for therapeutic drug monitoring of infliximab in inflammatory bowel disease." Aliment Pharmacol Ther 38(5): 447-459.	
15.	Lazebnik, L. B. and V. E. Sagynbaeva (2013). "[Level of adalimumab and its antibody titers define the effectiveness of the biological (anticytokine) therapy in Crohn's disease]." Eksperimental'Naia i Klinicheskaia Gastroenterologiia(7): 18-22.	Non-English
16.	Lichtenstein, G. R. (2013). "Comprehensive review: antitumor necrosis factor agents in inflammatory bowel disease and factors implicated in treatment response." Therapeutic Advances in Gastroenterology 6(4): 269-293.	SR without MA
17.	Malickova, K., D. Duricova, M. Bortlik, N. Machkova, I. Janatkova and M. Lukas (2011). "Serum infliximab trough levels and induction of antibodies to infliximab during the biological treatment of patients with inflammatory bowel diseases. [Czech]Serove hladiny infliximabu a indukce tvorby protilatek proti infliximabu pri biologicke lecbe nemocnych s idiopatickymi strevnimi zanety." Alergie 13(3): 216-222.	Non-English
18.	Rivero Marcotegui, A., R. Ibanez Bosch, A. Zuniga Vera, A. Arin Letamendia and M. J. Burusco Paternain (2014). "Clinical usefulness in measuring infliximab and human anti-chimeric antibodies. [Spanish]Utilidad clinica de la cuantificacion de infliximab y anticuerpos antiquimericos humanos." Revista del Laboratorio Clinico 7(2): 68-72.	patients >50% RA
19.	Roblin, X., M. Rinaudo, E. Del Tedesco, J. M. Phelip, C. Genin, L. Peyrin-Biroulet and S. Paul (2014). "Development of an algorithm incorporating pharmacokinetics of adalimumab in inflammatory bowel diseases." American Journal of Gastroenterology 109(8): 1250-1256.	C insufficient data M no algorithm specified / acted on
20.	Rutgeerts, P., G. D'Haens, S. Targan, E. Vasiliauskas, S. B. Hanauer, D. H. Present, L. Mayer, R. A. Van Hogezand, T. Braakman, K. L. DeWoody, T. F. Schaible and S. J. Van Deventer (1999). "Efficacy and safety of retreatment with anti-tumor necrosis factor antibody (infliximab) to maintain remission in Crohn's disease." Gastroenterology 117(4): 761-769.	C insufficient data
21.	Sono, K., A. Yamada, Y. Yoshimatsu, N. Takada and Y. Suzuki (2012). "Factors associated with the loss of response to infliximab in patients with Crohn's disease." Cytokine 59(2): 410-416.	C insufficient data
22.	Steenholdt, C., M. Svenson, K. Bendtzen, O. O. Thomsen, J. Brynskov and M. A. Ainsworth (2011). "Severe infusion reactions to infliximab: aetiology, immunogenicity and risk factors in patients with inflammatory bowel disease." Aliment Pharmacol Ther 34(1): 51-58.	C insufficient data
23.	Ungar, B., Y. Chowers, M. Yavzori, O. Picard, E. Fudim, O. Har-Noy, U. Kopylov, R. Eliakim, S. Ben-Horin and A. consortium (2014). "The temporal evolution of antidrug antibodies in patients with inflammatory bowel disease treated with infliximab." Gut 63(8): 1258-1264.	C insufficient data
24.	Van Assche, G., C. Magdelaine-Beuzelin, G. D'Haens, F. Baert, M. Noman, S. Vermeire, D. Ternant, H. Watier, G. Paintaud and P. Rutgeerts (2008). "Withdrawal of immunosuppression in Crohn's disease treated with scheduled infliximab maintenance: a randomized trial." Gastroenterology 134(7): 1861-1868.	C insufficient data
25.	Vermeire, S., M. Noman, G. Van Assche, F. Baert, G. D'Haens and P. Rutgeerts (2007). "Effectiveness of concomitant immunosuppressive therapy in suppressing the formation of antibodies to infliximab in Crohn's disease." Gut 56(9): 1226-1231.	C insufficient data
26.	Yamada, A., K. Sono, N. Hosoe, N. Takada and Y. Suzuki (2010). "Monitoring functional serum antitumor necrosis factor antibody level in Crohn's disease patients who maintained and those who lost response to anti-TNF." Inflamm Bowel Dis 16(11): 1898-1904.	C insufficient data
27.	Yanai H, Hanauer SB. Assessing response and loss of response to biological therapies in IBD. Am JGastroenterol. 2011;106(4):685-98	Review without

	MA

Excluded abstracts with reason

Reference	Reason for exclusion
28. Abraham, B. and M. Chiorean (2012). "False positive infliximab levels detected in patients treated with adalimumab for inflammatory bowel disease." American Journal of Gastroenterology 107: S627	C insufficient data
29. Afif, W., E. V. Loftus, W. A. Faubion, K. A. Hanson and W. J. Sandborn (2009). "Clinical utility of measuring infliximab and human anti-chimeric antibody levels in patients with inflammatory bowel disease." Gastroenterology 1): A147.	Superseded by full text
30. Anonymous (2012). "New Assay Can Detect Infliximab Levels and Anti-Infliximab Antibodies From a Single Serum Sample." Clinical Advances in Hematology and Oncology 10 (10): 27.	Editorial no original data
31. Armbruster, S., M. Ally, C. Maydonovitch, J. Betteridge and G. Veerappan (2012). "The use of human anti-chimeric antibody (HACA) and infliximab levels in the management of inflammatory bowel disease." American Journal of Gastroenterology 107: S641.	M no algorithm specified / acted on
32. Arranz, M. D. M., E. M. Arranz, D. P. Salcedo, C. De Diego, S. G. Senent, J. P. Cordon, B. B. Garcia and J. M. S. Parga (2014). "Infliximab trough levels and antibodies: Relationship with infusion reaction, immunomodulators and biological parameters." Gastroenterology 1): S-243.	C insufficient data
33. Baert, F. J., D. Drobne, V. Ballet, I. Cleynen, G. Compernolle, P. J. Rutgeerts, G. A. Van Assche, A. Gils and S. Vermeire (2011). "Early trough levels and antibodies predict safety and success of restarting infliximab after long drug holiday." Gastroenterology 1): S62.	C insufficient data
34. Baert, F. J., S. Lockton, S. Hauenstein, S. Singh, A. Gils and S. Vermeire (2014). "Antibodies to adalimumab predict inflammation in crohn's patients on maintenance adalimumab therapy." Gastroenterology 1): S-242	C insufficient data
35. Ben-Bassat, O., S. Hauenstein, A. Iacono, S. P. Irwin, S. Singh and G. R. Greenberg (2013). "Serum adalimumab and immunogenicity in IBD patients after 80mg biweekly maintenance therapy." Gastroenterology 1): S771.	C insufficient data
36. Ben-Horin, S., B. Ungar, Y. Chowers, M. Yavzori, O. Picard, E. Fudim and R. Eliakim (2013). "The temporal evolution of anti-drug antibodies in IBD patients treated with infliximab." Journal of Gastroenterology and Hepatology 28: 145.	C insufficient data
37. Bodini, G., V. Savarino, P. Dulbecco, I. Baldissarro and E. Savarino (2014). "TNF-alpha levels strongly correlated with disease activity based on HBI and CDEIS in patients with crohn's disease in maintenance treatment with adalimumab." Gastroenterology 1): S-238.	C insufficient data
38. Bodini, G., V. Savarino, P. Dulbecco, I. Baldissarro and E. Savarino (2014). "The influence of anti-adalimumab antibodies on adalimumab trough levels, TNF-alpha levels and clinical outcome." Journal of Crohn's and Colitis 8: S42.	C insufficient data
39. Bodini, G., V. Savarino, P. Dulbecco, I. Baldissarro and E. V. Savarino (2014). "Elisa vs. HMSA: A comparison between two different methods for measuring adalimumab serum concentration and anti-adalimumab antibodies-preliminary data." Digestive and Liver Disease 46: S67.	Duplicate
40. Bodini, G., V. Savarino, P. Dulbecco, L. Assandri, L. Bruzzone, F. Mazza, V. Fazio, E. Giambruno, L. Gemignani and E. Savarino (2013). "Correlation between adalimumab trough serum concentration, anti-adalimumab antibodies and TNF-alpha levels with clinical outcome in patients affected by crohn's disease." Gastroenterology 1): S780.	C insufficient data

41. Bodini, G., V. Savarino, V. Fazio, L. Assandri, L. Gemignani, P. Dulbecco, E. Giambruno and E. Savarino (2012). "Relationship between drug serum concentration and clinical activity in patients with Crohn's Disease who achieved remission with adalimumab." Digestive and Liver Disease 44: S69-S70.	Duplicate
42. Bodini, G., V. Savarino, V. Fazio, L. Assandri, P. Dulbecco, L. Gemignani and E. Savarino (2012). "Relationship between drug serum concentration and clinical activity in patients with crohn disease who achieved remission with adalimumab-a prospective study." Gastroenterology 1): S388.	C insufficient data
43. Bortlik, M., D. Duricova, K. Malickova, A. Komarek, N. Machkova, E. Bouzkova, L. Hrdlicka and M. Lukas (2012). "Infliximab trough levels may predict sustained response to infliximab in patients with Crohn's disease: A single cohort study." Journal of Crohn's and Colitis 6: S153.	Superseded by full text
44. Cardile, S., A. Costa, I. Loddo, G. Morabito, C. Pidone and C. Romano (2013). "Impact of measurement of infliximab and anti-infliximab antibodies levels in pediatric inflammatory bowel disease." Digestive and Liver Disease 45: e294-e295.	C insufficient data
45. Chauhan, U., U. Dutta, D. Armstrong, E. Greenwald, J. Marshall, F. Tse, T. Xenodemetropoulos and H. Smita (2012). "Does measuring infliximab and human anti-chimeric antibody concentrations in patients with inflammatory bowel disease impact clinical management? A canadian experience." Inflamm Bowel Dis 18: S82-S83	M no algorithm specified / acted on
46. Chauhan, U., U. Dutta, D. Armstrong, J. Marshall, F. Tse, E. Greenwald, T. Xenodemetropoulos and S. Halder (2013). "Does measuring IFX and human anti-chimeric antibody concentrations in patients with inflammatory bowel disease impact clinical management? A Canadian experience." Journal of Crohn's and Colitis 7: S228.	Duplicate
47. Chollet-Martin, S., P. Nicaise-Roland, L. De Chaisemartin, S. Grootenboer-Mignot, G. Hayem, A. L. Pelletier, A. Amiot, V. Descamps, Y. Bouhnik and O. Meyer (2013). "Simultaneous determination of anti-infliximab antibodies and residual infliximab levels to monitor anti-TNF therapy." Annals of the Rheumatic Disease 71.	Not M, C or ATC
48. Church, P., J. Guan, K. Frost, A. Muise, T. Walters and A. Griffiths (2013). "Infliximab treatment for paediatric Crohn's disease: Long-term ouCTomes at a single centre." Journal of Crohn's and Colitis 7: S198.	Not M, C or ATC
49. Church, P., J. Guan, L. Salz, K. Frost, A. Muise, T. Walters and A. Griffiths (2012). "Long-term outcomes with infliximab treatment in children with crohn's disease at a single centre." Inflamm Bowel Dis 18: S72-S7	C insufficient data
50. Cornillie, F., S. Hanauer, R. Diamond, J. Wang, D. Zelinger, Z. Xu, S. Vermeire and P. Rutgeerts (2011). "Early serum infliximab trough level, Clinical disease activity and crp as markers of sustained benefit of infliximab treatment in crohn's disease: A post-HOC analysis of the accent1 trial." American Journal of Gastroenterology 106: S462-S463	C insufficient data
51. Corstjens, P. L., K. Wiesmeijer, G. J. Wolbink, J. Tanke, D. W. Hommes and H. Fidder (2011). "A rapid test for quantitative determination of infliximab trough levels in blood." Gastroenterology 1): S276-S277	C insufficient data
52. Daperno, M., A. Lavagna, M. Fracchia, C. Guiotto, L. Germano, C. Rigazio, E. Ercole, M. Migliardi, R. Pellerito and R. Rocca (2013). "Infliximab trough levels (IFX-TL) are higher in patients with inflammatory bowel disease (IBD)treated with immunosuppressives: Clinical correlations of IFX-LT and antibodies to infliximab (ATI) in IBD." Gastroenterology 1): S781	C insufficient data
53. Daperno, M., A. Lavagna, M. Fracchia, C. Guiotto, L. Germano, C. Rigazio, E. Ercole, M. Migliardi, R. Pellerito and R. Rocca (2013). "Clinical correlations of infliximab trough levels (IFX-TL) and antibodies to infliximab (ATI) in inflammatory bowel disease." Journal of Crohn's and Colitis 7: S239.	C insufficient data

54. Daperno, M., F. Frigerio, C. Guiotto, G. Laura, E. Ercole, A. Lavagna, C. Rigazio, S. Arico, M. Migliardi, R. Pellerito and R. Rocca (2013). "Comparison of the performance of two commercially available tests for determination of infliximab trough levels (IFX-TL) and antibodies to infliximab (ATI), promonitor and immundiagnostik, in inflammatory bowel disease." Digestive and Liver Disease 45: S109.	C insufficient data
55. Daperno, M., F. Frigerio, C. Guiotto, L. Germano, E. Ercole, S. Arico, M. Fracchia, C. Rigazio, A. Lavagna, R. Pellerito, M. Migliardi and R. Rocca (2013). "Identical diagnostic performance of two commercially available tests for infliximab trough levels (IFX-TL) and antibodies to infliximab (ATI) titration in inflammatory bowel disease (IBD): Promonitor and immunodiagnostik tests." Gastroenterology 1): S780.	Duplicate
56. Daperno, M., M. Fracchia, C. Guiotto, L. Germano, E. Ercole, C. Rigazio, A. Lavagna, M. Migliardi, R. Pellerito and R. Rocca (2013). "Clinical implications and stability of determination of infliximab trough levels (IFX-TL) and antibodies to infliximab (ATI) in inflammatory bowel disease." Digestive and Liver Disease 45: S145.	C insufficient data
57. De Bruyn, M., T. Bessissow, T. Billiet, I. Cleynen, R. Kirkland, X. Liu, S. Hauenstein, K. Drake, S. Singh, M. Ferrante, P. Rutgeerts, G. Van Assche, I. Arijs, G. Opdenakker and S. Vermeire (2014). "Biomarker panel for prediction of mucosal healing in patients with Crohn's disease under infliximab therapy." Journal of Crohn's and Colitis 8: S45-S46.	C insufficient data
58. De Bruyn, M., T. Bessissow, T. Billiet, I. Cleynen, R. Kirkland, X. Liu, S. Hauenstein, K. Drake, S. Singh, M. Ferrante, P. J. Rutgeerts, G. A. Van Assche, I. Arijs, G. Opdenakker and S. Vermeire (2014).	Duplicate
59. Dotan, I., H. Yanai, Y. Ron, R. Kariv, S. Fishman, L. Yahav, M. Ben-Yehoyada, E. Santo and D. R. Mould (2014). "Population pharmacokinetic evaluation of adlimumab reveals patient factors that increase adalimumab clearance and shorten half-life in inflammatory bowel disease patients." Gastroenterology 1): S-243.	C insufficient data
60. Dotan, I., Y. Ron, H. Yanai, S. A. Becker, S. Fishman, L. Yahav, M. B. Yehoyada and D. R. Mould (2013). "Population pharmacokinetic evaluation of infliximab reveals patient factors that increase infliximab clearance and shorten half-life in inflammatory bowel disease patients." Gastroenterology 1): S774.	C insufficient data
61. Drastich, P., J. Kozeluhova, M. Jaresova and J. Spicak (2011). "Infliximab serum trough levels and deep remission in patients with IBD." Gastroenterology 1): S292	C insufficient data
62. Drobne, D., P. Bossuyt, C. Breynaert, N. Vande Casteele, G. Compernolle, M. Juergens, V. Ballet, I. Cleynen, W. Wollants, A. Gils, P. Rutgeerts, S. Vermeire and G. Van Assche (2011). "Long term evolution and impact of immunomodulator co-treatment and withdrawal on infliximab trough levels in 223 patients with Crohn's disease." Journal of Crohn's and Colitis 5 (1): S10-S11.	M no algorithm specified / acted on
63. Drobne, D., P. J. Bossuyt, C. Breynaert, N. V. Casteele, G. Compernolle, M. Jurgens, V. Ballet, W. J. Wollants, I. Cleynen, P. J. Rutgeerts, S. Vermeire, A. Gils and G. A. Van Assche (2011). "Crohn's disease: Infliximab trough levels and CRP during infliximab-immunomodulator combination treatment are associated with clinical ouCTome after immunomodulator withdrawal." Gastroenterology 1): S62.	C insufficient data
64. Duricova, D., K. Malickova, M. Bortlik, N. Machkova, V. Komarek, E. Bouzkova and M. Lukas (2011). "Predictors of sustained response to infliximab in patients with Crohn's disease: A single cohort study." Gastroenterology 1): S593.	C insufficient data
65. Echarri, A., R. Ferreiro, R. Fraga-Iriso, M. Barreiro-De Acosta, J. Cid, L. De-Castro, S. Pereira, A. Fernandez-Villaverde, S. Soto, D. Carpio, B. Gonzalez, E. Castro, V. Ollero and A. C. Campos (2014). "Drug trough levels and primary nonresponse to antitnf therapy in moderate-severe crohn disease. Results of the optimiza study." Gastroenterology 1): S-247.	C insufficient data

66. Eser, A., C. Primas, R. Shringarpure, S. Hauenstein, S. L. Wang and W. Reinisch (2012). "Detection of anti infliximab antibodies in patients with inflammatory bowel disease (IBD) in the presence of infliximab by homogeneous liquid phase anti infliximab mobility shift assay." American Journal of Gastroenterology 107: S657	Duplicate
67. Fasanmade, A. A., C. Wagner, H. Davis, M. Graham, D. Everitt and A. Gottlieb (2002). "Comparison of the pharmacokinetics of infliximab in patients with psoriasis or Crohn's disease not receiving concomitant immunosuppressants or corticosteroids." Journal of Investigative Dermatology 119(1): 243-243.	Not C, M or ATC
68. Fasanmade, A. A., P. Marsters, E. Munsanje, M. A. Graham, H. M. Davis and S. Van Deventer (2003). "Infliximab pharmacokinetics and improvement in fistulizing Crohn's Disease." Gastroenterology 124(4): A61-A61.	C insufficient data
69. Fasanmade, A. A., Y. W. Zhu, C. Wagner, C. Pendley and H. M. Davis (2002). "Population pharmacokinetics of single dose infliximab in patients with Crohn's disease." Clinical Pharmacology & Therapeutics 71(2): P66-P66.	Not C, M or ATC
70. Fasanmade, A., A. Olson, W. Bao, C. Pendley, H. Davis and L. Mayer (2002). "Relationship between infliximab pharmacokinetics and improvement in Crohns disease." Gastroenterology 122(4): A617-A618.	C insufficient data
71. Garces, S., J. Demengeot and E. Benito-Garcia (2012). "Clinical impact of immunogenicity of infliximab, adalimumab and etanercept: A systematic review of the literature with a meta-analysis." Annals of the Rheumatic Disease 71	Superseded by full paper
72. Garces, S., J. Demengeot, G. J. Wolbink, L. Aarden and E. Benito-Garcia (2011). "The immunogenicity of infliximab, adalimumab and etanercept in rheumatoid arthritis, ankylosing spondylitis, psoriatic arthritis, crohn's disease and ulcerative colitis a quantitative and a qualitative review." Arthritis and Rheumatism 1).	Superseded by full paper
73. Garces, S., J. Demengeot, J. C. Da Silva and L. Aarden (2011). "Bridging elisa as a screening assay to monitor immunogenicity in routine clinical practice." Arthritis and Rheumatism 1).	<50% CD
74. Garces, S., J. Demengeot, J. Canas-da-Silva and L. Aarden (2013). "Bridging ELISA as a secreening assay to monitor immunogenicity in routine clinical practice." Annals of the Rheumatic Disease 71.	Duplicate
75. Garces, S., J. Freitas, J. Canas-Silva, L. Aarden and J. Demengeot (2013). "The impact of immunogenicity on drug safety profile." Annals of the Rheumatic Diseases 72.	C insufficient data
76. Garimella, T. S., J. Z. Peng, K. Beck, P. A. Noertersheuser, K. G. Lomax, S. K. Paulson and P. F. Pollack (2006). "Pharmacokinetics of adalimumab in a long-term investigation of the induction and maintenance of remission in patients with Crohn's disease (CLASSIC I and CLASSIC II)." Gastroenterology 130(4): A481-A481.	C insufficient data
77. Guan, J., L. Salz, K. Frost, A. Muise, T. Walters and A. Griffiths (2012). "Long-term outcomes with infliximab treatment in children with crohn's disease at a single centre peter church." Inflamm Bowel Dis 18: S5-S6.	Duplicate
78. Guilday, C., D. Eastwood, Y. Zadvornova, D. Stein, A. S. Naik, K. Best, S. Skaros and L. P. Perera (2013). "Concomitant use of immunomodulator therapy results in higher serum infliximab levels compared to monotherapy without lowering serum haca levels." Gastroenterology 1): S429.	C insufficient data
79. Guiotto, C., L. Germano, M. Vizzini, R. Cerruti, F. Frigerio, M. Daperno, R. Rocca and M. Migliardi (2013). "Determination of infliximab trough levels (IFX-TL) and antibodies to Infliximab (ATI) in inflammatory bowel disease." Biochimica Clinica 37: S475.	Superseded
80. Hadigan, C. B. R., C. P. Braegger, E. Vasilauskis, J. C. Escher, M. Sinaasappel, G. D. Ferry, B. Kirschner, C. Wagner, R. Livingston, K. DeWoody and H. S. Winter (1999) Pharmacokinetics of infliximab (Anti-	C insufficient data

TNFx) in children with Crohn's disease: A multicenter trial. Journal of Pediatric Gastroenterology and Nutrition 29, 525	
81. Hauenstein, S., J. Salbato, S. Lockton and S. Singh (2013). "Characterization of neutralizing anti-drug antibody response in patients with loss of response to anti-TNF therapy." Gastroenterology 1): S418	C insufficient data
82. Hayes and Inc (2013) Use of anti-infliximab antibody levels to monitor infliximab treatment in patients with inflammatory bowel disease (IBD) (Structured abstract). Health Technology Assessment Database	Not available
83. Hester, K. D., X. Liu, J. Salbato, S. Lockton, S. Hauenstein and S. Singh (2014). "Improved homogeneous mobility shift assay (HMSA) for the detection of neutralizing antibodies (NAB) in ibd patients treated with infliximab or adalimumab." Gastroenterology 1): S-248	Not M, C or ATC
84. Hibi, T., A. Sakuraba, M. Watanabe, S. Motoya, H. Ito, N. Sato, T. Yoshinari, K. Motegi, K. Yoshitaka, M. Takazoe, Y. Suzuki, T. Matsumoto, K. Kawakami, I. Hirata, S. Tanaka, T. Ashida and T. Matsui (2012). "Decrease in serum infliximab level precedes loss of clinical response and can be easily detected by the elevation of C-reactive protein in crohn's disease." Gastroenterology 1): S388.	C insufficient data
85. Hoekman, D. R., J. F. Brandse, T. De Meij, T. Hummel, M. Lowenberg, M. A. Benninga, G. R. D'Haens and A. Kindermann (2014). "Large variation in infliximab trough levels is associated with disease activity in paediatric inflammatory bowel disease." Gastroenterology 1): S-782.	C insufficient data
86. Hoekman, D., H. Brandse, T. De Meij, T. Hummel, M. Lowenberg, M. Benninga, G. D'Haens and A. Kindermann (2014). "Large variation in infliximab trough levels is associated with disease activity in paediatric inflammatory bowel disease." Journal of Crohn's and Colitis 8: S35.	C insufficient data
87. Huang, V. W., C. Prosser, C. Shalapay, D. K. Fedorak, N. Dhami, H. Wang, K. I. Kroeker and R. N. Fedorak (2014). "In IBD outpatients knowledge of fecal calprotectin and infliximab trough levels significantly enhances infliximab dose escalation decision making." Journal of Crohn's and Colitis 8: S255.	M no algorithm specified / acted on
88. Huang, V. W., N. Dhami, D. K. Fedorak, C. Prosser, C. Shalapay, K. I. Kroeker and R. N. Fedorak (2014). "Disparity between infliximab trough level and infliximab associated adverse events." Journal of Crohn's and Colitis 8: S282	C insufficient data
89. Huang, V., K. I. Kroeker, H. Wang, C. Prosser, S. Carol, N. Dhami, D. K. Fedorak and R. N. Fedorak (2014). "In IBD outpatients knowledge of fecal calprotectin and infliximab trough levels significantly alters clinical decision making." Gastroenterology 1): S241-S242.	Duplicate
90. Huang, V., N. Dhami, D. K. Fedorak, C. Prosser, S. Carol, K. I. Kroeker and R. N. Fedorak (2014). "Infliximab trough levels are correlated with infliximab-associated adverse events." Gastroenterology 1): S-1.	Not M, C or ATC
91. Imaeda, H., A. Andoh, H. Ban, S. Bamba, M. Sasaki, T. Tsujikawa and Y. Fujiyama (2012). "The new immunoassay for the accurate determination of antibodies to infliximab, and relationship between its serum level and inflammatory values in crohn's disease." Gastroenterology 1): S349.	Superseded by full text
92. Imaeda, H., A. Andoh, K. Takahashi, T. Fujimoto, H. Ban, S. Bamba and Y. Fujiyama (2012). "Serum infliximab trough levels above 1.0 mg/ml are required to obtain clinical efficacy in patients with crohn's disease." Inflamm Bowel Dis 18: S59-S60.	C insufficient data
93. Imaeda, H., A. Andoh, S. Bamba, T. Tsujikawa and Y. Fujiyama (2011). "Development of a new Immunoassay for the accurate determinationof anti-infliximab antibodies in Crohn's Disease." Inflamm Bowel Dis 17: S42.	Duplicate
94. Imaeda, H., K. Takahashi, T. Fujimoto, S. Bamba, M. Sasaki, T. Tsujikawa, Y. Fujiyama and A. Andoh (2013). "Accurate determination of serum adalimumab and anti-adalimumab antibodies levels during maintenance therapy for crohn's disease." Gastroenterology 1): S431.	Not M, C or ATC

95. Irving, P. M., Z. Arkir, J. Duncan, M. Sastrillo, S. Anderson and J. Sanderson (2012). "Initial experience with infliximab levels in a tertiary IBD centre." Gut 61: A238.	C insufficient data
96. Jauregui-Amezaga, A., I. Ordas, M. Gallego, A. Ramirez, S. Pino, M. C. Masamunt, M. Juan, E. Ricart, J. Yague and J. Panes (2013). "Impact of serum drug level and human anti-drug antibody measurement on management of biologic drugs in inflammatory bowel disease." Journal of Crohn's and Colitis 7: S202-S203.	C insufficient data
97. Juan, G., A. Alvarino, L. Oltra, N. Maroto, N. Cano, I. Ferrer and J. Hinojosa (2014). "Utility of "trough levels" determination and anti-infliximab antibodies in patients with inflammatory bowel disease. Estimation of individual pharmacokinetic parameters (PK) through population pharmacokinetic model." Journal of Crohn's and Colitis 8: S190	C insufficient data M no algorithm specified / acted on
98. Karmiris, K., G. Paintaud, D. Degenne, M. Ferrante, A. C. Duveau, M. Noman, G. A. Van Assche, S. Vermeire and P. Rutgeerts (2008). "Adalimumab trough serum levels and clinical response in a single-center cohort of inflammatory bowel disease patients: Can trough serum levels serve as a predictor for future loss of response?" Gastroenterology 134(4): A68-A68.	C insufficient data
99. Karmiris, K., M. Noman, G. Paintaud, M. Ferrante, A. C. Duveau, D. Degenne, G. A. Van Assche, S. Vermeire and P. Rutgeerts (2008). "A 3-week course of 80 mg weekly administered adalimumab as a rescue therapy for patients with Crohn's disease who lost response to 40 mg weekly: Relationship with adalimumab trough serum levels." Gastroenterology 134(4): A640-A640.	C insufficient data
100. Karsan, S. S., E. R. Cohen, S. R. Targan, A. Ippoliti, D. Q. Shih, E. A. Vasiliauskas, M. Dubinsky, D. Berel and D. P. McGovern (2012). "Analysis of clinical and serological associations, and the clinical consequences of the development of human anti-chimeric antibodies (HACAS), and low serum infliximab (IFX) levels in inflammatory bowel disease (IBD)." Gastroenterology 1): S264	C insufficient data M no algorithm specified / acted on
101. Kerr, J., A. Nair and R. Hinds (2013). "Variable practice in children with inflammatory bowel disease requiring infliximab infusions across Australia." Journal of Gastroenterology and Hepatology 28: 141.	M no algorithm specified / acted on
102. Kong, J. Y., C. Bundell, J. Pawlik, P. Hollingsworth and G. Forbes (2013). "Low Trough serum infliximab and antibodies to infliximab in smokers." Inflamm Bowel Dis 19(3): E35-E36.	C insufficient data
103. Kong, J. Y., C. S. Bundell, J. Pawlik, P. N. Hollingsworth and G. M. Forbes (2011). "Smoking is associated with low trough serum infliximab levels and presence of anti-infliximab antibody in maintenance treatment of inflammatory bowel disease (IBD)." Journal of Gastroenterology and Hepatology 26: 59	C insufficient data
104. Lamblin, C., A. Aubourg, D. Ternant, L. Picon, T. Lecomte and G. Paintaud (2012). "Concentration effect relationship of infliximab in Crohn's disease: Results of a cohort study." Journal of Crohn's and Colitis 6: S142-S143.	C insufficient data
105. Leclerc, M., H. Marotte, S. Paul, E. Del Tedesco, P. Gonzalo, J. M. Phelip, L. Peyrin Biroulet and X. Roblin (2014). "Persistence of antibodies to infliximab for more than two months strongly predicts loss of response to infliximab in inflammatory bowel diseases." Journal of Crohn's and Colitis 8: S226-S227.	C insufficient data
106. Li, J. L., S. K. Paulson, Y. L. Chiu, A. Robinson, K. G. Lomax and P. F. Pollack (2010). "Evaluation of potential correlations between serum adalimumab concentration and remission in patients with Crohn's disease in classic I and II." Gastroenterology 1): S101.	C insufficient data
107. Lowenberg, M., J. Brandse, L. Vos, C. Ponsioen, G. Van Den Brink and G. D'Haens (2014). "High infliximab trough levels are associated with impaired quality of life in IBD patients in clinical and biochemical remission on maintenance IFX therapy." Journal of Crohn's and Colitis 8:	C insufficient data

S262-S263	
108. Lowenberg, M., J. F. Brandse, L. M. Vos, C. Ponsioen, G. R. Van Den Brink and G. R. D'Haens (2014). "High infliximab trough levels are associated with impaired quality of life in IBD patients in clinical and biochemical remission on maintenance IFX therapy." Gastroenterology 1): S-450.	C insufficient data
109. Lukas, M., K. Malickova, M. Bortlik and D. Duricova (2009). "Anti-infliximab antibodies in routine clinical practice - Is it worth to assess them?" Gastroenterology 1): A679.	C insufficient data
110. Malickova, K., I. Janatkova, D. Duricova, M. Bortlik and M. Lukas (2011). "Serum infliximab levels, antibodies to infliximab and albumin concentrations during infliximab treatment in patients with inflammatory bowel disease." Clinical and Experimental Rheumatology 29(1): 213-213.	C insufficient data
111. Martin Arranz, M. D., E. Martin Arranz, D. Pascual-Salcedo, C. De Diego, M. Jaquotot, S. Gomez Senent, J. Poza and J. M. Suarez Parga (2014). "Infliximab trough levels and antibodies: Relationship with infusion reaction, immunomodulators and biological parameters." Journal of Crohn's and Colitis 8: S251	C insufficient data
112. Mazor, Y., U. Koplov, D. Ben Hur, R. Almog, M. Waterman, S. Ben-Horin and Y. Chowers (2013). "Evaluating Adalimumab drug and antibody levels as predictors of clinical and laboratory response in Crohn's disease patients." Journal of Crohn's and Colitis 7: S217.	Cinsufficient data
113. McTigue, M., W. Sandborn, B. Levesque and D. Patel (2013). "Infliximab therapeutic drug monitoring in clinical practice: Indications and utility." American Journal of Gastroenterology 108: S512.	C insufficient data
114. Morgenstern, J., E. Baestlein, L. Leifeld, P. Nguyen, J. Stein and W. Kruis (2012). "Infliximab drug levels in Crohn's disease responding to the treatment." Journal of Crohn's and Colitis 6: S125	C insufficient data
115. Noman, M., F. Baert, S. Vermeire, G. Van Assche, G. D'Haens, A. Carbonez and P. Rutgeerts (2002). "Post infusion infliximab levels determine duration of response in Crohn's disease and are directly related to infusion reactions." Gastroenterology 122(4): A100-A100.	C insufficient data
116. O'Donnell, S., J. M. Stempak and M. S. Silverberg (2014). "Is there a higher rate of infliximab dose optimization in initial responders between UC and CD cases?" Gastroenterology 1): S462-S463.	C insufficient data M no algorithm specified / acted on
117. Papamichail, K., N. V. Casteele, S. Hauenstein, F. Princen, S. Singh, M. Ferrante, G. A. Van Assche, P. J. Rutgeerts, A. Gils and S. Vermeire (2014). "Prediction of sustained remission after discontinuation of infliximab in patients with crohn's disease." Gastroenterology 1): S-457.	C insufficient data
118. Pariente, B., G. P. De Chambrun, M. Desroches, C. De Cassan, J. M. Gornet, P. Desreumaux, R. Krzysiek, D. Emilie, J. F. Colombel and M. Allez (2011). "Clinical value of measuring trough levels and human anti-chimeric antibodies in patients with inflammatory bowel disease who lost response to infliximab therapy." Gastroenterology 1): S277.	Superseded by full text
119. Pariente, B., G. Pineton De Chambrun, M. Desroches, C. De Cassan, J. Gornet, P. Desreumaux, R. Krzysiek, D. Emilie, J. Colombel and M. Allez (2011). "Clinical value of measuring trough levels and human anti-chimeric antibodies in patients with inflammatory bowel disease who lost response to infliximab therapy." Journal of Crohn's and Colitis 5 (1): S111-S112.	Duplicate
120. Paul, S., E. Del Tedesco, H. Marotte, L. Clavel, J. M. Phelip, L. Peyrin-Biroulet and X. Roblin (2013). "Therapeutic drug monitoring of infliximab and mucosal healing in inflammatory bowel disease: A prospective study." Gastroenterology 1): S92.	Superseded by full text
121. Paul, S., E. Del Tedesco, H. Marotte, M. Rinaudo-Gaujous, J. M. Phelip, L. Peyrin-Biroulet and X. Roblin (2013). "Infliximab concentration is associated with mucosal healing in intestinal bowel disease (IBD)."	C insufficient data

Journal of Crohn's and Colitis 7: S199-S200.	
122. Paulson, S. K., P. Nocrtersheuser, P. F. Pollack and R. S. Hoffman (2005). "Pharmacokinetics of adalimumab from classic, a randomized phase 3 trial for the induction of clinical remission in patients with Crohn's." Gastroenterology 128(4): A585-A585.	C insufficient data
123. Pradhan, R. S., S. Sharma, R. Thakkar, A. Robinson, J. S. Hyams, J. R. Rosh, F. Ruemmele and W. M. Awni (2013). "Relationship between adalimumab concentration and efficacy for the induction of clinical remission in pediatric patients with moderate to severe crohn's disease." Gastroenterology 1): S230-S231.	Duplicate
124. Pradhan, R., S. Sharma, R. Thakkar, A. Robinson, J. Hyams, J. Rosh, F. M. Ruemmele and W. Awni (2013). "Relationship between adalimumab concentration and efficacy for the induction of clinical remission in pediatric patients with moderate to severe Crohn's disease." Journal of Crohn's and Colitis 7: S164.	C insufficient data
125. Rai, T., U. Navaneethan, D. Dalal, B. Lashner and B. Shen (2012). "Clinical implications of measuring infliximab levels and human anti-chimeric antibodies in patients with inflammatory bowel disease." American Journal of Gastroenterology 107: S634.	M no algorithm specified / acted on
126. Rekvig, M., M. Gedde Dahl, J. Bratlie, N. Bolstad, B. Moum, J. Jahnsen and K. E. A. Lundin (2014). "Anti-TNFalpha drug level measurements in IBD patients." Journal of Crohn's and Colitis 8: S301-S302.	C insufficient data
127. Roblin, X., H. Marotte, E. Del Tedesco, M. Rinaudo-gaujous, J. M. Phelip and S. Paul (2013). "Residual adalimumab trough levels are associated with clinical remission and mucosal healing in IBD." Gastroenterology 1): S778.	C insufficient data M no algorithm specified / acted on
128. Roblin, X., M. Rinaudo, E. Del Tedesco, J. M. Phelip, L. Peyrin Biroulet and S. Paul (2014). "Development of an algorithm incorporating pharmacokinetics of adalimumab in inflammatory bowel diseases." Journal of Crohn's and Colitis 8: S41.	C insufficient data M no algorithm specified / acted on
129. Roblin, X., M. Rinaudo-gaujous, E. Del Tedesco, J. M. Phelip, C. Genin, L. Peyrin-Biroulet and S. Paul (2014). "Development of an algorithm incorporating pharmacokinetics of adalimumab in inflammatory bowel diseases." Gastroenterology 1): S-150.	Duplicate
130. Rosenthal, C., G. Melmed, B. Tripuraneni, J. Gebbia, S. Callejas, S. Farrior, S. Rabizadeh and M. Dubinsky (2012). "Early infliximab trough levels predict remission at one year in pediatric IBD patients." Inflamm Bowel Dis 18: S81.	C insufficient data
131. Rubin, D., S. Hauenstein and S. Singh (2013). "Post-marketing review of serum adalimumab and antibodies to adalimumab using the mobility shift assay platform." American Journal of Gastroenterology 108: S532.	C insufficient data
132. Scaldaferri, F., S. Pecere, V. Petito, G. Cammarota, M. C. Campanale, G. L. Rapaccini, A. Armuzzi, L. Guidi, A. Poscia, A. Amato, A. Sgambato, A. Gasbarrini and A. Papa (2012). "Infliximab and TNF ALFA measurement in intestinal mucosa of IBD patients: A new tool for the clinic?" Digestive and Liver Disease 44: S189-S190.	C insufficient data
133. Schatz, S. B., C. Prell, F. Freudenberg, T. Schwerd, P. Bufler and S. Koletzko (2011). "Correlation of infliximab levels and antibodies with clinical outcome in children with IBD." Journal of Pediatric Gastroenterology and Nutrition 52: E45.	C insufficient data
134. Semmler, J. M., A. Pilch, F. P. Armbruster, A. Dignass, W. Kruis and J. Stein (2014). "Development of a new immunoassay for the accurate determination of anti-infliximab antibodies in inflammatory bowel disease." Journal of Crohn's and Colitis 8: S41.	Duplicate
135. Semmler, J., A. Pilch, F. P. Armbruster, A. Dignass and J. Stein (2014). "Development of a new immunoassay for the accurate determination of anti-infliximab antibodies in inflammatory bowel disease." Clinical Chemistry and Laboratory Medicine 52: S1008.	Duplicate

