HEALTH TECHNOLOGY ASSESSMENT

VOLUME 20 ISSUE 83 NOVEMBER 2016
ISSN 1366-5278

Clinical effectiveness and cost-effectiveness of use of therapeutic monitoring of tumour necrosis factor alpha (TNF-α) inhibitors [LISA-TRACKER® enzyme-linked immunosorbent assay (ELISA) kits, TNF-α-Blocker ELISA kits and Promonitor® ELISA kits] versus standard care in patients with Crohn's disease: systematic reviews and economic modelling

Karoline Freeman, Martin Connock, Peter Auguste, Sian Taylor-Phillips, Hema Mistry, Deepson Shyangdan, Rachel Court, Ramesh Arasaradnam, Paul Sutcliffe and Aileen Clarke

National Institute for Health Research

DOI 10.3310/hta20830

Clinical effectiveness and cost-effectiveness of use of therapeutic monitoring of tumour necrosis factor alpha (TNF-α) inhibitors [LISA-TRACKER® enzyme-linked immunosorbent assay (ELISA) kits, TNF-α-Blocker ELISA kits and Promonitor® ELISA kits] versus standard care in patients with Crohn's disease: systematic reviews and economic modelling

Karoline Freeman,[1] Martin Connock,[1] Peter Auguste,[1] Sian Taylor-Phillips,[1] Hema Mistry,[1] Deepson Shyangdan,[1] Rachel Court,[1] Ramesh Arasaradnam,[2] Paul Sutcliffe[1]* and Aileen Clarke[1]

[1]Warwick Evidence, Division of Health Sciences, Warwick Medical School, University of Warwick, Coventry, UK
[2]Clinical Sciences Research Institute, University of Warwick, University Hospitals Coventry and Warwickshire NHS Trust, Coventry, UK

*Corresponding author

Declared competing interests of authors: Aileen Clarke is a member of the National Institute for Health Research, Health Technology Assessment and Efficacy and Mechanism Evaluation Editorial Boards. Aileen Clarke and Sian Taylor-Phillips are partly supported by the National Institute for Health Research Collaboration for Leadership in Applied Health Research and Care West Midlands at the University Hospitals Birmingham NHS Foundation Trust.

Published November 2016
DOI: 10.3310/hta20830

This report should be referenced as follows:

Freeman K, Connock M, Auguste P, Taylor-Phillips S, Mistry H, Shyangdan D, *et al*. Clinical effectiveness and cost-effectiveness of use of therapeutic monitoring of tumour necrosis factor alpha (TNF-α) inhibitors [LISA-TRACKER® enzyme-linked immunosorbent assay (ELISA) kits, TNF-α-Blocker ELISA kits and Promonitor® ELISA kits] versus standard care in patients with Crohn's disease: systematic reviews and economic modelling. *Health Technol Assess* 2016;**20**(83).

Health Technology Assessment is indexed and abstracted in *Index Medicus*/MEDLINE, *Excerpta Medica*/EMBASE, *Science Citation Index Expanded* (SciSearch®) and *Current Contents®*/ Clinical Medicine.

Health Technology Assessment

NICE TAR and DAR

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 4.058

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the ISI Science Citation Index.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: nihredit@southampton.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: http://www.nets.nihr.ac.uk/programmes/hta

This report

The research reported in this issue of the journal was commissioned and funded by the HTA programme on behalf of NICE as project number 14/69/03. The protocol was agreed in January 2015. The assessment report began editorial review in May 2015 and was accepted for publication in December 2015. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

Abstract

Clinical effectiveness and cost-effectiveness of use of therapeutic monitoring of tumour necrosis factor alpha (TNF-α) inhibitors [LISA-TRACKER® enzyme-linked immunosorbent assay (ELISA) kits, TNF-α-Blocker ELISA kits and Promonitor® ELISA kits] versus standard care in patients with Crohn's disease: systematic reviews and economic modelling

Karoline Freeman,[1] Martin Connock,[1] Peter Auguste,[1] Sian Taylor-Phillips,[1] Hema Mistry,[1] Deepson Shyangdan,[1] Rachel Court,[1] Ramesh Arasaradnam,[2] Paul Sutcliffe[1]* and Aileen Clarke[1]

[1]Warwick Evidence, Division of Health Sciences, Warwick Medical School, University of Warwick, Coventry, UK
[2]Clinical Sciences Research Institute, University of Warwick, University Hospitals Coventry and Warwickshire NHS Trust, Coventry, UK

*Corresponding author p.a.sutcliffe@warwick.ac.uk

Background and objectives: Systematic reviews and economic modelling of clinical effectiveness and cost-effectiveness of therapeutic monitoring of tumour necrosis factor alpha (TNF-α) inhibitors [using LISA-TRACKER® enzyme-linked immunosorbent assay (ELISA) kits (Theradiag, Marne La Vallee, France, or Alpha Laboratories, Heriot, UK), TNF-α-Blocker ELISA kits (Immundiagnostik AG, Bensheim, Germany) and Promonitor® ELISA kits (Proteomika, Progenika Biopharma, Bizkaia, Spain)] versus standard care for Crohn's disease (CD).

Methods: Multiple electronic databases were searched from inception to December 2014 in order to identify primary studies and meta-analyses.

Population: Patients with moderate to severe active CD treated with infliximab (IFX) (Remicade®, Merck Sharp & Dohme Ltd, Kenilworth, NJ, USA) or adalimumab (ADA) (Humira®, AbbVie Inc., North Chicago, IL, USA).

Intervention: Monitoring of serum anti-TNF-α (IFX or ADA) and/or of anti-drug antibody levels using test assays with a test–treatment algorithm.

Comparator: Standard care.

Outcomes: Any patient-related outcome, test agreement and cost-effectiveness estimates. The quality assessments used recognised checklists (Quality Assessment of Diagnostic Accuracy Studies-2, Cochrane, Philips and Consolidated Health Economic Evaluation Reporting Standards). Evidence was synthesised using narrative review and meta-analysis. A Markov model was built in TreeAge Pro 2013 (TreeAge Software, Inc., Williamstown, MA, USA). The model had a 4-week cycle and a 10-year time horizon, adopted a NHS

and Personal Social Services perspective and used a linked evidence approach. Costs were adjusted to 2013/14 prices and discounted at 3.5%.

Results: We included 68 out of 2434 and 4 out of 2466 studies for the clinical effectiveness 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 [two randomised controlled trials (RCTs) and one retrospective observational study] investigated outcomes following implementation of a test algorithm. None used the specified commercial ELISA immunoassay test kits. Neither of the two RCTs demonstrated clinical benefit of a test–treatment regimen. A meta-analysis of 31 studies to estimate test accuracy for predicting clinical status indicated that 20–30% of test results are likely to be inaccurate. The four cost-effectiveness studies suggested that testing results in small cost reductions. In the economic analysis the base-case analysis showed that standard practice (no testing/therapeutic monitoring with the intervention tests) was more costly and more effective than testing for IFX. Sensitivity and scenario analyses gave similar results. The probabilistic sensitivity analysis indicated a 92% likelihood that the 'no-testing' strategy was cost-effective at a willingness to pay of £20,000 per quality-adjusted life-year.

Strengths and limitations: Rigorous systematic reviews were undertaken; however, the underlying evidence base was poor or lacking. There was uncertainty about a linked evidence approach and a lack of gold standard for assay comparison. The only comparative evidence available for economic evaluation was for assays other than the intervention assays.