136. Settesoldi, A., M. Giannotta, M. Milla, S. Genise, A. Santini, S. Bagnoli, V. Annese, A. Matucci, A. Vultaggio, G. Petroni, S. Pratesi, F. Nencini and E. Maggi (2012). "Loss of efficacy and adverse drug reactions during infliximab therapy in IBD patients are related to the appearance of anti-infliximab antibodies." Journal of Crohn's and Colitis 6: S151.	Duplicate
137. Settesoldi, A., M. Giannotta, S. Genise, A. Santini, S. Bagnoli, M. Milla, V. Annese, A. Matucci, A. Vultaggio, G. Petroni, S. Pratesi, F. Nencini and E. Maggi (2012). "Loss of efficacy and adverse drug reactions during infliximab therapy in IBD patients are related to the appearance of anti-infliximab antibodies." Digestive and Liver Disease 44: S194.	C insufficient data
138. Sharma, S., R. Pradhan, R. Thakkar, A. Robinson, J. Hyams, J. Rosh, F. M. Ruemmele and W. Awni (2013). "Relationship between adalimumab concentration and efficacy for the maintenance of clinical remission in pediatric patients with moderate to severe Crohn's disease." Journal of Crohn's and Colitis 7: S163.	C insufficient data
139. Sharma, S., R. S. Pradhan, R. Thakkar, A. Robinson, J. S. Hyams, J. R. Rosh, F. Ruemmele and W. M. Awni (2013). "Relationship between adalimumab concentration and efficacy for the maintenance of clinical remission in pediatric patients with moderate to severe crohn's disease." Gastroenterology 1): S231.	C insufficient data
140. Sorrentino, D., S. Hauenstein, M. Marino, S. Lockton, D. Zarifi, T. Del Bianco and S. Singh (2013). "Low dose infliximab for prevention of postoperative recurrence of crohn's disease: Long term follow-up and impact of infliximab trough levels and antibodies to infliximab." Gastroenterology 1): S777	C insufficient data
141. Steenholdt, C., J. Brynskov, O. Thomsen, L. K. Munck, J. Fallingborg, L. A. Christensen, G. Pedersen, J. Kjeldsen, K. Bendtzen and M. A. Ainsworth (2013). "Secondary infliximab treatment failure in crohn's disease: Therapeutic implications of measuring drug and anti-drug antibodies by three different binding assays." Gastroenterology 1): S773	Duplicate
142. Steenholdt, C., J. Brynskov, O. Thomsen, L. K. Munck, J. Fallingborg, L. A. Christensen, G. Pedersen, J. Kjeldsen, B. A. Jacobsen, A. S. Oxholm, J. Kjellberg, K. Bendtzen and M. A. Ainsworth (2013). "Treatment of secondary infliximab failure in crohn's disease based on serum levels of infliximab and antibodies against infliximab: The danish study of optimizing infliximab therapy in crohn's disease (do it crohn) randomized clinical trial." Gastroenterology 1): S22.	Superseded by full paper
143. Steenholdt, C., J. Brynskov, O. Thomsen, L. K. Munck, J. Fallingborg, L. A. Christensen, G. Pedersen, J. Kjeldsen, K. Bendtzen and M. A. Ainsworth (2013). "Secondary infliximab treatment failure in crohn's disease: Therapeutic implications of measuring drug and anti-drug antibodies by three different binding assays." Gastroenterology 1): S773.	Duplicate
144. Steenholdt, C., J. Brynskov, O. Thomsen, L. Munck, J. Fallingborg, L. Christensen, G. Pedersen, J. Kjeldsen, K. Bendtzen and M. Ainsworth (2013). "Secondary infliximab treatment failure in Crohn's disease: Therapeutic implications of measuring drug and anti-drug antibodies by three different binding assays." Journal of Crohn's and Colitis 7: S159.	Superseded by full text
145. Steenholdt, C., J. Brynskov, O. Thomsen, L. Munck, J. Fallingborg, L. Christensen, G. Pedersen, J. Kjeldsen, K. Bendtzen, S. Lockton, S. Hauenstein, R. Shringarpure, E. Chuang, S. Singh and M. Ainsworth (2012). "Comparison of techniques for monitoring infliximab and antibodies to infliximab in Crohn's disease patients with infliximab treatment failure." American Journal of Gastroenterology 107: S622.	Superseded by full text
146. Steenholdt, C., K. Bendtzen, J. Brynskov, O. O. Thomsen and M. A. Ainsworth (2014). "Clinical implications of measuring drug and anti-drug antibodies by different assays when optimizing infliximab treatment failure in Crohn's disease." Journal of Crohn's and Colitis 8: S291.	Duplicate
147. Steenholdt, C., K. Bendtzen, J. Brynskov, O. Thomsen and M. A. Ainsworth (2014). "Clinical implications of measuring drug and anti-drug antibodies by different assays when optimizing infliximab treatment	Superseded by full text

failure in crohn's disease." Gastroenterology 1): S-240.	
148. Steenholdt, C., K. Bendtzen, O. O. Thomsen, J. Brynskov and M. Ainsworth (2010). "Discriminating between response types in infliximab-treated patients with Crohn's disease: Sensitivity and specificity of combined assessment of infliximab trough levels and antidrug antibodies." Scandinavian Journal of Gastroenterology 45: 59-59.	Superseded by full text
149. Steenholdt, C., M. Svenson, K. Bendtzen, O. Thomsen, J. Brynskov and M. A. Ainsworth (2011). "Can measurements of anti-infliximab antibodies predict acute severe infusion reactions to infliximsab?" Gastroenterology 1): S774	C insufficient data
150. Steenholdt, C., M. Svenson, M. A. Ainsworth, O. Thomsen, J. Brynskov and K. Bendtzen (2012). "Comparison of techniques for monitoring infliximab bioavailability and immunogenicity in crohn's disease." Gastroenterology 1): S781	ATC insufficient data
151. Steenholdt, C., M. Svenson, M. A. Ainsworth, O. Thomsen, J. Brynskov and K. Bendtzen (2012). "Comparison of techniques for monitoring infliximab bioavailability and immunogenicity in crohn's disease." Gastroenterology 1): S781.	Duplicate
152. Steenholdt, C., O. O. Thomsen, J. Brynskov, K. Bendtzen and M. A. Ainsworth (2010). "Discriminating between response types in infliximab-treated patients with Crohn's disease: Sensitivity and specificity of combined assessment of infliximab trough levels and anti-drug antibodies." Gastroenterology 1): S687-S688.	Insufficient data
153. Steenholdt, C., Y. Palarasah, K. Bendtzen, A. Teisner, B. Teisner, J. Brynskov and C. Nielsen (2013). "Pre-existing IgG antibodies to the Fab region of infliximab predict efficacy and safety in IBD patients naive to anti-TNF agents." Journal of Crohn's and Colitis 7: S6.	C insufficient data
154. Steenholdt, C., Y. Palarasah, K. Bendtzen, A. Teisner, J. Brynskov, B. Teisner and C. H. Nielsen (2013). "Pre-existing IGG antibodies to the fab region of infliximab predict efficacy and safety in ibd patients naive to anti-TNF agents." Scandinavian Journal of Immunology 77 (4): 333.	C insufficient data
155. Szepes, Z., E. Kunstar, K. Farkas, F. Nagy, R. Gyulai, R. Kui, A. Kinyo, A. Balint, M. Szucs, T. Wittmann and T. Molnar (2013). "Clinical utility of measuring serum TNF alpha level, anti TNF alpha levels and antibody titers in critical situations in inflammatory bowel disease and in psoriasis." Journal of Crohn's and Colitis 7: S118-S119.	C insufficient data
156. Tang, J., X. Gao, M. Zhi, H. Zhou, H. Chen, M. Zhang, Q. Yang and Z. Liang (2014). "Serum infliximab levels and early mucosal healing in Crohn's disease." Journal of Crohn's and Colitis 8: S209-S210.	C insufficient data
157. Turon, J., A. Langseder, R. Irizarry, K. Ahuja and J. R. Rosh (2013). "Clinical outcome of pediatric IBD patients after measurement of infliximab drug and anti-drug antibody levels." Gastroenterology 1): S531.	M no algorithm specified / acted on
158. Ungar, B., A. Anafy, M. Yavzori, O. Picard, E. Fudim, U. Kopylov, Y. Ron, H. Yanai, I. Dotan, Y. Chowers, R. Eliakim and S. Ben-Horin (2014). "The clinical and immunological significance of low level of infliximab in the absence of anti-infliximab antibodies in patients with IBD." Journal of Crohn's and Colitis 8: S113.	Duplicate
159. Ungar, B., U. Kopylov, M. Yavzori, E. Fudim, O. Picard, A. Lahat, B. Avidan, A. Lang, B. Weiss, Y. Chowers, R. Eliakim and S. Ben-Horin (2014). "Predictors of formation of antibodies to infliximab (ATI) and secondary loss of response in IBD patients treated with infliximab." Journal of Crohn's and Colitis 8: S45.	C Insufficient data
160. Ussia, V., L. Ceccarelli, S. Maltinti, G. Di Fluri, M. G. Mumolo, V. Bolognesi, A. Ricchiuti, M. Bellini, S. Marchi and F. Costa (2014). "A prospective assessment of antidrug antibody response over time by a new ELISA in patients with IBD treated with infliximab." Journal of Crohn's and Colitis 8: S298-S299.	C insufficient data
161. Van Der Woude, C. J., E. Bultman, J. Deuring, R. West, Z. Zelinkova	C insufficient data

and M. Peppelenbosch (2013). "Adalimumab trough levels in a prospective cohort of Crohn's disease patients." Journal of Crohn's and Colitis 7: S250.	
162. Van Der Woude, C. J., J. J. Deuring, R. West, Z. Zelinkova and M. P. Peppelenbosch (2013). "Adalimumab trough levels in a prospective cohort of crohn's disease patients." Gastroenterology 1): S567.	Duplicate
163. Van Moerkercke, W., C. Ackaert, G. Compernolle, M. Jurgens, I. Cleynen, G. A. Van Assche, P. J. Rutgeerts, A. Gils and S. Vermeire (2010). "High infliximab trough levels are associated with mucosal healing in Crohn's disease." Gastroenterology 1): S60.	C insufficient data
164. Van Moerkercke, W., G. Compernolle, C. Ackaert, A. Gils, S. Vermeire, M. Jurgens, I. Cleynen, G. Van Assche and P. Rutgeerts (2010). "Mucosal healing in Crohn's disease is associated with high infliximab trough levels." Journal of Crohn's and Colitis Supplements 4 (1): 30-31.	C insufficient data
165. Vande Casteele, N., A. Gils, G. Compernolle, V. Ballet, M. Peeters, K. Van Steen, S. Simoens, G. Van Assche, M. Ferrante, S. Vermeire and P. Rutgeerts (2013). "Drug level versus clinically based dosing of infliximab maintenance therapy in IBD: Final results of the randomized controlled taxit trial." Inflamm Bowel Dis 19: S2-S3.	Superseded by full text
166. Vande Casteele, N., G. Compernolle, V. Ballet, G. Van Assche, A. Gils, S. Vermeire and P. J. Rutgeerts (2012). "Results on the optimisation phase of the prospective controlled trough level adapted infliximab treatment (TAXIT) trial." Gastroenterology 1): S211-S212	Superseded by full text
167. Vande Casteele, N., G. Compernolle, V. Ballet, G. Van Assche, A. Gils, S. Vermeire and P. Rutgeerts (2012). "Individualised infliximab treatment using therapeutic drug monitoring: A prospective controlled Trough level Adapted infliXImab Treatment (TAXIT) trial." Journal of Crohn's and Colitis 6: S6.	Duplicate
168. Vande Casteele, N., K. Drake, S. Hauenstein, B. G. Levesque, S. Singh and W. Sandborn (2014). "Infliximab and antibody to infliximab concentrations in 7,613 patients shows indication for testing, association with loss of response and provides new insights into binding characteristics of anti-drug antibodies." Gastroenterology 1): S-242.	C insufficient data M no algorithm specified / acted on
169. Vande Casteele, N., L. Cuypers, S. Singh, L. Ohrmund, S. Hauenstein, G. Van Assche, P. J. Rutgeerts, A. Gils and S. Vermeire (2012). "Antibodies to infliximab can either be persistent or transient: A retrospective case-control study in ibd patients treated with infliximab maintenance therapy." Gastroenterology 1): S114.	M no algorithm specified / acted on
170. Vande Casteele, N., L. Cuypers, S. Singh, S. Hauenstein, L. Ohrmund, E. Chuang, P. Rutgeerts, A. Gils and S. Vermeire (2012). "Transient versus sustained antibodies to infliximab: Possibility to overcome low titer antibody responses by dose optimisation." Journal of Crohn's and Colitis 6: S110.	M no algorithm specified / acted on
171. Vande Casteele, N., M. Peeters, M. Ferrante, G. Compernolle, G. Van Assche, S. Vermeire and A. Gils (2014). "Functional cellular based assay reveals neutralising anti-drug antibodies in IBD patients treated with maintenance adalimumab." Journal of Crohn's and Colitis 8: S268-S269.	Duplicate
172. Vaughn, B. P., M. Martinez-Vazquez, V. Patwardhan, A. C. Moss, W. J. Sandborn and A. S. Cheifetz (2014). "A pilot study of optimized monotherapy with infliximab for patients with inflammatory bowel disease." Gastroenterology 1): S-55.	Superseded by full paper
173. Vaughn, B. P., M. Martinez-Vazquez, V. Patwardhan, A. C. Moss, W. J. Sandborn and A. S. Cheifetz (2014). "Prospective therapeutic drug monitoring to optimizing infliximab (IFX) maintenance therapy in patients with inflammatory bowel disease (IBD)." Gastroenterology 1): S-54.	Superseded by full paper
174. Vaughn, B., M. Matinez-Vazquez and A. Cheifetz (2013). "Infliximab dosing changes based on trough levels in a cohort of IBD patients in clinical remission." Inflamm Bowel Dis 19: S59.	M no algorithm specified / acted on

175. Velayos, F. S., S. Sheibani, S. Lockton, S. Hauenstein, S. Singh, J. P. Terdiman and U. Mahadevan (2013). "Prevalence of antibodies to adalimumab (ATA) and correlation between ATA and low serum drug concentration on CRP and clinical symptoms in a prospective sample of IBD patients." Gastroenterology 1): S91.	C insufficient data
176. Veres, G., J. L. Kaplan, E. De Greef, E. Chuang, D. Szabo, K. Molnar, L. Ohrmund, S. Hauenstein, S. Singh, A. Arato, G. Veereman and H. S. Winter (2012). "New assay to detect infliximab levels and anti-infliximab antibodies from a single serum sample is useful in measuring efficacy of treatment with infliximab in children with IBD." Gastroenterology 1): S386.	C insufficient data
177. Wang, S. L., L. Ohrmund, S. Hauenstein, J. Salbato, R. Reddy, P. Monk and S. Lockton (2011). "Evaluation of a novel homogeneous mobility shift assay for the measurement of human antibodies-to-infliximab and infliximab levels in patient serum." Arthritis and Rheumatism 1).	Duplicate
178. Wang, S. L., S. Hauenstein, L. Ohrmund, R. Shringarpure, D. C. Wolf, I. A. Diab, J. Salbato, R. Reddy, K. McCowen, S. Shah, S. Lockton, E. Chuang and S. Singh (2012). "Influence of trough serum drug level and immunogenicity on the lack of response to adalimumab therapy in inflammatory bowel disease patients." Arthritis and Rheumatism 64: S819-S820.	C insufficient data
179. Wang, S. L., S. Hauenstein, L. Ohrmund, R. Shringarpure, D. Wolf, I. Diab, J. Salbato, R. Reddy, K. McCowen, S. Shah, S. Lockton, E. Chuang and S. Singh (2012). "Influence of trough serum drug level and immunogenicity on the lack of response to adalimumab therapy in IBD patients presidential poster." American Journal of Gastroenterology 107: S680.	Duplicate
180. Wolf, D. C., S. Hauenstein, S. Lockton and S. Singh (2013). "Mechanisms of loss of response to adalimumab in crohn's disease." Gastroenterology 1): S775	C insufficient data
181. Wolf, D. C., S. Lockton, S. Hauenstein, S. Carroll, S. Singh and E. Chuang (2013). "A multi-center observational study in community gastroenterology practices evaluating the clinical usage of testing for serum levels of infliximab and antibodies to infliximab." Gastroenterology 1): S423.	M no algorithm specified / acted on
182. Wolf, D., R. Shringarpure, S. Lockton, R. Corey, S. Woods, H. Aguilar and E. Chuang (2012). "Clinical experience with measurement of serum infliximab and antibodies to infliximab using a new homogenous mobility shift assay: Results of a multi-center observational study." American Journal of Gastroenterology 107: S658.	C insufficient data
183. Yamada, A., K. Sono, K. Takeuchi and Y. Suzuki (2013). "Clinical and basic studies to understand factors associated with the loss of response to infliximab in patients with Crohn's disease." Journal of Crohn's and Colitis 7: S239.	C insufficient data
184. Yanai, H., L. Lichtenstein, A. Assa, Y. Mazor, B. Weiss, A. Levine, Y. Ron, U. Kopylov, Y. Bujanover, Y. Rosenbach, B. Ungar, A. R. Eliakim, Y. Chowers, R. Shamir, G. Fraser, I. Dotan and S. Ben-Horin (2014). "Anti-TNF and anti-drug antibodies levels predict the ouCTomes of interventions after loss of response to adalimumab and infliximab." Gastroenterology 1): S-381.	C insufficient data M no algorithm specified / acted on
185. Yarur, A., J. P. Trivella, D. A. Sussman, K. Drake, J. S. Barkin, S. Hauenstein, A. R. Deshpande, M. A. Quintero, S. Singh and M. T. Abreu (2014). "Anti-tumor necrosis factor drug levels and anti-bodies are associated with crohn's disease recurrence at the level of the ileo-colonic anastomosis after ileal resection." Gastroenterology 1): S243-S244.	C insufficient data
186. Yarur, A., K. Drake, M. Kubiliun, R. M. Dauer, D. A. Sussman, S. Hauenstein, M. A. Quintero, S. Singh, J. S. Barkin and M. T. Abreu (2014). "Anti-tumor necrosis factor levels are not associated with intestinal extent of mucosal inflammation in patients with inflammatory bowel diseases." Gastroenterology 1): S-244.	Not M, C or ATC

187. Zelinkova, Z., M. P. Peppelenbosch, A. Van Liere-Baron, C. De Haar and C. J. Van Der Woude (2011). "Naturally-occurring autoantibodies against TNF-alpha are present in sera of inflammatory bowel disease patients and influence the response to adalimumab." Gastroenterology 1): S62.	C insufficient data

M – Management type study; C – Correlation type study; ATC – Assay type comparison study, CD - Crohn's disease; RA rheumatoid arthritis

8.7 Appendix 7 Ongoing trials

On-going trials using an algorithm to determine treatment change according to test results

Title (Acronym)	**Status** **Start date** **Estimated completion date**	**URL**
Pediatric Crohn's Disease AdalImumab Level-based Optimization Treatment (PAILOT)	Ongoing - not yet open for participant recruitment Start: Nov 2014 Primary completion due: July 2018	http://clinicaltrials.gov/ct2/show/NCT02256462
A randomized controlled trial investigating tailored treatment with infliximab for active luminal crohn's disease (TAILORIX)	Ongoing Start: Mar 2012 Primary completion due: June 2015	https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2011-003038-14
Adjusting infliximab dose in IBD patients in remission, based on infliximab trough levels: the study on Infliximab Levels in IBD patients Steering Treatment, the ILIST pilot (ILIST)	Ongoing Start: Oct 2013 Primary completion due: Dec 2014	http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=4067

On-going correlation studies

Title (Acronym)	**Status** **Start date** **Expected completion date**	**URL**
Anti-TNF-alpha Trough Level Measurements in Inflammatory Bowel Disease	Ongoing Study start: May 2013 Study primary completion due: May 2016	http://clinicaltrials.gov/ct2/show/NCT02073526
Improving Treatment of Inflammatory Bowel Diseases Through Better Understanding Infliximab Drug and Antibody Levels (OPTIMIZE)	Ongoing Study start: Mar 2014 Study primary completion due: March 2015	https://clinicaltrials.gov/ct2/show/NCT01787786
Personalised Anti-TNF Therapy in Crohn's Disease (PANTS)	Ongoing – currently recruiting	http://public.ukcrn.org.uk/Search/StudyDetail.aspx?StudyID=14175
Utilising drug levels and anti-drug antibodies to predict response to treatment in patients with Inflammatory Bowel Disease	Ongoing Start: Oct 2012 Primary completion due:	https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2011-006084-22

8.8 Appendix 8 Excluded assay type comparison studies

Reference	Reason for exclusion
1. Bodini, G., V. Savarino, P. Dulbecco, I. Baldissarro and E. Savarino (2014). "ELISA vs. HMSA: A comparison between two different methods for the evaluation of adalimumab serum concentration and anti-adalimumab antibodies Preliminary data." Journal of Crohn's and Colitis 8: S278.	Irrelevant comparison
2. Corstjens PL, Fidder HH, Wiesmeijer KC, de Dood CJ, Rispens T, Wolbink GJ, et al. A rapid assay for on-site monitoring of infliximab trough levels: a feasibility study. Analytical & Bioanalytical Chemistry. 2013;405(23):7367-75.	Irrelevant comparison
3. Greathead, L., P. Kelleher and A. Steel (2014). "Development and validation of ELISA to measure serum anti TNFa levels." Journal of Crohn's and Colitis 8: S97-S98.	Irrelevant comparison
4. Imaeda, H., A. Andoh and Y. Fujiyama (2012). "Development of a new immunoassay for the accurate determination of anti-infliximab antibodies in inflammatory bowel disease." Journal of Gastroenterology 47(2): 136-143.	Irrelevant comparison
5. Imaeda, H., K. Takahashi, T. Fujimoto, S. Bamba, T. Tsujikawa, M. Sasaki, Y. Fujiyama and A. Andoh (2014). "Clinical utility of newly developed immunoassays for serum concentrations of adalimumab and anti-adalimumab antibodies in patients with Crohn's disease." Journal of Gastroenterology 49(1): 100-109.	Irrelevant comparison
6. Kopylov, U., Y. Mazor, M. Yavzori, E. Fudim, L. Katz, D. Coscas, O. Picard, Y. Chowers, R. Eliakim and S. Ben-Horin (2012). "Clinical utility of antihuman lambda chain-based enzyme-linked immunosorbent assay (ELISA) versus double antigen ELISA for the detection of anti-infliximab antibodies." Inflamm Bowel Dis 18(9): 1628-1633.	Irrelevant comparison
7. McTigue, M., W. Sandborn, B. Levesque and D. Patel (2013). "Clinical utility of next generation infliximab and antibodies to infliximab assay." American Journal of Gastroenterology 108: S527.	Irrelevant comparison
8. Semmler, J., A. Pilch, F. Armbruster, A. Dignass and J. Stein (2013). "Development of a new immunoassay for the accurate determination of anti-infliximab antibodies in inflammatory bowel disease." Clinical Chemistry and Laboratory Medicine 51 (10): eA27-eA28.	Irrelevant comparison
9. Steenholdt, C., M. A. Ainsworth, M. Tovey, T. W. Klausen, O. O. Thomsen, J. Brynskov and K. Bendtzen (2013). "Comparison of techniques for monitoring infliximab and antibodies against infliximab in Crohn's disease." Therapeutic Drug Monitoring 35(4): 530-538.	Irrelevant comparison
10. Ungar, B., A. Anafy, U. Kopylov, Y. Ron, H. Yanai, I. Dotan, Y. Chowers, A. R. Eliakim and S. Ben-Horin (2014). "The clinical and immunological significance of low level of infliximab in the absence of anti-infliximab antibodies in patients with IBD." Gastroenterology 1): S-245.	Irrelevant comparison
11. Vande Casteele, N., M. Peeters, G. Compernolle, M. Ferrante, G. A. Van Assche, S. Vermeire and A. Gils (2014). "TNF-responsive cellular based assay reveals neutralizing capacity of anti-adalimumab antibodies in crohn's disease and ulcerative colitis patients." Gastroenterology 1): S-242.	Irrelevant comparison

8.9 Appendix 9 Summary of studies evaluating the clinical utility of measuring levels of anti-TNFα and its antibodies

This section summarises the studies by Afif et al. (2010),[55] Pariente et al. (2012),[58] Roblin et al. (2014)[56] and Paul et al. (2013)[57] and details the studies' proposed algorithms.

Table 45 Overview of study characteristics of studies evaluating the clinical utility of measuring levels of anti-TNFα and its antibodies

	Afif et al. (2010) [55]	**Pariente et al. (2012)[58]**	**Roblin et al. (2014)[56]**	**Paul et al. (2013)[57]**
Patients N	155	76	82	52
Condition	IBD (CD 78%)	IBD (CD 72%)	IBD (CD 58%)	IBD (CD 65%)
Study design	Retrospective	Retrospective	Prospective	Prospective
Drug	Infliximab	Infliximab	Adalimumab	Infliximab
Assay type	ELISA (Prometheus laboratories)	LISA TRACKER Premium infliximab ELISA kits	LISA TRACKER Premium adalimumab ELISA kits	LISA TRACKER Premium infliximab ELISA kits
Time of assessment of clinical response following treatment change	Unclear	8 weeks and 6 months	6 and 12 months	8 weeks
Outcome: definition	Complete response: cessation of diarrhea and abdominal pain, complete closure of all fistulas	Clinical response: i) decrease of at least 2 points of HBI Or ii) overall assessment by treating clinician	Clinical remission: CDAI<150, FC<250µg/g	Clinical remission: CDAI<150, Mucosal healing: FC<250µg/g
Reason for testing	At discretion of clinician (LOR or partial response: 71%)	LOR	LOR	LOR
Treatment change	Various according to treating clinician	No change: n=31 (41%) Intensification: n=39 (51%) Switch to adalimumab: n=5	1st) Adalimumab 40mg eow to adalimumab 40mg ew 2nd) Switch to infliximab 5mg/kg	Infliximab 5mg/kg every 8 weeks to infliximab 10mg/kg every 8 weeks

		(7%) Surgery: n=1 (1%)	at 0, 2, and 6 weeks	
Algorithm proposed	yes	no	yes	yes

Abbreviations: IBD inflammatory bowel syndrome; CD Crohn's disease; ELISA enzyme linked immunosorbent assay; HBI Harvey Bradshaw index; CDAI Crohn's disease activity index; FC faecal calprotectin; LOR loss of response; eow every other week; ew every week

The two retrospective studies[55, 58] assessed response to any treatment change that was prescribed by the treating clinician in response to treatment failure and evaluated the relationship between clinical outcomes and test outcomes. The prospective studies[56, 57] tested IBD patients once treatment failure occurred and before a fixed treatment change (dose intensification) was applied. The response to the treatment change was then correlated with the test outcome. Afif et al. 2010[55] concluded that measurement of drug and anti-drug antibodies impacts management and is clinically useful, Paul et al. (2013) concluded that measurement of drug and anti-drug antibodies predicts clinical remission and may guide clinical decisions in practice, Roblin et al. (2014)[56] concluded that knowledge of drug and antidrug antibody levels may have a strong impact on the management of IBD patients with LOR, but Pariente et al. (2012)[58] concluded that clinical response to drug intensification cannot be accurately predicted by measurement of drug and anti-drug antibody levels. The authors reported that there was a considerable number of patients (70%) that showed a clinical response to dose intensification even though the drug test result before dose intensification was positive.

The patients included in the studies were different between and within studies in terms of disease (CD, UC), treatment duration before testing, co-treatment with immunomodulators, disease duration and time of disease assessment following anti-TNFα optimisation. The proposed algorithms varied considerably in terms of drug and anti-drug antibody levels used to predict response but the proposed treatment changes were comparable but differed in detail.

Algorithm proposed by Afif et al. (2010)[55]

On the basis of the study results the following treatment algorithm for IBD patients with drug and anti-drug antibody testing results was suggested:

- Detectable anti-drug antibodies: switch to another anti-TNFα agent and switch class of drug if symptoms persist
- Therapeutic infliximab concentrations (>12µg/ml at 4 weeks or detectable trough level > 1.4 µg/ml): for active disease on endoscopy switch class of drug and for inactive disease on endoscopy investigate for other causes of symptoms

- Sub-therapeutic infliximab concentrations (<12µg/ml at 4 weeks or undetectable trough level <1.4 µg/ml): infliximab intensification or switch within class followed by switch of class if symptoms persist

In summary the study showed that clinical response depends on whether patients are anti-drug antibodies positive or if they had sub-therapeutic or therapeutic infliximab levels and how they are managed according to their serum levels of drugs and antibodies. Those who were anti-drug antibodies positive responded better if switched to a different anti-TNFα drug, while those with therapeutic infliximab level responded if they stayed on the same dose of infliximab. Patients with sub-therapeutic infliximab levels responded to dose intensification of infliximab. The study concluded that use of infliximab and anti-drug antibody tests can potentially avoid inappropriate management.

Algorithm proposed by Paul et al. (2013)[57]

On the basis of their study results the authors suggested the following treatment algorithm for IBD patients with drug and anti-drug antibody testing results:

- Sub-therapeutic infliximab levels (<2µg/ml) and anti-drug antibodies >200ng/ml: switch to another anti-TNFα agent, or optimise IFX and introduce immunomodulator
- Sub-therapeutic infliximab levels (<2µg/ml) and anti-drug antibodies <200ng/ml: infliximab intensification
- Therapeutic infliximab levels (>2µg/ml) and anti-drug antibodies <10ng/ml: infliximab intensification
- Therapeutic infliximab levels (>2µg/ml) and anti-drug antibodies >10ng/ml: switch class of drug

The study found that therapeutic monitoring of drug can help predict response defined as mucosal healing in patients with IBD following infliximab dose intensification.

Algorithm proposed by Roblin et al. (2014)[56]

On the basis of the study results the study suggested the following treatment algorithm for IBD patients with drug and anti-drug antibody testing results:

- Low trough adalimumab concentrations (<4.9µg/ml) and detectable anti-drug antibodies (>10ng/ml): switch to another anti-TNFα agent
- High trough adalimumab concentrations (>4.9µg/ml): switch class of drug
- Low trough adalimumab concentrations (<4.9µg/ml) and no detectable anti-drug antibodies (<10ng/ml): adalimumab intensification (40mg every week)

The findings of the study suggested that those with low trough levels of anti-TNFα drug and undetectable levels of antibodies or high trough levels of anti-TNFα drug had the greatest chance of achieving clinical remission following anti-TNFα drug optimisation whereas those with low levels of anti-TNFα drug levels and high levels of antibodies had the lowest chance of achieving clinical remission.

8.10 Appendix 10 Quality appraisal of included management studies

Cochrane Collaboration's tool for assessing risk of bias for a randomised controlled trial (adapted from Higgins et al. (2011)[69])

First author surname and year of publication: Steenholdt 2014[122] and 2015[123]

Name of first reviewer: *Paul Sutcliffe* ***Name of second reviewer:*** *Martin Connock*

Domain	Description	Review authors' judgment
Selection bias: Sequence generation	The author's state: "Randomisation was performed centrally by an independent person (block randomisation in blocks of 20; sequentially numbered opaque envelopes)" Using a block size of 20 may not be appropriate. There are potential concerns about whether the allocation sequence was adequately generated	Unclear risk of bias
Selection bias: Allocation concealment	No further details are provided (see above). Allocation appears to be appropriately concealed	Low risk of bias
Performance bias: Blinding of participants, personnel	The author's state: "Patients were blinded to randomisation group and results of serum analyses. Physicians were blinded to IFX and IFX antibodies test results from patients in the IFX escalation group. Physicians were not completely blinded because they had to use the results of analyses of serum IFX and IFX antibodies in the treatment of those patients who were randomised to the algorithm group". Overall, the patient blinding appears appropriate and physician knowledge of the allocated intervention was acknowledged; physician knowledge in the algorithm arm has probably resulted in treatment selection not conforming to algorithm for a significant proportion of patients	High risk of bias
Detection bias: Blinding of outcome assessors	See above; no further details were provided related to blinding of outcome assessors	Unclear risk of bias
Attrition bias: Incomplete outcome data	The completeness of outcome data for each main outcome, including attrition and exclusions from the analysis was appropriate. Reasons for attrition/exclusions were reported. Incomplete outcome data appears to be adequately addressed	Low risk of bias
Reporting bias: Selective	The study appears to be free of any	Low risk of bias

reporting of the outcome, subgroups, or analysis	selective outcome reporting of outcome, subgroups, or analysis	
Other sources of bias: Funding source, adequacy of statistical methods used, type of analysis [ITT/PP], baseline imbalance in important characteristics	We note that 42% of patients were not treated in accordance with the algorithm resulting in patient's crossing over to treatment more similar to the "control" group	High risk of bias

Summary assessment of the risk of bias across domains (please highlight overall risk of bias rating)

Risk of bias across key domains	**Interpretation**	**Summary risk of bias**
Low risk of bias for all key domains	Plausible bias unlikely to seriously alter the results	Low risk of bias
Unclear risk of bias for one or more key domains	Plausible bias that raises some doubt about the results	Unclear risk of bias
High risk of bias for one or more key domains	**Plausible bias that seriously weakens confidence in the results**	**High risk of bias**

First author surname and year of publication: Vande Casteele 2015[72]

Name of first reviewer: *Paul Sutcliffe* ***Name of second reviewer:*** *Martin Connock*

Domain	Description	Review authors' judgement
Selection bias: Sequence generation	The author's state: "Randomization was performed by one person (VB) not in charge of the clinical care of patients using a computer-generated randomization schedule, with random block sizes". The range of block sizes is not presented	Low risk of bias
Selection bias: Allocation concealment	No further details are provided (see above). Allocation adequately appears to be appropriately concealed	Low risk of bias
Performance bias: Blinding of participants, personnel	The author's state: "Both patients and treating physicians were blinded to individual infliximab trough and ATI concentrations". This is unclear. No further information is provided; this limits our rating of whether the knowledge of the allocated intervention was adequately prevented during the study	Unclear risk of bias
Detection bias: Blinding of outcome assessors	The author's state: "Stable clinical response was assessed by the treating physician"; no further details were provided related to blinding of outcome assessors	Unclear risk of bias
Attrition bias: Incomplete outcome data	The completeness of outcome data for each main outcome, including attrition and exclusions from the analysis was appropriate. Patients who discontinued the optimization phase due to personal reasons (non-compliant to treatment algorithm or consent withdrawal) were described; these were excluded from the analysis. Attrition and exclusions were reported. Incomplete outcome data appears to be adequately addressed	Low risk of bias
Reporting bias: Selective reporting of the outcome, subgroups, or analysis	The study appears to be free of any selective outcome reporting	Low risk of bias
Other sources of bias: Funding source, adequacy of statistical methods used, type of analysis [ITT/PP], baseline imbalance in important characteristics	It is noted that the duration of the randomized maintenance phase was only one year which prevents the analysis of long term clinical and pharmaco-economical outcomes	Low risk of bias

Summary assessment of the risk of bias across domains (please highlight overall risk of bias rating)

Risk of bias across key domains	Interpretation	Summary risk of bias
Low risk of bias for all key domains	**Plausible bias unlikely to seriously alter the results**	**Low risk of bias**
Unclear risk of bias for one or more key domains	Plausible bias that raises some doubt about the results	Unclear risk of bias
High risk of bias for one or more key domains	Plausible bias that seriously weakens confidence in the results	High risk of bias

Downs and Black checklist[70] for non-randomised primary clinical studies

First author (year) study ID: Vaughn 2014[127]

Name of first reviewer: *Paul Sutcliffe* ***Name of second reviewer:*** *Martin Connock*

Reporting	**Rating**
1. Is the hypothesis/aim/objective of the study clearly described? **(Yes/No)**	Yes
2. Are the main outcomes to be measured clearly described in the Introduction or Methods section? **(Yes/No)** *If the main outcomes are first mentioned in the Results section, the question should be answered "No"*	Yes
3. Are the characteristics of the patients included in the study clearly described? **(Yes/No)** *In cohort studies and trials, inclusion and/or exclusion criteria should be given. In case-control studies, a case-definition and the source for controls should be given*	Yes
4. Are the interventions of interest clearly described? **(Yes/No)** *Treatments and placebo (where relevant) that are to be compared should be clearly described*	Yes
5. Are the distributions of principal confounders in each group of subjects to be compared clearly described? **(Yes/Partially/No)** *A list of principal confounders is provided*	Partially – no list of principal confounders
6. Are the main findings of the study clearly described? **(Yes/No)** *Simple outcome data (including denominators and numerators) should be reported for all major findings so that the reader can check the major analyses and conclusions (This question does not cover statistical tests which are considered below)*	Yes
7. Does the study provide estimates of the random variability in the data for the main outcomes? **(Yes/No)** *In non-normally distributed data the inter-quartile range of results should be reported. In normally distributed data the standard error, standard deviation or confidence intervals should be reported. If the distribution of the data is not described, it must be assumed that the estimates used were appropriate and the question should be answered "Yes"*	Yes
8. Have all important adverse events that may be a consequence of the intervention been reported? **(Yes/No)** *This should be answered "Yes" if the study demonstrates that there was a comprehensive attempt to measure adverse events. (A list of possible adverse events is provided)*	Yes
9. Have the characteristics of patients lost to follow-up been described? **(Yes/No)** *This should be answered "Yes" where there were no losses to follow-up or where losses to follow-up were so small that findings would be unaffected by their inclusion. This should be answered "No" where a study does not report the number of patients lost to follow-up*	Yes
10. Have actual probability values been reported (e.g., 0.035 rather than <0.05) for the main outcomes except where the probability value is less than 0.001? **(Yes/No)**	Yes
External validity	**Rating**
11. Were the subjects asked to participate in the study representative of the entire population from which they were recruited? **(Yes/No/Unable to determine)** *The study must identify the source population for patients and describe how the patients were selected. Patients would be representative if they comprised the entire source population, an unselected sample of consecutive patients, or a random sample. Random sampling is only feasible where a list of all members of the relevant*	Unable to determine – insufficient information is provided
12. Were those subjects who were prepared to participate representative of the entire population from which they were recruited? **(Yes/No/Unable to determine)** *The proportion of those asked who agreed should be stated. Validation that the sample was representative would include demonstrating that the distribution of the main confounding factors was the same in the study sample and the source population*	Unable to determine – insufficient information is provided
13. Were the staff, places, and facilities where the patients were treated, representative of the treatment the majority of patients receive? **(Yes/No/Unable to determine)** *For the question to be answered "Yes" the study should demonstrate that the intervention was representative of that in use in the source population. The question should be answered "No" if, for example, the intervention was undertaken in a specialist centre unrepresentative of the hospitals most of the source population would attend*	Unable to determine – insufficient information is provided
Internal validity – bias	**Rating**

14. Was an attempt made to blind study subjects to the intervention they have received? **(Yes/No/Unable to determine)** *For studies where the patients would have no way of knowing which intervention they received, this should be answered "Yes"*	No
15. Was an attempt made to blind those measuring the main outcomes of the intervention? **(Yes/No/Unable to determine)**	No
16. If any of the results of the study were based on "data dredging", was this made clear? **(Yes/No/Unable to determine)** *Any analyses that had not been planned at the outset of the study should be clearly indicated. If no retrospective unplanned subgroup analyses were reported, then answer "Yes"*	Yes – no data dredging
17. In trials and cohort studies, do the analyses adjust for different lengths of follow-up of patients, or in case-control studies, is the time period between the intervention and outcome the same for cases and controls? **(Yes/No/Unable to determine)** *Where follow-up was the same for all study patients the answer should "Yes". If different lengths of follow-up were adjusted for by, for example, survival analysis the answer should be "Yes". Studies where differences in follow-up are ignored should be answered "No"*	Yes
18. Were the statistical tests used to assess the main outcomes appropriate? **(Yes/No/Unable to determine)** *The statistical techniques used must be appropriate to the data. For example nonparametric methods should be used for small sample sizes. Where little statistical analysis has been undertaken but where there is no evidence of bias, the question should be answered "Yes". If the distribution of the data (normal or not) is not described it must be assumed that the estimates used were appropriate and the question should be answered "Yes"*	Yes
19. Was compliance with the intervention/s reliable? **(Yes/No/Unable to determine)** *Where there was non-compliance with the allocated treatment or where there was contamination of one group, the question should be answered "No". For studies where the effect of any misclassification was likely to bias any association to the null, the question should be answered "Yes"*	Yes
20. Were the main outcome measures used accurate valid and reliable? **(Yes/No/Unable to determine)** *For studies where the outcome measures are clearly described, the question should be answered "Yes". For studies which refer to other work or that demonstrates the outcome measures are accurate, the question should be answered as "Yes"*	Yes
Internal validity - confounding (selection bias)	**Rating**
21. Were the patients in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited from the same population? **(Yes/No/Unable to determine)** *For example, patients for all comparison groups should be selected from the same hospital. The question should be answered "Unable to determine" for cohort and case-control studies where there is no information concerning the source of patients included in the study*	Yes
22. Were study subjects in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited over the same period of time? **(Yes/No/Unable to determine)** *For a study which does not specify the time period over which patients were recruited, the question should be answered as "Unable to determine"*	Unable to determine – insufficient information is provided
23. Were the subjects randomised to intervention groups? **(Yes/No/Unable to determine)** *Studies which state that subjects were randomised should be answered "Yes" except where method of randomisation would not ensure random allocation. For example alternate allocation would score "No" because it is predictable*	No
24. Was the randomised intervention assignment concealed from both patients and health care staff until recruitment was complete and irrevocable? **(Yes/No/Unable to determine)** *All non-randomised studies should be answered "No". If assignment was concealed from patients but not from staff, it should be answered "No"*	No
25. Was there adequate adjustment for confounding in the analyses from which the main findings were drawn? **(Yes/No/Unable to determine)** *This question should be answered "No" for trials if: the main conclusions of the study were based on analyses of treatment rather than intention to treat; the distribution of known confounders in the different treatment groups was not described; or the distribution of known confounders differed between the treatment groups but was not taken into account in the analyses. In nonrandomised studies if the effect of the main confounders was not investigated or confounding was demonstrated but no adjustment was made in the final analyses the question should be answered as "No"*	No – However, the study reports continued use of IFX for 700 weeks in therapeutic concentration

	monitoring group for ~90% of patients; this seems implausible
26. Were losses of patients to follow-up taken into account? **(Yes/No/Unable to determine)** *If the numbers of patients lost to follow-up are not reported, the question should be answered as "Unable to determine". If the proportion lost to follow-up was too small to affect the main findings, the question should be answered "Yes"*	Yes
Power	**Rating**
27. Did the study have sufficient power to detect a clinically important effect where the probability value for a difference being due to chance is less than 5%? **(Yes/No/Unable to determine)***	No

QUADAS-2[68] tool with index questions adapted to the review for studies comparing performance of different tests: Steenholdt [121] and [122]

Name of first reviewer: Sian Taylor-Phillips Name of second reviewer: Martin Connock

Phase 1: State the review question

What is the level of concordance between the index tests and reference standard tests for measurement of drug and antibody levels?