Conclusions: Our finding that testing is not cost-effective for IFX should be viewed cautiously in view of the limited evidence. Clinicians should be mindful of variation in performance of different assays and of the absence of standardised approaches to patient assessment and treatment algorithms.

Future work recommendations: There is substantial variation in the underlying treatment pathways and uncertainty in the relative effectiveness of assay- and test-based treatment algorithms, which requires further investigation. There is very little research evidence on ADA or on drug monitoring in children with CD, and conclusions on cost-effectiveness could not be reached for these.

Study registration: This study is registered as PROSPERO CRD42014015278.

Funding: The National Institute for Health Research Health Technology Assessment programme.

Contents

List of tables

List of figures

List of boxes

List of abbreviations

5-ASA	5-aminosalicylic acid
ACCENT	A Crohn's disease Clinical trial Evaluating infliximab in a New long-term Treatment regimen trial
ADA	adalimumab
AUC	area under the curve
BNF	*British National Formulary*
CD	Crohn's disease
CDAI	Crohn's Disease Activity Index
CHEERS	Consolidated Health Economic Evaluation Reporting Standards
CI	confidence interval
CRP	C-reactive protein
DOR	diagnostic odds ratio
ELISA	enzyme-linked immunosorbent assay
EQ-5D	European Quality of Life-5 Dimensions
Fc	fragment crystallising
FC	faecal calprotectin
HBI	Harvey–Bradshaw Index
HMSA	homogeneous mobility shift assay
HTA	*Health Technology Assessment*
IBD	inflammatory bowel disease
IBDQ	Inflammatory Bowel Disease Questionnaire
ICER	incremental cost-effectiveness ratio
IFX	infliximab
IgG	immunoglobulin gamma
IgG1	immunoglobulin gamma type 1
IPD	individual patient data
IQR	interquartile range
ITT	intention to treat
LOR	loss of response
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NPV	negative predictive value
OR	odds ratio
PDAI	Perianal Disease Activity Index
PP	per protocol
PPV	positive predictive value
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PSA	probabilistic sensitivity analysis
PSS	Personal Social Services
QALY	quality-adjusted life-year
QoL	quality of life
QUADAS-2	Quality Assessment of Diagnostic Accuracy Studies-2
RCT	randomised controlled trial
RIA	radioimmunoassay
ROC	receiver operating characteristic
RR	relative risk
SD	standard deviation
SEM	standard error of mean
sROC	summary receiver operating characteristic
TAXIT	Trough level Adapted infliXImab Treatment
TNF-α	tumour necrosis factor alpha
UC	ulcerative colitis

Note

This monograph is based on the Technology Assessment Report produced for NICE. The full report contained a considerable number of data that were deemed confidential. The full report was used by the Appraisal Committee at NICE in its deliberations. The full report with each piece of confidential data removed and replaced by the statement 'confidential information (or data) removed' is available on the NICE website: www.nice.org.uk.

The present monograph presents as full a version of the report as is possible while retaining readability, but some sections, sentences, tables and figures have been removed. Readers should bear in mind that the discussion, conclusions and implications for practice and research are based on all the data considered in the original full NICE report.

Plain English summary

Crohn's disease is a serious chronic inflammatory condition of the digestive tract. It currently affects about 115,000 patients in the UK. Severely ill patients can be treated with drugs called infliximab and adalimumab.

These are expensive drugs for the NHS. Some patients improve on them, whereas others improve initially but then lose response. One cause of lost response is that the patient develops antibodies against the drug which cancel out the effect of treatment.

Tests have been developed to measure both the level of drug and the level of antibodies against these drugs in the patient's blood. The idea is that treatment can be adapted in response to the test outcomes to ensure that the patient is on the best treatment for them.

In this assessment we systematically reviewed the literature for three of these new tests and combined the results to obtain our own estimates of clinical effectiveness and cost-effectiveness.

We found that no test accurately measures levels of drugs or antibodies to drugs and that tests disagree, which means that it is difficult to assess the effectiveness of new tests. The model drew mainly on evidence from two randomised controlled trials using infliximab and showed that, compared with standard care, testing appeared to be more costly and less effective.

We conclude that more evidence is required to tell us how the tests and the treatment options prescribed by the test results can benefit the management of patients with severe Crohn's disease.

Scientific summary

Introduction

Crohn's disease (CD) is a chronic fluctuating inflammatory condition of the digestive tract. It is currently estimated to affect approximately 115,000 patients in the UK, with 3000 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 are monoclonal antibodies that inactivate tumour necrosis factor alpha (anti-TNF-α). The two anti-TNF-α agents considered here are infliximab (IFX) (Remicade®, Merck Sharp & Dohme Ltd, Kenilworth, NJ, USA) and adalimumab (ADA) (Humira®, AbbVie Inc., North Chicago, IL, USA).

Response to anti-TNF-α agents is variable. Loss of response (LOR) is thought to be caused by subtherapeutic drug levels or the development of anti-drug antibodies that neutralise anti-TNF-α and hasten clearance from the circulation. This idea has led to the development of test kits able to measure circulatory levels of anti-TNF-α drugs and the antibodies directed against them, and to the use of test results in treatment algorithms to bring the anti-TNF-α levels into the therapeutic range and to prevent continuing use of ineffective agents.

Decision problem

The decision problem for this assessment is:

- Does testing of TNF-α inhibitor levels and antibodies to TNF-α inhibitors (IFX or ADA) represent a clinically effective and cost-effective use of NHS resources in patients with moderate or severe CD whose disease responds to treatment or who have lost response to treatment with TNF-α inhibitors?

The comparator for testing is standard care with an appropriate anti-TNF-α.

Three commercially available test kits for estimation of serum anti-TNF-α agents and anti-drug antibodies have been identified as interventions. These are LISA-TRACKER® enzyme-linked immunosorbent assay (ELISA) kits (Theradiag, Marne La Vallee, France, or Alpha Laboratories, Heriot, UK), TNF-α-Blocker ELISA kits (Immundiagnostik AG, Bensheim, Germany) and Promonitor® ELISA kits (Proteomika, Progenika Biopharma, Bizkaia, Spain).

Objectives

Objective A: review of comparative performance of tests

To review and critique studies:

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

Objective B: description of algorithms

To describe algorithms used in studies that 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 C1).

Objective C1: review of clinical effectiveness of test algorithm combinations

To systematically review the literature comparing the clinical effectiveness of an intervention or other assays for anti-TNF-α agents and/or for anti-drug antibodies used in conjunction with a treatment algorithm in patients with CD treated with IFX or ADA with the clinical effectiveness of standard care (no tests or test-informed algorithm used) in patients with CD treated with the same anti-TNF-α agent.

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

To analyse correlation studies that investigate the relationship between test results for anti-TNF-α and anti-drug antibody levels, and clinical outcome measured as clinical response. This objective was added post protocol because of the paucity of studies which address the decision question.

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

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

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

Multiple electronic databases were searched from inception up to the point of searching, during October to December 2014. Supplementary searches were used to identify additional published and unpublished studies. Reference lists and citation searches of included studies and review articles were undertaken. Further information was provided by the companies.