Patients (setting, intended use of index test, presentation, prior testing): Crohn's disease patients (adults and children) receiving infliximab or adalimumab, either whose disease responds to treatment with TNF inhibitor, or who experience secondary loss of response during maintenance treatment with TNF inhibitor.
Index test(s): ELISA (LISA-TRACKER or Promonitor or Immundiagnostik)
Reference standard: Spiked drug levels. Where this is not available tests for which we have a prospective link to outcomes using a pre-specified algorithm may be used (these are HPLC, RIA, Prometheus ELISA, or Leuven in-house ELISA).

Phase 2: Draw a flow diagram for the primary study

Phase 3: Risk of bias and applicability judgements

Domain 1: Patient selection

A. Risk of bias	
Describe methods of patient selection: Study[121] included 66 CD patients with secondary loss of response to infliximab, which were all part of an RCT. The RCT paper[122] described 69 patients, recruited from six Danish centres. Inclusion criteria stated but not selection method	
Was a consecutive or random sample of patients enrolled?	Unclear
Was a case-control design avoided?	Yes
Did the study avoid inappropriate exclusions?	No
Could the selection of patients have introduced bias? **Risk: High**	
B. Concerns regarding applicability	
Describe included patients (prior testing, presentation, intended use of intervention test and setting): CD patients with secondary loss of response to infliximab Range of drug / antibody concentrations:	
Is there concern that the included patients or range of drug / antibody concentrations do not match the review question?	

Concern: Low

Domain 2: Index test(s)

A. Risk of bias	
Describe the intervention test and how it was conducted and interpreted: HMSA, Prometheus ELISA and RGA.	
Were the number of failed results and measurement repeats reported?	No
Was the threshold pre-specified?	Yes
Were index tests interpreted without knowledge of reference standard?	Yes
Could the conduct or interpretation of the intervention test have introduced bias? **Risk: Low**	
B. Concerns regarding applicability They are comparing presence of drug at limit of quantisation of each test, rather than therapeutic drug levels. The anti-drug part does not differ from the review question. Describe the preparation and storage of the sample before the intervention test was applied:	
Is there concern that the intervention test, its conduct, or interpretation differ from the review question? **Concern: High**	

Domain 3: Reference standard

A. Risk of bias	
Describe the reference standard and how it was conducted and interpreted: Radioimmunoassay on samples stored at room temperature.	
	Yes
	No
Could the comparison test, its conduct, or its interpretation have introduced bias? **Risk: High**	
B. Concerns regarding applicability	

Same test and threshold as use in RCT.
Is there concern that the comparison test does not match that used in studies assessing the link to outcomes? **Concern: Low**

Domain 4: Flow and timing

A. Risk of bias	
Describe any patients who did not receive the intervention test and/or comparison test(s) or who were excluded from the correlation calculations: Three patients received the reference standard in the RCT, but were not given the index test. This is not described in this paper. Describe the time interval and any interventions between intervention test and comparison test(s): Comparator was conducted on samples stored at room temperature at the time, index tests were performed on frozen samples at a later stage.	
Was there an appropriate interval between intervention test and comparison test(s)?	No
Were both intervention test and reference standard conducted on all samples?	No
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	No
Could the patient flow have introduced bias? **Risk: High**	

QUADAS-2[68] tool with index questions adapted to the review for studies comparing performance of different tests: Vande Casteele 2013[125]

Name of first reviewer: Sian Taylor-Phillips Name of second reviewer: Martin Connock

Phase 1: State the review question

What is the level of concordance between the index tests and reference standard tests for measurement of drug and antibody levels?

Patients (setting, intended use of index test, presentation, prior testing): Crohn's disease patients (adults and children) receiving infliximab or adalimumab, either whose disease responds to treatment with TNF inhibitor, or who experience secondary loss of response during maintenance treatment with TNF inhibitor.
Index test(s): ELISA (LISA-TRACKER or Promonitor or Immundiagnostik)
Reference standard: Spiked drug levels. Where this is not available tests for which we have a prospective link to outcomes using a pre-specified algorithm may be used (these are HPLC, RIA, Prometheus ELISA, or Leuven in-house ELISA).

Phase 2: Draw a flow diagram for the primary study

Phase 3: Risk of bias and applicability judgements

QUADAS-2 is structured so that four key domains are each rated in terms of the risk of bias and the concern regarding applicability to the review question (as stated in Phase 1). Each key domain has a set of signalling questions to help reach the judgements regarding bias and applicability.

Domain 1: Patient selection

A. Risk of bias	
Describe methods of patient selection: Selected from biobank based on index test results for antidrug levels	
Was a consecutive or random sample of patients enrolled?	No
Was a case-control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Unclear
Could the selection of patients have introduced bias? **Risk: High**	
B. Concerns regarding applicability	
Describe included patients (prior testing, presentation, intended use of intervention test and setting): Crohns and UC patients selected on basis of index test results.	

Range of drug / antibody concentrations:
Is there concern that the included patients or range of drug / antibody concentrations do not match the review question? **Concern: High**

Domain 2: Index test(s)

A. Risk of bias	
Describe the intervention test and how it was conducted and interpreted: Leuven in-house ELISA administered in same manner as described in VC 2012	
Were the number of failed results and measurement repeats reported?	No
Was the threshold pre-specified?	Yes
Were index tests interpreted without knowledge of reference standard?	Yes
Could the conduct or interpretation of the intervention test have introduced bias? **Risk: Low**	
B. Concerns regarding applicability Describe the preparation and storage of the sample before the intervention test was applied: Tested at point of trough levels, no further details given	
Is there concern that the intervention test, its conduct, or interpretation differ from the review question? **Concern: Low**	

Domain 3: Reference standard

A. Risk of bias	
Describe the comparison test and how it was conducted and interpreted: HMSA at Prometheus labs from biobanked samples	
Is the comparison test likely to correctly classify the target condition? (only matters if doing more than correlation studies)	No
	Unclear
Could the comparison test, its conduct, or its interpretation have introduced bias? **Risk: High**	
B. Concerns regarding applicability	

Is there concern that the comparison test does not match that used in studies assessing the link to outcomes? Correct HMSA test but using threshold of 7.95 U/ml, authors suggest it subsequently changed to 3.13 U/ml **Concern: Low**

Domain 4: Flow and timing

A. Risk of bias	
Describe any patients who did not receive the intervention test and/or comparison test(s) or who were excluded from the correlation calculations: Unclear. Describe the time interval and any interventions between intervention test and comparison test(s): HMSA was from biobanked samples, Leuven ELISA was conducted at the time	
Was there an appropriate interval between intervention test and comparison test(s)? (ideally conducted at same time so samples cant deteriorate)	No
Were both intervention test and reference standard conducted on all samples?	Unclear
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Unclear
Could the patient flow have introduced bias? **Risk: High**	

QUADAS-2[68] tool with index questions adapted to the review for studies comparing performance of different tests: Vande Casteele 2012[66]

Name of first reviewer: Sian Taylor-Phillips Name of second reviewer: Martin Connock

Phase 1: State the review question

What is the level of concordance between the index tests and reference standard tests for measurement of drug and antibody levels?

Patients (setting, intended use of index test, presentation, prior testing): Crohn's disease patients (adults and children) receiving infliximab or adalimumab, either whose disease responds to treatment with TNF inhibitor, or who experience secondary loss of response during maintenance treatment with TNF inhibitor.
Index test(s): ELISA (LISA-TRACKER or Promonitor or Immundiagnostik)
Reference standard: Spiked drug levels. Where this is not available tests for which we have a prospective link to outcomes using a pre-specified algorithm may be used (these are HPLC, RIA, Prometheus ELISA, or Leuven in-house ELISA).

Phase 2: Draw a flow diagram for the primary study

Phase 3: Risk of bias and applicability judgements

QUADAS-2 is structured so that four key domains are each rated in terms of the risk of bias and the concern regarding applicability to the review question (as stated in Phase 1). Each key domain has a set of signalling questions to help reach the judgements regarding bias and applicability.

Domain 1: Patient selection

A. Risk of bias	
Describe methods of patient selection: Unclear. Combination of spiked samples and samples from departments of gastroenterology and rheumatology.	
Was a consecutive or random sample of patients enrolled?	No
Was a case-control design avoided?	Yes
Did the study avoid inappropriate exclusions?	Unclear
Could the selection of patients have introduced bias? **Risk: High**	
B. Concerns regarding applicability	
Describe included patients (prior testing, presentation, intended use of intervention test and setting): No details given except departments of origin. Those from rheumatology may not be applicable, those	

from gastroenterology with diseases other than CD may not be applicable. Range of drug / antibody concentrations:
Is there concern that the included patients or range of drug / antibody concentrations do not match the review question? **Concern: High**

Domain 2: Index test(s)

A. Risk of bias	
Describe the intervention test and how it was conducted and interpreted: LISA TRACKER according to manufacturer's guidelines	
Were the number of failed results and measurement repeats reported?	Yes
Was the threshold pre-specified?	Yes
Were index tests interpreted without knowledge of reference standard?	Yes
Could the conduct or interpretation of the intervention test have introduced bias? **Risk: Low**	
B. Concerns regarding applicability Describe the preparation and storage of the sample before the intervention test was applied: Unclear	
Is there concern that the intervention test, its conduct, or interpretation differ from the review question? **Concern: Low**	

Domain 3: Reference standard

A. Risk of bias	
Describe the comparison test and how it was conducted and interpreted: Leuven in-house ELISA.	
Is the comparison test likely to correctly classify the target condition? (only matters if doing more than correlation studies)	No
	Yes
Could the comparison test, its conduct, or its interpretation have introduced bias? **Risk: High**	
B. Concerns regarding applicability	

Is there concern that the comparison test does not match that used in studies assessing the link to outcomes? **Concern: Low**

Domain 4: Flow and timing

A. Risk of bias	
Describe any patients who did not receive the intervention test and/or comparison test(s) or who were excluded from the correlation calculations: Describe the time interval and any interventions between intervention test and comparison test(s):	
Was there an appropriate interval between intervention test and comparison test(s)? (ideally conducted at same time so samples cant deteriorate)	Unclear
Were both intervention test and reference standard conducted on all samples?	Yes
Did patients receive the same reference standard?	No
Were all patients included in the analysis?	No
Could the patient flow have introduced bias? **Risk: High**	

QUADAS-2[68] tool with index questions adapted to the review for studies comparing performance of different tests: Wang 2012[130]

Name of first reviewer: Sian Taylor-Phillips Name of second reviewer: Martin Connock

Phase 1: State the review question

What is the level of concordance between the index tests and reference standard tests for measurement of drug and antibody levels?

Patients (setting, intended use of index test, presentation, prior testing): Crohn's disease patients (adults and children) receiving infliximab or adalimumab, either whose disease responds to treatment with TNF inhibitor, or who experience secondary loss of response during maintenance treatment with TNF inhibitor.
Index test(s): ELISA (LISA-TRACKER or Promonitor or Immundiagnostik)
Reference standard: Spiked drug levels. Where this is not available tests for which we have a prospective link to outcomes using a pre-specified algorithm may be used (these are HPLC, RIA, Prometheus ELISA, or Leuven in-house ELISA).

Phase 2: Draw a flow diagram for the primary study

Phase 3: Risk of bias and applicability judgements

Domain 1: Patient selection

A. Risk of bias	
Describe methods of patient selection: Controls were from blood bank donors in California, and cases were left over blood samples from tests carried out at Prometheus Laboratories No information is given on how they selected samples from these sources.	
Was a consecutive or random sample of patients enrolled?	No
Was a case-control design avoided?	No
Did the study avoid inappropriate exclusions?	No
Could the selection of patients have introduced bias? **Risk: High**	
B. Concerns regarding applicability	
Describe included patients (prior testing, presentation, intended use of intervention test and setting): 100 inflammatory bowel patients as defined by index test and 100 healthy controls. No details of split between CD and Ulcerative colitis. Range of drug / antibody concentrations:	

Is there concern that the included patients or range of drug / antibody concentrations do not match the review question? **Concern: High**

Domain 2: Index test(s)

A. Risk of bias	
Describe the intervention test and how it was conducted and interpreted: Prometheus bridging ELISA. Threshold not specified at all.	
Were the number of failed results and measurement repeats reported?	No
Was the threshold pre-specified?	Unclear
Were index tests interpreted without knowledge of reference standard?	Yes
Could the conduct or interpretation of the intervention test have introduced bias? **Risk: High**	
B. Concerns regarding applicability Threshold not given so applicability unclear. Describe the preparation and storage of the sample before the intervention test was applied:	
Is there concern that the intervention test, its conduct, or interpretation differ from the review question? **Concern: High**	

Domain 3: Reference standard

A. Risk of bias	
Describe the reference standard and how it was conducted and interpreted: HMSA at Prometheus Labs.	
	Unclear
	Unclear
Could the comparison test, its conduct, or its interpretation have introduced bias? **Risk: Unclear**	
B. Concerns regarding applicability	
Is there concern that the comparison test does not match that used in studies assessing the link	

to outcomes? **Concern: Low**

Domain 4: Flow and timing

A. Risk of bias	
Describe any patients who did not receive the intervention test and/or comparison test(s) or who were excluded from the correlation calculations: This is unclear from the report. Describe the time interval and any interventions between intervention test and comparison test(s):	
Was there an appropriate interval between intervention test and comparison test(s)?	Unclear
Were both intervention test and reference standard conducted on all samples?	Unclear
Did patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Unclear
Could the patient flow have introduced bias? **Risk: High**	

8.11 Appendix 11 Parametric modelling for Vaughn and TAXIT

Parametric models were fit to reconstructed IPD of time to treatment failure for proactive drug monitoring and control patients in remission who commenced maintenance infliximab at the start of 2009.[127] This was done so that treatment failure could be modelled to 10 years (the time horizon of the economnic model) with potentiall for use in the model.

According to AIC and BIC information criteria best fit to the data was provided by lognormal loglogistic and Weibull models (Table 46); A gamma model could not be fit for the standard care arm.

Table 46 AIC and BIC values for parametric models for time to treatment failure

Model	Obs	ll(model)	df	AIC	BIC
Standard care arm					
exponential	68	-54.9247	1	111.8493	114.0688
weibull	68	-52.2516	2	108.5031	112.9422
gompertz	68	-53.618	2	111.2359	115.6749
lognormal	68	-51.55	2	107.0999	111.5389
loglogistic	68	-51.7864	2	107.5728	112.0118
Proactive drug monitoring arm					
exponential	39	-19.2462	1	40.49236	42.15592
weibull	39	-19.2257	2	42.45134	45.77846
gompertz	39	-19.2243	2	42.44856	45.77569
lognormal	39	-18.8709	2	41.74174	45.06886
loglogistic	39	-19.1771	2	42.35412	45.68124

Lognormal and Weibull models are shown in Figure 36.

TAXIT TRIAL time to relapse.

Parametric modelling of time to relapse based on the TAXIT study is shown below in Figure 37. Again this was done because of potential relevance to the economic model.

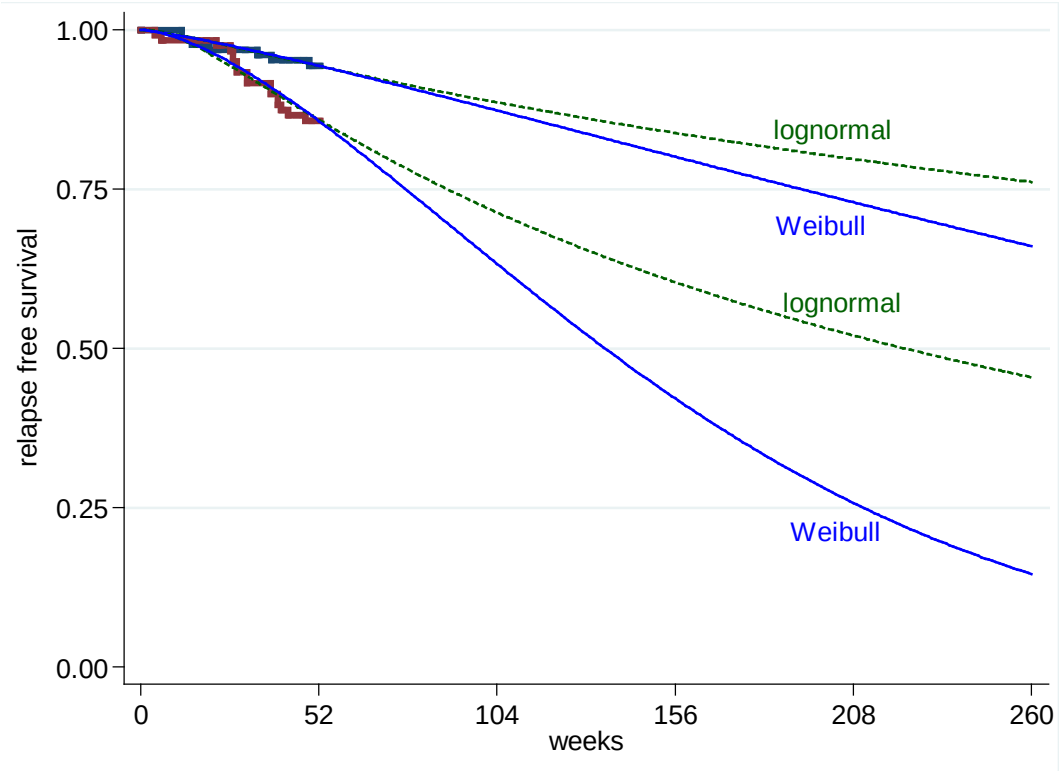


Figure 37 Parametric modelling of time to relapse based on the TAXIT study

8.12 Appendix 12 Meta-analysis results

Correlation studies that permitted extraction of a two by two table for test result (positive or negative and clinical status (response or loss/lack of response) were carried forward for hierarchical meta-analysis. The major features of these studies are summarised in Table 47. Forest plots of sensitivity and specificity for prediction of loss or lack response and summary ROC plots are presented below according to the test applied.

Table 47 Major features of studies included for hierarchical meta-analysis

STUDY	DRUG	DIAGNOSIS	RESP/2L	TEST	RES-DEF
IFX trough level as predictor of loss or lack of response					
Ainsworth 2008[46]	IFX	CD	2L	RIA	PJ
Ben-Basset 2013[77] abs	IFX	IBD ~.93 CD	Resp	HMSA	HBI
Bortlik 2013[81]	IFX	CD	Resp	ELISA	PJ
Cornillie 2014[84]	IFX	CD	Resp	ELISA	CDAI
Hibi 2014[97]	IFX	CD	Resp	ELISA	CDAI
Imaeda 2012[98]	IFX	CD	Resp	ELISA	CDAI
Kopylov 2012[102]	IFX	CD	Resp	ELISA	PJ
Maser 2006[37]	IFX	CD	Resp	Unclear	HBI
Steenholdt 2011[119]	IFX	CD	Resp	RIA	PJ
Steenholdt 2014[122]	IFX	CD	2L	RIA	CDAI
Yanai 2012[133] abs	IFX	CD	Resp	ELISA	PJ
Trough antibodies to IFX as predictor of loss or lack of response					
Ainsworth 2008[46]	IFX	CD	2L	RIA	PJ
Baert 2014[76]	IFX	IBD ~0.8 CD	2L	HMSA	PJ
Ben-Horin 2011[78]	IFX	IBD ~.82 CD	Resp	NR	ST
Ben-Horin 2012[79]	IFX ADA	IBD ~0.9 CD	2L	ELISA	PJ
Bodini 2014[80] abs	IFX	CD	Resp	HMSA	HBI
Candon 2006[82]	IFX	CD	2L	ELISA]	UC
Dauer 2013[87] abs	IFX	CD ~.83 CD	Resp	NR	PJ
Farrell 2003[91]	IFX	CD	Resp	ELISA	PJ
Hanauer 2004[39]	IFX	CD	Resp	ELISA	CDAI
Imaeda 2012[98]	IFX	CD	Resp	ELISA	CDAI

Kong[101] abs	IFX	IBD ~.83 CD	Resp	ELISA	PJ
Kopylov 2012[102]	IFX	CD	Resp	ELISA	PJ
Marzo[105]	IFX	NR	Resp	ELISA	CDAI
Nagore[109] abs	IFX	IBD ~.86 CD	Resp	ELISA	PJ
Pariente 2012[58]	IFX	CD & UC	2L	ELISA	PJ or HBI
Steenholdt 2011[119]	IFX	CD	Resp	RIA	PJ ST
Steenholdt 2013[51]	IFX	CD	Resp	ELISA	PJ
Steenholdt 2014[122]	IFX	CD	2L	RIA	CDAI
Vande Casteele 2013[125]	IFX	IBD ~.70 CD	2L	HMSA	CRP TC
Vande Casteele 2013[125]	IFX	IBD ~.70 CD	Resp	HMSA	CRP TC
Adalimumab trough level as predictor of loss or lack of response					
Chiu 2013[83]	ADA	CD	2L	ELISA	CDAI
Frederiksen 2014[93]	ADA	IBD	Resp	RIA	PJ BM
Imaeda 2014[99]	ADA	CD	Resp	ELISA	CRP
Mazor 2014[107]	ADA	CD	Resp	ELISA	PJ
Roblin 2014[114]	ADA	IBD ~.55 CD	Resp	ELISA	CDAI
Trough antibodies to adalimumab as predictor of loss or lack of response					
Frederiksen 2014[93]	ADA	IBD	Resp	RIA	PJ BM
Imaeda 2014[99]	ADA	CD	Resp	ELISA	CRP
Mazor 2014[107]	ADA	CD	Resp	ELISA	PJ
West 2008[132]	ADA	CD	Resp	RIA	PJ
Ben-Horin 2012[79]	IFX ADA	IBD ~0.9 CD	2L	ELISA	SA
Roblin 2014[114]	ADA	CD	Resp	ELISA	CDAI

DIAGNOSIS = study patient population; 2L = patients with secondary loss of response ; RESP = responding patients; RES-DEF = method used for defining clinical response; ADA = adalimumab; IFX = infliximab; CD = Crohn's disease; IBD = inflammatory bowel disease; ELISA = enzyme linked immunoassay; RIA = radioimmunoassay; CDAI = Crohn's disease activity index score; CRP = C reactive protein level; PJ = physicians' judgement ; PJ BM = physicians' judgement and biological measure ; HBI = Harvey Bradshaw Index score. SA = switch anti-TNF; ST = stop anti-TNF.

Appendix 12.1 Infliximab trough level tests for loss of response or lack of regaining response

Eleven studies were included, of which two were reported as abstracts (Ben Basset et al., 2013[77] and Yanai et al., 2012[133]). Sensitivity and specificity pairs are summarised in Figure 38.

Study	TP	FP	FN	TN	assay	RES	POP	Sensitivity (95% CI)	Specificity (95% CI)
Ainsworth 2008	7	1	1	18	RIA	PJ	2L	0.88 [0.47, 1.00]	0.95 [0.74, 1.00]
Ben-Bassat 2013	50	29	10	145	HMSA	HBI	R	0.83 [0.71, 0.92]	0.83 [0.77, 0.89]
Bortlik 2013	16	23	7	38	ELISA	PJ	R	0.70 [0.47, 0.87]	0.62 [0.49, 0.74]
Cornillie 2014	14	11	8	38	ELISA	CDAI	R	0.64 [0.41, 0.83]	0.78 [0.63, 0.88]
Hibi 2014	8	4	7	22	ELISA	CDAI	R	0.53 [0.27, 0.79]	0.85 [0.65, 0.96]
Imaeda 2012	9	8	8	33	ELISA	CDAI	R	0.53 [0.28, 0.77]	0.80 [0.65, 0.91]
Kopylov 2012	21	2	9	31	ELISA	PJ	R	0.70 [0.51, 0.85]	0.94 [0.80, 0.99]
Maser 2006	22	11	13	44	UC	HBI	R	0.63 [0.45, 0.79]	0.80 [0.67, 0.90]
Steenholdt 2011	18	7	3	41	RIA	PJ	R	0.86 [0.64, 0.97]	0.85 [0.72, 0.94]
Steenholdt 2014	10	7	21	31	RIA	CDAI	2L	0.32 [0.17, 0.51]	0.82 [0.66, 0.92]
Yanai 2012	7	10	7	16	ELISA	PJ	R	0.50 [0.23, 0.77]	0.62 [0.41, 0.80]

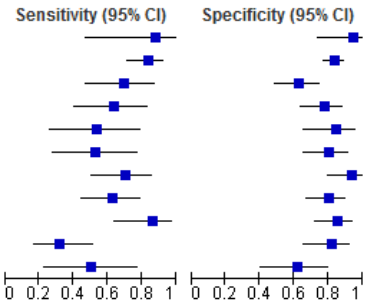


Figure 38 Trough infliximab for predicting LOR or failure to regain response
RES = method for estimating clinical response, POP = study patient population, RIA = radioimmunoassay, HMSA = homogeneous mobility shift assay, ELISA = enzyme linked immunoassay, UC = unclear, PJ = physicians' judgement, HBI = Harvey Bradshaw Index score, CDAI = Crohn's disease activity index score, R= responders, 2L = loss of response

Hierarchical meta-analysis yielded the test accuracy results summarised in Table 48 and Figure 39. Subgroup analyses examining responder populations only, and ELISA test studies only, had little effect on pooled estimates.

Table 48 Test accuracy results from hierarchical meta-analysis

Studies included	parameter	Point estimate	95% LCI	95% UCI
all 11 studies	Sens	0.657232	0.546288	0.753299
all 11 studies	Spec	0.80625	0.744166	0.85618
all 11 studies	DOR	7.978975	4.119972	15.45254
all 11 studies	LR+	3.392169	2.35152	4.893351
all 11 studies	LR-	0.425139	0.305104	0.592398
all 11 studies	1/LR-	2.352175	1.688056	3.277573
responder populations only	Sens	0.681452	0.592117	0.759178
responder populations only	Spec	0.790873	0.723301	0.845468
responder populations only	DOR	8.090128	4.353039	15.03551
responder populations only	LR+	3.258549	2.287802	4.641198
responder populations only	LR-	0.402781	0.298559	0.543385
responder populations only	1/LR-	2.482739	1.840315	3.349423
ELISA studies only	Sens	0.652104	0.564027	0.730877
ELISA studies only	Spec	0.789041	0.691592	0.861849
ELISA studies only	DOR	7.010794	3.450232	14.24578
ELISA studies only	LR+	3.091133	1.959085	4.877331
ELISA studies only	LR-	0.440911	0.329778	0.589495
ELISA studies only	1/LR-	2.268033	1.696367	3.032348

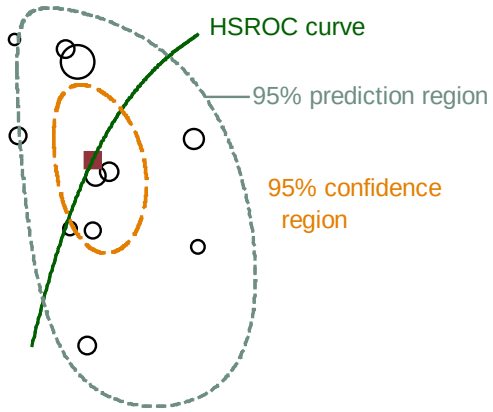


Figure 39 Trough infliximab levels for predicting LOR; hierarchical meta-analysis of test accuracy
Left = all 11 studies, right = responder studies only. The red square represents the summary point estimate on the HSROC curve

The random effects pooled estimate for the prevalence of loss or lack of response was 0.335 (95% CI: 0.289 to 0.382). If responder populations only were considered this changed slightly to 0.325 (95% CI: 0.278 to 0.372). Given the meta-analysis values sensitivity specificity and prevalence (P) values the point estimate for the probability of positive and negative test results is as shown in Table 49.

Table 49 Probability of a positive and negative test result (range based 95% CI prevalence)

Probability of positive test result	[P * Sens] + ([1-P]*[1-Spec])	0.349 (0.328 to 0.371)
Probability of negative test result	([1-P] * Spec) + (P*[1-Sens])	0.651 (0.629 to 0.672)

Appendix 12.2 Antibodies to infliximab tests for loss of response or lack of regaining response

Twenty studies were included, of which five were reported as abstracts (Bodini et al., 2014,[80] Dauer et al., 2013,[87] Kong et al., 2011,[101] Marzo et al., 2014[105] and Nagore et al., 2015[109]). Sensitivity and specificity pairs are summarised in Figure 40.

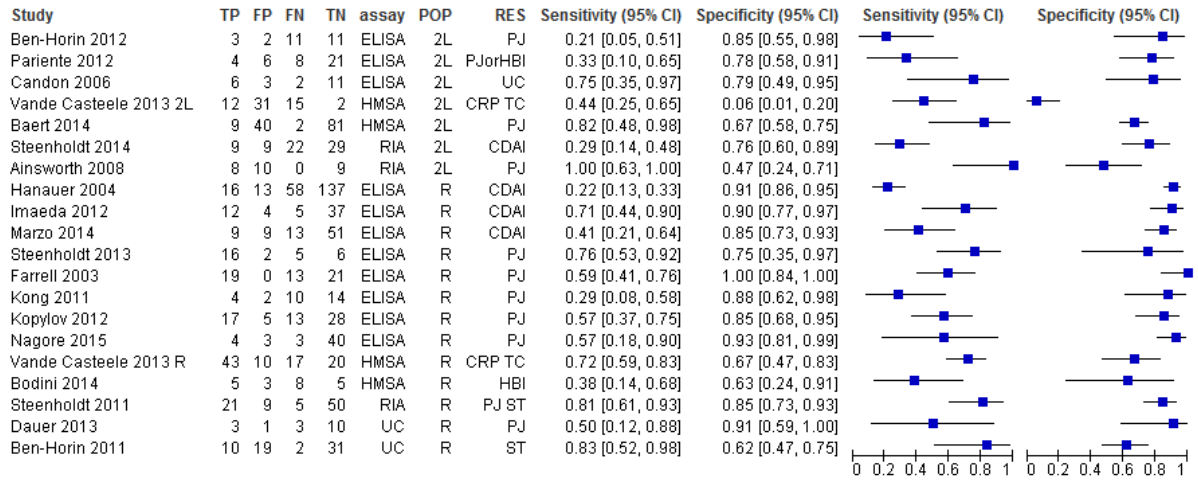


Figure 40 Sensitivity and specificity of tests of antibodies to infliximab for prediction of LOR or failure to regain response

POP = study patient population, RES = method for estimating clinical response, RIA = radioimmunoassay, HMSA = homogeneous mobility shift assay, ELISA = enzyme linked immunoassay, UC = unclear, PJ = physicians' judgement, HBI = Harvey Bradshaw Index score, CDAI = Crohn's disease activity index score, TC treatment change, ST = stop IFX therapy, CRP = CRP level, R= responders, 2L = loss of response. Note: Bodini, Dauer, Kong, Marzo and Nagore were published in brief as abstracts or conference proceedings.

Hierarchical meta-analysis yielded test accuracy results summarised in Table 50 and Figure 41.

Subgroup analyses removing two outlier studies, examining responder populations only, and ELISA test studies only, had little effect on pooled summary point estimates.

Table 50 Test accuracy results from hierarchical meta-analysis

Studies included	parameter	Point estimate	95% LCI	95% UCI
all 20 studies	Sens	0.559745	0.444812	0.668611
all 20 studies	Spec	0.792243	0.688105	0.868267
all 20 studies	DOR	4.848283	2.519589	9.329239
all 20 studies	LR+	2.694226	1.72293	4.213088
all 20 studies	LR-	0.555707	0.426575	0.72393
all 20 studies	1/LR-	1.799509	1.38135	2.344251
all studies minus outliers	Sens	0.597	0.477	0.707
all studies minus outliers	Spec	0.807	0.742	0.859
all studies minus outliers	DOR	6.183	3.805	10.050

all studies minus outliers	LR+	3.088	2.311	4.127
all studies minus outliers	LR-	0.500	0.381	0.655
all studies minus outliers	1/LR-	2.002	1.528	2.623
responder populations only	Sens	0.570	0.445	0.687
responder populations only	Spec	0.849	0.787	0.896
responder populations only	DOR	7.460	4.544	12.250
responder populations only	LR+	3.778	2.722	5.244
responder populations only	LR-	0.506	0.388	0.660
responder populations only	1/LR-	1.974	1.514	2.574
ELISA studies only	Sens	0.482	0.355	0.611
ELISA studies only	Spec	0.880	0.841	0.911
ELISA studies only	DOR	6.830	3.872	12.050
ELISA studies only	LR+	4.022	2.805	5.768
ELISA studies only	LR-	0.589	0.459	0.755
ELISA studies only	1/LR-	1.698	1.324	2.178

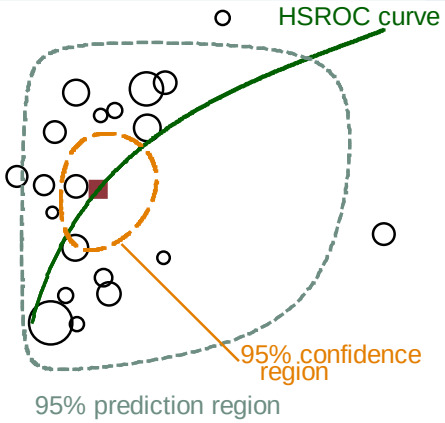


Figure 41 Antibodies to infliximab for predicting LOR; hierarchical meta-analysis of test accuracy
Left upper= all 20 studies, right upper = ELISA studies only, left lower = all studies excluding two influential outliers, right lower = responder populations only. The red square represents the summary point estimate on the HSROC curve.

The random effects pooled estimate for the prevalence of lack of response was 0.390 (95% CI: 0.302 to 0.477). If responder populations only were considered this changed slightly 0.411 (95% CI: 0.312 to 0.511). Given the meta-analysis values sensitivity specificity and prevalence values the point estimate for the probability of positive and negative test results is as shown in Table 51.

Table 51 Probability of a positive and negative test result, all studies (range based 95% CI prevalence)

Probability of positive test result	[P * Sens] + ([1-P]*[1-Spec])	0.345 (0.324 to 0.365)
Probability negative test result	([1-P] * Spec) + (P*[1-Sens])	0.655 (0.635 to 0.686)

Appendix 12.3 Adalimumab trough level test for loss of response or lack of regaining response

Four studies of responders were included. The study of Roblin et al. (2014)[114] included 18 UC and 22 CD patients. Mazor et al. (2014)[107] reported results by test rather than by patients (there were 118 tests in 71 patients; authors stated using the first test result for each patient did not alter the results). Sensitivity and specificity pairs are summarised in Figure 42.

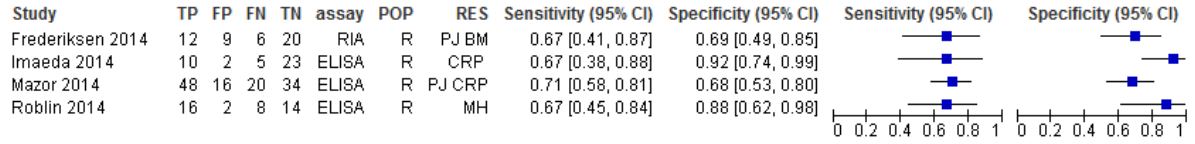


Figure 42 Trough adalimumab for predicting LOR in responders
RES = method for estimating clinical response, POP = study patient population, RIA = radioimmunoassay, ELISA = enzyme linked immunoassay, CRP = CRP level > 3mg/mL, PJ CRP = physicians' judgement and CRP level, PJ BM = physicians' judgement and biological measure, R= responders

A single study of patients with loss of response was identified (Chiu 2013;[83] as shown in Figure 43, this study appeared to be an outlier and meta-analysis was restricted to responder populations.

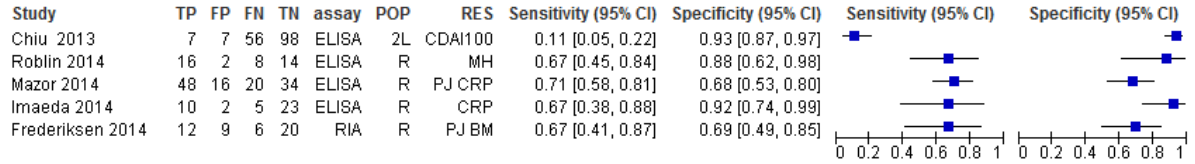


Figure 43 Trough adalimumab for predicting LOR or failure to regain response
RES = method for estimating clinical response, POP = study patient population, RIA = radioimmunoassay, ELISA = enzyme linked immunoassay, CDAI100 = CDAI score reduction of 100, CRP = CRP level > 3mg/mL, PJ CRP = physicians' judgement and CRP level, PJ BM = physicians' judgement and biological measure, R= responders, 2L = loss of response

Hierarchical meta-analysis yielded test accuracy results summarised in Table 52 and Figure 44.

Table 52 Test accuracy results from hierarchical meta-analysis (4 studies)

Parameter	Point estimate	95% LCI	95% UCI
Sens	0.684	0.591	0.764
Spec	0.786	0.643	0.883
DOR	7.971	3.646	17.428
LR+	3.201	1.822	5.623
LR-	0.402	0.297	0.542
1/LR-	2.490	1.844	3.363

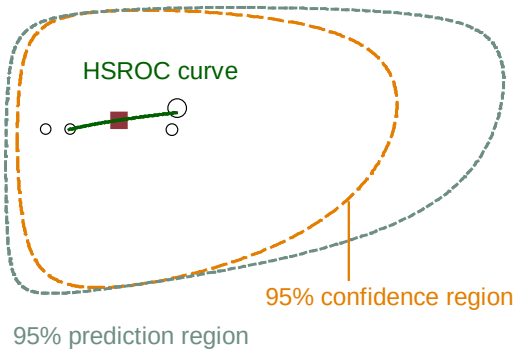


Figure 44 Trough adalimumab levels for predicting LOR; hierarchical meta-analysis of test accuracy
(only responder studies are included)

The random effects pooled estimate for the prevalence of lack of response was 0.489 (95% CI: 0.372 to 0.606); this is likely to be an over estimate due to double counting patients from the Mazor et al. (2014)[107] study. Given the meta-analysis values sensitivity specificity and prevalence values the point estimate for the probability of positive and negative test results is 0.444 (range 0.389 to .499) and 0.556 (range 0.501 to 0.611) respectively.

Appendix 12.4 Antibodies to adalimumab as test for loss of response or lack of regaining response

Six studies of responders or secondary starters were included. Mazor et al. (2014)[107] reported results by test rather than by patients (there were 118 tests in 71 patients; authors stated using the first test result for each patient did not alter the results). Sensitivity and specificity pairs are summarised in Figure 45.

Figure 45 Antibodies to adalimumab for predicting LOR
RES = method for estimating clinical response, POP = study patient population, RIA = radioimmunoassay, ELISA = enzyme linked immunoassay, CDAI = CDAI score, CRP = CRP level, PJ BM = physicians' judgement and biological measure , SA = stop anti-TNF, R= responders, RS = restarters on adalimumab

Hierarchical meta-analysis yielded test accuracy results summarised in Table 53 and Figure 46.

Table 53 Test accuracy results from hierarchical meta-analysis (5 studies)

Parameter	Point estimate	95% LCI	95% UCI
Sens	0.471206	0.2903357	0.66
Spec	0.915467	0.7939073	0.968
DOR	9.65022	4.387759	21.22
LR+	5.574189	2.646268	11.74
LR-	0.577623	0.4208713	0.793
1/LR-	1.731233	1.261422	2.376

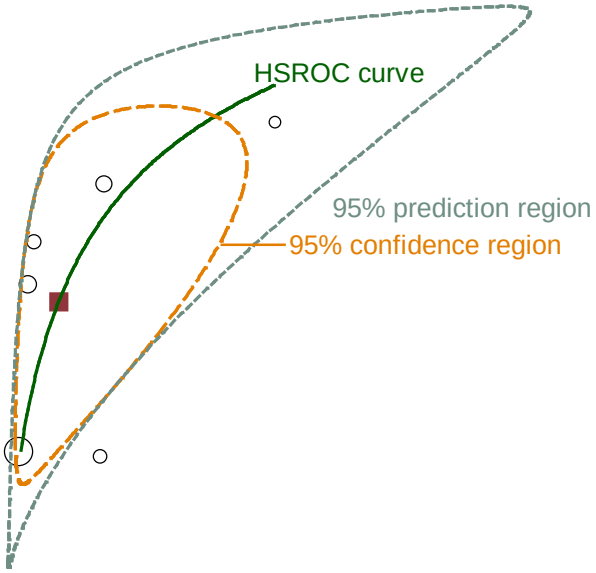


Figure 46 Antibodies to adalimumab for predicting LOR; hierarchical meta-analysis of test accuracy

The random effects pooled estimate for the prevalence of LOR was 0.435 (95% CI: 0.330 to 0.540); this is likely to be an over estimate due to double counting patients from the Mazor et al. (2014)[107] study. Given the meta-analysis values sensitivity specificity and prevalence values the point estimate for the probability of positive and negative test results is 0.253 (range 0.212 to 0.293) and 0.747 (range 0.707 to 0.788) respectively.