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

- population – adults and children with moderate to severe active CD treated with IFX or ADA
- intervention – monitoring of serum anti-TNF-α (IFX or ADA) and/or anti-drug antibody levels using intervention tests or other tests implemented using a test–treatment algorithm
- comparator – standard care (no anti-TNF-α or antibody monitoring)
- outcomes – any patient-related outcome, test agreement, cost-effectiveness estimates
- study design – any primary study design, systematic reviews with meta-analyses.

Study quality assessments were undertaken using an adapted Quality Assessment of Diagnostic Accuracy Studies-2 checklist, the Cochrane risk-of-bias tool, the Downs and Black checklist, Philips' checklist and the Consolidated Health Economic Evaluation Reporting Standards. Data were extracted by one reviewer and checked by a second reviewer. Disagreements were resolved by consensus or with a third reviewer. Evidence was synthesised using narrative review and statistical methods when appropriate. Individual patient data were reconstructed from available Kaplan–Meier plots using the method of Guyot *et al.* (Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan–Meier survival curves. *BMC Med Res Methodol* 2012;**12**:9). Meta-analyses were undertaken in Stata version 11 (StataCorp LP, College Station, TX, USA) or using 'MetaAnalyst' software (Tufts University, Medford, MA, USA) and RevMan version 5.3 (The Cochrane Collaboration, The Nordic Cochrane Centre, Copenhagen, Denmark). The Harbord and Whiting (Harbord R, Whiting P. metandi: meta-analysis of diagnostic accuracy using hierarchical logistic regression. *Stata J* 2009;**9**:211–29) method of hierarchical meta-analysis was used for diagnostic studies.

Cost-effectiveness model

A de novo Markov model was built in TreeAge Pro 2013 (TreeAge Software, Inc., Williamstown, MA, USA) to evaluate the cost-effectiveness of a test algorithm strategy-based treatment versus standard care. Two populations were considered: (1) patients responding to treatment and (2) patients who had lost response to treatment. Two test strategies were assessed: (1) concurrent and reflex testing of drugs and (2) antibodies to the drugs (i.e. simultaneous or sequential drug and antibody testing). Concurrent testing yields four possible outcomes: drug+/antibody–, drug+/antibody+, drug–/antibody+ or drug–/antibody–. Reflex testing (antibody testing if drug tests are negative) yields three outcomes: drug+, drug–/antibody– or drug–/antibody+. The model structure was informed by the literature search and expert clinical advice. The model had a 4-week cycle, a 10-year time horizon and adopted a NHS and Personal Social Services perspective. Costs were adjusted to 2013/14 prices and discounted at 3.5% per annum. The starting point was a hypothetical cohort of patients aged 30 years. Outcomes are reported as incremental cost-effectiveness ratios (ICERs), expressed in terms of cost per quality-adjusted life-years (QALYs) gained. A linked evidence approach was adopted (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 4-week cycle lengths and different transition probabilities for LOR. Probabilistic sensitivity analysis was also undertaken (10,000 model runs).

Results

The searches identified 2434 studies of clinical effectiveness and 2466 studies of cost-effectiveness, of which 68 and four studies, respectively, 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 [homogeneous mobility shift assay (HMSA), radioimmunoassay (RIA), PROMETHEUS® ELISA (Prometheus Laboratories Inc., San Diego, CA, USA) or Leuven in-house ELISA (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* 2015;**148**:1320–9)].

Three studies – two randomised controlled trials (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 IFX. Neither of the RCTs found evidence of clinical benefit for a test–algorithm–treatment regimen. In the Trough level Adapted infliXImab Treatment trial, which investigated the effectiveness of drug monitoring following dose optimisation in patients with response to IFX treatment, 131 out of 178 (73.59%) patients with CD were in clinical remission before dose optimisation and 138 out of 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 the control group ($p = 0.353$). Both RCTs estimated cost savings in drug expenditure with a test–treatment algorithm compared with normal care. The retrospective observational study compared a proactive test–treatment algorithm with normal care and reported greater retention on IFX treatment for the intervention group. However, the algorithm was ill-defined. Much of the evidence comes from studies, including this retrospective study, that investigated mixed groups of patients with inflammatory bowel disease [CD and ulcerative colitis (UC)].

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/LOR using drug and/or anti-drug antibody levels.

Among these there were three studies that reported results from both drug and anti-drug antibody tests for individual patients (one for IFX-treated responders, one for IFX-treated patients with LOR and one for ADA-treated responders). However, the patients in these studies did not receive treatment according to a test–treatment algorithm, therefore no outcomes data from the studies were available, and outcomes data from the RCTs 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 ADA test-based treatment strategy. The model therefore addressed IFX therapy only.

In the base case, the de novo Markov model showed that for IFX reflex testing dominates concurrent testing (that means that it is less costly and produces more QALYs); however, no-testing is more costly and produces more QALYs than reflex testing, with an ICER of approximately £50,800.

However, the probabilistic sensitivity analysis indicated a 92% likelihood that the 'no-testing' strategy was cost-effective at a willingness to pay of £20,000 per QALY.

Discussion and conclusions

Main findings

The 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 which 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, albeit at a higher 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.

The main limitation relates to the availability of relevant high-quality evidence. Although we undertook extensive systematic searches 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 LOR) lack standardisation, which impact on the classification of patients in different studies. Consensus on treatment algorithms 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 were very limited concordance data from studies comparing test performance of different assays. Evidence on ADA was lacking.

Populating the economic model with information from the literature was problematic because of the small size and short duration of the studies and their use of 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 UC; 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 finding that testing for levels of IFX and its antibodies is 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 about which clinical algorithms to follow as a result of testing. There is very little research on ADA or the use of testing strategies and algorithms in children. The key questions for future research consideration are:

1. What is the relative performance of methods of measuring anti-TNF-α drug and their antibodies by ELISA kits compared with other methods, such as RIA and HMSA, and are any potential differences clinically significant? For example, is there a validated drug threshold that is a useful predictor of clinical outcome?
2. What are the best criteria for estimating response, non-response and LOR in CD?
3. At what time should assessments of drug and antibody take place?
4. What is the effectiveness of clinical algorithms for disease management in response to testing in the UK?
5. What is the clinical effectiveness and cost-effectiveness of monitoring patients with CD on ADA and for paediatric patients with CD?
6. What is the relevance of cotreatment with immunosuppressants in the monitoring of anti-TNF-α agents and their antibodies?
7. Is there a benefit of measuring total drug/antibodies compared with measurements of free drug/antibody alone?

Study registration

This study is registered as PROSPERO CRD42014015278.

Funding

Funding for this study was provided by the Health Technology Assessment programme of the National Institute for Health Research.

Chapter 1 Introduction

Overview

Anti-tumour necrosis factor alphas (TNF-αs), including infliximab (IFX) (Remicade®, Merck Sharp & Dohme Ltd, Kenilworth, NJ, USA) and adalimumab (ADA) (Humira®, AbbVie Inc., North Chicago, IL, USA), are given to patients with inflammatory bowel disease (IBD), including Crohn's disease (CD), as a second- or third-line therapy. Response to anti-TNF-α treatment varies among patients treated for inflammatory chronic conditions. Although some patients stay in response over a long period of time, some are weaned off the drug because it is no longer needed and others may lose response at some stage during treatment. Loss of response (LOR) can occur for various reasons, the most common being (1) formation of antibodies against the drug, which neutralise the drug's action, rendering it ineffective; and (2) ongoing illness as a result of inflammation that is not meditated by TNF-α. It has been proposed that measurement of serum levels of anti-TNF-α and its antibodies can aid the management of patients with chronic diseases on anti-TNF-α drugs.

Measurement of anti-TNF-α levels and its antibodies can 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 that was adopted in this review is a methodology to handle shortcomings in the evidence for medical test evaluations.[1] The idea is to link evidence from other relevant research to the expected benefits of the test in question when direct evidence from the test and its effects on patient outcomes is absent. The decision-analytic model for the cost-effectiveness analysis 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 of the populations, tests and outcomes across the linkages.[1]

This report contains reference to confidential information provided as part of the National Institute for Health and Care Excellence (NICE) appraisal process. This information has been removed from the report and the results, discussions and conclusions of the report do not include the confidential information. These sections are clearly marked in the report.

Descriptions of the health problem: Crohn's disease

Crohn's disease 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,[2] 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]

Aetiology and pathology

Crohn's disease can affect adults, adolescents or children. CD manifests mainly during late adolescence or early adulthood. The first onset most commonly occurs between the ages of 16 and 30 years, with a second peak between the ages of 60 and 80 years. 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 of European descent.[4]

Crohn's disease 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 constant.

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 CD 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. Extraintestinal 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. A response is defined as a reduction in symptoms.

Measurement of disease activity

Crohn's disease can be difficult to diagnose because symptoms overlap with those of other gastrointestinal disorders, such as ulcerative colitis (UC) and irritable bowel syndrome. Investigations to aid diagnosis include taking the patient's medical history, physical examination, blood and stool tests and, finally, endoscopy to confirm diagnosis. As the treatment for CD depends on the location and severity of disease, an 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 'reference standard' for the assessment of disease severity, which has important implications for this assessment. For example, 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 that emphasises the less subjective laboratory parameters has been developed.[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.[5] However, these tools have been primarily developed for clinical trials rather than clinical practice. In clinical practice, the use of endoscopy to assess mucosal healing as an indicator of response and remission is becoming increasingly important, and the potential of objective laboratory markers, such as C-reactive protein (CRP) and faecal calprotectin (FC), for the assessment of disease activity, risk of complications and prediction of relapse, and for monitoring the effect of therapy has been recognised.[11]

The variables measured by the CDAI measures include number of liquid stools, abdominal pain, general well-being, extraintestinal 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.[2]

Although the CDAI uses a symptom diary of the patient over 7 days for the assessment, the HBI uses only a 1-day 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. HBI scores 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 technology appraisal 187:[6]

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

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 CD was defined for the purpose of the indication of IFX or ADA treatment as:

> *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.*
>
> *NICE (2010) technology appraisal number 187.* Infliximab (Review) and Adalimumab for the Treatment of Crohn's Disease. *London: NICE. Available from www.nice.org.uk/guidance/TA187.*[6] *Reproduced with permission. This information is accurate at time of publication*

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

Mild to 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.*
>
> *Hanauer and Sandborn*[7]

Moderate to 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.*
>
> *Hanauer and Sandborn*[7]

Severe to 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.*
>
> *Hanauer and Sandborn*[7]

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'.*
>
> *Hanauer and Sandborn. Reprinted by permission of Nature Publishing Group from Macmillan Publishers Ltd: The American Journal of Gastroenterology, Hanauer SB, Sandborn W. Management of Crohn's disease in adults.* Am J Gastroenterol *2001;**96**:635–43,*[7] *copyright 2001*

Management and care pathway

The treatment of CD is complex; in general, it aims at (1) reducing symptoms through induction and maintenance of remission, (2) minimising drug-related toxicity and (3) reducing the risk of surgery.[15] The management options for CD include drug therapy [e.g. glucocorticosteroids, 5-aminosalicylic acid (5-ASA), antibiotics, immunosuppressants, TNF-α inhibitors], enteral nutrition, smoking cessation and, in severe or chronic active disease, surgery. The choice of treatment among the available drugs is influenced by patient age, site and activity of disease, previous drug tolerance and response to treatment, and the presence of extraintestinal 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 extraintestinal manifestations, antibiotics, corticosteroids or immunosuppressant therapy. Between 50% and 80% of people with CD require surgery because of complications such as strictures causing symptoms of obstruction, fistula formation, perforation or failure of medical therapy.[2]

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]

Induction of remission according to the National Institute for Health and Care Excellence Clinical Guideline number 152[2]

Usually, at first presentation, patients with active CD are recommended to receive monotherapy treatment 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-ASA or enteral nutrition may be offered for patients who do not choose to take or who are intolerant of glucocorticosteroid therapy.

The addition of an immunosuppressant (azathioprine, mercaptopurine or methotrexate) to a conventional glucocorticosteroid or budesonide is recommended as an add-on therapy for inducing remission in patients who have active CD and who have experienced two or more inflammatory exacerbations in a 12-month period, or in whom glucocorticosteroid doses cannot be tapered. As advised in the current online version of the *British National Formulary* (BNF)[18] or *BNF for Children*,[19] the effects of azathioprine, mercaptopurine and methotrexate, as well as levels of neutropenia (in patients on azathioprine or mercaptopurine), should be monitored.[15]

In adults and children aged 6–17 years with severe active CD who fail to respond to the first line of treatment with conventional therapy (e.g. immunosuppressants or corticosteroids), or who are intolerant to or who have contraindications to conventional therapy, anti-TNF-α agents (IFX and ADA) are recommended as treatment options within their licensed indications. The administration of anti-TNF-α agents is recommended until 12 months after the start of treatment or until treatment failure (including the need for surgery), 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. IFX), considering drug administration costs, dose and product price per dose. The use of anti-TNF-α drugs for the treatment of CD is covered in the 2010 NICE technology appraisal guidance 187 [IFX (review) and ADA for the treatment of CD], which is summarised in NICE guidelines.[6]

Surgery should be considered early in the course of the disease for patients 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.[2]

Maintenance of remission according to the National Institute for Health and Care Excellence clinical guideline 152[2]

Patients 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. Patients 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 or general ill-health) to ensure that they keep their disease appropriately under review with their health-care professionals.

Patients 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 patients who are not able to tolerate, or who have contraindications to, azathioprine or mercaptopurine. Treatment with 5-ASA can be used to maintain remission after surgery.

If remission has been achieved with anti-TNF-α medication, then maintenance with anti-TNF-α with or without an immunosuppressant can be used. Continuation of treatment with IFX or ADA 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 patients who relapse after anti-TNF-α treatment may start it again.[15]

National Institute for Health and Care Excellence guidelines

The NICE guideline technology appraisal 187[6] describes when IFX or ADA should be used to treat patients with severe active or fistulising CD in the NHS in England and Wales.