Appendix 12.5 Predictive values for drug and anti-drug antibodies tests for LOR or failure to regain response.

Figure 47 summarises PPVs and NPVs according to prevalence of the clinical state of interest. The dashed vertical lines indicate the pooled prevalence and 95% CI and what is probably a meaningful clinical range across which the tests might be employed.

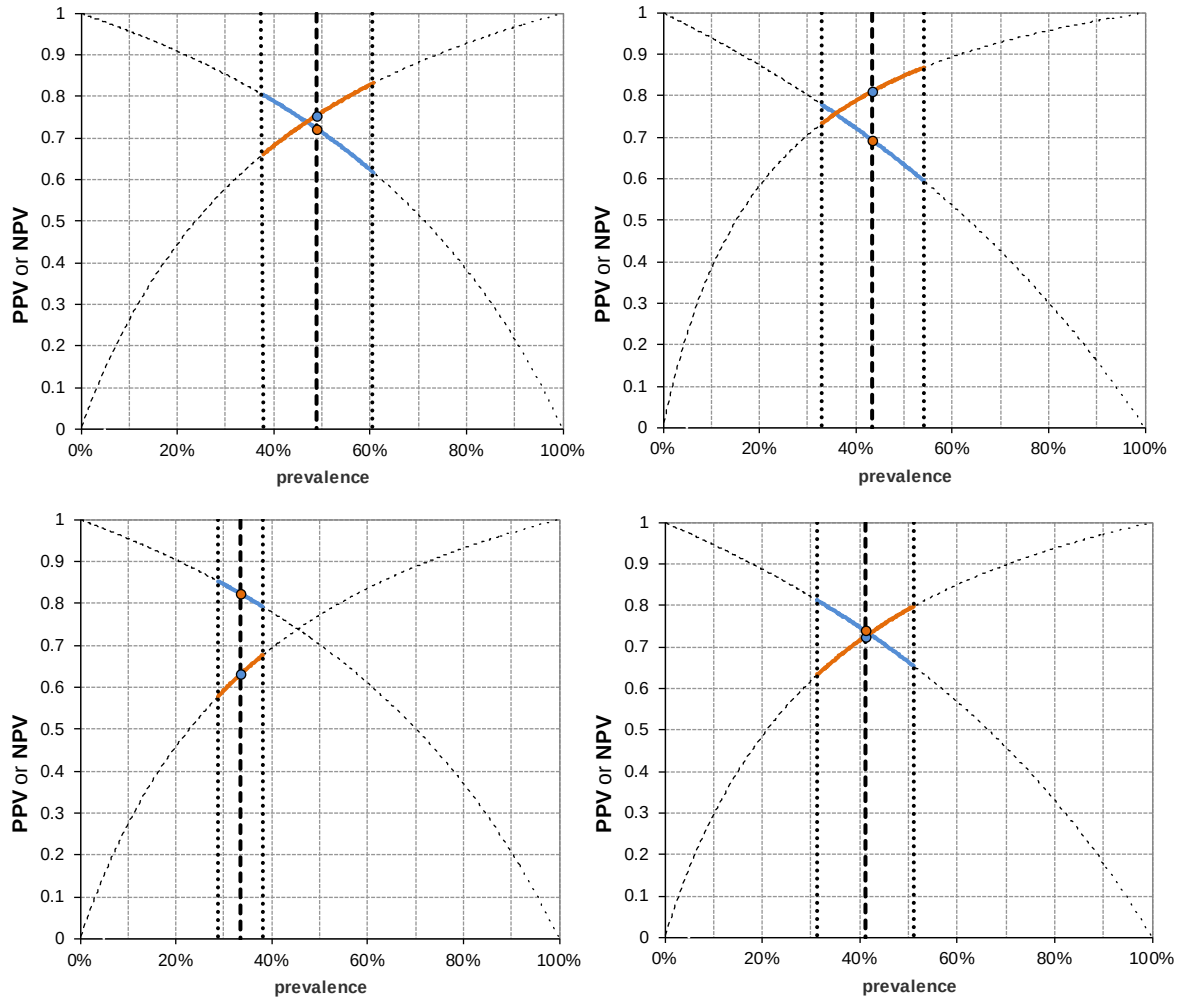


Figure 47 PPV and NPV according to prevalence of LOR (or inability to regain response) at the sROC model estimate of sensitivity and specificity; Top: adalimumab (left) and antibodies to adalimumab (right). Bottom: infliximab (left) and antibodies to infliximab (right)
As prevalence increases PPV increases and NPV decreases. Data points = PPV and NPV at sROC sensitivity and specificity and pooled prevalence. Dashed vertical lines = pooled prevalence and 95% CI. Thick curves = PPV and NPV at hierarchical model sensitivity and specificity across at pooled prevalence and 95% CI

The predictive values are indicative of moderate test accuracy so that between about 20 to 30% of positive and negative test results are likely to be incorrect.

8.13 Appendix 13 List of excluded cost-effectiveness studies with reason

Table 54 List of excluded studies from the literature review

Reference	Reason(s) for exclusion
1. Blackhouse G, Assasi N, Xie F, Marshall J, Irvine EJ, Gaebel K, et al. Canadian cost-utility analysis of initiation and maintenance treatment with anti-TNF-alpha drugs for refractory Crohn's disease. Journal of Crohn's & colitis. 2012;6(1):77-85.	No testing kits used to monitor anti-TNFα or antibodies to anti-TNFα levels
2. Bodger K, Kikuchi T, Hughes D. Cost-effectiveness of biological therapy for Crohn's disease: Markov cohort analyses incorporating United Kingdom patient-level cost data. Aliment Pharmacol Ther. 2009;30(3):265-74.	No testing kits used to monitor anti-TNFα or antibodies to anti-TNFα levels
3. Buchanan J, Wordsworth S, Ahmad T, Perrin A, Vermeire S, Sans M, et al. Managing the long term care of inflammatory bowel disease patients: The cost to European health care providers. Journal of Crohns & Colitis. 2011;5(4):301-16.	No testing kits used to monitor anti-TNFα or antibodies to anti-TNFα levels
4. Dretzke J, Edlin R, Round J, Connock M, Hulme C, Czeczot J, et al. A systematic review and economic evaluation of the use of tumour necrosis factor-alpha (TNF-α) inhibitors, adalimumab and infliximab, for Crohn's disease. Health Technol Assess. 2011;15(6):1-244.	No testing kits used to monitor anti-TNFα or antibodies to anti-TNFα levels
5. Kaplan GG, Hur C, Korzenik J, Sands BE. Infliximab dose escalation vs. initiation of adalimumab for loss of response in Crohn's disease: a cost-effectiveness analysis. Aliment Pharmacol Ther. 2007;26(11-12):1509-20.	No testing kits used to monitor anti-TNFα or antibodies to anti-TNFα levels

8.14 Appendix 14 Data extraction sheets of included health economic studies

Name of first reviewer: Hema Mistry
Name of second reviewer: Peter Auguste

Study details	
Study title	A Test-based Strategy Is More Cost Effective Than Empiric Dose Escalation for Patients With Crohn's Disease Who Lose Responsiveness to Infliximab
First author	Fernando S Velayos
Co-authors	Ames G Kahn, William J Sandborn, and Brian G Feagan
Source of publication Journal yy;vol(issue):pp	Clinical Gastroenterology and Hepatology 2013;11:654–666
Language	English
Publication type	Journal article
Inclusion criteria/study eligibility/PICOS	
Population	Patients with Crohn's disease who become unresponsive to therapy with tumour necrosis factor antagonists - infliximab
Intervention(s)	Testing-based strategy.
Comparator(s)	Empiric dose escalation strategy
Outcome(s)	Cost per quality-adjusted life-year (QALY) gained
Study design	Cost-effectiveness analysis
Methods	
Target population and subgroups	Patients with moderate-severe active Crohn's disease
Setting and location	Not reported
Study perspective	Third party payer
Time horizon	1 year time horizon with a 4 week cycle duration
Discount rate	Not reported
Measurement of effectiveness	Quality-adjusted life years
Measurement and valuation of preference based outcomes	Not reported
Resource use and costs	Direct medical costs included: cost of the interventions – infliximab, adalimumab, certolizumab, natalizumab, and surgery; and the cost of diagnostics: anti-infliximab antibody/serum infliximab measurement, CT enterography and colonoscopy
Currency, price date and conversion	US $
Model type	Decision analytical model
Assumptions	Adverse side effects causing discontinuation of medical therapy were considered to not have a significant effect on QALYs. The overall rate of response to infliximab dose escalation was assumed to be equal to that of adalimumab switching. The presence of drug antibody, drug concentration, and inflammation accurately categorises the mechanism for loss of response and the proposed interventions represent the best approach to remedy a given mechanism.
Analytical methods	Incremental cost-effectiveness ratios were presented. Extensive one-way sensitivity analyses were conducted and probabilistic sensitivity analyses using 10,000 simulations determined uncertainty in model results.
Results	
Study parameters	Proportion with mild/minimal inflammation with symptoms; Initial response – switching to adalimumab, anti-infliximab antibody

	present, subtherapeutic and therapeutic infliximab concentrations, infliximab increase to 10 mg/kg, anti-infliximab antibody present, subtherapeutic and therapeutic infliximab concentration Sustained response at 1 year - adalimumab switch, infliximab increase to 10 mg/kg, adalimumab increase to 40 mg every week, infliximab 5 mg/kg maintenance, surgery switch, sustained responders in remission, restart biological for postoperative recurrence, and proportion sustained responders in remission Mortality - after biological therapy and after surgery
Incremental costs and outcomes	The testing strategy yielded similar QALYs compared with the empiric strategy (0.801 vs 0.800, respectively) but was less expensive ($31,870 vs $37,266, respectively). The testing strategy dominated the empiric strategy.
Characterising uncertainty	One-way sensitivity analysis – key observations: the testing strategy was superior with regard to cost in almost every circumstance and the empiric strategy was less expensive when the cost of surgery was tested at 5-fold more than the base case. Probabilistic sensitivity analyses of the base case showed that 68.9% of results were within quadrant 4 (testing strategy was both less costly and more effective).
Discussion	
Study findings	The results showed that the testing strategy was cheaper and more effective than the empiric strategy.
Limitations	A prospective trial is needed to provide more precise estimates for the data such as data on the efficacy of biological therapy in the minimal/mild inflammation subgroup, data on the efficacy of biological therapy after failing standard and high dose biological therapy, as well data on efficacy of TNFα switching and infliximab dose escalation in the setting of the various drug antibody and drug level subgroups.
Generalisability	The model was defined a priori and does not reflect all possible permutations of managing loss of response.
Other	
Source of funding	Supported by an investigator-initiated research grant from Prometheus Laboratories.
Conflicts of interest	Disclosed
Comments	None
Authors conclusion	
The results support the hypothesis that a testing-based strategy is a more cost-effective alternative than the currently advocated strategy of empiric dose escalation. The basis for this difference is lower cost at similar outcomes.	
Reviewer's conclusion	
The authors used appropriate modelling techniques to demonstrate the cost-effectiveness of a testing based strategy compared with empiric strategy.	

Name of first reviewer: Hema Mistry
Name of second reviewer: Peter Auguste

Study details	
Study title	Individualised therapy is more cost-effective than dose intensification in patients with Crohn's disease who lose response to anti-TNFα treatment: a randomised, controlled trial
First author	Casper Steenholdt
Co-authors	Jørn Brynskov, Ole Østergaard Thomsen, Lars Kristian Munck, Jan Fallingborg, Lisbet Ambrosius Christensen, Gitte Pedersen, Jens Kjeldsen, Bent Ascanius Jacobsen, Anne Sophie Oxholm, Jakob Kjellberg, Klaus Bendtzen, Mark Andrew Ainsworth
Source of publication Journal yy;vol(issue):pp	Gut 2014;63: 919-927
Language	English
Publication type	Journal article
Inclusion criteria/study eligibility/PICOS	
Population	Eligible adult patients with Crohn's disease.
Intervention(s)	Receive treatment based on serum concentrations of infliximab and infliximab antibodies at the time of infliximab treatment failure in accordance with the algorithm
Comparator(s)	Receive infliximab at an increased dose frequency of 5 mg/kg every 4 weeks
Outcome(s)	Cost per intention-to-treat and cost per-protocol population
Study design	Randomised, controlled, single-blind, clinical trial
Methods	
Target population and subgroups	All patients had secondary infliximab treatment failure on infliximab maintenance therapy defined as recurrence of active disease with a Crohn's Disease Activity Index (CDAI) ≥ 220 and/or a minimum of one draining perianal fistula. Subgroup analyses included: proposed mechanisms for therapeutic failure, assessment of co-primary end points in patients stratified for C -reactive protein (CRP) level at inclusion, disease phenotype and grouping in algorithm.
Setting and location	Six Danish hospitals
Study perspective	Not reported
Time horizon	12 weeks with scheduled visits at weeks 0, 4, 8 and 12
Discount rate	Not applicable
Measurement of effectiveness	Clinical response rates – loss of response to infliximab maintenance therapy
Measurement and valuation of preference based outcomes	Not applicable
Resource use and costs	All costs of inpatient and outpatient contacts in hospitals recorded in the National Patient Registry (NPR) relating to treatment of Crohn's disease such as diagnoses and diagnostic and treatment procedures were recorded, as well as standardised infliximab doses. Expenses related to Crohn's disease in the 12 months before inclusion were comparable between randomisation groups.
Currency, price date and conversion	Danish kroner (DKK) and converted into Euros. Price date 1 January 2012.
Model type	Not applicable
Assumptions	Not applicable
Analytical methods	Costs were compared using arithmetic means and were assessed by non-parametric bootstrap analysis to determine statistical significance. Data

	were analysed by intention-to-treat and per protocol population. One-way sensitivity analyses of key primary and secondary endpoints conducted.
Results	
Study parameters	Primary endpoints: Costs of Crohn's disease and clinical response Secondary endpoints included: Crohn's Disease Activity Index (CDAI) 100 response, clinical remission, CDAI decrease, Perianal Disease Activity Index decrease and Inflammatory Bowel Disease Questionnaire increase
Incremental costs and outcomes	Costs were significantly lower in the algorithm group than in the infliximab intensification group in both the intention-to-treat population (mean difference per patient €-3,141 and the per-protocol population €-5,116. Response rates in the intention-to-treat population were 58% in the algorithm group and 53% in the infliximab intensification group: relative risk (RR) 1.091 (95% CI 0.713–1.673). The difference between response rates was 5% in favour of the algorithm group. In the per-protocol population, 47% in the algorithm group and 53% in the IFX intensification group showed a clinical response: RR 0.898 (95% CI 0.510–1.580). Incremental cost-effectiveness ratios not reported.
Characterising uncertainty	One-way sensitivity analyses included (1) estimated administrative costs for biological drugs, (2) use of actual infliximab dosing and (3) price reductions on biological agents. Findings were similar to the base-case analysis.
Discussion	
Study findings	The present clinical trial testing of whether a personalised patient treatment based on IFX bioavailability and immunogenicity at the time of therapeutic failure proved more cost-effective than standard IFX intensification. That is, the interventions based on the algorithm achieved similar clinical, biological and life quality outcomes to dose intensification, but at a lower cost. Findings were also robust and consistent in subgroups.
Limitations	Small numbers
Generalisability	Only reported in terms of costs
Other	
Source of funding	Disclosed
Conflicts of interest	Disclosed
Comments	None
Authors conclusion	
Managing secondary infliximab treatment failure by an algorithm based on serum infliximab and infliximab antibodies to define the mechanistic basis and corresponding interventions is more cost-effective than an intensified infliximab regimen.	
Reviewer's conclusion	
Although patient numbers were small, the authors used appropriate trial evidence to demonstrate the cost-effectiveness of algorithm based strategy compared with intensified dose strategy.	

Name of first reviewer: Hema Mistry
Name of second reviewer: Peter Auguste

Study details	
Study title	Trough Concentrations of Infliximab Guide Dosing for Patients with Inflammatory Bowel Disease
First author	Niels Vande Casteele
Co-authors	Marc Ferrante, Gert Van Assche, Vera Ballet, Griet Compernolle, Kristel Van Steen, Steven Simoens, Paul Rutgeerts, Ann Gils, Séverine Vermeire
Source of publication Journal yy;vol(issue):pp	Gastroenterology (in press)
Language	English
Publication type	Journal article
Inclusion criteria/study eligibility/PICOS	
Population	Patients with a diagnosis of moderate-to-severe Crohn's disease or ulcerative colitis
Intervention(s)	Concentration-based infliximab dosing
Comparator(s)	Clinically based infliximab dosing
Outcome(s)	Cost per quality-adjusted life year (QALY)
Study design	Randomised controlled trial
Methods	
Target population and subgroups	Cohort of Crohn's (and ulcerative colitis) responder patients. Patients needed to be treated with maintenance infliximab therapy for at least 14 weeks and needed to be in stable clinical response.
Setting and location	Tertiary referral centre, Belgium
Study perspective	Third party payer
Time horizon	1 year
Discount rate	Not applicable
Measurement of effectiveness	Quality-adjusted life years
Measurement and valuation of preference based outcomes	EuroQol (EQ-5D) used to calculate QALYs.
Resource use and costs	Drug costs per patient per year Resource use and costs not reported in detail
Currency, price date and conversion	Euros, price year 2012
Model type	Not applicable
Assumptions	Not applicable
Analytical methods	QALYs were adjusted for differences in baseline utility scores using a multiple regression approach. Incremental cost-effectiveness ratios were presented. Uncertainty in incremental QALYs and costs was determined by non-parametric bootstrapping consisting of 1,000 iterations and plotted onto a cost-effectiveness plane.
Results	
Study parameters	Primary endpoints: Clinical and biochemical remission at 1 year after the optimisation phase (increasing and maintaining remission) Secondary endpoints: Durable remission, relapse, infliximab trough concentration within the optimal interval, antibodies to infliximab positivity, total cost of infliximab treatment, and quality adjusted life years (QALY)
Incremental costs and outcomes	Concentration-based dosing: QALY = 0.8227; Costs = €20,723 Clinically based dosing: QALY = 0.8421; Costs = €21,023 Incremental QALYs = -0.0193

	Incremental costs = -€300 Incremental cost-effectiveness ratio = €15,525
Characterising uncertainty	Cost-effectiveness plane showing probabilistic sensitivity analyses found that 58.4% of simulations were in quadrant 3 where concentration-based dosing was less costly and less effective
Discussion	
Study findings	Concentration-based dosing was slightly less effective and less costly than clinically based dosing, but overall differences were small
Limitations	Duration of randomised treatment was 1 year.
Generalisability	Not reported
Other	
Source of funding	Disclosed
Conflicts of interest	Disclosed
Comments	None
Authors conclusion	
Concentration-based dosing was slightly less effective and less costly than clinically based dosing	
Reviewer's conclusion	
The authors used appropriate trial evidence to demonstrate the cost-effectiveness of concentration-based dosing compared with clinically based dosing	

Name of first reviewer: Hema Mistry
Name of second reviewer: Peter Auguste

Study details	
Study title	Individualized Therapy Is a Long-Term Cost-Effective Method Compared to Dose Intensification in Crohn's Disease Patients Failing Infliximab
First author	Casper Steenholdt
Co-authors	Jørn Brynskov, Ole Ø. Thomsen, Lars K. Munck, Jan Fallingborg, Lisbet A. Christensen, Gitte Pedersen, Jens Kjeldsen, Bent A. Jacobsen, Anne Sophie Oxholm, Jakob Kjellberg, Klaus Bendtzen, Mark A. Ainsworth
Source of publication Journal yy;vol(issue):pp	Digestive Diseases and Sciences or Gut 2015; DOI 10.1007/s10620-015-3581-4
Language	English
Publication type	Journal article
Inclusion criteria/study eligibility/PICOS	
Population	Eligible adult patients with Crohn's disease
Intervention(s)	Receive treatment based on serum concentrations of infliximab and infliximab antibodies at the time of infliximab treatment failure in accordance with the algorithm
Comparator(s)	Receive infliximab at an increased dose frequency of 5 mg/kg every 4 weeks
Outcome(s)	Cost per intention-to-treat and cost per-protocol population
Study design	Randomised, controlled, single-blind, clinical trial
Methods	
Target population and subgroups	All patients had secondary infliximab treatment failure on infliximab maintenance therapy defined as recurrence of active disease with a Crohn's Disease Activity Index (CDAI)≥ 220 and/or a minimum of one draining perianal fistula.
Setting and location	Six Danish hospitals
Study perspective	Not reported
Time horizon	1 year with cost evaluations at 20 weeks and 1 year
Discount rate	Not applicable
Measurement of effectiveness	Clinical response was defined as ≥70 point reduction in CDAI from baseline in luminal disease and a reduction in active fistulas of≥50 % from baseline in fistulising disease. Clinical remission was defined as CDAI ≤150 and complete closure of all fistulas despite gentle pressure.
Measurement and valuation of preference based outcomes	Not applicable
Resource use and costs	All costs of inpatient and outpatient contacts in hospitals recorded in the National Patient Registry (NPR) relating to treatment of Crohn's disease such as diagnoses and diagnostic and treatment procedures were recorded, as well as standardised infliximab and anti-infliximab doses.
Currency, price date and conversion	Danish kroner (DKK) and converted into US $. Price date 1 January 2012.
Model type	Not applicable
Assumptions	Not applicable
Analytical methods	Costs were analysed using arithmetic means and were compared by non-parametric bootstrap analysis to determine statistical significance. Data were analysed by intention-to-treat, per protocol population, per-protocol completion at end of trial week 12, and per-protocol completion at end of follow-up week 20. One-way sensitivity analyses of key primary and secondary endpoints conducted
Results	
Study parameters	Endpoints: Costs of Crohn's disease, clinical response and

	clinical remission
Incremental costs and outcomes	Incremental costs in favour of the algorithm group – that is costs were substantially and highly significantly lower in the algorithm group than in the infliximab intensification group: 20 weeks Intention to treat: $-5,296 Per protocol: $-8,494 Per protocol end of trial week 12: $-8,546 Per protocol end of follow-up week 20: $-10,720 1 year Intention to treat: $-7,006 Per protocol: $-13,383 Per protocol end of trial week 12: $-13,265 Per protocol end of follow-up week 20: $-16,618
Characterising uncertainty	One-way sensitivity analyses at both 20 weeks and 1 year included (1) estimated administrative costs for biological drugs, (2) use of actual infliximab dosing and (3) price reductions on biological agents. Findings were similar to the 20 weeks and 1 year time frames.
Discussion	
Study findings	The algorithm group had significantly lower costs than in the infliximab intensification group at the 20 week follow-up and this was maintained throughout the 1 year.
Limitations	Small sample size for the study
Generalisability	Compared findings with other studies and some studies have used their algorithm
Other	
Source of funding	Disclosed
Conflicts of interest	Disclosed
Comments	None
Authors conclusion	
Clinical interventions at infliximab treatment failure based on monitoring of infliximab and anti-infliximab antibodies are long-term cost-effective method compared to infliximab dose intensification.	
Reviewer's conclusion	
Although patient numbers were small, the authors used appropriate trial evidence to demonstrate the cost-effectiveness of algorithm based strategy compared with intensified dose strategy over a 1 year time period.	

Name of first reviewer: Hema Mistry
Name of second reviewer: Peter Auguste

Study details	
Study title	A systematic review and economic evaluation of the use of tumour necrosis factor-alpha (TNF-a) inhibitors, adalimumab and infliximab, for Crohn's disease
First author	J Dretzke
Co-authors	R Edlin, J Round, M Connock, C Hulme, J Czeczot, A Fry-Smith, C McCabe and C Meads
Source of publication Journal yy;vol(issue):pp	Health Technology Assessment 2011;15(6)
Language	English
Publication type	Monograph
Inclusion criteria/study eligibility/PICOS	
Population	Adult patients with moderate to severe Crohn's disease
Intervention(s)	Anti-TNFα therapy for CD – infliximab and adalimumab
Comparator(s)	Standard care for CD
Outcome(s)	Cost-per quality adjusted life year (QALY) gained
Study design	Cost-effectiveness analysis
Methods	
Target population and subgroups	Adult patients with moderate to severe Crohn's disease where response was defined as remission within 8 weeks
Setting and location	Not reported
Study perspective	NHS and PSS perspective
Time horizon	1 year time horizon with a 4 week cycle duration
Discount rate	Not reported
Measurement of effectiveness	Quality-adjusted life years
Measurement and valuation of preference based outcomes	Choice based time-trade off measure providing utility value
Resource use and costs	Cost of anti-TNFα treatment for both induction and maintenance therapy, plus administration costs. Type specific health state costs were also included: costs for surgery were modelled as the cost of inpatient IBD interventions, while moderate and severe relapse costs were modelled as the cost of IBD outpatient major and intermediate interventions. Post-surgery remission costs were based on outpatient surgical gastrointestinal follow-up. Relapse costs were based on a gastrointestinal admission to hospital. Remission costs were modelled using literature. Unit costs were obtained from the NHS reference costs.
Currency, price date and conversion	Price year 2005-2006
Model type	Markov model
Assumptions	Model did not take into account mortality. Used Silverstein et al for all transition probabilities in the intervention arm.
Analytical methods	Incremental cost-effectiveness ratios and cost-effectiveness acceptability curves were presented. One-way sensitivity analyses and probabilistic sensitivity analyses using 10,000 simulations were conducted to characterise uncertainty in the model.
Results	
Study parameters	For the three arms: standard care, induction and maintenance the parameters included transition probabilities, costs and utilities for the following health states: remission, relapse (moderate and severe), surgery and post-surgery

Incremental costs and outcomes	For induction therapy for severe Crohn's disease, both adalimumab and infliximab dominated standard care (i.e. cheaper and more effective). For maintenance therapy for severe Crohn's disease, neither drug was cost-effective (well above NICE thresholds). For moderate Crohn's disease, for maintenance therapy for both drugs and induction therapy for infliximab, these were not cost-effective (well above NICE thresholds); however, for induction therapy for adalimumab dominated standard care.
Characterising uncertainty	Patients who had severe disease, infliximab induction treatment was found to be cost-effective relative to maintenance treatment and standard care in over 99% of cases at all points up to £100,000 per QALY. Likewise, adalimumab induction treatment was found to be cost-effective relative to maintenance treatment and standard care for thresholds up to £100,000 per QALY.
Discussion	
Study findings	The results for induction, both adalimumab and infliximab were cost-effective (dominant relative to standard care) for severe Crohn's disease and that adalimumab was cost-effective (dominant relative to standard care) for moderate Crohn's disease. Induction therapy with infliximab was not cost-effective for moderate Crohn's disease. Neither drug was cost-effective as maintenance therapy for moderate or severe disease .
Limitations	Exclusion of death from the model. A 1 year time horizon. No RCT data available for maintenance therapy. Silverstein et al data had its own problems i.e. surgery rates are higher and relapse rates much lower than in routine practice.
Generalisability	Not reported
Other	
Source of funding	Disclosed
Conflicts of interest	Not reported
Comments	None
Authors conclusion	
Infliximab is not likely to be cost-effective in the management of moderate Crohn's disease. While adalimumab may be cost-effective, there is uncertainty regarding the incremental cost-effectiveness ratio value. Neither of these therapies is likely to be cost-effective as maintenance therapy for moderate or severe disease. Both treatments are highly cost-effective, with no meaningful uncertainty, as induction therapy in severe disease.	
Reviewer's conclusion	
The authors used appropriate modelling techniques to demonstrate the cost-effectiveness of the interventions for two anti-TNFα drug therapies compared with standard care; although there are some limitations in terms of how the transition probabilities and utility values were estimated.	

8.15 Appendix 15 Quality assessment of included health economic studies

CHEERS quality assessment checklist for economic evaluation studies

Assessment	Velayos et al. (2013)	Steenholdt et al. (2014)	Vande Casteele et al. (2015)	Steenholdt et al. (2015)	Dretzke at al. (2011)
Title	Y	Y	N	Y	Y
Abstract	Y	Y	Y	Y	Y
Introduction					
Background and objectives	Y	Y	Y	Y	Y
Methods					
Target population and subgroups	P	Y	Y	Y	Y
Setting and location	N	Y	Y	Y	N
Study perspective	Y	N	Y	N	Y
Comparators	Y	Y	Y	Y	Y
Time horizon	Y	Y	Y	Y	Y
Discount rate	N	N/A	N/A	N/A	N/A
Choice of health outcomes	Y	Y	Y	Y	Y
Measurement of effectiveness	Y	Y	Y	Y	Y
Measurement and valuation of preference-based outcomes	N	N	Y	N	Y
Estimating resources and costs	Y	Y	UNC	Y	Y
Currency, price date, and conversion	P	Y	P	Y	Y
Choice of model	Y	N/A	N/A	N/A	Y
Assumptions	Y	N	N	N	Y
Analytical methods	Y	Y	Y	Y	Y
Results					
Study parameters	Y	Y	Y	Y	Y
Incremental costs and outcomes	Y	Y	Y	Y	Y
Characterising uncertainty	Y	Y	Y	Y	Y
Discussion					
Study findings	Y	Y	Y	Y	Y
Limitations	Y	Y	Y	Y	Y
Generalizability	P	P	N	P	N
Other					
Source of funding	Y	Y	Y	Y	Y
Conflicts of interest	Y	Y	Y	Y	N
N- No; N/A- Not Applicable; P – Partial; Y- Yes; UNC-Unclear					

Philips' quality assessment checklist for studies that included an economic model

Philips' criteria		Studies	
		Velayos et al 2013	Dretzke et al 2011
Structure			
1.	Is there a clear statement of the decision problem?	Y	Y
2.	Is the objective of the model specified and consistent with the stated decision problem?	Y	Y
3.	Is the primary decision maker specified?	N	Y
4.	Is the perspective of the model stated clearly?	Y	Y
5.	Are the model inputs consistent with the stated perspective?	Y	Y
6.	Has the scope of the model been stated and justified?	Y	Y
7.	Are the outcomes of the model consistent with the perspective, scope and overall objective of the model?	Y	Y
8.	Is the structure of the model consistent with a coherent theory of the health condition under evaluation?	Y	Y
9.	Are the sources of the data used to develop the structure of the model specified?	Y	Y
10.	Are the causal relationships described by the model structure justified appropriately?	UNC	Y
11.	Are the structural assumptions transparent and justified?	Y	Y
12.	Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Y	Y
13.	Is there a clear definition of the options under evaluation?	Y	Y
14.	Have all feasible and practical options been evaluated?	Y	Y
15.	Is there justification for the exclusion of feasible options?	Y	Y
16.	Is the chosen model type appropriate given the decision problem and specified casual relationships within the model?	Y	Y
17.	Is the time horizon of the model sufficient to reflect all important differences between the options?	Y	Y
18.	Are the time horizon of the model, the duration of treatment and the duration of treatment described and justified?	Y	Y
19.	Do the disease states (state transition model) or the pathways (decision tree model) reflect the underlying biological process of the disease in question and the impact of interventions?	Y	Y
20.	Is the cycle length defined and justified in terms of the natural history of disease?	Y	Y
Data			
21.	Are the data identification methods transparent and appropriate given the objectives of the model?	Y	Y
22.	Where choices have been made between data sources are these justified appropriately?	UNC	Y
23.	Has particular attention been paid to identifying data for the important parameters of the model?	Y	Y
24.	Has the quality of the data been assessed appropriately?	Y	Y
25.	Where expert opinion has been used are the methods described and justified?	UNC	N/A

Philips' criteria		Studies	
		Velayos et al 2013	Dretzke et al 2011
26.	Is the data modelling methodology based on justifiable statistical and epidemiological techniques?	UNC	Y
27.	Is the choice of baseline data described and justified?	UNC	Y
28.	Are transition probabilities calculated appropriately?	UNC	N
29.	Has a half-cycle correction been applied to both costs and outcomes?	N	N
30.	If not, has the omission been justified?	N	N
31.	If relative treatment effects have been derived from trial data, have they been synthesised using appropriate techniques?	N/A	Y
32.	Have the methods and assumptions used to extrapolate short-term results to final outcomes been documented and justified?	UNC	Y
33.	Have alternative extrapolation assumptions been explored through sensitivity analysis?	Y	Y
34.	Have assumptions regarding the continuing effect of treatment once treatment is complete been documented and justified?	UNC	Y
35.	Have alternative assumptions regarding the continuing effect of treatment been explored through sensitivity analysis	UNC	Y
36.	Are the costs incorporated into the model justified?	Y	Y
37.	Has the source for all costs been described?	Y	Y
38.	Have discount rates been described and justified given the target decision maker?	N	N
39.	Are the utilities incorporated into the model appropriate?	Y	Y
40.	Is the source of utility weights referenced?	N	Y
41.	Are the methods of derivation for the utility weights justified?	N	Y
42.	Have all data incorporated into the model been described and referenced in sufficient detail?	Y	Y
43.	Has the use of mutually inconsistent data been justified (i.e. are assumptions and choices appropriate?)	Y	Y
44.	Is the process of data incorporation transparent?	Y	Y
45.	If data have been incorporated as distributions, has the choice of distributions for each parameter been described and justified?	Y	Y
46.	If data have been incorporated as distributions, is it clear that second order uncertainty is reflected?	UNC	Y
47.	Have the four principal types of uncertainty been addressed?	N	N
48.	If not, has the omission of particular forms of uncertainty been justified?	N	N
49.	Have methodological uncertainties been addressed by running alternative versions of the model with different methodological assumptions?	N	Y
50.	Is there evidence that structural uncertainties have been addressed via sensitivity analysis?	N	Y
51.	Has heterogeneity been dealt with by running the model separately for different sub-groups?	Y	Y
52.	Are the methods of assessment of parameter uncertainty appropriate?	Y	Y

Philips' criteria		Studies	
		Velayos et al 2013	Dretzke et al 2011
53.	If data are incorporated as point estimates, are the ranges used for sensitivity analysis stated clearly and justified?	Y	Y
54.	Is there evidence that the mathematical logic of the model has been tested thoroughly before use?	N	Y
55.	Are any counterintuitive results from the model explained and justified?	Y	Y
56.	If the model has been calibrated against independent data, have any differences been explained and justified?	N	UNC
57.	Have the results been compared with those of previous models and any differences in results explained?	N	N
N- No; N/A- Not Applicable; Y- Yes; UNC-Unclear			

8.16 Appendix 16 Decision tree structure for the responders' model

This Appendix summarises the underlying decision tree structure of the model for responders to anti-TNFα therapy in several Figures:

- For concurrent testing see Figure 48 to Figure 53
- For no testing see **Figure** 54 and Figure 55
- For reflexing testing see Figure 56 to Figure 59

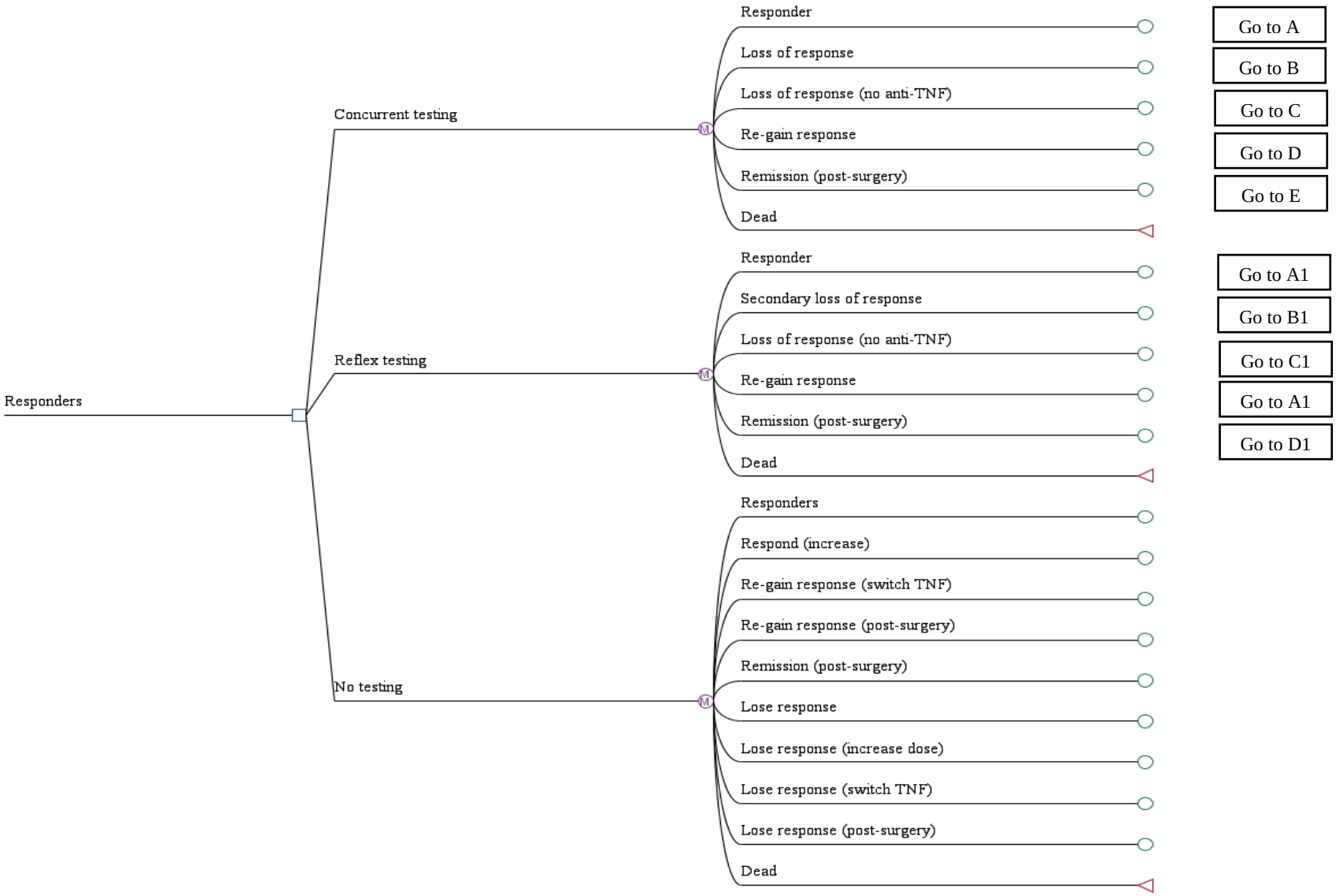


Figure 48 Decision tree structure for the responders' model for concurrent testing

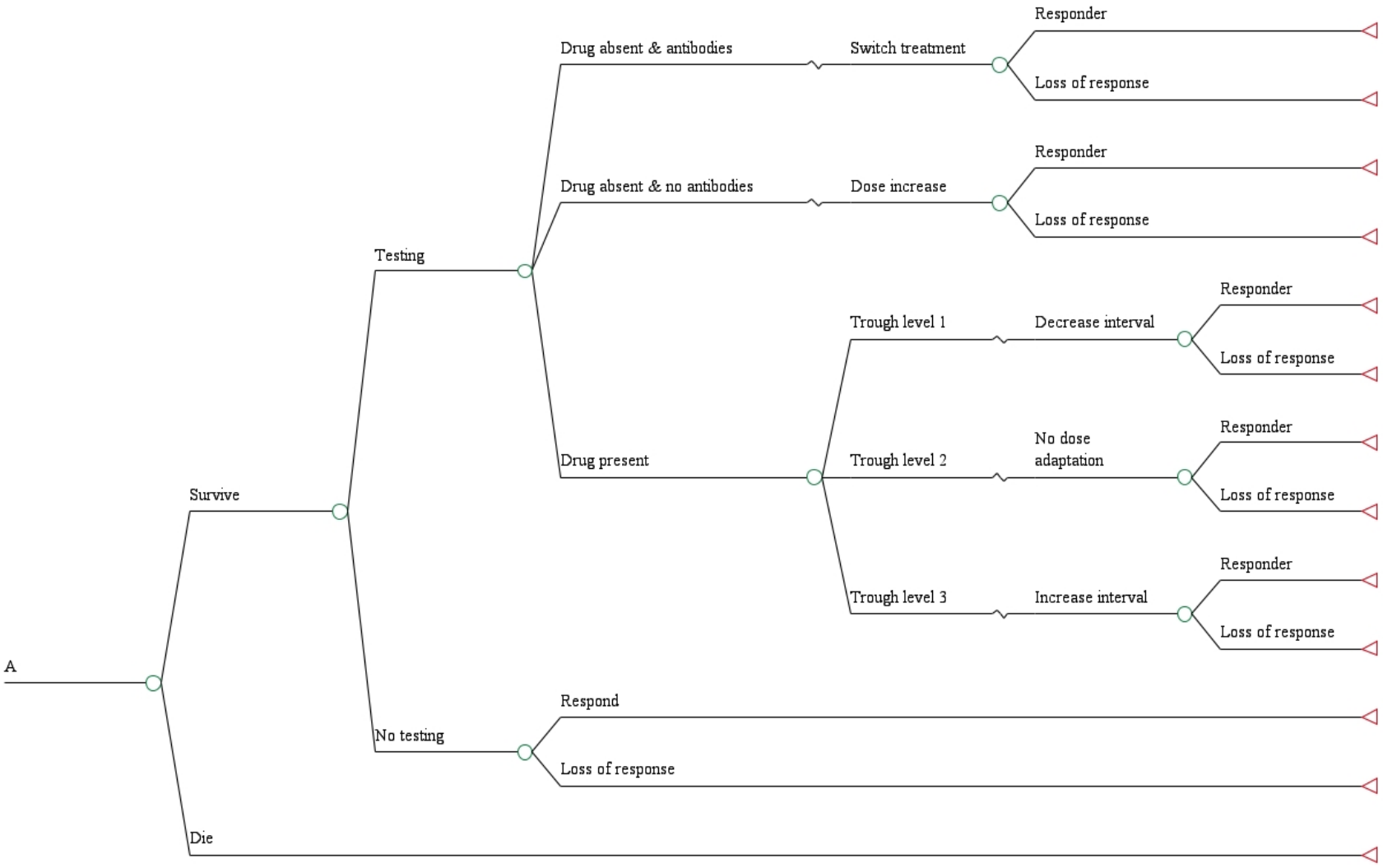


Figure 49 Decision tree structure for the responders' model for concurrent testing (A)