Infliximab

The guideline states:

> *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.*
>
> *NICE (2010) technology appraisal number 187.* Infliximab (Review) and Adalimumab for the Treatment of Crohn's Disease. *London: NICE. Available from www.nice.org.uk/guidance/TA187.*[6]
> *Reproduced with permission. This information is accurate at time of publication*

Administration of IFX should follow this pattern:

> *. . . 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).*
>
> *NICE (2010) technology appraisal number 187.* Infliximab (Review) and Adalimumab for the Treatment of Crohn's Disease. *London: NICE. Available from www.nice.org.uk/guidance/TA187.*[6]
> *Reproduced with permission. This information is accurate at time of publication*

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

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 ADA should follow this pattern:

> *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.*
>
> *NICE (2010) technology appraisal number 187.* Infliximab (Review) and Adalimumab for the Treatment of Crohn's Disease. *London: NICE. Available from www.nice.org.uk/guidance/TA187.*[6]
> *Reproduced with permission. This information is accurate at time of publication*

Anti-tumour necrosis factor alpha agents

Crohn's disease is associated with elevated levels of the immune regulatory protein, TNF-α. The reasons for this elevation in CD are still largely unknown. TNF-α is a small cell signalling protein (cytokine) involved in

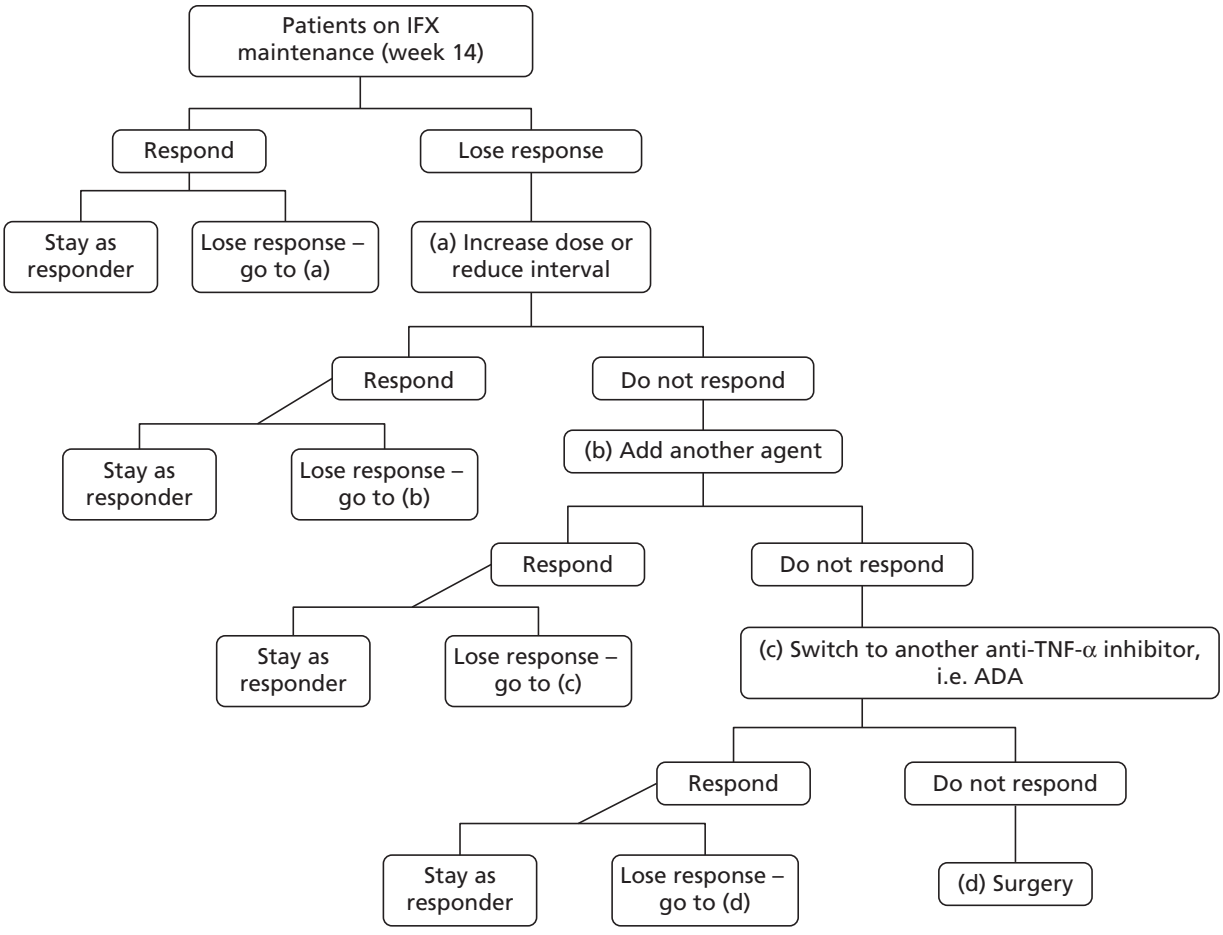


FIGURE 1 Patient pathway of patients with CD on IFX therapy.

inflammatory responses primarily by influencing regulation of various effector cells of the immune system. TNF-α has been shown to have a role in several inflammatory diseases including CD, UC, rheumatoid arthritis and ankylosing spondylitis. Anti-TNF-α agents bind to cell surface TNF-α and free TNF-α, and block their activity. Blocking of TNF-α with anti-TNF-α drugs has been shown to be successful for some patients with inflammatory diseases, including CD. Anti-TNF-α agents recommended by NICE for the treatment of CD are IFX and ADA. These monoclonal antibodies are introduced into the human body to bind and block TNF-α. 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 that are all of identical structure to and affinity for TNF-α.[15]

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. IFX belongs to the immunoglobulin gamma type 1 (IgG1) 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 IFX is generated from cultured mouse cells, the carbohydrate part of the molecule corresponds to that of mouse rather than human glycoproteins.[15]

Infliximab is composed of human IgG1 heavy-chain constant regions and human kappa light-chain constant regions (together representing 70% of the genetic make-up of the molecule), plus mouse-derived heavy- and light-chain variable regions (30% of the genetic make-up, 4/12 domains), which carry the binding sites with high affinity and specificity to TNF-α (see *Figure 2*). IFX was the first anti-TNF-α agent that was approved and licensed for treating severe active CD and active fistulising CD in adults and children aged > 6 years. It is administered intravenously over 1–2 hours.

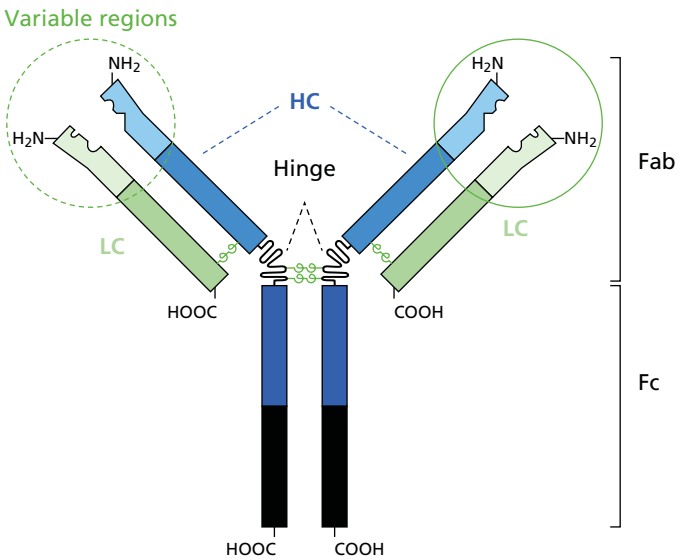


FIGURE 2 Diagrammatic representation of the structure of an IgG1 antibody molecule. The molecule has two heavy chains and two light chains; the heavy chains are joined by disulphide bonds (S–S) and each light chain is joined to a heavy chain by S–S bonding. The light and heavy chains have a variable 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 heavy and light chains are identical to other IgG1 antibodies and are called constant regions. The proteolytic enzymes papain and pepsin cut the molecule just above or below the S–S bonds holding the heavy chains together. When the split is below the heavy chain S–S bond, this generates a fragment crystallising (Fc) product and a fragment antigen-binding (Fab) product. When the split is above the heavy chain S–S bond, two antigen-binding fragments are formed [$F(ab)_2$]. COOH, carboxylic acid; HC, heavy chain; LC, light chain.

Side effects of IFX include:

- allergic reaction to the infusion (or IFX) 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 aged > 65 years
- 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]

Adalimumab

Adalimumab is a purely human IgG1 monoclonal antibody. ADA is a more recent anti-TNF-α 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 ADA include:

- reactions to the injection including pain, swelling, redness, bruising and itching
- allergic reaction to ADA including:
  - rashes or hives
  - swollen face, hands and feet
  - trouble breathing.
- greater susceptibility to infections such as colds, influenza, 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]

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 (QoL) fluctuates through time and, unsurprisingly, has been found to be better during remission. Studies using various disease-specific health-related QoL measures (such as McMaster's IBDQ, IBDQ-36, short IBDQ, Rating Form of IBD Patient Concerns, Cleveland Clinic questionnaire and Gastrointestinal Quality of Life Index) show a clear correlation between health-related QoL and symptoms. These measures in patients with CD have allowed utility estimates to be developed for patients in various clinical states.[20] QoL has been found to be somewhat worse in CD than in UC, 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.[21] 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]

Patients with CD can be cared for in primary or secondary care depending on symptom severity. Although general practitioners 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 immunosuppressants or anti-TNF-αs, or requiring surgery. It has been estimated that about 50% of patients with CD experience at least one flare per year. Of these, 20% of patients will require hospitalisation.[22] Disease flares have been found to be associated with a two- to threefold increase in hospitalisation and a 20-fold increase in cost compared with managing patients in remission.[23] Audit data show that anti-TNF-α agents are potentially cost-saving by successfully maintaining patients in remission and reducing hospital admissions. Cost reductions of £138 per patient at 6 months[24] and £2750 per patient at 12 months (excluding IFX costs) have been demonstrated in a before-and-after study of IFX therapy.[25] However, in that study both non-responders (£3608) as well as responders (£1656) incurred a considerably higher annual cost than the mean annual costs of long-term care in CD of £631/£762 (UK) and £838/£796 (Europe) estimated in another study using decision modelling.[26] Using 2008 prevalence figures, the total annual cost to the NHS for the approximate 60,000 patients with CD was estimated at £38M. Updating this figure with more recent prevalence data (115,000), but still using 2008 prices, would double that cost to £73M. This, however, might be a modest estimate when considering the wide range of measured 6-month costs for individual patients with CD (£73–£33,254).[23] The main drivers of costs were hospital admission, surgery and anti-TNF-α treatment.[26] In the 2003 *Health Technology Assessment* (HTA) publication by Clark *et al.*,[3] the average cost of a single 5-mg/kg IFX infusion for a 70-kg patient was reported to be £1804.80. No comparable data for ADA 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 lifetime care costs for patients with CD are now comparable to the costs of caring for patients with other major chronic diseases, such as diabetes mellitus and cancer.[22] This argues not only clinically but also economically for interventions that keep patients in remission and out of hospital. This review focuses on whether or not monitoring anti-TNF-α agents and their antibodies with enzyme-linked immunosorbent assays (ELISAs) could potentially contribute to this aim.

Rationale for measuring anti-tumour necrosis factor alpha drug and anti-drug antibody levels

Responders and non-responders definitions and incidence rates

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

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 that the assessment of response is undertaken. For example, A Crohn's disease Clinical trial Evaluating infliximab in a New long-term Treatment regimen (ACCENT) I, a randomised controlled trial (RCT) that investigated the benefit of maintenance IFX therapy in 573 active CD patients, assessed response after a single IFX infusion at 2 weeks.[27] The ACCENT II study was a post hoc analysis to determine the efficacy and safety of IFX therapy in patients with fistulising CD in which patients were assessed at 10 weeks.[28] The Crohn's trial of the fully Human Antibody adalimumab for Remission Maintenance RCT assessed response to ADA 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.[29] However, Ben-Horin and Chowers[30] 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 found that non-response ranged from 20% to 40% in clinical trials and from 10% to 20% in 'real-life' series.[30] In contrast, lack of remission at week 4 in patients with luminal CD was reported to be as high as 67% for IFX and 64% for ADA.[31,32] The true magnitude of the rate of primary non-response is therefore difficult to determine.

Factors associated with non-response are believed to include:[30,33,34]

- severity of disease
- duration of disease
- smoking
- drug elimination
- drug binding
- anti-drug antibodies
- alternative non-TNF-α-mediated disease pathways
- concomitant treatment with immunosuppressants
- prior failure of other anti-TNF-α.

Loss of response

Patients with an initial response to anti-TNF-α treatment can lose response 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,[35] in their review, reported a range of LOR to IFX of 11–48% (mean 37%) for varying lengths of follow-up, and de Boer *et al.*[33] reported a range of 21–46% for LOR to ADA. For this reason the incidence of LOR is better expressed as the annual risk of LOR per patient-year (13% for IFX[35] and 20.3% for ADA[36]). LOR to ADA and IFX did not differ significantly in a retrospective study of 375 patients; however, patients treated with ADA required more dose optimisation intervention than patients on IFX.[37]

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

- pre-medication with steroids
- concomitant immunosuppressants
- maintenance therapy as opposed to episodic treatment.

Mechanisms of LOR to anti-TNF-α agents are still unclear. The next section describes some of the possible mechanisms in more detail.

Anti-drug antibodies

Anti-drug antibodies can be elicited by IFX or ADA during therapy as a response by the human immune system to these foreign proteins; this is termed immunogenicity of anti-TNF-α agents. These anti-drug antibodies bind to the anti-TNF-α 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%[27,38,39] and 3–17%[39] for IFX and ADA, respectively. The similar rates for IFX and ADA might initially appear counterintuitive as ADA is a fully human recombinant protein, whereas IFX is partly human and partly mouse protein and, therefore, 'more' foreign. However, ADA, similar to IFX, is a foreign protein that 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 (i.e. formation of antibodies to a foreign protein).[39]

Levels of antibodies have been found to be higher during episodic treatment, at 36–61%,[39–41] than levels found during maintenance therapy. This indicates that other factors may influence immunogenicity.

The true incidence of antibodies in anti-TNF-α-treated patients is, therefore, unknown. The ability to mount an immune response and measurement of that response depends on a number of factors, including the method of measuring antibody levels and age, and also depends on concomitant treatment with immunosuppressants.[14,27,40–44] For that reason, concomitant immunosuppressants might be given to patients to prevent or reduce the formation of antibodies. This effect was not observed in one study for ADA,[14] and Vermeire *et al.*[45] reported that increasing anti-TNF-α above antibody-binding capacity might have similar effects to immunosuppressants, by neutralising free antibodies. Vande Casteele *et al.*[46] 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:[39]

- 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 or 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.[27,38,40,43,45,47,48] This can be explained by antibodies binding to the epitope of anti-TNF-α and neutralising the drug (i.e. making it unable to bind to TNF-α and inhibiting the working mechanism of anti-TNF-α) or by forming immune complexes with the drug (non-neutralising antibodies), which are subsequently cleared from the circulation (reducing the drug's bioavailability).[49]

Although the importance of the neutralising antibodies has been universally acknowledged,[14,34,49,50] reviewers seem to disagree in their conclusion about the importance of non-neutralising antibodies.[14,34] Over 90% of antibodies to IFX and ADA are neutralising.[51] In a meta-analysis, Garces *et al.*[49] estimated that detectable antibodies can decrease response to anti-TNF-α by as much as 80%.