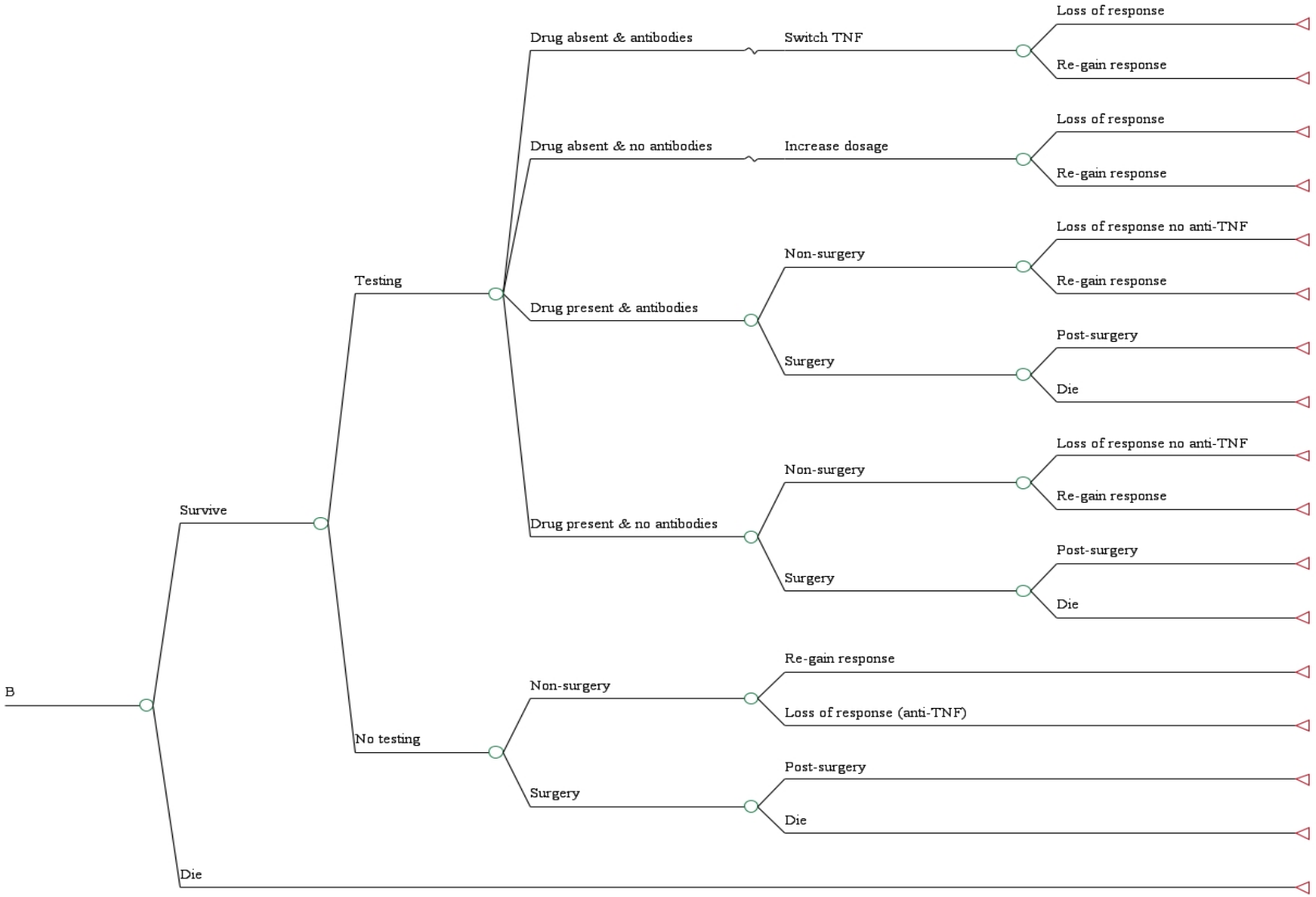


Figure 50 Decision tree structure for the responders' model concurrent testing (B)

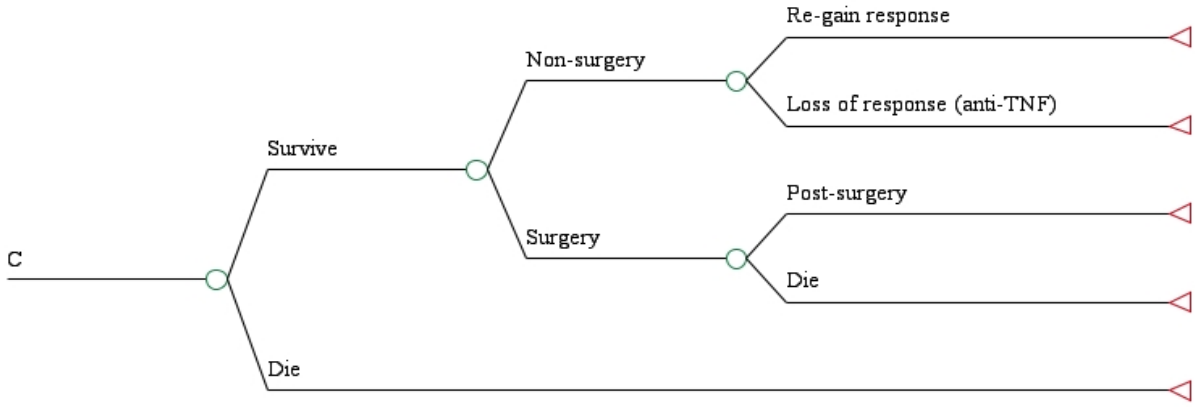


Figure 51 Decision tree structure for the responders' model concurrent testing (C)

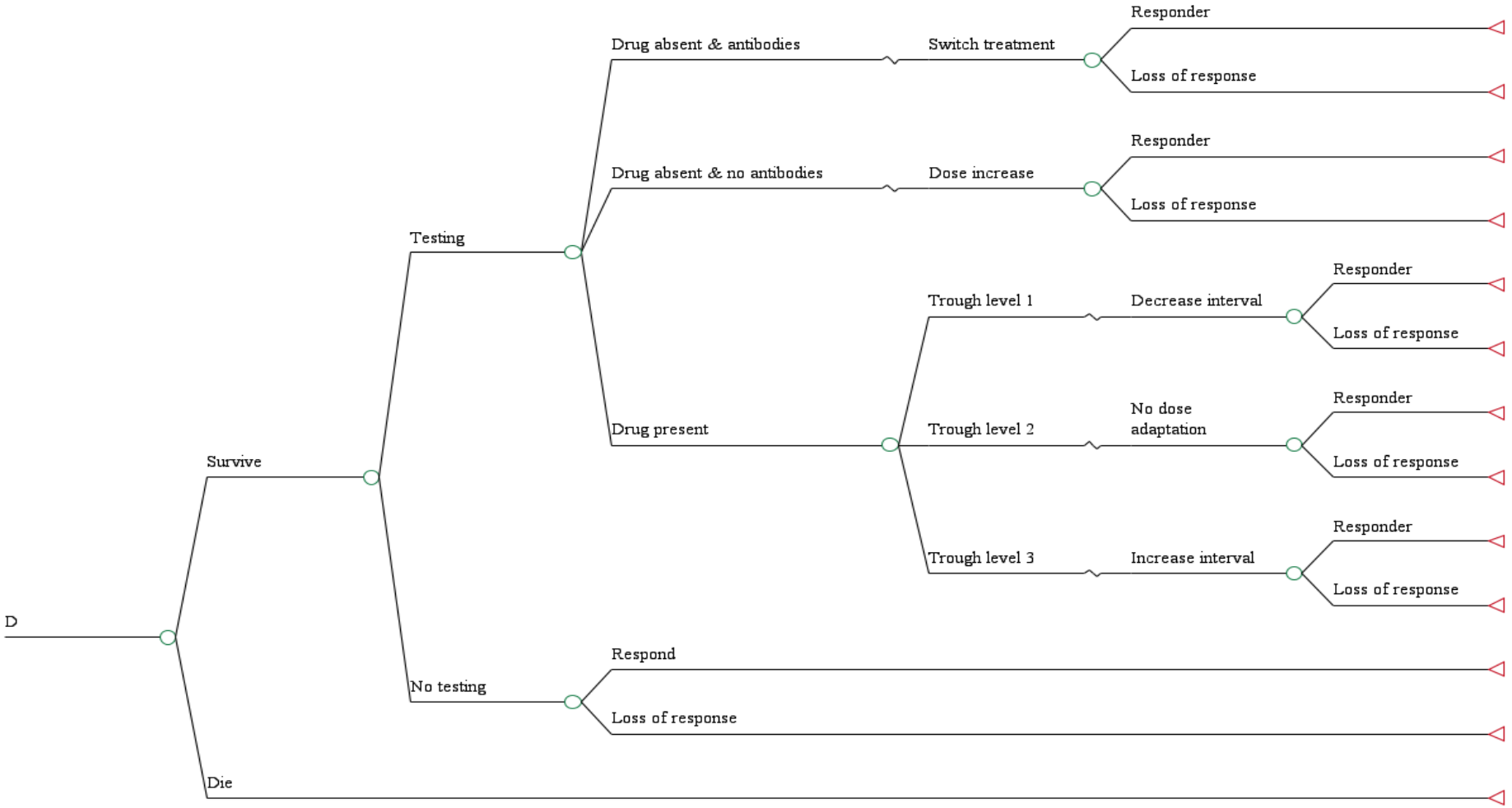


Figure 52 Decision tree structure for the responders' model concurrent testing (D)

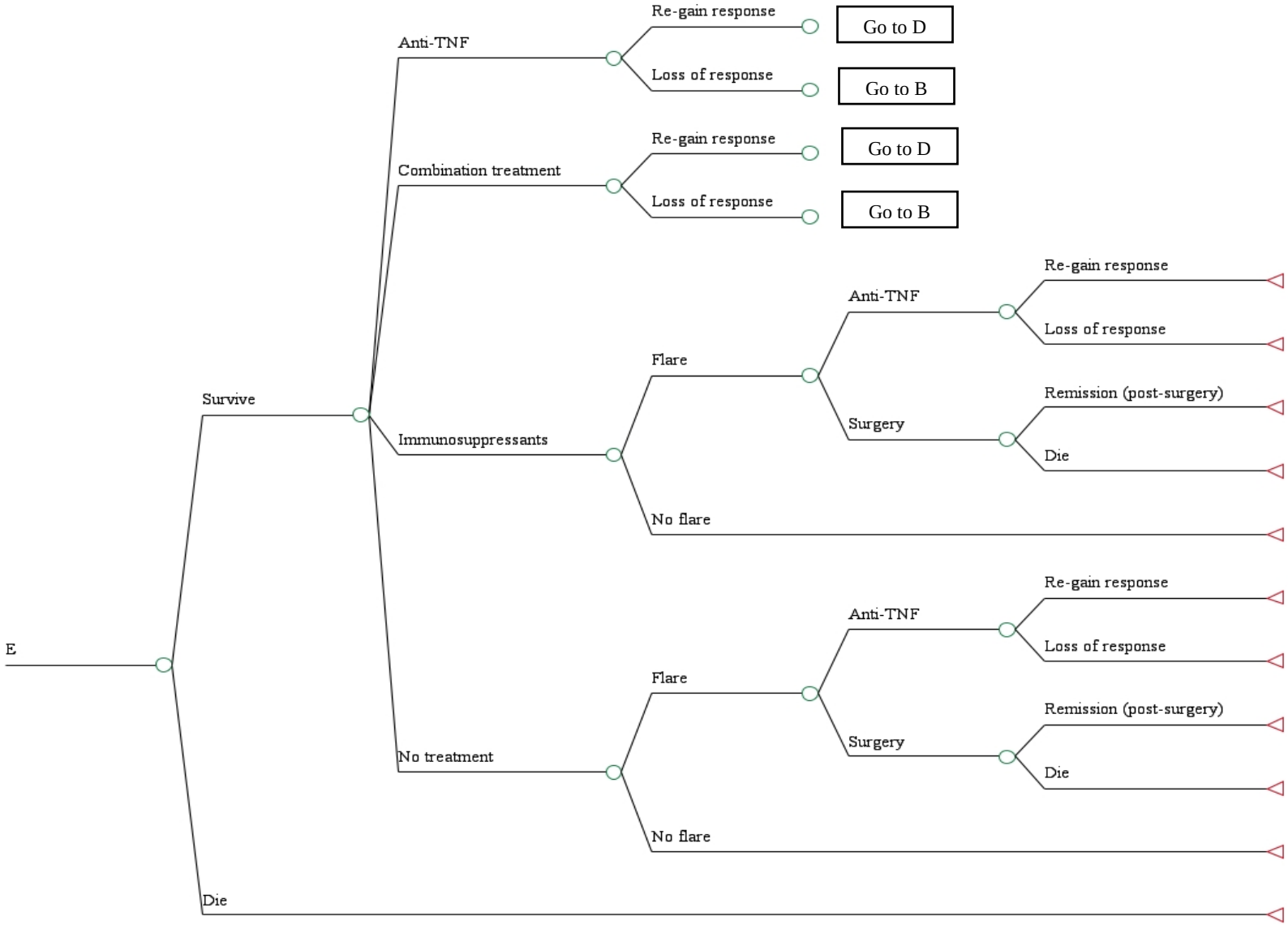


Figure 53 Decision tree structure for the responders' model concurrent testing (E)

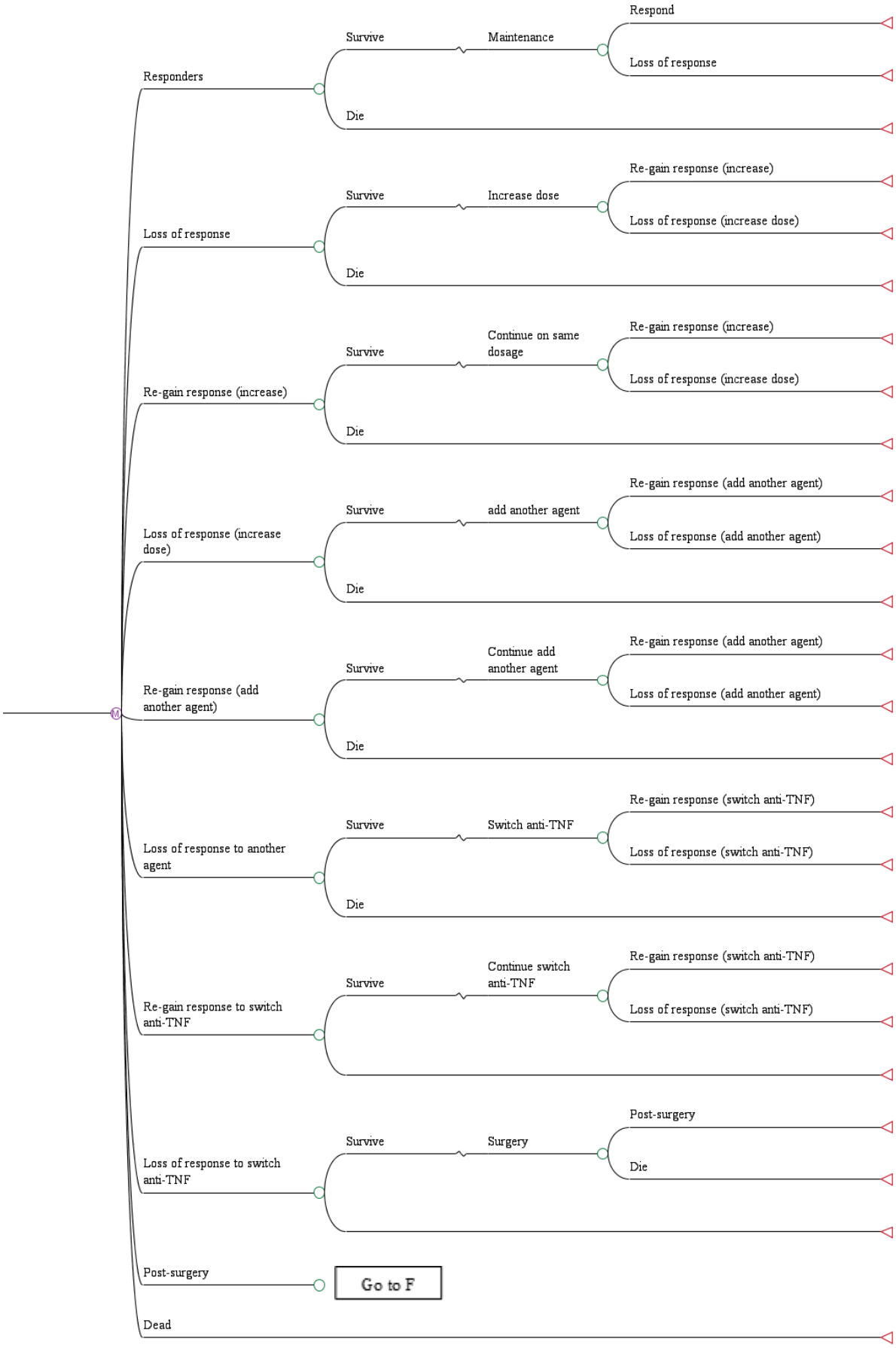


Figure 54 Decision tree structure for the no testing strategy

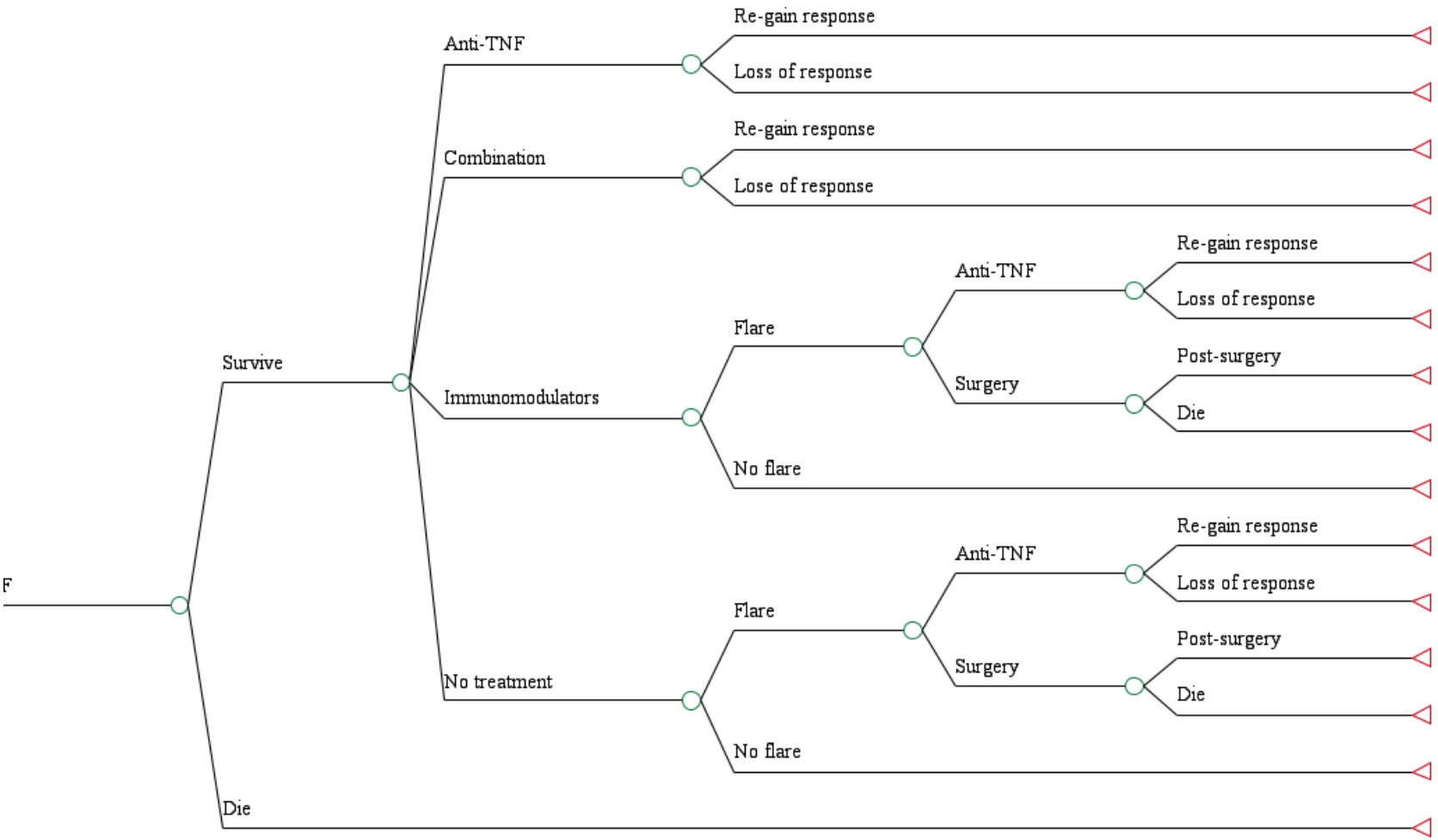


Figure 55 Patient pathway for people in the post-surgery health state

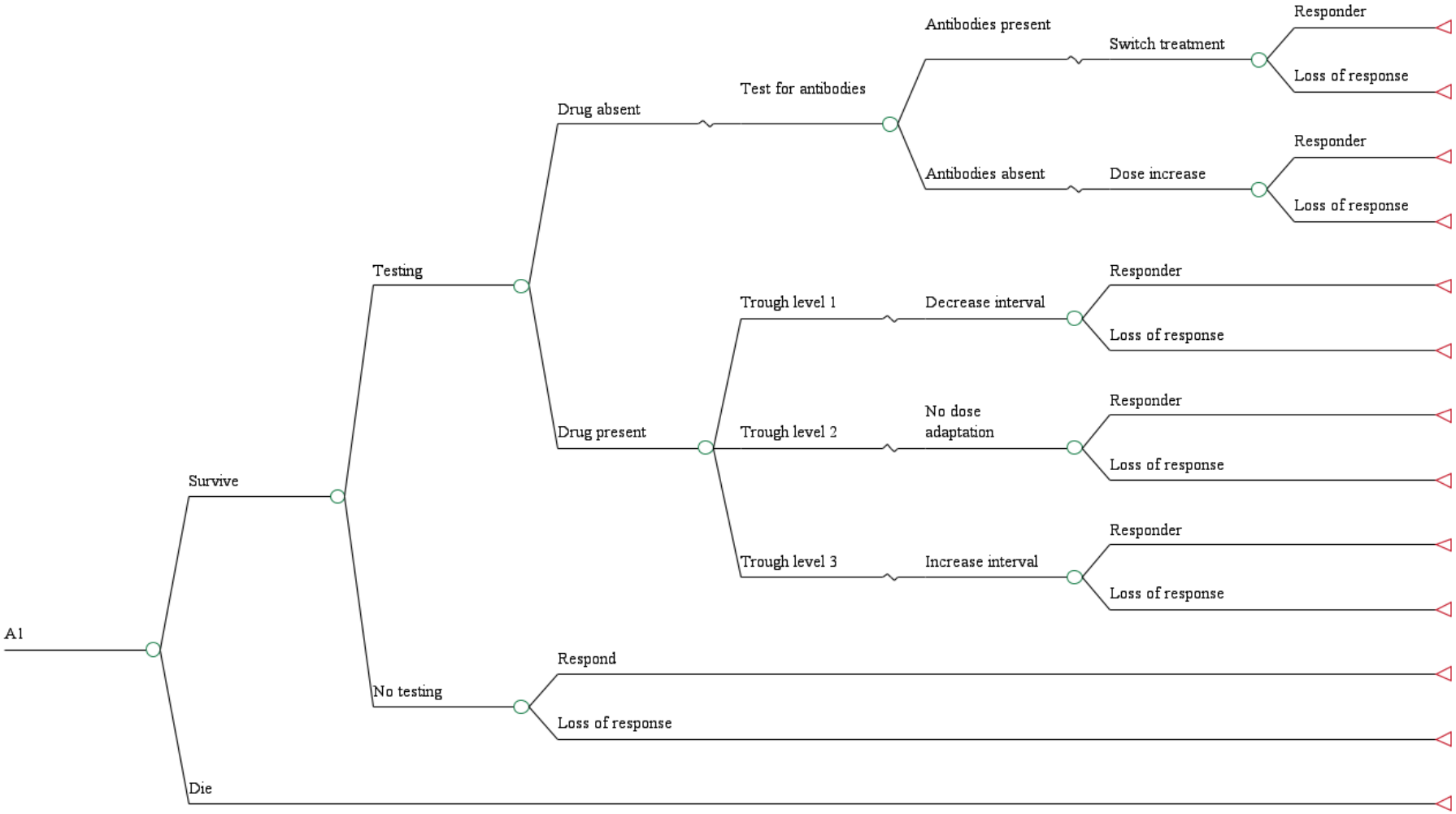


Figure 56 Decision tree structure for responders' model for reflex testing (A1)

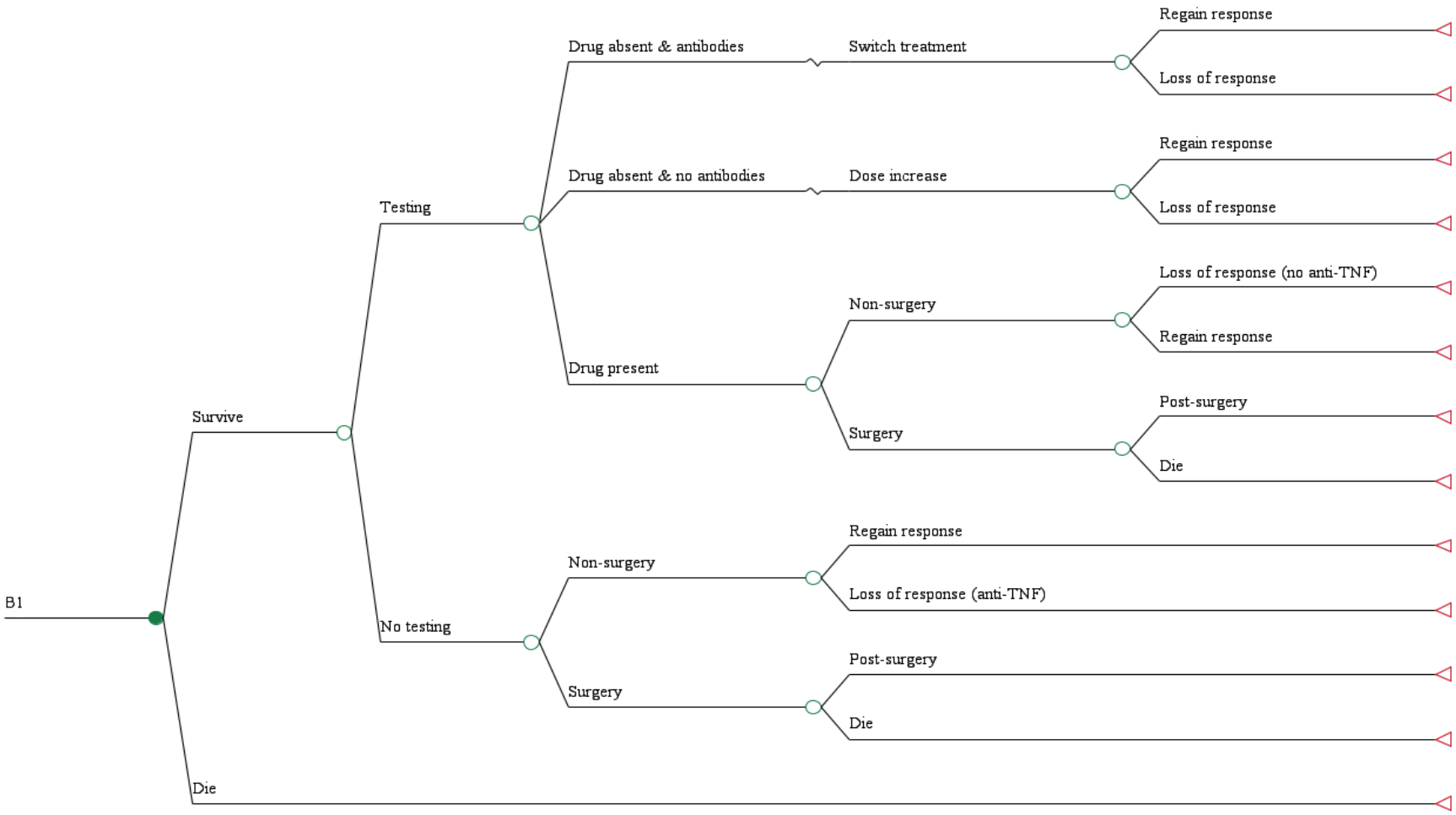


Figure 57 Decision tree structure for responders' model for reflex testing (B1)

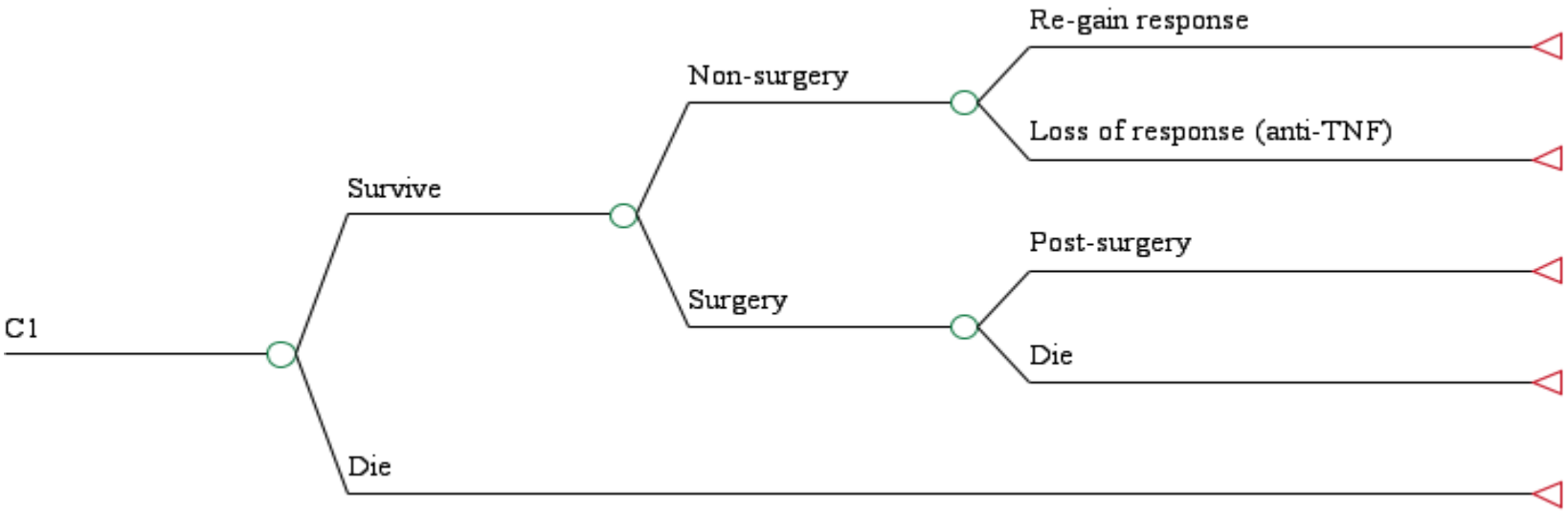


Figure 58 Decision tree structure for responders' model for reflex testing (C1)

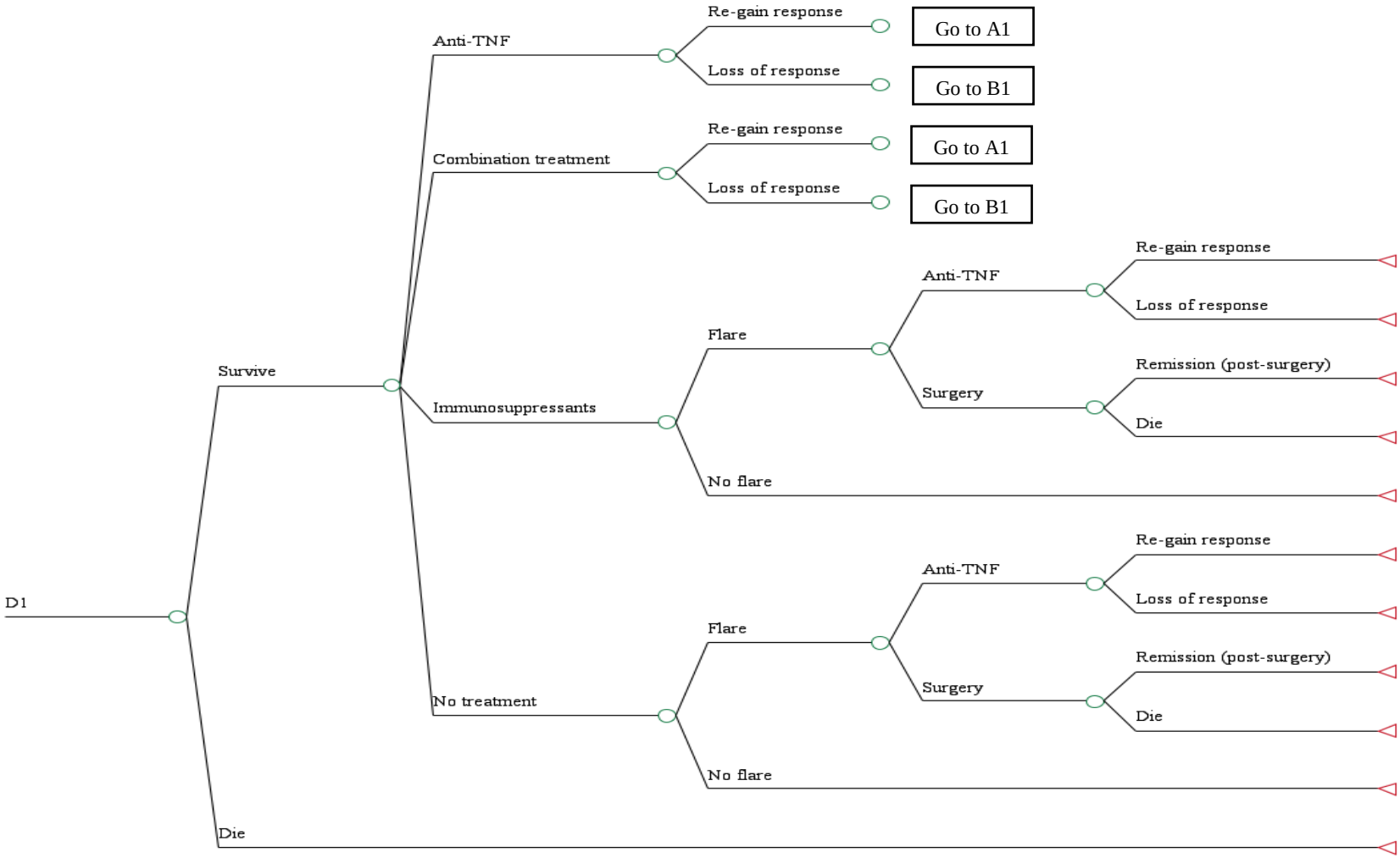


Figure 59 Decision tree structure for the responders' model for reflext testing (D1)

8.17 Appendix 17 Transition probabilities derived from published studies

Transition from response to infliximab to loss of response in primary responders

There was insufficient published information to model an adalimumab test-based treatment strategy. The model therefore addresses patients responding to infliximab maintenance therapy (the transition probabilities used are summarised in Table 34). It should be emphasised that there were no prospective or other test-directed management studies describing outcomes for infliximab responders followed from maintenance treatment through to treatments subsequent to loss of response to maintenance. Therefore, by necessity, model structure for the intervention arm is based on the algorithms used in the two identified randomised controlled trials describing test-based patient management, specifically TAXIT[72] for responders and Steenholdt et al. (2014)[122] for patients with loss of response to maintenance infliximab (section 3.2.3); we aimed to use as much data from these RCTs as possible to populate the model. Unfortunately the control arm in TAXIT does not provide information for the model's standard care arm (a no-test management strategy) because all patients in TAXIT were dose-optimised according to test results prior to randomisation; consequently the model structure for the standard care arm is based on expert clinical advice and alternative studies were examined for model input.

Standard care arm; loss of response to infliximab maintenance

For the standard care arm three studies which reported reasonable quality data for time to loss of response or to cessation of infliximab treatment for patients on maintenance treatment with infliximab were identified (Juillerat et al., 2015,[151] Bortlik et al., 2013[81] and Vaughn et al., 2014[127]). Reconstructed Kaplan Meier plots with candidate parametric models are shown in Figure 60. For the Juillerat study, a Weibull model provided the best fit to 130 cycles.

Figure 60 Reconstructed Kaplan Meier plots for time to loss of response or to cessation of treatment of responders on maintenance infliximab therapy by four-week cycle
Left, Bortlik N=84; centre, Juillerat N=1014; right, Vaughn N = 68

These three studies generate quite different transition probabilities. Because of its size, the availability of observed data to 130 cycles (model time horizon), and the inclusion of only CD patients, the Juillerat study was selected for model inputs. In Juillerat, 21% of patients received dose escalation but the time to escalation was not reported. However Ma et al. (2014)[153] have reported the time to loss of response requiring dose escalated for CD patients on infliximab maintenance therapy; Weibull and Gompertz models provided best fits to the Ma et al. (2014)[153] data. Figure 61 shows both Juillerat and Ma data with Weibull parametric models. Transition probabilities generated by these Weibull models were used for economic model input. These allow estimates of the percentage of time over 130 cycles spent in each of the following conditions: (i) untreated with infliximab, (ii) in standard dose treatment with infliximab, (iii) in escalated dose treatment with infliximab; the resulting percentages were 35.6%, 24.0% and 40.4% respectively.

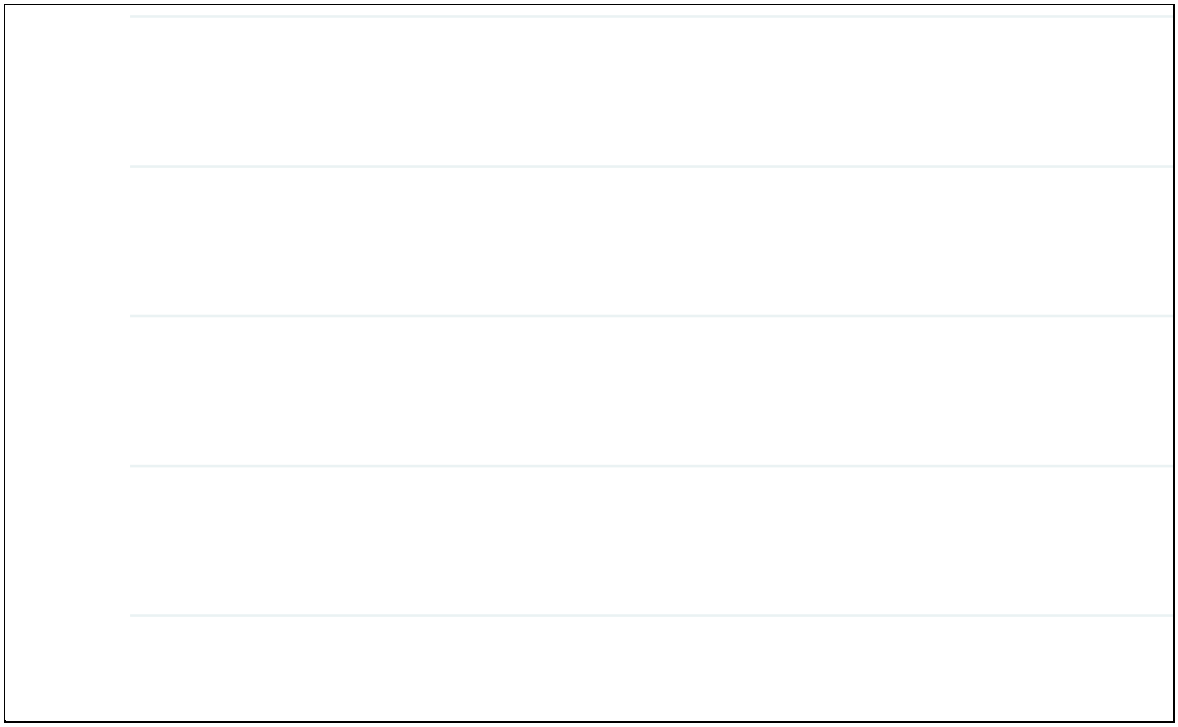


Figure 61 Reconstructed Kaplan-Meier plots and Weibull fits for time to cessation of infliximab treatment and time to loss of response requiring dose escalation of infliximab by four-week cycle (studies of Juillerat and Ma)

An alternative approach, used in sensitivity analysis, retained the Juillerat Weibull shape parameter, applied this for time to dose escalation and found the lambda parameter that generated 21% of the 130 cycle time spent in dose escalation (Figure 62).

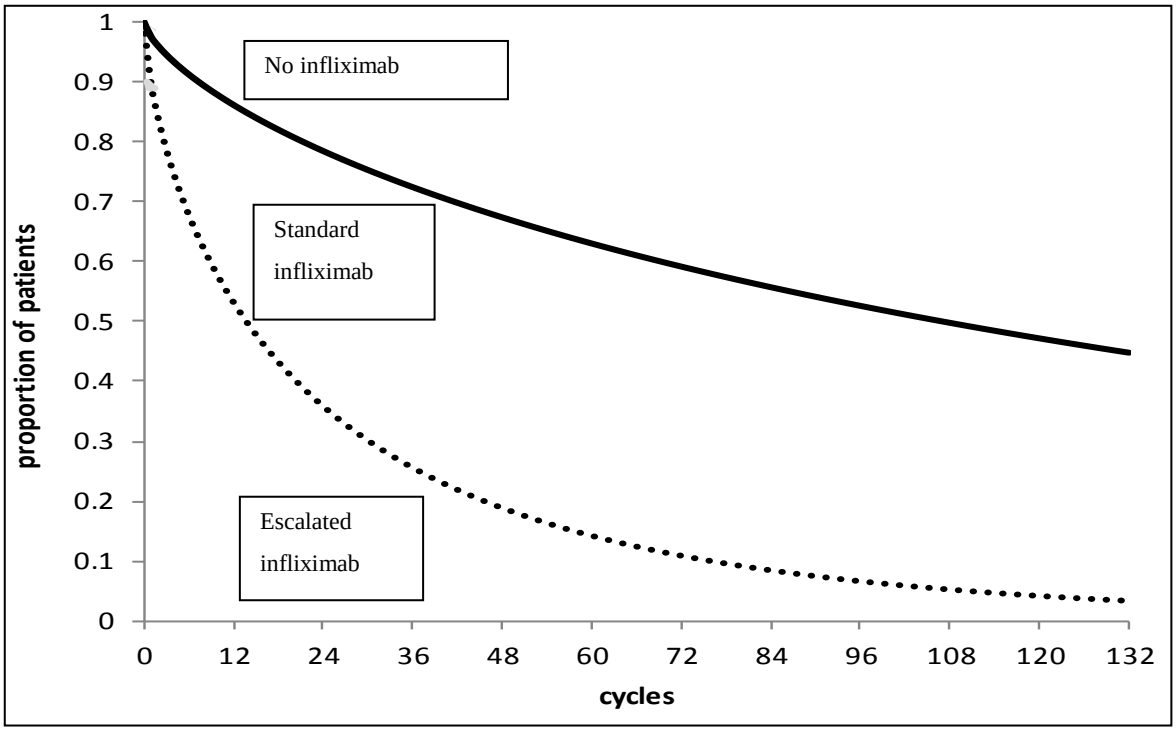


Figure 62 Partition of time over 132 cycles occupied by patients who stopped infliximab, continued on standard dose, and who dose escalated by four-week cycle
(Distributions based on Weibull fit to Juillerat data (shape parameter) and scale parameter for dose escalation that generates 21% of time in dose escalation)

Standard care arm treatment after loss of response to infliximab

On failure of response to infliximab maintenance (with or without dose escalation) it is assumed patients are switched to adalimumab induction therapy followed by maintenance on adalimumab for those responding to induction. We classify those that fail induction as patients that have lost response during the first cycle of treatment. We have taken this from the GAIN RCT (Sandborn et al., 2007)[154] that investigated adalimumab for patients that had failed infliximab. This provides a first cycle transition probability of 0.484. Thereafter the transition probability for loss of response to adalimumab was derived from the study by Karmaris et al. (2009)[47] of 152 CD responders receiving adalimumab followed up prospectively (Figure 63). Exponential and Weibull distributions provided a best fit to re-constructed individual patient data (IPD). The former generates a transition probability of 0.032369/cycle and this was employed in the economic model.

Figure 63 Reconstructed Kaplan-Meier and Weibull model for time to loss of response for CD patients on maintenance therapy with adalimumab by four-week cycle (Karmaris et al., 2009)[47]

After failure of adalimumab we have assumed patients remain in a loss of response state until such time that they receive surgery. This assumption was necessitated by lack of data and was based on advice of clinical experts. The transition to surgery was based on a large Canadian study[155] and is described below.

Time to surgery

No data was found for time to surgery for patients who experience loss of response or a failure to regain response after a treatment switch aimed to reinstate a response. We identified three studies[155, 170, 171] that provided time from diagnosis to surgery for recent cohorts of CD patients (i.e. coincident with the era of anti-TNFα therapies for CD). Vester-Anderson et al. (2014)[170] reported surgical relapse rates of 6%, 18% and 23% at 1, 5 and 7 years (91 cycles) after diagnosis; similarly a UK study[171] that included 137 patients observed approximately 24% primary surgery five years after diagnosis (Figure 64); a larger Canadian study[155] included >1000 patients and also data for recurrent surgery. Figure 64 shows the time to primary surgery was similar in the UK and Canadian studies and that Weibull distributions fitted the reconstructed IPD well. Because of reduced uncertainty in the large Canadian study this was used in the economic modelling for both primary and recurrent surgery.

Figure 64 Time from diagnosis of CD to primary surgery (upper) and to recurrent surgery (lower) by four-week cycle
Upper left UK study; upper right Canadian study

CD patients in the TAXIT and Steenholdt management studies varied considerably in the time from diagnosis to study entry and also in whether they had experienced previous surgery (e.g. in TAXIT patients on average were diagnosed 13.7 years prior to entry and 70% had received previous surgery; in Steenholdt patients were diagnosed an average of nine years before entry). Surgery was not a primary or secondary outcome measure in these studies but each reported that one patient received surgery (1/69 by week 20 in Steenholdt, and 1/251 by week 52 in TAXIT). It appears that during the short follow up periods observed the use of surgery was a relatively rare event. In the absence of more appropriate data we assumed the incidence of primary surgery was described by the Weibull distribution for the Canadian data and that the relevant patients had been diagnosed ten years before entering the model; thus the transition probabilities were calculated from the Weibull fit for the Canadian study between 10 years (130 cycles) and 20 years (260 cycles) post diagnosis. Patients with loss of response were assumed at risk of primary surgery irrespective of whether they had experienced surgery at some unspecified time previously (most likely soon after diagnosis according to the three studies described above). After primary surgery patients were assumed to be at risk of recurrent surgery. The transition probability for recurrent surgery was based the Gompertz fit to the Canadian recurrent surgery data between cycles 130 and 260. It is recognised that these selections are somewhat

arbitrary and that modelling extends beyond the observed data; therefore in sensitivity analysis we explored the effect of using the transition probabilities from cycle one after diagnosis.

Maintenance of surgery-induced remission

The scant evidence about maintenance of surgically induced remission in CD was reviewed by Gordon et al. (2014)[156] in a Cochrane systematic review. It should be noted that the authors' rated the included studies to be at high risk of bias for these outcomes. At two years across three studies there was no difference in risk of clinical relapse between patients receiving purine analogues and those receiving 5-ASA (fixed effects pooled RR = 1.01; 95% CI 0.81 to 1.24). The total events were 146 amongst 265 patients. Assuming a constant hazard the estimated transition probability to post-surgical clinical relapse is 0.023971/cycle (95% CI 0.025398 to 0.035624). In the economic model this was taken to apply for both therapies (ASA and purine analogues). Relative to purine analogues, the review data suggest patients receiving no therapy (placebo group in two studies) were at 1.35 (95% CI 1.06 to 1.72) greater risk of clinical relapse. Assuming a constant hazard provided an estimated transition probability of 0.050961/cycle (95% CI 0.033248 to 0.108412); this was used in the model for the group given no therapy. One study[172] included in the Gordon et al. (2014)[156] review found a relative risk for clinical relapse of 0.5 for infliximab versus purine analogues; this study observed only 3 events amongst 22 patients giving, on assumption of constant hazard, a transition probability to clinical relapse for infliximab treated patients of 0.0119855 / cycle.

In view of the considerable uncertainty necessarily associated with this estimate of response loss with infliximab, and the lack of information on timing of events, we looked for alternative data. Baert et al. (2014)[76] reported time to event data for reintroduction of infliximab following at least 15 months after loss of response despite dose optimization. During the $\geq$ 15 month infliximab holiday some patients received surgery. Time to loss of response after infliximab reintroduction is shown in Figure 65 together with the exponential fit used to estimate transition probabilities for the economic model. Due to lack of data we have assumed the same transition probabilities for patients receiving anti-TNFα in combination with immunomodulators to be the same as that for infliximab alone.

Figure 65 Reconstructed Kaplan Meier plot and Weibull fit for time to loss of response after reintroduction of infliximab after surgery by four-week cycle (based on data from Baert et al. (2014)[76])

Intervention arm: loss of response to test-directed infliximab maintenance

Only two management studies of infliximab responders were found and one of these, Vaughn et al. (2014)[127] was a retrospective study at considerable risk of selection bias such that the large reported advantage for the poorly-defined test-based strategy lacks face validity. Data from this study was used in sensitivity analysis (see below). The TAXIT[72] randomised controlled management study of responders to infliximab maintenance did not report time to loss of response. "Durable remission" amongst TAXIT IBD patients at week 52 post randomisation (13 cycles) was reported to be almost the same for test-algorithm strategy patients who were dose escalated, or who received no dose adjustment, or whose dose was reduced (28.6%, 26.4% and 25% respectively). On this basis we have assumed that loss of response was also unlikely to differ significantly between these groups. The P-value for the comparison of test-based dosing with clinically based dosing was 0.88. Of CD patients in the TAXIT intervention arm 79.77% were in clinical remission at randomisation and 62.6% in clinical and biological remission at week 52. There was no time to event data for clinical remission, however if a constant hazard is assumed for loss of remission the resulting transition probability is 0.018477165/cycle; Figure 67). This represents a very severe test for loss of response since patients without clinical remission are likely to be retained in anti-TNFα treatment because of a partial response. Therefore this was used only for sensitivity analysis and we looked at alternative data sources. The retrospective management study of Vaughn et al. (2014)[127] (Figure 66) reported vastly superior performance for 39 IBD patients receiving a test-algorithm strategy relative to 68 patients given clinically-based dosing strategy; when time to treatment cessation for these 39 patients was fitted with an exponential distribution a transition probability of only 0.003928414/cycle is generated

(Figure 67). These transition probabilities are very different and it is doubtful that either generates an appropriate input for the economic model.

Figure 66 Lognormal models for retention in infliximab maintenance therapy for infliximab responders (based on Vaughn et al., 2014[127] and used in sensitivity analysis)

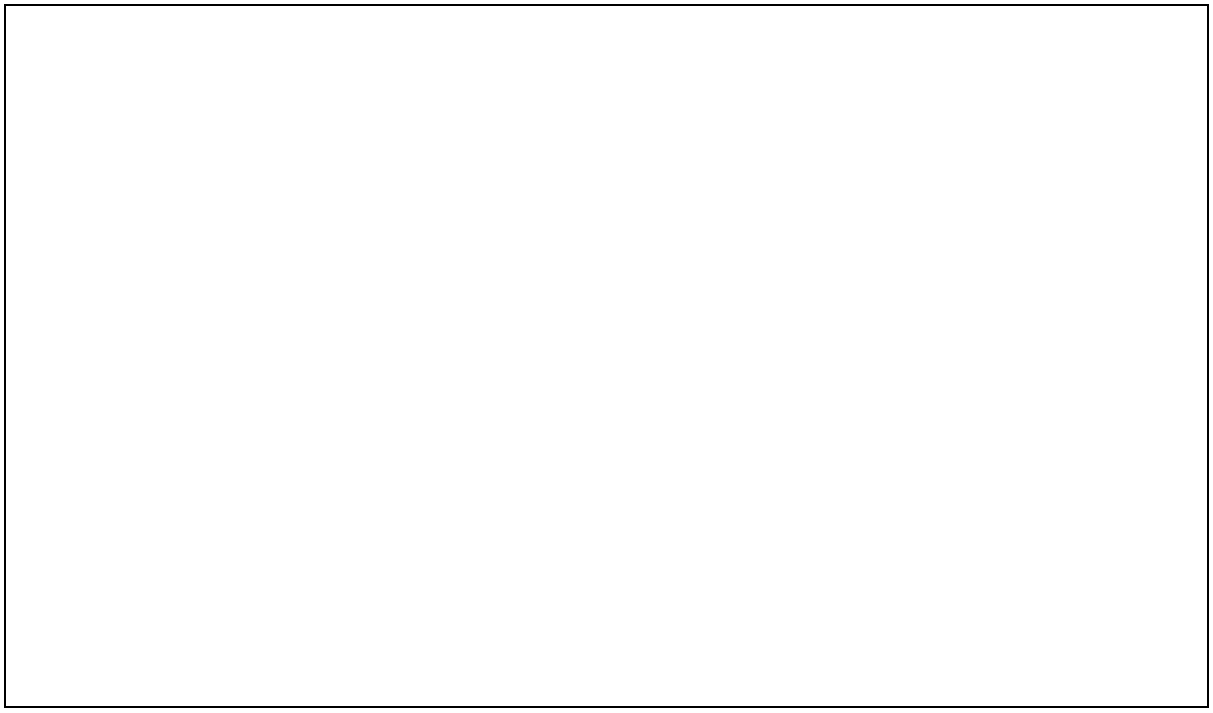


Figure 67 Time to event for Responders receiving a test-algorithm strategy. Time to clinical remission in TAXIT and retention in treatment in Vaughn et al. (2014)[127] by four-week cycle

In the TAXIT study before dose optimisation 131/178 (73.59%) CD patients were in clinical remission; after dose optimisation with a test-directed dose adjustments 138/173 (79.77%) were in remission (5 CD patients could not be optimised to target trough level). According to intention-to-treat analysis this represents a 3.9% improvement. With continued test-directed dosing post randomisation 62.6% of CD patients were in remission (clinical and biological) at 52 weeks, whereas 54.9% were in remission with clinically based post-randomisation dosing, implying a small advantage for the testing strategy (approximately 7.7%); P = 0.353 for comparison between groups. These small differences (3.9% and 7.7%) can be explained by the play of chance and are obviously associated with considerable uncertainty. We found no other evidence of clinical benefit from a test-algorithm based strategy. In the absence of other evidence demonstrating an advantage for a test-algorithm based strategy the model uses the same probability for loss of response to infliximab as used for the standard care arm (Weibull distribution fit to Juillerat data). In sensitivity analysis a 3.9% advantage for a testing strategy was implemented based on TAXIT data (see below).

Intervention arm: loss of response to test-directed infliximab maintenance sensitivity analysis

The retrospective study by Imaeda et al. (2012)[98] offers an additional data source (Section 3.2.5.2.3). In this, 58 patients received concurrent testing for infliximab and anti-drug antibodies and were classified according to maintenance or loss of response to infliximab, but test results did not inform patient management (patients received standard 5mg/kg every 6 to 8 weeks). The median time of follow up was not reported. For the whole group (n = 58) 70.69% retained response. When this proportion is anchored on the largest long term dataset for retention in infliximab treatment of responders (i.e. Juillerat et al. (2015)[151], see above) the cycles corresponding to this percent equal 39.74 cycles. This time is similar to that calculated from data in Imaeda et al. (2012)[98] which reports the mean number of infusions (= 22.52) with mean gap between infusions of 7 weeks (= 1.75 cycles), providing a mean follow up of 39.4 cycles (22.52 x 1.75). Using an assumed advantage of 3.9% for the test-directed strategy raises the proportion of responders by simple addition to 74.59%. The NICE Decision Support Unit[173] recommends that the same parametric form should be used for modelling intervention and comparator arms. We therefore adopted the Weibull shape parameter from Juillerat (standard care arm) and found the required scale parameter that delivered the 74.59% retained response for test strategy patients at the assumed follow up 39.4 cycles. These Weibull distribution transition probabilities were used in the economic model; for sensitivity analysis we used exponential models. Figure 68 Time to loss of response to maintenance infliximab in standard care and test-directed strategiesFigure 68 shows these distributions and compares them with those used for the standard care arm.

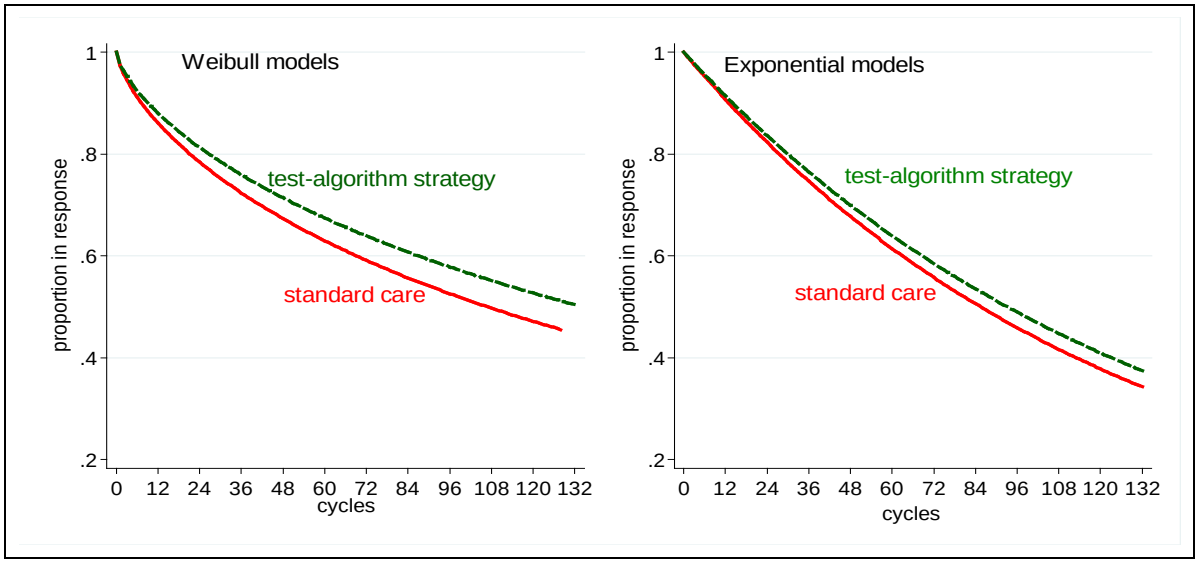


Figure 68 Time to loss of response to maintenance infliximab in standard care and test-directed strategies
Models are based on data from Juillerat[151], TAXIT and Imaeda[98] studies using Weibull and exponential distributions

Intervention arm: regain of response with test-directed treatments following response loss to maintenance infliximab

The treatments for patients with loss of response to maintenance infliximab were informed by the management study of Steenholdt et al. (2014)[122] (section 3.2.5). Patients enrolled in this study had failed infliximab maintenance in which patients received "regular infusions of 5mg/kg". It is recognised that this regimen does not exactly correspond to the dose being received by patients during the 52 weeks of the TAXIT trial in which dose was variously adjusted to bring trough infliximab to a target range. In Steenholdt, patients received concurrent testing at the time of infliximab failure and subsequent treatment followed an algorithm based on test results and was aimed at regaining response.

Concurrent testing identified four groups of intervention patients in the following proportions: [1] infliximab – / antibodies + , n = 5 (15.15%); [2] infliximab – / antibodies – , n = 1 (3.03%); [3] infliximab + / antibodies – , n = 26 (78.79%); [4] infliximab + / antibodies +, n = 1 (3.03%). The study reported the proportion that regained a response by 12 weeks but time to event data was not reported. We have assumed those who had not regained response by week 12 have lost response at a rapid rate over 3 cycles and remained in the non-response state (until surgery was implemented) and those that were in a response state at week 12 then proceeded to lose response at a given rate dependent on their algorithm-directed treatment regimen. The number of patients in all groups except group three was small so that outcomes are associated with great uncertainty. We have assumed that the single group four individual (positive test results for both infliximab and antibodies to infliximab) had the test results confirmed and was subsumed according to the treatment algorithm into group

three, which then accounts for 27/33 (81.8%) of intervention patients. Unfortunately, the various treatments used for the group three patients were insufficiently prescribed to be usable (e.g. surgery "should be considered").

For intervention group [1] patients (15.15% of infliximab failures), the algorithm-prescribed treatment was a switch to maintenance therapy with adalimumab; at 12 weeks 2/5 had regained response. This is a poor response rate but is based on only 5 patients and is uncertain. We have therefore used the same transition probabilities for these patients as for adalimumab-treated patients in the standard care arm (based on the GAIN RCT and on the study by Karmaris et al. (2009)[47] described above).

The single group [2] patient (3.03% of intervention patients) received infliximab intensification and failed to regain response by week 12. However all control arm patients in the trial also received infliximab intensification and at 12 weeks 19/36 had regained response; when combined with the single group two patient this provides an estimate of 19/37 (51.3%) in response at week 12 on infliximab intensification and 11/37 at 12 weeks in loss of response state. We assume the latter patients move to non-response at constant hazard over the first 3 cycles (12 weeks) providing a TP of 0.19948/cycle. Thereafter the rate of loss of regained response was assumed to be the same as that for dose escalated infliximab patients described by Ma et al. (2014)[153] (see above for the Weibull model based on data from Ma et al., 2014[153]).

In groups [3+4] (81.81% of infliximab failures) 16/27 had regained a response at 12 weeks and 11/27 were in a state of non-response. We assume that the latter group lost response at constant hazard over the 12 weeks providing a transition probability of 0.16004/cycle. Because the treatment for group [3] patients was not prescribed, other than that it lacked anti-TNF, we have assumed that after cycle 4 (12 weeks) loss of response occurs at constant hazard based on the Rutgeerts (1999) RCT[157]placebo arm (background therapies including purine analogues, steroids, methotrexate and 5-ASA) in which about half of patients had previously received previous anti-TNF therapy. This suggested a transition probability of 0.08617343/cycle.

In the Steenholdt study[122] about half of group 3] patients likely received infliximab in contradiction to the specified treatment according to the algorithm. The footnote[1] indicates the various treatments

[1] Of these, 12 patients continued IFX (9 patients were in group 3, and 1patient was in group 4). The applied infliximab (IFX) regimen was (all received 5 mg/kg):
- IFX q8 regimen (2 infusions during the trial, i.e. week 0 and 8): n=5
- IFX q4 regimen (4 infusions during the trial, i.e. week 0,4,8,12): n=2
- IFX q4 regimen but not throughout the entire trial (3 infusions during the trial): n=1
- IFX q4 regimen but not throughout the entire trial (2 infusions during the trial): n=2
- IFX q4 regimen but not throughout the entire trial (1 infusions during the trial): n=2

The remaining 2 patients had been switched to ADL due to misinterpretation of test results (Figure 2). Both patients were in group 3.
The applied ADL regimen was:
- ADL induction (160-80-40) and followed by 40 mg every other week.
- ADL induction (80-40) and followed by 40 mg every other week.

received by the 14 patients in the intervention arm who did not receive algorithm-directed treatments. In view of these difficulties we undertook a sensitivity analysis by redeploying patients in group 3] to groups 1] and 2]. The proportion in group 3+4] was reduced from 81.8% and set at 20%, and the proportions in groups 1] and 2] raised to 66.67% and 13.33% respectively.

The people with LOR from all groups remain on palliative care in a loss of response state until surgery. It is possible that some of these patients (and also those people with LOR after adalimumab in the standard care arm), at some time may be reintroduced to infliximab (or possibly adalimumab) prior to surgery and may regain response, however lack of evidence precluded modelling this. We have assumed that after surgery various treatments are administered in attempts to maintain post-surgical remission and that these are the same as for the standard care arm (see above).

8.18 Appendix 18 Resource use data

In this Appendix, we report on the unit costs derived for monitoring infliximab and antibodies to infliximab, treatment costs for people receiving infliximab maintenance therapy and cost of a surgical procedure.

Table 55 Unit costs for monitoring infliximab and antibodies to infliximab

Resource use	Quantity	Description	Unit costs (£, 2014)	Source
LISA TRACKER for monitoring infliximab and antibodies to infliximab (concurrent testing)				
Assay kit used to monitor infliximab and antibodies to infliximab (concurrent testing)	1	Total cost of kits for monitoring infliximab and antibodies to infliximab is £1568. Number of patient samples per kit is 42	37.33	NICE
Lab technician	1	Assay takes three hours to perform in the lab. Based on a clinical support worker as a proxy (£21 per hour)	1.50	Curtis 2014
LISA TRACKER for monitoring infliximab and antibodies to infliximab (reflex testing)				
Assay kit used to monitor infliximab	1	Total cost of kit for monitoring infliximab is £850. Number of patient samples per kit is 42	20.24	NICE
Assay kit used to monitor antibodies to infliximab	1	Total cost of kit for monitoring antibodies to infliximab is £850. Number of patient samples per kit is 42	20.24	NICE
Lab technician	1	Assay takes three hours to perform in the lab. Based on a clinical support worker as a proxy (£21 per hour)	1.50	Curtis 2014
Estimated total cost for monitoring infliximab and antibodies per person (concurrent testing)				38.83
Estimated total cost for monitoring infliximab and antibodies to infliximab per				43.48

person (reflex testing)	
Estimated total cost for monitoring infliximab per person	21.74

Table 56 Treatment of Crohn's disease with infliximab and adalimumab

Resource use	Quantity	Description	Unit costs (£,2014)	Source
Infliximab treatment				
Infliximab (Remicade)	1	5-mg/kg intravenous infusion over a 2-hour period every 8-weeks 100mg/vial = £419.62 Four vials required 4*£419.62 = £1678.48	1678.48	BNF 2013/14[158]
Administration cost	1		287.93	Curtis 2014[160]
Estimated cost per individual receiving infliximab maintenance therapy every eight weeks				**1966.41**
Adalimumab treatment				
Adalimumab (Humira)	1	40mg every two weeks	704.28	BNF 2013/14
Estimated cost per individual receiving infliximab maintenance therapy				
Estimated cost per individual receiving infliximab maintenance therapy				
[a] People on maintenance therapy receiving infliximab treatment is given 5-mg/kg intravenous infusion over a 2-hour period every 8-weeks. We assumed that people are on average weighing 70kg				

Table 57 C**ost of a surgical procedure**

Resource use	Quantity	Description	Unit costs (£,2014)	Source
Investigations				
Laparoscopic ilecolic resection	1	FZ74F Elective inpatients - complex large intestine procedures, 19 years +, with CC score 0-2	6803	NHS reference costs 2013/14[159]
Outpatient visits (follow-up consultation)	1	WF01A Colorectal surgery - consultant led outpatient attendance non-admitted	105	NHS reference costs 2013/14[159]
Cost of laparoscopic ileocolic resection				**£6908**

Table 58 Additional costs associated with occupying health states

Health state	Quantity	Description	Unit costs (£,2014)	Source
Responder				
Outpatient visits	2	WF01A Colorectal surgery - consultant led outpatient attendance non-admitted	105	NHS reference costs 2013/14[159] and expert opinion
Colonoscopy	1	Weighted average of NHS reference cost outpatient for FZ51Z diagnostic colonoscopy without biopsy, or FZ52Z diagnostic colonoscopy with biopsy	370.69	
MRI	1	Outpatient RA01A MRI scan	145	
Cost for the responder health state				**725.69**
Regain response				
Outpatient visits	2	WF01A Colorectal surgery - consultant led outpatient attendance non-admitted	105	NHS reference costs 2013/14[159] and expert opinion
Colonoscopy	1	Weighted average of NHS reference cost outpatient for FZ51Z diagnostic colonoscopy without biopsy, or FZ52Z diagnostic colonoscopy with biopsy	370.69	
MRI	1	Outpatient RA01A MRI scan	145	
Cost for the regain response health state				**725.69**
Loss of response				
Outpatient visits	2	WF01A Colorectal surgery - consultant led outpatient attendance non-admitted	105	NHS reference costs 2013/14[159] and expert opinion
Colonoscopy	2	Weighted average of NHS reference cost	370.69	

Health state	Quantity	Description	Unit costs (£,2014)	Source
		outpatient for FZ51Z diagnostic colonoscopy without biopsy, or FZ52Z diagnostic colonoscopy with biopsy		
MRI	2	Outpatient RA01A MRI scan	145	
Cost for the loss of response health state				**1241.38**
Post-surgery (remission)				
Outpatient visits	4	WF01A Colorectal surgery - consultant led outpatient attendance non-admitted	105	NHS reference costs 2013/14[159] and expert opinion
Colonoscopy	1	Weighted average of NHS reference cost outpatient for FZ51Z diagnostic colonoscopy without biopsy, or FZ52Z diagnostic colonoscopy with biopsy	370.69	
Cost for the post-surgery (remission) health state				**790.69**

Diagnostic Assessment Report commissioned by the NIHR HTA Programme on behalf of the National Institute for Health and Care Excellence – Additional sensitivity and scenario analyses

Title of project:

Clinical and cost-effectiveness of use of therapeutic monitoring of TNFα inhibitors (LISA-TRACKER ELISA kits, TNFα-Blocker ELISA kits, and Promonitor ELISA kits) versus standard care in people with Crohn's disease: systematic reviews and economic modelling

Name of External Assessment Group (EAG) and project lead:

Produced by: Warwick Evidence

Lead author: Karoline Freeman

Co-authors: Martin Connock
Peter Auguste
Sian Taylor-Phillips
Hema Mistry
Deepson Shyangdan
Rachel Court
Ramesh Arasaradnam
Paul Sutcliffe
Aileen Clarke

Correspondence to: Dr Paul Sutcliffe
Associate Professor
Deputy Director for Warwick Evidence
Populations, Evidence and Technologies
Division of Health Sciences
Warwick Medical School
University of Warwick
Coventry CV4 7AL

Tel: 02476 150189

Fax: 02476 528375

Email: p.a.sutcliffe@warwick.ac.uk

Additional sensitivity analyses requested by NICE

The base case in our model is the concurrent testing strategy for a) responders and b) people who have lost response with ongoing 3-monthly testing. NICE has requested that we investigate one-off testing. We therefore looked at the following scenarios:

a) In the responders model - 3 possible modes of one-off testing

1. One-off testing at three months followed by yearly retesting or
2. One-off testing at three months and one retest for those who regained response or
3. One-off testing at three months and no retesting for responders / regained response

b) In the loss of response model - 3 monthly testing for patients with loss of response (LOR); no testing of people who have regained response

c) In the loss of response model - probabilistic sensitivity analysis

d) We have also undertaken one further set of analyses changing time to event transition probabilities to exponential transition probabilities

a) Responders model:

The only data available for this question are from the TAXIT trial (evidence for responders on an anti-TNFα drug (infliximab), however neither arm of the trial represents the one-off testing strategy at 3-4 months, or a strategy with annual retesting as suggested by the decision question. In the TAXIT trial all included participants received testing for dose optimisation. Over 50% of people required more than one dose optimisation step to reach the target infliximab trough level. After dose optimisation patients were randomised either to receive clinically based or concentration-based dosing of infliximab.

1. Sensitivity analyses of one-off testing and 3-month testing followed by yearly retesting

In the responder model we have undertaken an analysis where people (responder/regain response) were tested at three months followed by 12-montly testing after commencing treatment. The results showed that reflex testing dominated concurrent testing being less cheaper and more effective. The no testing strategy was more expensive and produced more QALYs with a reported ICER of approximately £132,800 per QALY when compared to the reflex testing strategy (Table 1).

Table 1 Testing at 3 month followed by 12-montly re-testing for responders

Exponential

Testing strategy	Mean cost per strategy	Difference in costs	Effectiveness (QALYs)	Incremental QALYs	Incremental cost-effectiveness ratio (£) (ICER)
Testing at three months followed by 12-months after commencing treatment					
Reflex testing	113,400	-	6.2290	-	
Concurrent testing	113,800	800	6.2244	-0.0046	Dominated
No testing	150,500	37,100	6.5084	0.2794	132,800

2. One-off testing at three months for responders and those who regained response and three-monthly testing for those who lost response

In this analysis in the responder model, testing was undertaken once, at three months for responders and those who regained response whilst people who lost response were tested three-monthly. The results showed that testing had similar mean costs (102,000 vs 103,000) and QALYs (6.2255 vs 6.2390). The no testing strategy had an estimated ICER of approximately £176,300 per QALY (Table 2).

Table 2 Testing at 3 months for responders and one re-test for people who have regained response

Exponential					
Testing strategy	**Mean cost per strategy**	**Difference in costs**	**Effectiveness (QALYs)**	**Incremental QALYs**	**Incremental cost-effectiveness ratio (£) (ICER)**
Testing at three months for responders and regain response					
Concurrent testing	102,000	-	6.2255	-	-
Reflex testing	103,000	1000	6.2390	0.0135	74,100
No testing	150,500	47,500	6.5084	0.2694	176,300

3. One-off testing at three months and no retesting for responders / regained response

In this analysis in the responder model responders were tested once at three months and people who lost response were tested three-monthly. People who regained response were not re-tested. The results showed that testing had similar mean costs (£102,000 vs £102,900) and QALYs (6.2255 vs 6.2255). The no testing strategy had an estimated ICER of approximately £176,700 per QALY (Table 3).

Table 3 One-off testing for responders

Exponential					
Testing strategy	**Mean cost per strategy**	**Difference in costs**	**Effectiveness (QALYs)**	**Incremental QALYs**	**Incremental cost-effectiveness ratio (£) (ICER)**
Testing responders at three months only					
Concurrent testing	102,000	-	6.2255	-	-
Reflex testing	102,900	900	6.2255	0.0135	66,700
No testing	150,500	47,600	6.5084	0.2694	176,700

Interpretation

We found that the one-off testing strategies did not alter the findings from the base case as presented in the report i.e. a no testing strategy is more expensive generating more QALYs. A concurrent testing strategy remains generally the most cost effective.

As we have had to base these additional sensitivity analyses on data from the TAXIT trial, the findings should be treated with caution. The study states that "in the concentration-based dosing group, individual infliximab trough concentrations were *evaluated at each infusion…*" In fact this means that in the trial patients might be tested 8-weekly if they reached the target concentration after the initial dosing, 12-weekly if dose reduction was required more than once, or 4 weekly if dose increases were required (by shortening the infusion interval). We therefore assumed in our original analyses that three monthly testing is a fair estimate to reflect the trial without over testing. The one-off testing strategy does not correspond with any available data. We do not therefore know what clinical algorithm might be followed or what clinical consequences might then follow from this unknown algorithm for a responder who gets tested once.

b) LOR model: 3 monthly testing for patients with loss of response (LOR); no testing for people who have regained response

The only available comparative evidence for people with LOR comes from the RCT by Steenholdt et al. (2014) where people with LOR were tested once and followed up for 12 weeks. There is no available evidence that suggests the frequency of further testing. However, no further testing does not seem plausible in the treatment arm and with clinical advice we compromised on a three-monthly retesting strategy to avoid over testing. In this additional sensitivity analysis requested by NICE we investigated the option that only people with LOR receive testing. Patients with regained response were not retested. Table 4 summarises the results of this analysis.

Table 4 Regular testing of patients with LOR, with no testing for people who have regained response

Exponential					
Testing strategy	**Mean cost per strategy**	**Difference in costs**	**Effectiveness (QALYs)**	**Incremental QALYs**	**Incremental cost-effectiveness ratio (£) (ICER)**
Three monthly testing for people who lose response whilst being treated with anti-TNF					
Concurrent testing	96,200	-	6.1453	-	-
Reflex testing	97,700	1500	6.1630	0.0177	84,700
No testing	215,800	118,100	6.4961	0.3331	354,500

Interpretation

In the loss of response model we assumed three-monthly testing for people who lost response while being treated with anti-TNF, and no testing for people who subsequently regained response. These results showed that concurrent testing was both cheaper and less effective than the reflex testing

strategy with estimated mean costs of approximately £96,200 vs. £97,000 and QALYs (6.1453 vs. 6.1630) with a reported ICER of £84,700. The no testing strategy was more costly (215,800) and more effective (6.4961) with an ICER of approximately £354,500 per QALY compared to the reflex testing strategy.

c) Probabilistic sensitivity analysis for LOR model

As requested by NICE a PSA analysis has been undertaken for the LOR model with the following results:

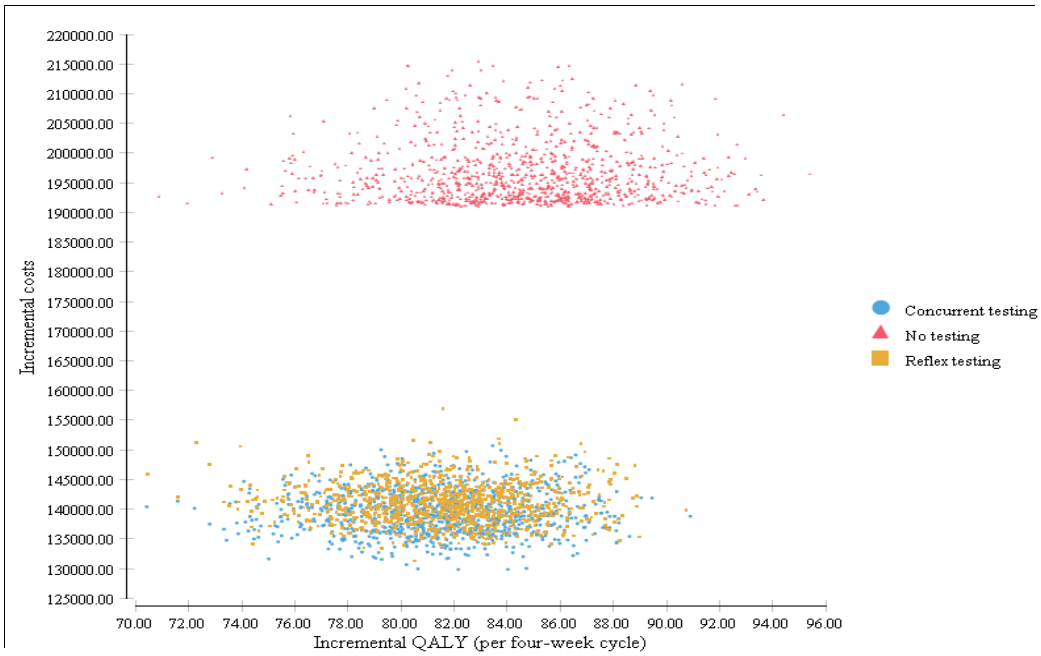


Figure 1: Probabilistic sensitivity analysis results for concurrent and reflex testing and no testing in the loss of response model. Scatterplot using distributions around model input parameters

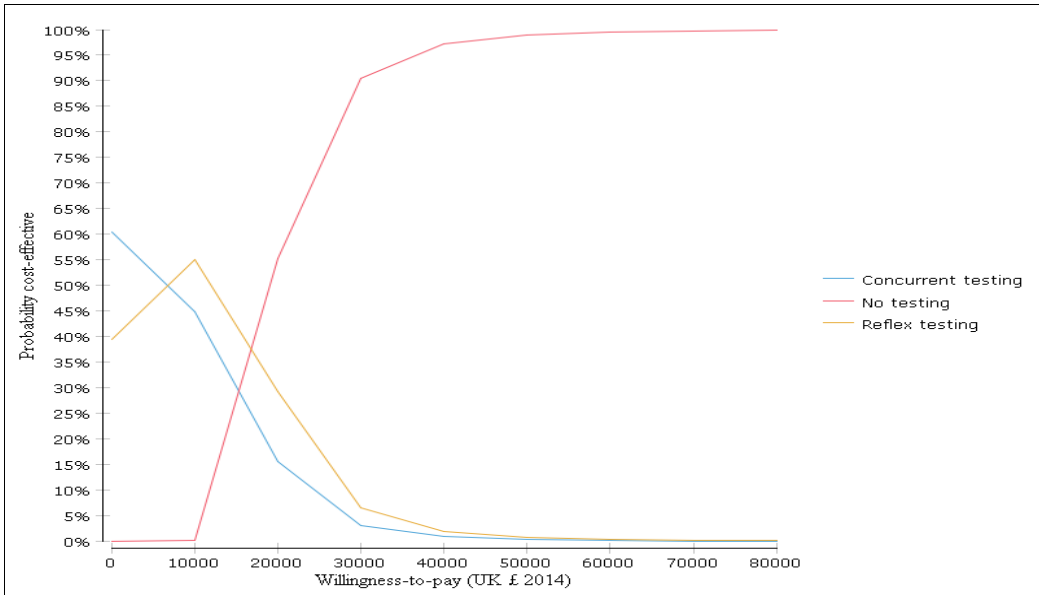


Figure 2: Cost-effectiveness acceptability curve using distributions around outcomes in the loss of response model

Results and interpretation

Figure 1 shows the Monte Carlo simulation for the loss of response model. The scatterplot shows the uncertainty in expected incremental costs and QALYs associated with testing (concurrent and reflex) and no testing. The scatterplot shows considerable uncertainty about the expected incremental costs and incremental QALYs. The results for the loss of response model are presented in the form of cost-effectiveness acceptability curves (Figure 2). These results suggest that that no testing is more expensive. At a willingness to pay of £20,000 per QALY no testing is 50% likely to be the most cost effective strategy.