An interesting additional observation was made by Steenholdt *et al.*,[52] who showed that immunoglobulin gamma (IgG) antibodies reacting with the fragment antigen-binding portion (see *Figure 2*) of IFX exist in IFX-naive IBD patients prior to treatment. The presence of pre-existing antibodies affected response and safety of IFX treatment in patients with CD, and the study concluded that the clinical utility of measuring pre-treatment antibodies should be assessed.

Drug levels

Although anti-drug antibodies have been shown to reduce anti-TNF-α drug levels, there are other known mechanisms that affect drug levels. These include dose and dosing interval, body mass index, sex, serum albumin levels (serum albumin transports drugs and can affect the half-life of drugs), concomitant immunosuppressants, severity of inflammation, mode of administration (intravenous vs. subcutaneous) and drug half-life.[14,44]

As a consequence, drug levels vary considerably between patients and within individuals over time.[34] Following administration of the anti-TNF-α 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[44] and it is largely unknown what the optimal drug levels would be at the different time points. Although a threshold trough level is thought to be needed for effectiveness, it is also known that supratherapeutic levels can cause infections and other adverse events.[14]

Anti-tumour necrosis factor alpha 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, immunosuppressant therapy and response in patients with CD on anti-TNF-α agents was the study by Baert *et al.*[43]

Baert *et al.*[43] was an early, and influential, study of the development of anti-drug antibodies to IFX in patients with CD; this study stimulated numerous subsequent investigations. The study enrolled 125 consecutive patients (38 with fistulising and 87 with luminal disease) who received 5 mg/kg of IFX at 0, 2 and 6 weeks. Responders (89/125, 71%) were retreated with this regimen if they required restart of IFX therapy according to clinical judgement. Mean treatment period was 10 months and median follow-up was 36 months. Anti-drug antibodies and IFX serum levels were measured before and at 4, 8 and 12 weeks after each infusion, using ELISA (PROMETHEUS® ELISA, Prometheus Laboratories Inc., San Diego, CA, USA). After five infusions, 76 out of 125 (61%) patients were classified as positive for anti-drug antibodies.

When a level of 8 µg of anti-drug antibodies/ml serum was selected, it was found that concentrations of anti-drug antibodies were > 8 µg/ml in 24 out of 56 (43%) patients taking immunosuppressants, compared with 52 out of 69 (75%) not taking immune suppressants. The relative risk (RR) of anti-drug antibodies concentration > 8 µg/ml in patients taking compared with those not taking suppressive therapy was 2.40 [95% confidence interval (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 in patients experiencing infusion reactions was higher than in those with no reactions ($p < 0.001$). Reactions were significantly more common in patients not taking immunosuppressant 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 levels > 8 µg/ml relative to those with levels < 8 µg/ml (median 35 days vs. 71 days; $p < 0.001$).

The level of IFX at 4 weeks after an infusion was correlated with the level of anti-drug antibodies prior to the infusion ($R^2 = 0.34$; $p < 0.001$) and was positively correlated with duration of response. IFX 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 IFX > 12 µg/ml was the use of immunosuppressant therapy. IFX 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 IFX therapy; anti-drug antibodies are associated with reduced IFX levels; duration of response is reduced by the presence of anti-drug antibodies; and production of anti-drug antibodies may be reduced when concomitant immunosuppressant therapy is employed.

Further evidence steadily accumulated from retrospective analyses of multiple clinical trials and case series,[38,53–55] and the observation that detectable trough levels of drug are associated with greater clinical efficacy is now well established.[44]

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

Without monitoring, the options for a clinician if the anti-TNF-α agent fails are to wait and see, to intensify drug treatment, to switch drug within its class or to switch to a different class of drugs.[34] 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.[56–59]

Drug and anti-drug antibody monitoring could be undertaken in good responders (i.e. those responding to an initial induction course of anti-TNF-α treatment), as well as in patients with LOR (i.e. those initially

responding to anti-TNF-α treatment but losing this response over time). The use of these technologies provides a clinician with potentially useful information that may guide an individual patient's future treatment. Such information may aid in anticipating the LOR in responders or allow drug optimisation, whereas for non-responders such analyses may help in estimating the likelihood of various candidate reasons for LOR. In non-responders with low levels of drug and high levels of anti-drug antibodies, for example, the loss or lack of response may be surmised to be because of rapid clearance of the drug as a result of the action of anti-drug antibodies. On the other hand, a low level of anti-TNF-α in the absence of anti-drug antibodies may be suggestive of non-immune mechanisms of rapid drug clearance, whereas high levels of drug in the absence of antibodies in non-responders may be suggestive of a pathology for the condition independent of TNF-α in a particular patient. Algorithms for future treatment based on anti-TNF-α and anti-drug antibody estimates have been published.[56–59]

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

- improve QoL 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-TNF-α 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]

Description of technology under assessment

Intervention technologies

Various assay procedures for anti-TNF-α agents and for anti-drug antibodies have been developed in the belief that the levels of circulating anti-TNF-α 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, Marne La Vallee, France, or Alpha Laboratories, Heriot, UK; the TNF-α-Blocker ELISA kits, Immundiagnostik AG, Bensheim, Germany; and the Promonitor® ELISA kits, Proteomika, Progenika Biopharma, Bizkaia, Spain) are the intervention technologies designed to measure IFX and ADA levels, and their antibodies and are investigated in this review.

These are all particular examples of solid-phase ELISAs. They estimate the following molecules in patient blood sera:

- IFX
- ADA
- anti-IFX antibodies
- anti-ADA antibodies.

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

Other ELISAs commercially available for measuring these molecules in sera include the SHIKARI® ELISA kits (Matriks Biotechnology Co. Ltd, Ankara, Turkey). These are not included as index tests in the NICE scope.

In the UK a number of non-commercial kits are also available, but these are not the focus of this report.

Other methodologies based on alternative principles of detection and measurement include (1) radioimmunoassays (RIAs) (liquid-phase assays using the radioisotope ^{125}I to label TNF); (2) cell reporter assays based on genetically engineered cells incubated in culture medium; and (3) mobility shift assays (liquid-phase assays using size-exclusion high-performance liquid chromatography with 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.

Enzyme-linked immunosorbent assays for infliximab and adalimumab

For details of the ELISA kits, refer to *Appendix 1*. All three specified ELISA methods employ similar principles in which, typically, microtitre 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-TNF-α 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 (3,3′,5,5′-tetramethylbiphenyl-4,4′-diamine)]. Peroxidase catalyses the following reaction:

$$\text{tetramethylbenzidine} + \text{hydrogen peroxide} \rightarrow \text{chromogen} + \text{water}. \quad (1)$$

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 IFX and for ADA are illustrated in *Figure 3*.