Changing time to event transition probabilities to exponential transition probabilities

As a further sensitivity analysis we undertook additional analyses are based on exponential transition probabilities (constant hazard over time) as opposed to time to event transition probabilities used in the final report. Further assessment of the model post-submission confirmed that constant hazards over time appear to be more appropriate for the model and the model has been re-run with the new transition probabilities. Table 5 below summarises the results in light of the changes, showing that concurrent testing remains in almost all cases the best strategy with very similar QALYs generated as in the time to event transition probabilities but higher costs in the no testing arms.

Table 5 Model results using time to event versus exponential transition probabilities

Time to event						**Exponential**					
Testing strategy	**Mean cost per strategy**	**Difference in costs**	**Effectiveness (QALYs)**	**Incremental QALYs**	**Incremental cost-effectiveness ratio (£) (ICER)**	**Testing strategy**	**Mean cost per strategy**	**Difference in costs**	**Effectiveness (QALYs)**	**Incremental QALYs**	**Incremental cost-effectiveness ratio (£) (ICER)**
Base case (responder)						**Base case (responder)**					
No testing	137,600	-	6.5146	-	-	Reflex testing	138,700		6.2761		
Concurrent testing	145,900	8300	6.3315	0.1831	Dominated	Concurrent testing	139,800	1100	6.2637	-0.0124	Dominated
Reflex testing	147,100	9500	6.3215	0.1931	Dominated	No testing	150,500	11,800	6.5084	0.2323	50,800
Base case (loss of response)						**Base case (loss of response)**					
Concurrent testing	139,200	-	6.2600	-	-	Concurrent testing	129,400	-	6.1807	-	-
Reflex testing	140,300	1100	6.2715	0.0115	95,700	Reflex testing	131,000	1600	6.1976	0.0169	94,700
No testing	199,900	59600	6.5031	0.2316	257,340	No testing	215,800	84,800	6.4961	0.2985	284,100
Annual testing in responder model						**Annual testing in responder model**					
Concurrent testing	116,300	-	6.2446	-	-	Concurrent testing	114,000	-	6.2201	-	-
Reflex testing	116,400	100	6.2508	0.0062	16,100	Reflex testing	114,100	100	6.2281	0.0080	12,500
No testing	137,600	21,200	6.5146	0.2638	80,400	No testing	150,500	36,400	6.5084	0.2803	129,900
Annual testing in loss of response model						**Annual testing in loss of response model**					
Concurrent testing	111,200	-	6.1877	-	-	Concurrent testing	106,900	-	6.1406	-	-
Reflex testing	112,100	900	6.1946	0.0069	130,400	Reflex testing	108,100	1200	6.1532	0.0126	95,200
No testing	199,900	87,800	6.5031	0.3085	284,600	No testing	215,800	107,700	6.4961	0.3429	314,100
One-year time horizon in responder model						**One-year time horizon in responder model**					
No testing	15,200	-	0.8269	-	-	No testing	14,900	-	0.7686	-	-
Concurrent testing	20,300	5100	0.8085	-0.0184	Dominated	Reflex testing	18,500	3600	0.7549	-0.0137	Dominated
Reflex	20,400	5200	0.8092	-0.0177	Dominated	Concurrent	19,200	4300	0.7543	-0.0143	Dominated

testing						testing					
One-year time horizon in loss of response model						**One-year time horizon in loss of response model**					
Concurrent testing	14,200	-	0.7531	-	-	Concurrent testing	12,000	-	0.6870	-	-
Reflex testing	14,600	400	0.7562	0.0031	129,000	Reflex testing	12,500	500	0.6915	0.0045	111,100
No testing	23,400	8800	0.8154	0.0592	148,600	No testing	23,500	11,000	0.7560	0.0645	170,500
Three monthly testing for people who lose response whilst being treated with anti-TNF						**Three monthly testing for people who lose response whilst being treated with anti-TNF**					
Concurrent testing	101,000	-	6.2092	-	-	Concurrent testing	96,200	-	6.1453	-	-
Reflex testing	102,200	1200	6.2223	0.0131	91,600	Reflex testing	97,700	1500	6.1630	0.0177	84,700
No testing	199,900	97,700	6.5031	0.2808	347,900	No testing	215,800	118,100	6.4961	0.3331	354,500
Testing at three months followed by 12-months after commencing treatment						**Testing at three months followed by 12-months after commencing treatment**					
Concurrent testing	116,000	-	6.2533	-	-	Reflex testing	113,400	-	6.2290	-	
Reflex testing	116,100	100	6.2597	0.0064	15,600	Concurrent testing	113,800	800	6.2244	-0.0046	Dominated
No testing	137,600	21,500	6.5143	0.2546	84,500	No testing	150,500	37,100	6.5084	0.2794	132,800
Testing at three months for responders and regain response						**Testing at three months for responders and regain response**					
Concurrent testing	105,500	-	6.2745	-	-	Concurrent testing	102,000	-	6.2255	-	-
Reflex testing	106,300	800	6.2856	0.0111	72,100	Reflex testing	103,000	1000	6.2390	0.0135	74,100
No testing	137,600	31,300	6.5143	0.2287	136,900	No testing	150,500	47,500	6.5084	0.2694	176,300
Testing responders at three months only						**Testing responders at three months only**					
Concurrent testing	105,500	-	6.2745	-	-	Concurrent testing	102,000	-	6.2255	-	-
Reflex testing	106,300	800	6.2856	0.0111	72,100	Reflex testing	102,900	900	6.2390	0.0135	66,700
No testing	137,600	31,300	6.5143	0.2287	136,900	No testing	150,500	47,600	6.5084	0.2694	176,700
No regain response following best supportive care (responders)						**No regain response following best supportive care (responders)**					
Concurrent testing	86,900	-	5.7472	-	-	No testing	150,550	-	6.5084	-	-
Reflex testing	87,700	800	5.7760	0.0288	27,800	Reflex testing	158,300	7750	6.4813	-0.02710	Dominated
No testing	137,600	49,900	6.5143	0.7383	67,600	Concurrent	160,800	10,250	6.4813	-0.00001	Dominated

						testing					
No regain response following best supportive care (loss of response)						**No regain response following best supportive care (loss of response)**					
Concurrent testing	49,700	-	5.4154	-	-	Concurrent testing	54,000	-	5.4649	-	-
Reflex testing	53,000	3300	5.4446	0.0292	113,000	Reflex testing	57,700	3700	5.4992	0.0343	107,900
No testing	199,900	164,900	6.5031	1.0585	155,800	No testing	215,700	158,000	6.4961	0.9969	158,500

Figures 3 to 6 present the probabilistic sensitivity analysis and cost-effectiveness acceptability curve for the responder and LOR model using constant hazard transition probabilities.

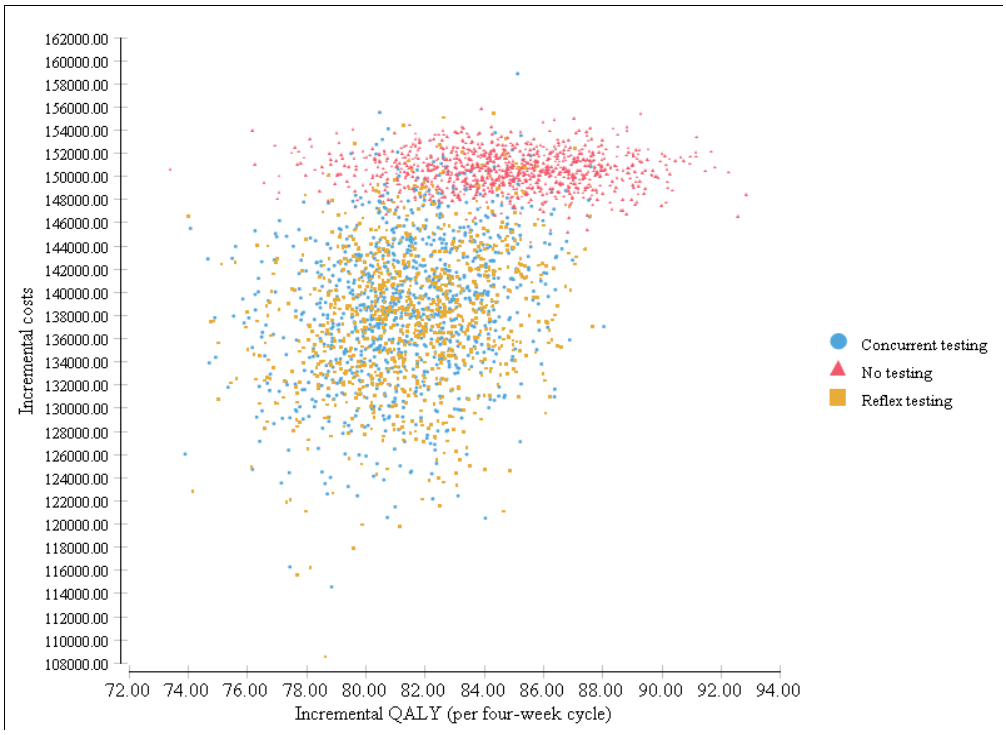


Figure 3: Probabilistic sensitivity analysis results for concurrent and reflex testing and no testing in the response model. Scatterplot using distributions around model input parameters (constant hazard transition probabilities)

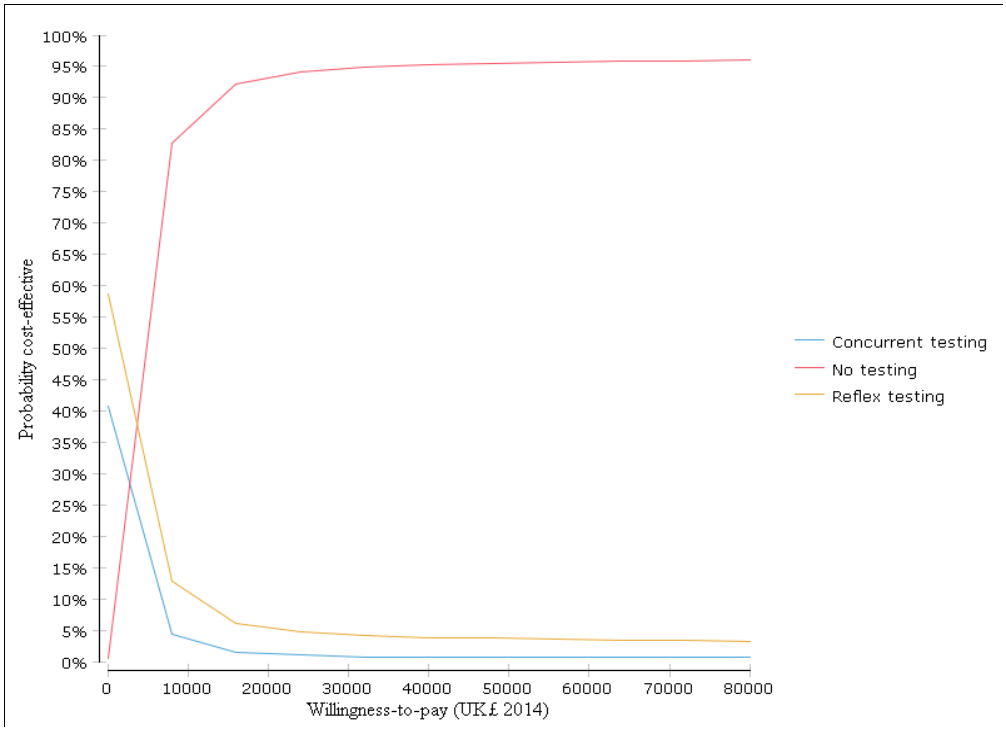


Figure 4: Cost-effectiveness acceptability curve using distributions around outcomes in the response model (constant hazard transition probabilities)

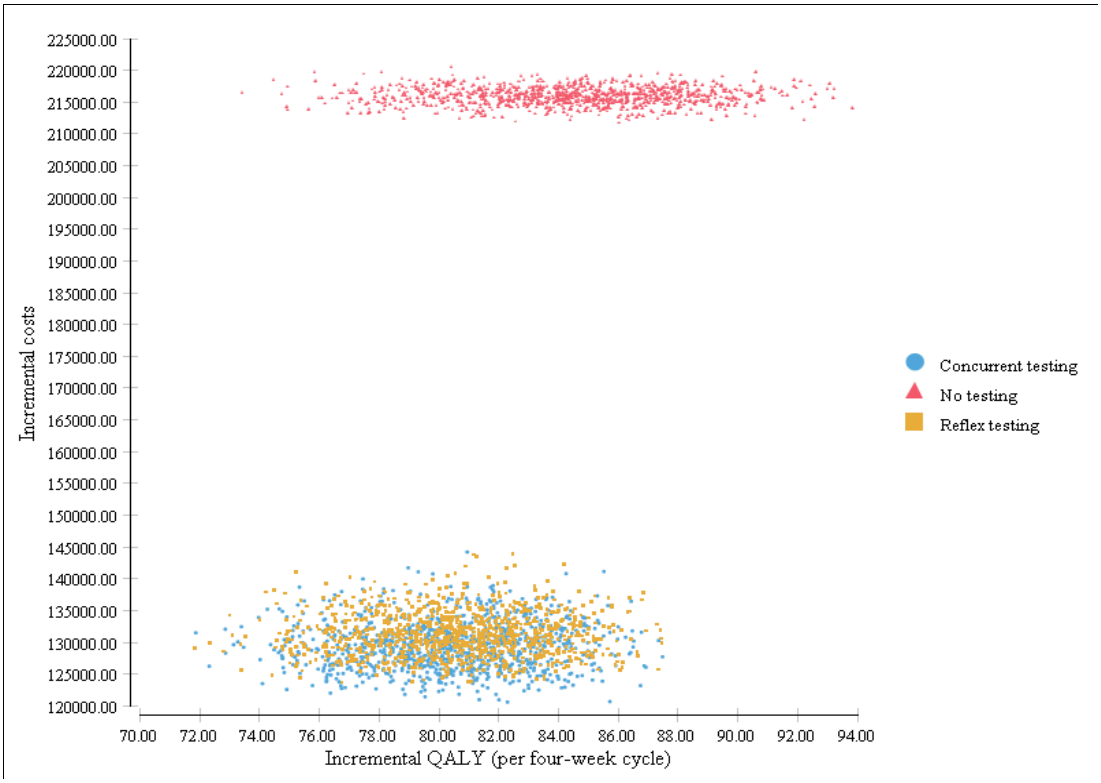


Figure 5: Probabilistic sensitivity analysis results for concurrent and reflex testing and no testing in the loss of response model. Scatterplot using distributions around model input parameters (constant hazard transition probabilities)

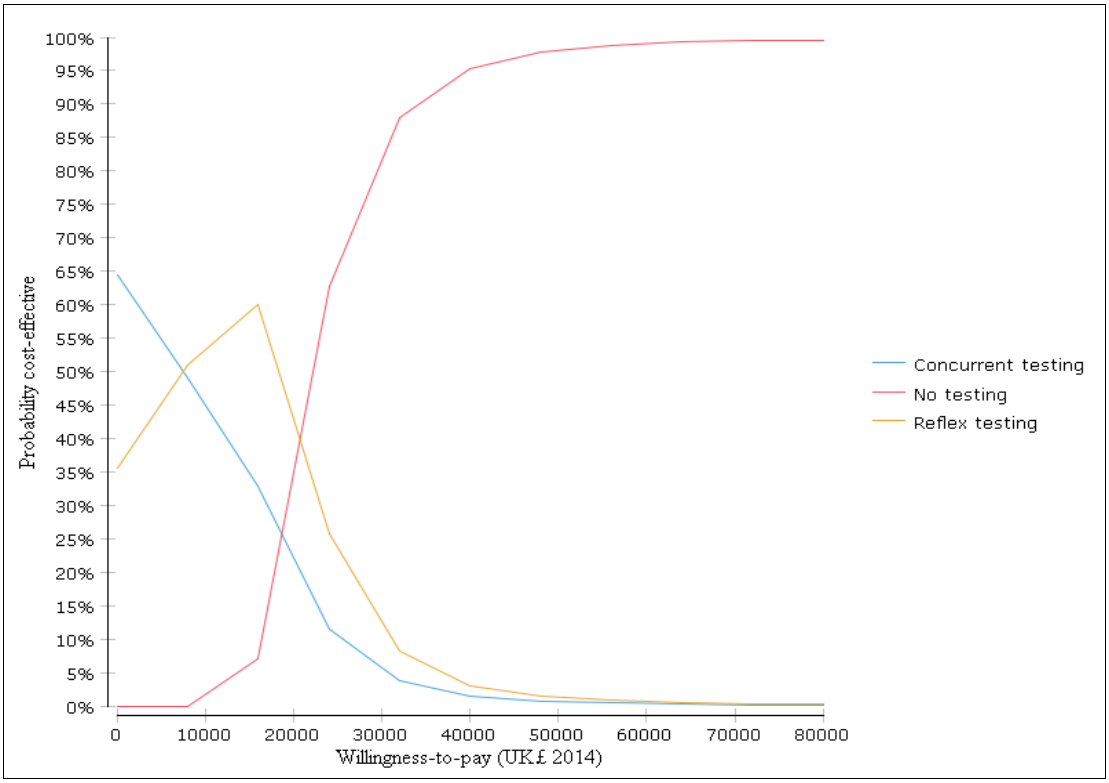


Figure 6: Cost-effectiveness acceptability curve using distributions around outcomes in the loss of response model (constant hazard transition probabilities)

Figure 3 shows the Monte Carlo simulation for the response model using constants hazard transition probabilities. The scatterplot shows the uncertainty in expected incremental costs and QALYs associated with testing (concurrent and reflex) and no testing. The scatterplot shows considerable uncertainty about the expected incremental costs and incremental QALYs. In figure 4 we present the results for the response model in the form of cost-effectiveness acceptability curves. The results suggest that there is a clear preference for no testing in the response. This is reflected in the CEAC which suggests that at a willingness-to-pay threshold of £20,000 per QALY no testing is 92% cost-effective when compared to testing.

Figures 5 and 6 show the scatterplot and CEAC, respectively, for the 10,000 Monte Carlo simulations of the loss of response model using constant hazard transition probabilities. Results from the CEAC shows that at a willingness-to-pay threshold of £20,000 per QALY there is no preference between no testing and reflex testing strategies. However, at higher willingness-to-pay thresholds (e.g. >£30,000 per QALY) no testing is likely to be the most cost-effectiveness strategy.

Diagnostic Assessment Report commissioned by the NIHR HTA Programme on behalf of the National Institute for Health and Care Excellence – Table of errata

Title of project:

Clinical and cost-effectiveness of use of therapeutic monitoring of TNFα inhibitors (LISA-TRACKER ELISA kits, TNFα-Blocker ELISA kits, and Promonitor ELISA kits) versus standard care in people with Crohn's disease: systematic reviews and economic modelling

Name of External Assessment Group (EAG) and project lead:

Produced by: Warwick Evidence

Lead author: Karoline Freeman

Co-authors: Martin Connock
Peter Auguste
Sian Taylor-Phillips
Hema Mistry
Deepson Shyangdan
Rachel Court
Ramesh Arasaradnam
Paul Sutcliffe
Aileen Clarke

Correspondence to: Dr Paul Sutcliffe
Associate Professor
Deputy Director for Warwick Evidence
Populations, Evidence and Technologies
Division of Health Sciences
Warwick Medical School
University of Warwick
Coventry CV4 7AL

Tel: 02476 150189

Fax: 02476 528375

Email: p.a.sutcliffe@warwick.ac.uk

Corrections to the main document are underlined.

DAR section	Page number	DAR text	Corrected text
4.3.1.2	147	‘administration costs were also included for adalimumab’	‘administration costs were also included for infliximab’
4.3.4.2.2	173	‘for people switching to adalimumab, we derived a cost of £1408.28 (2x £704.28, assuming 40mg of adalimumab is required every two weeks) per four-week cycle’	‘for people switching to adalimumab, we derived a cost of £704.28 (2 x £352.14; assuming 40mg of adalimumab is required every two weeks) per four-week cycle
4.3.4.2.2 Table 35	173	Base case value maintenance adalimumb: 704.28’	Base case value maintenance adalimumb: 352.14
8.18 Table 56	376	unit cost for adalimumab 40mg every other week: 704.28	unit cost for adalimumab 40mg every other week: 352.14

Addendum Table 5 Model results using time to event versus exponential transition probabilities

Time to event						**Exponential**					
Testing strategy	**Mean cost per strategy**	**Difference in costs**	**Effectiveness (QALYs)**	**Incremental QALYs**	**Incremental cost-effectiveness ratio (£) (ICER)**	**Testing strategy**	**Mean cost per strategy**	**Difference in costs**	**Effectiveness (QALYs)**	**Incremental QALYs**	**Incremental cost-effectiveness ratio (£) (ICER)**
No regain response following best supportive care (responders)						**No regain response following best supportive care (responders)**					
Concurrent testing	86,900	-	5.7472	-	-	No testing	150,550	-	6.5084	-	-
Reflex testing	87,700	800	5.7760	0.0288	27,800	Reflex testing	158,300	7750	6.4813	-0.02710	Dominated
No testing	137,600	49,900	6.5143	0.7383	67,600	Concurrent testing	160,800	10,250	6.4813	-0.00001	Dominated

Errata for Addendum Table 5 Model results using time to event versus exponential transition probabilities

Time to event						**Exponential**					
Testing strategy	**Mean cost per strategy**	**Difference in costs**	**Effectiveness (QALYs)**	**Incremental QALYs**	**Incremental cost-effectiveness ratio (£) (ICER)**	**Testing strategy**	**Mean cost per strategy**	**Difference in costs**	**Effectiveness (QALYs)**	**Incremental QALYs**	**Incremental cost-effectiveness ratio (£) (ICER)**
No regain response following best supportive care (responders)						**No regain response following best supportive care (responders)**					
Concurrent testing	86,900	-	5.7472	-	-	Reflex testing	87,900	-	5.7853	-	-
Reflex testing	87,700	800	5.7760	0.0288	27,800	Concurrent testing	89,900	2000	5.7838	-0.0015	Dominated
No testing	137,600	49,900	6.5143	0.7383	67,600	No testing	150,500	62,600	6.5084	0.7231	86,600

NICE National Institute for Health and Care Excellence

Therapeutic monitoring of TNF inhibitors in Crohn's disease (LISA-TRACKER ELISA kits, Immundiagnostik TNFα-Blocker ELISA kits, and Promonitor ELISA kits)

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
AbbVie Ltd	1.	30	1.2.3.1	It is stated that 'usually, treatment is initiated with the less expensive drug (i.e. infliximab), considering drug administration costs, dose and product price per dose.' Section 1.2.13 in NICE clinical guidelines 152 doesn't specify the treatment that is the less expensive. It is inaccurate to state infliximab as the less expensive drug as the cost depends on the required dose which is based on the weight of the patient. The cost also varies for year 1 or the subsequent years as indicated in the costing tool of NICE technology appraisal 187. The statement in the report should therefore be amended as follows, without reference to infliximab: 'usually, treatment is initiated with the less expensive drug, considering drug administration costs, required dose and product price per dose.'	We agree that often the least costly treatment is initiated taking into account drug acquisition and administration costs, dose and product price per dose.
	2.	33	1.2.3.3.1 and 1.2.3.3.2	In section 1.2.3.3.1 on infliximab, figure 1 presents a patient pathway of Crohn's disease patients on infliximab therapy. In section 1.2.3.3.2 on adalimumab, no pathway for patients on adalimumab therapy is presented. As the guidance from TA 187 indicates that both infliximab and adalimumab are recommended as treatment options for adults with severe active Crohn's disease, we suggest amending the first box of the	We agree with this comment after loss of response to adalimumab, patients would follow a pathway similar to that following infliximab failure (Figure 1) with patients eventually switching to infliximab.

Therapeutic monitoring of TNF inhibitors in Crohn's disease (LISA-TRACKER ELISA kits, Immundiagnostik TNFα-Blocker ELISA kits, and Promonitor ELISA kits)

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
				pathway in figure 1 so that it says 'patients on infliximab or adalimumab maintenance' and the box c) of the pathway to say 'c) Switch to another anti-TNF inhibitor i.e. adalimumab or infliximab. Also the reference and source of this pathway is missing. The information needs to be added.	
	3.	33	1.2.3.3.2	In line with both section 3.9 of TA187 (reference 6 of the DAR) and the SmPC of adalimumab, the paragraph on the administration of adalimumab should be amended as underlined. "The adalimumab induction treatment dose regimen for adults with severe Crohn's disease is 80mg via subcutaneous injection, followed by 40mg 2 weeks later. <u>In case there is a need for a more rapid response to therapy, the regimen 160mg at Week 0 followed by 80mg at Week 2, can be used with the awareness that the risk for adverse events is higher during induction.</u> After induction treatment the recommended dose is 40mg every other week. This can be increased to 40mg every week in people whose disease shows a decrease in response to treatment."	We agree that in case there is a need for a more rapid response to therapy, the regimen 160mg at Week 0 followed by 80mg at Week 2, could be used with consideration of the higher risk for adverse events during induction.
	4.	37	1.2.4	At the end of the first paragraph, it is stated that 'no comparable data for adalimumab are available'. However in a recent retrospective study, Choi et al. reported the total costs of care for adalimumab and infliximab in The Leeds Teaching Hospitals	We concur that for adalimumab little comparable data are available and that a recent small UK study of 70 matched patients indicated that the cost associated

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
				Inflammatory Bowel Diseases clinic. Total costs included outpatient, inpatient, surgery, radiology, endoscopy as well as drug costs. 72 matched patients receiving adalimumab (n=36) and infliximab (n=36) as first line anti-TNF therapies were studied. The total costs of care were £18165.57 for adalimumab and £24858.52 for infliximab. Costs were significantly lower with adalimumab (£6692.95 less per patient (95% confidence interval £1816.61–£11569.29), p=0.008) than with infliximab. This was largely driven by the drug costs and drug administration costs associated with infliximab. (Choi et al. Journal of Crohn's and Colitis (2014) 8, 375–383)	with adalimumab treatment of Crohn's disease might be less than that associated with infliximab treatment-matched patients".
	5.	38	1.3.1.2	It is stated that 'the incidence of LOR is better expressed as the annual risk for LOR per patient year (13% for infliximab[34] and 20.3% for adalimumab[35]).' In Billioud et al.[35] patients received different induction doses of 160mg/80mg and 80mg/40mg at weeks 0 and 2 respectively and some patients received adalimumab as a second anti-TNF after failure of infliximab. In Gisbert et al.[34] the majority of patients had received a consistent induction dose of 5mg/kg at 0, 2 and 6 weeks and were receiving infliximab as first line anti-TNF. As concluded by Echarria et al (European Journal of Gastroenterology & Hepatology 2015, 27:430–435),	Echarria et al 2015 (European Journal of Gastroenterology & Hepatology 2015, 27:430–435) post-dated our searches of electronic data bases and therefore could not be included in our report. Thank you for this clarification. This section of the introduction is meant to give an indication of the magnitude of LOR for both IFX and ADA and did not imply a comparison. This is followed by a statement from a study that directly compared LOR to IFX and ADA: *"LOR to adalimumab and infliximab did not differ significantly in a*

Therapeutic monitoring of TNF inhibitors in Crohn's disease (LISA-TRACKER ELISA kits, Immundiagnostik TNFα-Blocker ELISA kits, and Promonitor ELISA kits)

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
				the difference in LOR between a cohort of patients who are naive to anti-TNF and those who have been experienced in the past could account for a difference in these figures, Therefore, the statement on the annual risk of LOR needs to be followed by this underlined statement: 'It has to be noted that these two annual risks for LOR per patient year should not be compared because the difference in the line of therapy of the anti-TNF and the different induction doses could account for a difference in these figures.'	*retrospective study of 375 patients who had lost response to either infliximab or adalimumab…"*
	6.	147	4.3.1.2	At the top of the second paragraph, it is stated that 'administration costs were also included for adalimumab'. The statement is incorrect and should be amended as underlined: 'administration costs were also included for infliximab'. In page 225 of the HTA report Dretzke et al. (2011) used as reference it is stated: 'This includes the cost of administration in hospital or clinic in the case of infliximab.' and 'No administration costs were given for adalimumab on the grounds that it can be given subcutaneously.'	We agree that the corrected statement is correct.

Therapeutic monitoring of TNF inhibitors in Crohn's disease (LISA-TRACKER ELISA kits, Immundiagnostik TNFα-Blocker ELISA kits, and Promonitor ELISA kits)

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
	7.	173	4.3.4.2.2	It is stated that 'for people switching to adalimumab, we derived a cost of £1408.28 (2x £704.28, assuming 40mg of adalimumab is required every two weeks) per four-week cycle'. The summary of product characteristics indicates that 'After induction treatment, the recommended dose is 40mg every other week via subcutaneous injection.' The drug acquisition cost of adalimumab 40mg is not accurate. The cost is £352.14 for 40mg adalimumab. (BNF) The above statement should therefore be amended as underlined: 'for people switching to adalimumab, we derived a cost of <u>£704.28</u> (2x <u>£352.14</u>, assuming 40mg of adalimumab is required every two weeks) per four-week cycle.'	Thank you for picking this up. We agree that: for people switching to adalimumab, we derived a cost of £704.28 (2 x £352.14; assuming 40mg of adalimumab is required every two weeks) per four-week cycle. PLEASE NOTE: in the model the cost input for adalimumab was £704.28 per four week cycle.
	8.	173	4.3.4.2.2 table 35	The cost for the base-case value of maintenance adalimumab should be amended to £352.28.	The model used £352.14 / cycle and this changed price is noted.
	9.	21-22 176-182 187-188	Scientific summary 4.3.6, 4.3.7, 4.3.8 5.2.2	The results of the base case analyses, univariate sensitivity analyses and probabilistic sensitivity analyses need to be updated once the cost of adalimumab has been corrected.	No action required (see above).

Therapeutic monitoring of TNF inhibitors in Crohn's disease (LISA-TRACKER ELISA kits, Immundiagnostik TNFα-Blocker ELISA kits, and Promonitor ELISA kits)

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
	10.	376	8.18 Table 56	The unit cost for adalimumab 40mg every other week should be amended to £352.28 (BNF).	See point 8.
Immundiagnostik AG	1.	1	2	The name "TNFα-Blocker ELISA kits" was recently changed by Immundiagnostik to the new brand name "*IDK*monitor® ELISA kits".	No action required. This does not change the conclusions of our review.
	2.	58	Table 4	The individual kit names are now: • IDKmonitor® Infliximab free ADA ELISA (K 9650) • IDKmonitor® Adalimumab free ADA ELISA (K 9652) • IDKmonitor® Infliximab total ADA ELISA (K 9654) • IDKmonitor® Adalimumab total ADA ELISA (K 9651) • IDKmonitor® Infliximab drug level ELISA (K 9655) • IDKmonitor® Adalimumab drug level ELISA (K 9657)	No action required.
Healthcare professional	1.	General comment		The conclusions of this assessment are surprising and counterintuitive. In the area where I work, a combined test for Infliximab antibodies and levels costs 70 pounds. A single dose of Infliximab costs well over 1000 pounds, with an average annual cost per patient of 12,000 pounds. Another way of looking at it is that	These issues will doubtless be considered by the committee. The conclusion of our assessment includes: "*Our findings that testing anti-TNFα drugs and their antibodies are not cost effective*

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
				one year of treatment for a single patient with Inflximab at average dose would pay for about 170 combined Infliximab level and antibody tests. 15% of patients who appear to be in a good remission on Inflximab have no detectable drug when trough levels are measured, and hence continuing administration to this group of patients has no greater clinical effect over and above not giving the drug. Therefore it would appear that it is only necessary to test 6 or 7 patients to save drug costs equivalent to 170 tests, during the first year after testing. While I accept the findings of the study according to the methodology that has been used, it does seem that a general recommendation not to use these tests will prevent significant cost saving.	*should be viewed cautiously by clinicians and policy makers, in view of the linked-evidence approach required and the poor quality of the evidence available to us."*
Merck Sharpe and Dohme	1.	General	General	MSD notes that the Assessment Report finds that tests for therapeutic monitoring of TNF inhibitors are not cost-effective; however, we believe that it is important for the role of these tests in supporting optimal patient care to be recognised. The British Society of Gastroenterology guidelines state that "There is emerging evidence linking low serum trough levels of IFX to lack of sustained response. Further research is required, but it appears serum IFX levels	As we state above: These issues will doubtless be considered by the committee. The conclusion of our assessment includes: "*Our findings that testing anti-TNFα drugs and their antibodies are not cost effective should be viewed cautiously by clinicians and policy makers, in view of the linked-evidence approach required and the poor*

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
				are influenced by ATIs and other - probably pharmacokinetic - factors. At this stage, it is not known what the target trough level should be. In the UK this issue is at present academic because there is no available commercial resource for measuring either trough levels or antibody levels. *We think such a resource would be valuable*." (Mowat *et al*. 2011). In addition, identifying and understanding any loss of response that occurs allows clinicians to avoid unnecessary wastage of NHS resources. MSD would welcome NICE guidance supporting the use of these technologies in clinical practice.	*quality of the evidence available to us."*
	2.	General	General	The Assessment Report states that testing is not a cost-effective strategy. We believe that the External Assessment Group is correct to acknowledge the limitations in the data that underpin this assessment. MSD believe that the forthcoming data (PANTS, TAILORIX) should be considered, to the greatest extent that is possible at this stage, in any resulting recommendation by NICE.	We agree that the PANTS and TAILORIX studies will likely provide important and relevant information, however at the time of writing the assessment report no results from these studies were available in the public domain.
	3.	General	General	The External Assessment Group performed sensitivity analyses in which the frequency of testing is reduced from 3-monthly to annually. MSD notes that in these analyses, testing is found to be a cost-effective strategy. We believe that regular testing may not reflect likely clinical practice; rather opportunistic (and therefore less frequent) testing may represent a more	There was a lack of evidence about the clinical effectiveness of different testing frequencies. The model is based on the evidence available. 'Opportunistic' testing was not used in any of the two comparison studies.

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
				likely scenario.	
	4.	31	1.2.3.3.1	The marketing authorisation for infliximab (in adults) has been taken from NICE TA 187. The marketing authorisation for infliximab in adults has now been updated, and is as follows: "Adult Crohn's disease: Remicade is indicated for: • treatment of moderately to severely active Crohn's disease, in adult patients who have not responded despite a full and adequate course of therapy with a corticosteroid and/or an immunosuppressant; or who are intolerant to or have medical contraindications for such therapies. • treatment of fistulising, active Crohn's disease, in adult patients who have not responded despite a full and adequate course of therapy with conventional treatment (including antibiotics, drainage and immunosuppressive therapy)." The marketing authorisation is also described on p.35: "Infliximab was the first anti-TNFα agent that was approved and licensed for treating severe active Crohn's disease and active fistulising Crohn's disease in adults and children over the age of six. It is administered intravenously over 1–2 hours." This	The EAG notes the following. The EMEA has extended authorisation to "treatment of moderately to severely active disease". (Procedure No. EMEA/H/C/000240/II/00142)

Therapeutic monitoring of TNF inhibitors in Crohn's disease (LISA-TRACKER ELISA kits, Immundiagnostik TNFα-Blocker ELISA kits, and Promonitor ELISA kits)

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
				description does not acknowledge the inclusion of moderate patients in the present marketing authorisation for infliximab.	
	5.	33	1.2.3.3.1	Figure 1 shows one of the options with infliximab treatment to be "a) Increase dose or reduce interval". MSD would like to note that reducing the interval of treatment is not supported by the marketing authorisation for infliximab in adult patients.	We note this comment regarding the figure caption.
	6.	174	4.3.4.2.2	The footnote of Table 1 states "People receiving of Adalimumab during maintenance therapy every 40mg/kg every two weeks". The correct amount is 40 mg total not 40 mg/kg.	We note this.
Lay person	1.	18	Methods	Is it worth adding the names of the searched electronic databases in here?	Ideally the names should be mentioned, however, due to word count restrictions this has not been done as it is not a requirement for publication. The reader can refer to the full report for details of the databases searched. More importantly we mentioned that databases were searched, included the dates of the searches and reported that supplementary searches were undertaken.
	2.	22	Strengths and limitations	How can you be sure that the included studies can be used to determine clinical and cost effectiveness of the TNF alpha inhibitor kits I classifications of the patients in the studies is not standardised? Could the	We cannot be sure, which is why we have added this as one of the major limitations of the review. Disease activity scales were used in some but not all of the studies

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
				disease activity scales mentioned on page 27 of the report be used to classify the patients?	identified (two (TAXIT and Steenholdt) of the three comparative studies) even though they are superior to clinical assessment in terms of disease classification.
	3.	26	1.2	What other infections may cause Crohn's disease?	The true cause of CD is unknown. One theory is that in individuals that are genetically susceptible to develop Crohn's disease a previous infection (mycobacterium has been named) might lead to abnormal inflammation in the digestive tract. There is still a lot of uncertainty around this research area.
	4.	74	3.2.2.3.1	I believe that papers which patients with other medical conditions such as UC and those with unspecified medical conditions from other hospital departments should be excluded from the report as they may have an impact on the results of the literature search.	Ideally this would be the case. Searches were required to be sufficiently sensitive to identify articles mentioning CD or studies indexed as IBD in order not to miss any relevant studies. During sifting, studies on UC and RA patients were excluded from the review. Studies with a mix of patients >50% CD patients were included due to lack of evidence on CD patients only. One of the key trials included IBD patients and where available we reported results for CD patients only. We cautioned that we do not know

Therapeutic monitoring of TNF inhibitors in Crohn's disease (LISA-TRACKER ELISA kits, Immundiagnostik TNFα-Blocker ELISA kits, and Promonitor ELISA kits)

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
					what impact the mixed patient group has on the outcomes reported.
	5.	148	4.1.3.3.1	Was there a published cut of point at which a non-responder becomes a partial responder or is it just any decrease in symptoms or disease activity? Same question if this happens in other studies	To our knowledge no such cut off has been proposed.
	6.	159	Figure 28	What happens to partial responders in this care pathway? I'm not clear which pathway they fall into? I assume it's the do not respond pathway? What happens to the patients once they have had surgery if they have been non responders to all types of anti TNF tried?	Partial responders would be included as responders in the model.
	7.	160	Table 27	Where do the partial responders fit in as it's not defined?	See above.
	8.	167	4.4.3	Does the proposed model remain workable if the hypothetical cohort of 30 year olds are replaced with the second peak onset cohort at age 60 as these patients may have an increased risk of dying from age related diseases or complications?	This has not been investigated.
Viapath, Guy's and St Thomas' NHS Foundation Trust	1.	**19**		**When testing concurrently there are four possible test outcomes: 'drug + / antibody +'** This scenario is most likely to be due to interference in the drug assay if the technology used is measuring free drug and free antibody. In real patient samples, this is only observed when drug levels are <1 ug/mL	Thank you for this additional information. It needs to be considered that this is the experience of using one test. The question that follows on from that is what management / treatment should a likely algorithm suggest for individuals with a

Therapeutic monitoring of TNF inhibitors in Crohn's disease (LISA-TRACKER ELISA kits, Immundiagnostik TNFα-Blocker ELISA kits, and Promonitor ELISA kits)

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
				and there is dynamic changes in drug/antidrug antibody binding possibly due to presence of low affinity antibodies. See abstract attached. DDF15-1753	drug+ / antibody+ test outcome. The only evidence available was that generated by the Steenholdt algorithm, which used RIA for testing.
	2.	**22**		**'The evidence on assay performance was sparse and sometimes conflicting with lack of an agreed gold or reference standard for tests. There was very limited concordance data from studies comparing test performance of different assays.'** I entirely agree with this statement and also want to share our findings on the comparison of the three assays quoted in this review. Please refer to DDF15-2619 Further comments on technology in the context of assessing papers: - Assay comparisons: It is also important to consider how the ELISA was performed i.e. manual versus auto mated platform. Our experience suggests results may vary depending on the mode of analysis. Please refer to data provided on ROC study ADA (Promonitor additional info tab)	The comment and the additional evidence provided as unpublished abstract submitted to the Digestive Disorders Federation (British Society of Gastroenterology) conference in June 2015 confirms and strengthens our concerns about the applicability of outcomes to the intervention assays under assessment.

Therapeutic monitoring of TNF inhibitors in Crohn's disease (LISA-TRACKER ELISA kits, Immundiagnostik TNFα-Blocker ELISA kits, and Promonitor ELISA kits)

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
				- In UK, service providers are now obliged to provide traceability of their assays as part of accreditation. Unfortunately there is minimal information on all three assays kit inserts. This is further complicated by the fact that assay comparisons are challenging due to variation seen in these assays. In the context of patient monitoring, this is significant as results or clinical decision points are not transferrable from one assay to another. - It is hoped that NICE will consider recommendations on the standardisation of these assays and independent serum based reference materials to check accuracy.	
	3.	**45/46**		**Immundiagnostik TNFα-Blocker monitoring, infliximab drug level (e.g. Remicade®) ELISA (K9655).** The plate coating of this assay stated as 'Monoclonal anti- infliximab antibody' on page 207. This is the modified version of the assay to make it more specific for Infliximab. The plate coating for the same kit was	Thank you for this additional information. It confirms our finding that assay information provided by the assay kit inserts is insufficient. This information does not affect the outcome of our assessment as the comparative studies used in the economic model did not use the Immundiagnostik TNFα-Blocker

Therapeutic monitoring of TNF inhibitors in Crohn's disease (LISA-TRACKER ELISA kits, Immundiagnostik TNFα-Blocker ELISA kits, and Promonitor ELISA kits)

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
				previously TNFα. The impact of this change and date it is modified is unknown. Both kits were sold under identical order number and no notification of this change found in the reagent boxes. It will worth considering this modification when reviewing papers and when comparing assays. It is possible that this change was tested at manufacturer's lab and found to have no impact on results. Please also refer to data provided where Batch 1 Immunodiagnostik assay was performed using the old version of the assay and re-analysed with the new version. (ROC study IFX results)	monitoring, infliximab drug level (e.g. Remicade®) ELISA (K9655).
	4.	**50**	**1.4.2**	**'Current usage of assays in the NHS'** **However demand is low, analyses are often undertaken in batches, and it can be weeks (in some cases) before a clinician receives a result on which to act.** This is not entirely correct for our practice. Laboratories need to demonstrate assay verification on CE marked kits and validation (where in house assays are used) for accreditation purposes and	This is reflective of the experience of a clinical expert in the field and might be applicable to other parts of the country without strong links to Guy's and St Thomas' NHS Foundation Trust. We appreciate that Guy's and St Thomas' NHS Foundation Trust has considerable interest in this area and might not be representative of the rest of the country. Unfortunately this and other data about uptake was not available at the time of

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
				patient safety. We are not aware of any service offering routine in house assay except for one centre. TAT is not an issue as this is now significantly reduced. Also we regularly receive direct calls from clinicians and specialist nurses where test results are required for decision making. The uptake of anti-TNF testing has been enormous showing significant increase in the last year. Please refer to confidential data in further information and abstracts provided: Utility of test in real UK practice, TATs and demand for test. Please refer to following abstracts: DDF15-1753, DDF15-2523, DDF15-2502, DDF15-2089, ECCO2015	preparing our report
	5.	**80**		**There are no studies linking LISA-TRACKER to any of the comparator tests for detecting adalimumab or antibodies to adalimumab.** Please refer to DDF15-2619. Raw data also provided. Please refer to Excel worksheet (ROC study ADA; password roc01) (Confidential).	This provides very useful information comparing the index tests, but does not link to the comparator tests (e.g. Prometheus HMSA).

Therapeutic monitoring of TNF inhibitors in Crohn's disease (LISA-TRACKER ELISA kits, Immundiagnostik TNFα-Blocker ELISA kits, and Promonitor ELISA kits)

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
	6.	**P81**		**'In summary in one study using a range of clinical and spiked samples there is some evidence that LISA-TRACKER may give false positive results for infliximab in the presence of antibodies to infliximab or adalimumab, where the Leuven and Amsterdam in-house ELISAs do not.'** 'drug + / antibody +' scenario: This scenario is most likely to be due to interference in the drug assay where free drug and free antibody is measured and will always require interference studies. This is what happens in real practice. Please refer to abstracts: DDF15-1753 and AbtractESPHGAN2015.	These abstracts do not seem to add further information on interference. We agree that interference could be problematic.
	7.	**P75/76/77**		**'Comparisons between the index tests'** Method comparison presented requires further information. There are guidelines on method comparison which should clearly state the range covered, Altman and Bland plots and regression analysis (Deming or Passing Bablok). R2 values presented are misleading and do not provide information on the relationship between assays other than systematic correlation. Assay variations should	We agree that R^2 only gives information on correlation and we state clearly that we cannot reach conclusions from R^2 values but that is all that was available at the time of the review. The excel spreadsheets provide measurements of IFX, antibodies to IFX, ADA, and antibodies to ADA for 80 patients. For every patient there are measurements

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
				also be considered in the context of advice given to clinical users i.e. what level is considered adequate or therapeutic. Please refer to abstracts ADAECCO/ECCO poster Ward ADA poster and DDF15-2619. Raw data also provided (confidential). Clinical cohort discussed in the ECCO poster has measurements using Promonitor as well. Data was not presented at the time due to pending further checks on the Promonitor assay. Analytical comparison is presented in DDF15-2619. Clinical re-assessment due to take place. Raw data provided.	using Theradiag (Lisa Tracker) Promonitor and Immunodiagnostik. They present a range of plots plus raw data. If presented in time this would have been very useful and included in the report. It adds information to the comparison between the index tests, including Bland Altman plots, and analysis of systematic bias, and using the raw data comparisons at set thresholds could have been made. The ECCO data follows up patients and finds higher measured drug levels are associated with remission, and present some performance comparisons between assays. This is a very powerful and useful data set, currently presented as a series of abstracts and posters. In summary this data is more useful than that presented in the report, but there remain inconsistent results when comparing the index tests, without clear understanding of why, and therefore conclusions about assay comparisons may have remained difficult

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
					even with this extra information.
				'Additionally ten percent of samples were above the upper limit of quantification for LISA-TRACKER, and not for the Promonitor ELISA.' Please refer to ADA ROC study (tab all data). Out of 80 samples tested by all three assays % of samples requiring dilutions are as follows: Immunodiagnostik: 0/80 samples Lisatracker: 17/80 (21%) Promonitor: 41/80 (51%)	This is interesting and very informative. The majority of the information available to us for the review was from Promonitor.
				'In summary we have one abstract giving a Cohen's Kappa of 0.8 between LISA-TRACKER	

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
				and Promonitor, in tests for antibodies to adalimumab but it is not known how many samples and of which type were included in this analysis.' Please refer to raw data provided on antibody testing using three assays in IBD patients described in ADAECCO abstract (confidential)	
	8.	**P78**		**There was one abstract comparing Promonitor to LISA-TRACKER for infliximab (unpublished abstract provided by Proteomika (Nagore et al., 2015)).** We have evaluated three assays using patient specimens presented in the abstract below. Please refer to abstract: **DDF15-2619** The conclusion was assays show significant degree of variation and the cause of this require further investigation. Further testing were performed to identify the cause of differences observed between the assays and that is included in the raw data provided (ROC study IFX,	As above this is superior data to that which was available to us in the review, but due in part to the inconsistent results even if we were to have been able to include it, we consider that the conclusions may not be altered significantly.

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
				password roc01, confidential) Please refer to raw data provided on Infliximab testing using three assays in IBD patients described in ECCO abstract (confidential).	
	9.	**P82**		**'There are no studies linking Immundiagnostik to any of the comparator tests for adalimumab or antibodies to adalimumab.'** Please refer to abstracts ADAECCO/ECCO poster Ward ADA poster and DDF15-2619. Raw data also provided (ROC studies ADA, password roc01, confidential).	As above.
	10.	**P87**		**'They also describe the upper limits of the measurement range for LISA-TRACKER as low'** This statement is misleading. Out of the three drug assays, Immundiagnostik assay has the widest measurement range. Both Immundiagnostik and Lisatracker assays are optimised using single dilution. However, Promonitor assays uses two dilutions on single specimen to obtain drug measurement. This has two implications: - **Cost** : Unnecessary wastage of wells by using two dilutions per specimen, leading to increased cost	Same as for point 7 above. This is interesting and very informative. The majority of the information available to us for the review was from Promonitor.

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
				- **Impracticality:** If implemented this is deemed very unusual practice in a routine laboratory. Performing analysis using two dilutions on one specimen suggests need for further optimisation of the measurement range. Assays should be optimised to cover the most clinically appropriate range and dilutions should only be done if a result falls outside the dynamic measurement range of the assay. It is also possible to obtain two results on a specimen using different dilutions as recommended by IFU, potentially leading to errors. This approach undoubtedly raises the questions about the optimisation of the assay and potentially may be a contributory factor in variation in data presented on method comparisons. **Please refer to data provided on Excel worksheet.**	
	11.	**P124**	**3.2.44**	**'Summary of major findings from three management studies** ' None employed designated intervention tests (LISA-TRACKER, TNFα-Blocker, or Promonitor ELISAs) Please refer to abstract: DDF15-2089, DDF15-2502	These abstracts all post-date the completion of the report. It is not possible to implement changes to a report of this complexity with information which only appeared in the public domain in February 2015.

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
				and DDF15-2523	
Healthcare professionals	1.	**194**	**6.1**	**Definitions of response and remission** The lack of a gold standard definition of response, remission and loss of response and time at which this should be categorised limits any assessment of the relative performance of drug and antibody levels assays. A gold standard definition of remission must include mucosal healing defined at colonoscopy but this has not frequently been used as an end-point in correlation studies (and is not collected in the PANTS study) and has not been used in any comparative studies of drug and antibody levels assays. Response and loss of response are much more difficult to define. In the PANTS study we will use mathematical modelling to develop and refine clinically meaningful definitions using quantitative and qualitative data, (including quality of life data) collected at multiple time points. We aim to generate and test these definitions by October 2015 so that these may be utilised for comparative studies in the first quarter of 2016.	We agree with this comment. The PANTS study will add valuable knowledge in this area. However, results were not available for this assessment and the lack of treatment prescribed by a prescriptive algorithm in response to test outcomes would have limited the usefulness of the study in addressing the decision questions in this assessment.
	2.	**194**	**6.1**	**Relative performance of anti-TNF drug and antibody assays** In the PANTS study. Serum has been collected and stored on all patients at multiple time points. We will compare the relative performance of 4 ELISA kits	We agree that the PANTS study will add valuable knowledge in the area of assay type comparison.

Therapeutic monitoring of TNF inhibitors in Crohn's disease (LISA-TRACKER ELISA kits, Immundiagnostik TNFα-Blocker ELISA kits, and Promonitor ELISA kits)

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
				(Immunodiagnostic, LISA-TRACKER, Promonitor and the Leuven assay / R-Biopharm AG), and will consider comparing with other methods including RIA and HMSA. We will explore the benefit of measuring total vs. free drug / antibodies. These data will be available in the fourth quarter 2016	
	3.	**194**	**6.1**	**Effectiveness of clinical algorithms** PANTS is a correlation study and a specific algorithm is not being tested. However we will be able to observe the relationship between dose changes, drug and antibody levels and clinical outcome in a large cohort of patients	See previous comment about the PANTS study.
	4.	**194**	**6.1**	**Adalimumab and Paediatric data** The PANTS study includes patients treated with Adalimumab (to date more than 500 patients have been recruited) and children aged over 6 years.	See previous comment about the PANTS study.
	5.	**194**	**6.1**	**Impact of immunomodulators on monitoring of drug and antibody levels** Information on the use of immunomodulators (Azathioprine, Mercaptopurine and Methotrexate) is collected as part of the PANTS study. We will report the impact of immunomodulators on drug and antibody levels in the fourth quarter 2016	See previous comment about the PANTS study.

Therapeutic monitoring of TNF inhibitors in Crohn's disease (LISA-TRACKER ELISA kits, Immundiagnostik TNFα-Blocker ELISA kits, and Promonitor ELISA kits)

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
	6.	**194**	**6.1**	**PANTS study - Anticipated research output, timelines and alignment to research priorities** The PANTS study is a 3 year prospective correlation study which includes an investigation of the relationship between anti-TNF trough levels and anti-drug antibody levels. The study aims to recruit 1500 patients aged 6 and over treated with Infliximab or Adalimumab. Recruitment will close in December 2015. We anticipate the data generated from the PANTS study may help further explore the objectives of this Diagnostic Assessment Report.	See previous comment about the PANTS study.
	7.	**244**	**8.4.3 Information provided by Immundiagnostik**	vi) Published in Annuals of Biochemistry http://www.ncbi.nlm.nih.gov/pubmed/25780249 Ann Clin Biochem. 2015 Mar 16. pii: 0004563215580001. [Epub ahead of print] **Infliximab and adalimumab are stable in whole blood clotted samples for seven days at room temperature.** Perry M[1], Bewshea C[2], Brown R[3], So K[2], Ahmad T[2], McDonald T[3].	Thank you for this latest evidence which will help in answering questions about practical implications if tests are routinely implemented.
Theradiag	1.	80 257	3.2.2.3.3 8.5	Detection of infliximab levels and anti-infliximab antibodies Vande Casteele 2012: a comparison of three different assays: see Letter to the Editor AP&T	Theradiag clearly disagree with the Vande Casteele paper. This is one of the few papers to compare assays. This probably

Therapeutic monitoring of TNF inhibitors in Crohn's disease (LISA-TRACKER ELISA kits, Immundiagnostik TNFα-Blocker ELISA kits, and Promonitor ELISA kits)

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
				from Theradiag attached and explanations pdf file. Results clearly invalidate the data published by showing no false-positive result	would not affect our overall conclusions.
	2.		3.1.6.3	Ben Horin and Al. Undetectable anti-TNF drug levels in patients with long-term remission predict successful drug withdrawal. Conclusion: Incidental finding of undetectable anti-TNF drug levels in patients with stable long-term deep remission may identify a subset of patients whose clinical remission is no longer dependent on anti-TNF treatment, and who may be considered for therapy discontinuation after careful weighing of risks of drug stopping. Manuscript attached. Confidential information as article not already published	This is an interesting paper but does not appear to relate to the research questions we investigated. If published in time of the review it might have provided some useful parameters for the model.
	3.	84 85 86 87 4.88	3.2.2.3.4	Steenholdt, AMG 2014: Clinical Implications of Measuring Drug and Anti-Drug Antibodies by Different Assays When Optimizing Infliximab Treatment Failure in Crohn's Disease: Post Hoc Analysis of a Randomized Controlled Trial. Conclusion: despite variable analytical properties, common assays result in similar classifications and interventions in patients with IFX treatment failure, and comparable clinical outcomes. Publication attached	This paper is included in our review. It is the most useful paper for comparing tests, but unfortunately does not include any of our index tests.
	4.		3.1.6.3	A predictive model for relapse in CD patients	Thank you for this unpublished material. This was not submitted to NICE in time for the

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
				presenting clinical remission undergoing infliximab (IFX) treatment. Conclusions: In IFX-treated CD patients and in clinical remission, a combination of TLI (< 2µg/ml) and faecal calprotectin (>250µg/g of stools) enable the prediction of clinical recurrence within 6 months in 95% of cases. Intervention studies are needed to assess the impact of treatment modification in this group of patients. Pr. Xavier Roblin Presentation attached	review and could not be considered in the review.
	5.		3.1.6.2 3.1.6.3	Relationships between the adalimumab/infliximab levels, antibodies levels and clinical response, see 20150511 TDM presentation	Thank you for this eclectic collection of studies that relate test levels to clinical status. Without full referencing it is difficult to assess in the short time frame whether any additional information is presented in this collection of figures and citations.
	6.	282	8.6	Additional proof. Levels of drug and antidrug antibodies are associated with outcome of interventions after loss of response to Infliximab or Adalimumab (Yanai publication attached) reinforce the publication from Pr. Roblin (19. Trough levels of drugs or ADAs may guide therapeutic decision: Development of an algorithm Incorporating Pharmacokinetics of adalimumab in IBD) More over we disagree with the reason of exclusion of the publication by Roblun.	Thank you for the attached study by Yanai et al. This was not considered in the review as it was published too late for our review. The study of Roblin was interesting but a test algorithm did not inform treatment; rather, retrospective analysis of test categories was used to determine what treatment patients had received. Furthermore a suitable comparator group was unavailable.

Therapeutic monitoring of TNF inhibitors in Crohn's disease (LISA-TRACKER ELISA kits, Immundiagnostik TNFα-Blocker ELISA kits, and Promonitor ELISA kits)

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
	7.	142	4.1	COSTS SAVINGS OF ANTI-TNF THERAPY USING A TEST-BASED STRATEGY VERSUS AN EMPIRIC DOSE ESCALATION IN CROHN'S DISEASE PATIENTS LOOSING RESPONSE TO INFLIXIMAB. UEG Week 2014 abstract accepted attached; submitted article for publication attached (Confidential) RESULTS: Costs savings among the 10,000 Crohn's disease patients using a test-based strategy were 131.300.293 € at 5 years. At 5 years the mean costs savings were 13.130 euros per patient. The direct cost of the test had no impact on the results until a cost per test of 2.000 euros. CONCLUSION: A test-based strategy leads to major cost savings related to anti-TNF therapy in Crohn's disease.	Thank you for this abstract and submitted article itemising cost savings from a test strategy for Crohn's patients. There does not appear to be any comparison of clinical outcomes between the strategies. These were not available in time for the review.
	8.			Abirisk consortium will use Lisa Tracker kits to investigate the risks of immunogenicity generated by biotherapies. The European project ABIRISK is a collaboration between the European Federation of Pharmaceutical Industries, the Innovative Medicine Initiative (IMI) and the European Union. It counts 39 partners: large members of the pharmaceutical	Thank you for informing us of this interesting and useful collaborative effort. The results from this collaboration will be of considerable interest to all.

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
				industry (Pfizer, GSK, Sanofi-Aventis, Merck, UCB, Ipsen, Novartis, Novo Nordisk, and Bayer), academic labs and research institutes. LISA TRACKER kits have been validated for this project (press release attached)	
	9.		3.2.4	Prospective studies to be considered: Warman et al, 2015, Therapeutic drug monitoring of infliximab in inflammatory bowel disease patients in a teaching hospital setting: results of a prospective cohort study. Byron Vaughn et al, 2015 (abstract) Prospective Therapeutic Drug Monitoring and Optimization of Infliximab Maintenance Therapy in IBD (attached) – DDW	Vaughn et al. (2015): this is an abstract to the full text study by the same author which was considered in detail in the EAG report (see section 3.2.4.3.3). Thank you for the list of studies. The listed infliximab studies have entered the public domain very recently in 2015 and could not be included in our report; those for adalimumab were either too recent for inclusion or not relevant for inclusion.
	10.	89 90	3.2.2.5.1 3.2.2.5.2	Other Infliximab and Adalimumab cut off values publications may to be considered	We clearly stated that we considered all identified studies that reported ROC analysis for cut-offs which sufficiently showed that cut-offs vary considerably and are not readily transferable. Of the 14 provided publications 10 are included in our overview of cut-off thresholds, one was excluded as it details a

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
					review without ROC analysis of thresholds, one was excluded from the review as it is a study on UC patients and only two are additional studies which were published too late for inclusion in the review.
	11.	76, 77, 78, 79	3.2.2.3.2	We suggest that unpublished data provided by Proteomika should not be considered in this report	We are requested to review and consider all material that was provided in time by the companies.
Proteomika	1.	16 & 23	Abstract (conclusions) & Scientific Summary (Implications)	We think that the conclusion is not founded given the limited evidence available**.** Since far more evidence is available regarding infliximab, the conclusions should differentiate between infliximab and adalimumab. However, more patients receiving adalimumab are likely to require dose escalation based on the annual risks reported for LOR (page 38) which were 13 & 20.3% for IFX and ADA respectively. Therefore it would be reasonable to suggest that efficiency savings seen with IFX could be potentially greater with patients who are receiving ADA. We agree with the statement "…*testing is not cost effective should be viewed cautiously due to the limited evidence available*".However, although the evidence is limited, Steenholdt *et al. (Pg 111,112)* showed a significant cost reduction ($p<0.0001$) particularly for group 3 (table 14 p111). The majority of	We refer the commentators to the main body of our report. We used the evidence available to us at the time to draw our conclusions. There is very little peer-reviewed published information on adalimumab. But we note with interest this comment that saving in the case of adalimumab would be 'potentially greater.' However we found that immediate savings do not necessarily counteract the longer term differential costs of different strategies over the life time of the model. We agree that $p<0.0001$ appears highly significant – but even this level of significance as a one-off finding in a small study should be treated with caution.

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
				the costs were attributed to the 65.9% of patients who were dose intensified. This should at a minimum support monitoring trough levels. While concern was raised that RIA was utilised on samples stored in a biobank we can demonstrate that freeze thaw cycles do not affect trough level drug measurements. Our previous points were emphasised by Vande Casteele *et al*. pg 122 & 123 (bottom and top paragraphs respectively) which underpin a trough level strategy. Supporting these comments on page 124 are the authors' summary report, which attribute cost savings to less use of infliximab. In addition the TAXIT study demonstrated costs between the two strategies of 20,723 vs 21,023.These costs included a UC cohort (78 of 251), which may skew the cost effectiveness.	
	2.	19	Scientific Summary (Cost-effectiveness model)	"Drug + / antibody +" results with most ELISAs are likely false positives. ELISA is very susceptible to drug interference, meaning that antibody ELISA assays cannot detect antibodies in the presence of virtually any drug concentration present in the sample, therefore concurrent testing with most current ELISA tests (free antibodies) cannot include the option of "Drug + / antibody +". This is a recurrent problem with results provided by ELISA kits of two of the manufacturers included in this evaluation that give a high rate of false positives for both drug levels and antibodies (see Vande Casteele *et al*. (2012)).	Thank you for this comment; we believe the potential shortcomings of ELISAs due to the mutual interference between drug and anti-drug antibodies was adequately discussed in the EAG report.

Therapeutic monitoring of TNF inhibitors in Crohn's disease (LISA-TRACKER ELISA kits, Immundiagnostik TNFα-Blocker ELISA kits, and Promonitor ELISA kits)

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
	3.	20	Scientific Summary (Cost-effectiveness model)	Solid phase RIA is also susceptible to drug interference, therefore +/+ results must be interpreted with caution, since they will very likely represent a false positive for at least one of the analytes. Bloem *et al*., 2015.	Thank you for this comment; solid phase RIA was not considered in the report.
	4.	21	Scientific Summary (Clinical effectiveness)	In the absence of a gold standard test against which to compare the index tests the conclusion that between 20 and 30% of positive and negative test results are likely to be inaccurate does not, in our opinion, appear to be adequately supported.	Thank you for this comment. The 20 to 30% refers to the reliability of tests as predictors of clinical status. This was based on the predictive values derived from meta-analysis of 30 test studies described in Appendix 12 of the EAG report. The obverse will also hold; namely the reliability of clinical status as a predictor of test result will be moderate, (the balance between positive and negative predictive values will depend on which is taken as gold standard).
	5.	22	Scientific Summary (Discussion and Conclusions)	"*The limited RCT evidence from short term studies indicated that there is little or no benefit from a test-algorithm strategy although there may be some cost savings*". A recent abstract published at the Digestive Disease Week 2015 in Washington (American Gastroenterology Association) shows that dose optimisation of infliximab using therapeutic drug monitoring is more effective than dose optimisation based on clinical assessment alone in patients with	Thank you for indicating this interesting abstract. Regrettably its late availability has precluded its use in the EAG report. Relapse was somewhat better with testing in TAXIT, as described in the EAG report. For the economic model the relevant distinction is between response and loss of response; and TAXIT did not report difference for this outcome. As mentioned in the report the

Therapeutic monitoring of TNF inhibitors in Crohn's disease (LISA-TRACKER ELISA kits, Immundiagnostik TNFα-Blocker ELISA kits, and Promonitor ELISA kits)

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
				active inflammatory bowel disease (Abstract Tu1316, Kelly *et al*., 2015). This abstract has been provided to NICE for consideration in the evaluation. One of the significant clinical benefits demonstrated in the TAXIT trial (p=0.018 pg 122 bottom paragraph) was that more patients relapsed and require rescue therapy in the clinical based arm. Relapse free survival time was also superior (Figure 22) with a TDM strategy. Both of these outcomes alone are already part of routine clinical practice. As previously commented certain groups and strategies as demonstrated by Steenholdt *et al. (Pg 111,112)* showed a significant cost reduction (p<0.0001) particularly for group 3 (therapeutic infliximab and anti-drug antibody undetectable).	lower requirement for rescue in the test arm of the trial was probably the main reason for the slightly lower cost found for that arm relative to clinically based dosing (savings of € 300 per patients in one year). The Steenholdt study savings were included in the EAG report.
	6.	23	Scientific Summary (Research priorities)	If the benefit of measuring total antibodies is not clear, it does not seem logical to compare tests with different Intended Uses; that is, ELISA kits included in the evaluation measure free antibodies other comparator tests measure total antibodies. In our opinion, heterogeneity makes adequate comparison of results problematic.	…'heterogeneity makes adequate comparison of results problematic'…. Thank you. We have also have made this point in our report.
	7.	24	Plain English Summary	The statement "…*and that current tests disagree*…" we feel is not substantiated, please refer to comment 18. In addition a pending evaluation of 3 ELISA technologies in 120 IBD patients is currently underway	Because the plain English summary is so short we have to summarise a huge amount of information in a short space. Notwithstanding we consider that our Bland

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
				with the aim of answering this point. The results of this trial will be available early to mid-July.	Altman plots indicate that current tests disagree.
	8.	39	1.3.2 Anti-drug antibodies	Currently, no human standard for anti-drug antibodies exists. Therefore antibody results should be given as arbitrary units instead of mass units. The measurement of antibodies against infliximab and adalimumab provided in ng/mL by one of the manufacturers included in this evaluation is questionable. The immune response of patients is polyclonal, and the standard provided by this manufacturer cannot reflect all antibody specificities. Therefore test results should not be expressed as mass units unless the means of measurement and the nature of the standard are provided.	Thank you for drawing our attention to this.
	9.	43	1.3.4 Anti-TNFα and antibody level monitoring in Crohn's disease	One of bullet the points listed as potential benefits for testing is to *"allow earlier de-escalation of therapy, leading to a reduction in the overall drug used*". We believe this to be potentially the single greatest impact that TDM can have on demonstrating cost effectiveness. While we accept there is currently limited data to support this assertion, our concern is that the model does not incorporate drug de-escalation for responders if they lie within the target range of the treatment algorithm (4.3.3 pg 167 bullet 6). The model input (page 163, 4.3.2.2.1 bullet 3)	The intention of this introductory part of the EAG report was to provide background and set the scene including a description of the suggested potential benefits of testing. The model input was constrained by the available evidence. There are many potentially useful algorithms based on testing results but these have not been investigated to date.

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
				allows for de-escalation only when the trough drug level is above the higher target range outlined within the study. For the modelling, de-escalation (increase in dose interval only) was attributed to the basecase for the responders. Bullet point 6 (page 167, 4.3.3) within the model assumptions, suggests no dose adaptation took place if the responders had trough drug concentrations within the defined trough range. We would suggest that patients within that range could potentially be de-escalated and still achieve response. This group of patients represented almost 50% (48.21%) of the model population, and represented the largest patient group used in the model. We believe a more cost effective strategy would be to de-escalate based on response and on a patient-by-patient basis, thus providing a personalised therapeutic approach.	
	10.	45, Figure 4	1.4.1.1 ELISAs for infliximab and adalimumab	The format description provided for the Promonitor assays are stated incorrectly, please review the new IFU and previous product insert provided. The wells of the infliximab assay are coated with an anti-TNF human antibody fragment bound to human TNF (not "*a reagent … able to bind to the TNF binding site of infliximab*"). The conjugate is an anti-idiotipic anti-infliximab antibody fragment (not a "*peroxidase labelled antibody able to bind the Fc region of the anti-*	Thank you for this clarification

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
				TNFα")).	
	11.	49	1.4.1.3 Brief overview of identified non-ELISA assay methods	Liquid phase radioimmunoassays are referred to but not solid phase RIA – the omission is not commented on and appears arbitrary.	In the brief overview we only considered types of tests that featured in the main part of the report. Solid phase RIA did not feature in any included study.
	12.	60	3.1.5.1 Objective A	The statement "*However, where spiked samples were unavailable, one of the comparator tests used in the management studies and considered for a linked evidence approach was used. If the reference standard was one of the four comparator tests then it was classified as unlikely to correctly classify the target condition. This is because due to the lack of evidence they constitute an imperfect reference standard."*, would appear to undermine the justification for the use of a linked evidence approach that is widely relied on in comparing test performances.	The linked evidence approach is widely used, but considered inappropriate under certain well defined conditions. Where the reference standard in a linked evidence approach is imperfect, a conservative approach is necessary. (see Merlin, Intl. J. of technology assessment in health care 29:3, 2013) This is because when the index test and reference standard do not agree, we have no way of knowing which is better, and whether the index test would result in better or worse clinical outcomes for these patients where the two tests disagree.
	13.	72, 74 and 80	3.2.2.3 Results of assay	We do not understand why the "*Amsterdam Sanquin in-house ELISA (and RIA for ADAbs)*" is not considered a comparator test alongside HMSA,	Thank you for this comment. The NICE scope prescribed the assays included as interventions in the assessment.

Therapeutic monitoring of TNF inhibitors in Crohn's disease (LISA-TRACKER ELISA kits, Immundiagnostik TNFα-Blocker ELISA kits, and Promonitor ELISA kits)

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
			type comparison studies Figures 7 and 8 3.2.2.3.3 Comparisons between index tests and comparator tests	Biomonitor RIA and the Leuven in-house ELISA? Aside from figures 7 and 8 the first mention of the Amsterdam assay on page 80 states "*These tests from the Amsterdam group are not included as comparators…*". Given that the Sanquin ELISA and RIA tests have been used in a significant number of studies on biological drug monitoring, that they were the first in Europe to offer such testing as a service and that both the laboratory and the scientists involved are of international standing, this omission appears unjustified as it skews much of the subsequent analysis based on a linked evidence-based approach.	Comparator tests were considered those which had evidence linking them to clinical outcomes. Ideally test treat RCT evidence.
	14.	70, 71, 75, 76, 77, 78, 79, 87	3.2.2.3 Results of assay type comparison studies 3.2.2.3.2 Comparison between	The abstract Nagore *et al*. (2015) describing a comparison between Promonitor and LISA-TRACKER for antibodies to infliximab has been accepted for publication in the British Digestive Disease Federation 2015 meeting to be held in London in Jun 2015 (Abstract DDF15-1162). This work has been provided to NICE to be considered in the evaluation. Dichotomous analysis with Cohen's Kappa provided a value of 0.8 for antibodies against infliximab. Quantification of antibodies using the LISA TRACKER Duo IFX test	We agree that the abstract clearly states how many patients in total were in the study, but it is unclear if they are all included in the analysis for antibodies.

Therapeutic monitoring of TNF inhibitors in Crohn's disease (LISA-TRACKER ELISA kits, Immundiagnostik TNFα-Blocker ELISA kits, and Promonitor ELISA kits)

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
			the index tests Antibodies to infliximab Antibodies to adalimumab Summary	was impaired by the limited measurement range of the kit (10-200 ng/mL), with 75% of measurements falling outside the upper concentration range thereby necessitating retesting. "*In summary we have one abstract giving a Cohen's Kappa of 0.8 between LISA-TRACKER and Promonitor, in tests for antibodies to adalimumab but it is not known how many samples and of which type were included in this analysis*". The abstract clearly states that 109 samples (69 infliximab and 40 adalimumab) from 71 patients with IBD were included in the study.	
	15.	70	3.2.2.3 Rationale	As explained above (comment 2), drug + / anti-drug antibody + results based on the use of solid phase assays should be interpreted with caution.	
	16.	79	3.2.2.3.2 Antibodies to infliximab	A trial will of 120 IBD patients will be available early to mid-July.	Thank you for this information.
	17.	80	3.2.2.3.3 Comparisons between index	On page 80 it is stated *"In detecting infliximab LISA-TRACKER gave positive results for 11 samples which were negative using either Amsterdam or Sanquin in-house ELISAs."* Should this read "…Amsterdam or Leuven…"?	Yes thank you for pointing this out.

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
			tests and comparator tests		
	18.	81	3.2.2.3.3 Comparisons between index tests and comparator tests Figure 10	The reconstructed Bland-Altman plots unambiguously show the high concordance between the Amsterdam and Leuven tests (panel a), whereas the comparisons between either of the above tests with LISA-TRACKER (panels b and c) show significantly lower concordance.	We agree that this study found lower concordance with LISA-TRACKER, we describe this in the report.
	19.	82	3.2.2.3.3 Comparisons between index tests and comparator tests Promonitor	Units for Promonitor antibody tests are incorrect ("*Analytical sensitivity of the Promonitor assay was higher than that of Amsterdam Sanquin RIA, 4ng/mL and 20AU/mL for Infliximab respectively, 2ng/mL and 30AU/mL for adalimumab respectively*"). Mass units are given in the report whereas, as for the Amsterdam tests, values should be expressed in units (*c.f.* comment 8).	Thank you for pointing this out. This has no material effect as units for antibodies are not comparable between the different tests.
	20.	82, 83	3.2.2.3.3 Comparisons between	The report says there is no evidence of concordance between tests at clinically relevant thresholds. This statement, implies that clinically relevant cut-offs are known, which as the report frequently mentions, is not	We agree and reported in our review that agreed cut-off levels for clinical use are not available. The point is that the tests are not far enough developed/tested to have

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
			index tests and comparator tests Promonitor Immundiagnostik	the case. However, we would emphasise the following statement taken from the study of Steenholdt *et al*. (2014) that argues "*despite variable analytical properties, common assays result in similar classifications and interventions in patients with infliximab treatment failure, and with comparable clinical outcomes*".	performance known at clinically relevant thresholds. The statement in the Steenholdt paper does not change this fact.
	21.	83	3.2.2.3.3 Comparisons between index tests and comparator tests Immundiagnostik	The final paragraph notes the high concordance between the Amsterdam and Immundiagnostik tests and then states that the link to the Leuven ELISA is not known in terms of their agreement (*c.f*. comment 18). To us this further emphasises the potential bias introduced by excluding the Amsterdam test from the comparators.	The NICE scope prescribed to only include as comparators those tests with a good link to clinical outcomes so that they could be used in a linked evidence approach. We have already used a very relaxed threshold for this due to the scarcity of the evidence. Adding the Amsterdam tests to the comparators would not have changed the conclusions.
	22.	88	3.2.2.3.3 Comparisons between	Why is the study by Ruiz-Argüello *et al*. (2013) excluded from this analysis since it provides a good link between the Amsterdam assays and Promonitor, and also an estimation of test accuracy?	It was included in the first phase, but not in the second phase as it did not make a link at a set threshold that could be used in a linked evidence approach. The Amsterdam assay

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
			index tests and comparator tests Summary		was not used in any of the 3 included comparative studies reporting patient outcomes while following a test-treatment algorithm.
	23.	124	3.2.4.4 Summary of major findings from three management studies	The report states "*For 64% of patients in the algorithm arm (those with therapeutic infliximab but no anti-drug antibodies) the algorithm recommended cessation of infliximab therapy. That cessation of infliximab therapy was not associated with reduced disease control suggests that infliximab may not be useful for most CD patients with LOR…*". While cessation of anti-TNF therapy may be considered for patients in remission, it maybe is inaccurate to state that an anti-TNF may not be useful for most patients with LOR, a point that is borne out in many of the studies referenced in the report.	We note this point.
	24.	124	3.2.4.4 Summary of major findings from three management	In the final paragraph the retrospective observational study of Vaughn *et al.* (2014) is noted as observing that "*in the trough-monitoring group some dose changes were dose reductions and dose escalations were considerably more moderate than in the clinically-managed group*". As stated previously (comment 9) alongside many other potential patient benefits the economic case for patient monitoring will only be made once thorough studies are available	We agree that 'the economic case for patient monitoring will only be made once thorough studies are available showing that drug usage, and therefore treatment costs, can be reduced by informed dosing based on trough levels of anti-TNF therapies on a patient-by-patient basis' except that these reductions need to translate into cost effectiveness over the life time of the of the model for the

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
			studies	showing that drug usage, and therefore treatment costs, can be reduced by informed dosing based on trough levels of anti-TNF therapies on a patient-by-patient basis. In addition the findings of Vaughn *et al*. that "*relative to clinical monitoring trough monitoring was associated with far superior retention on infliximab treatment…*" (page 125 and also referred to in section, 3.3 Summary of clinical effectiveness findings, page 139), notwithstanding possible weaknesses in the study design, is also an area where further research is called for.	population under consideration, and for the health system under consideration – in this case the NHS. Simple cost savings may not be sufficient. …..We agree that the 'findings of Vaughn et al.' are of interests and that further research in this area is warranted.
	25.	138	3.2.5.3.4 Summary	The paragraph headed "*Representativeness of available evidence*" would appear to conclude that despite negligible data the partial correlation of patient level data with the meta-analyses is sufficient to justify their inclusion in the study. To us this logic seems questionable.	Thank you for this comment. We are unclear about its precise meaning.
	26.	140	3.3 Summary of clinical effectiveness findings	The bullet point "*ELISA assays are susceptible to interference to a greater extent than other assays such as RIA and HMSA*", ignores the fact that solid phase RIA is also quite susceptible to interference (*c.f.* comment 3). While it is true that ELISA is less drug-tolerant, solid-phase RIA is not much better in terms of drug-tolerance. Besides, clinical decisions based on both methods are the same (Steenholdt et al.(2014).	Thank you for this suggestion for more precise wording. Because solid phase RIA was not an assay type used in the included studies and because Steenholdt 2014 employed fluid phase RIA (and not solid phase RIA) only fluid phase RIA was intended by the term "RIA".
	27.	147	4.1.3.2 Summar	"*Type specific health state costs included: costs for surgery which were modelled as the cost of inpatient*	Thank you for this. Actual surgery costs were included in the report as

Therapeutic monitoring of TNF inhibitors in Crohn's disease (LISA-TRACKER ELISA kits, Immundiagnostik TNFα-Blocker ELISA kits, and Promonitor ELISA kits)

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
			y of the Health Technology Assessment report by Dretzke *et al*. (2011)	*IBD interventions and post-surgery remission costs were based on outpatient surgical gastrointestinal follow-up*." Is this saying the actual surgery costs were omitted?	'Laparoscopic ileocolic resection: £6908 (NHS reference costs 2013/14) Ref 159 and expert opinion with further costs break down in appendices) in Table 35 (p174 of submitted version).
	28.	156	4.1.3.5 Quality assessment	In the concluding sentence of the section the Velayos economic model is criticised for "*the omission of half cycle correction*". Given that the use of half cycle correction in Markov modelling is itself controversial it might be more appropriate for the study authors to justify why they used this approach in their own modelling.	NICE DSU guidance currently recommends the use of a half cycle correction.
	29.	160	4.3.2 Developing the model structure	The report states "*In the models we compared concurrent and reflex testing conducted every three months compared to standard care for responders and those who experience loss of response*:". However in the project decision questions (see box, page 51) the report states "*Testing will be carried out:* *a) 3 to 4 months after start of treatment or* *b) 3 to 4 months and every 12 months from start of treatment*". Thus the question to be answered appears to have been changed, *ad hoc* with a different	We have addressed this in our addendum delivered to NICE.

Diagnostics Assessment Report (DAR) - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	Response by the EAG
				question,	
	30.	185	5.2.1 Clinical Effectiveness	The report casts doubts on the findings of Steenholdt *et al*. 2014 and 2015 for assuming that standard care would treat LOR with dose escalation, stating *"…dose escalation was the most expensive treatment option in the standard care arm and might not be representative for UK clinical practice*.". Our understanding is that since the IBD and UC market place has limited number of biological therapies available dose escalation is representative of UK clinical practice.	This is a moot point. Because of the lack of published evidence on this we relied for this information on our excellent UK clinical advisors. We believe that dose escalation in 100% of patients with LOR is not representative for UK clinical practice and represents the most costly standard care to compare drug monitoring against.