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

TNF-α-Blocker and Promonitor assays for anti-TNF-α drugs differ from the LISA-TRACKER assay in that the well coat is not TNF-α, but rather a reagent (antibody or antibody fragment) able to bind specifically to the TNF-α binding site of IFX or of ADA that is added to the microtitre well in the patient's sample (or calibrator). After washing, the second reagent is a peroxidase-labelled antibody able to bind the fragment crystallising (Fc) region of the anti-TNF-α antibody (*Figure 4*). Thus, fewer steps and a single reagent are used to detect well-bound anti-TNF-α drug. *Table 1* summarises the information describing the mechanisms underlying these assays.

Enzyme-linked immunosorbent assays for anti-drug antibodies

The LISA-TRACKER assays for antibodies to IFX and to ADA are illustrated in *Figure 5*.

This assay quantitatively estimates only free antibodies to anti-TNF-α. Therefore, 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 IFX, 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-TNF-α

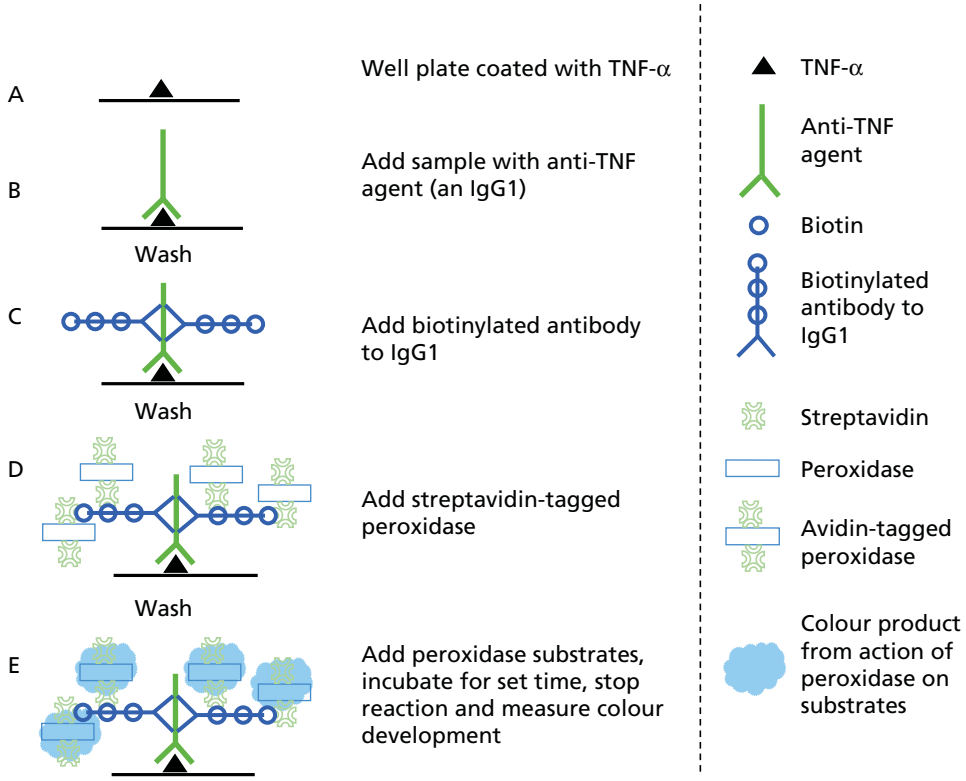


FIGURE 3 Diagrammatic representation of the LISA-TRACKER assay for IFX and ADA. Procedural steps C and D are detection steps that function to detect the anti-TNF-α that is bound to the microtitre well surface via TNF-α, ensuring a quantitative relationship between anti-TNF-α and peroxidase. Step E quantifies the amount of peroxidase (and, therefore, anti-TNF-α) in the microtitre well (streptavidin has four very high-affinity binding sites for biotin).

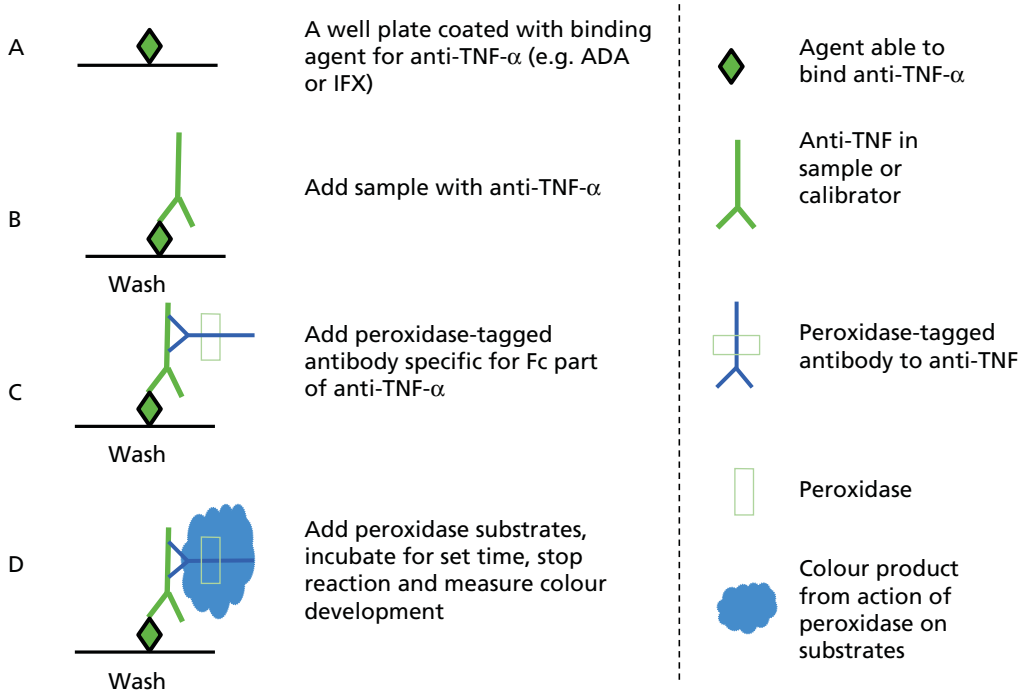


FIGURE 4 Diagrammatic representation of TNF-α-Blocker and Promonitor assays for IFX and ADA. Procedural steps C and D are detection steps that function to detect the anti-TNF-α (e.g. drugs IFX or ADA) that has been bound to the well surface via the coating (diamond). This ensures a quantitative relationship between anti-TNF-α (IFX or ADA) in the sample and peroxidase.

TABLE 1 Summary of ELISAs to be considered in this review for detection of IFX and ADA

Manufacturer and kit	Microtitre plate pre-coat	Detection reagent(s)
LISA-TRACKER indirect ELISA	Recombinant human TNF-α	Biotinylated mouse monoclonal IgG antibody directed to IgG Fc fragment
		Avidin-tagged peroxidase
TNF-α-Blocker ELISA[a]	Monoclonal anti-TNF-α antibody[b]	Peroxidase-labelled antibody[c]
Promonitor ELISA[a]	Monoclonal anti-TNF-α antibody[d]	Peroxidase-labelled monoclonal anti-TNF-α antibody[e]

a Further details supplied were labelled commercial in confidence.
b Confidential information has been removed.
c Confidential information has been removed.
d Confidential information has been removed.
e Confidential information has been removed.

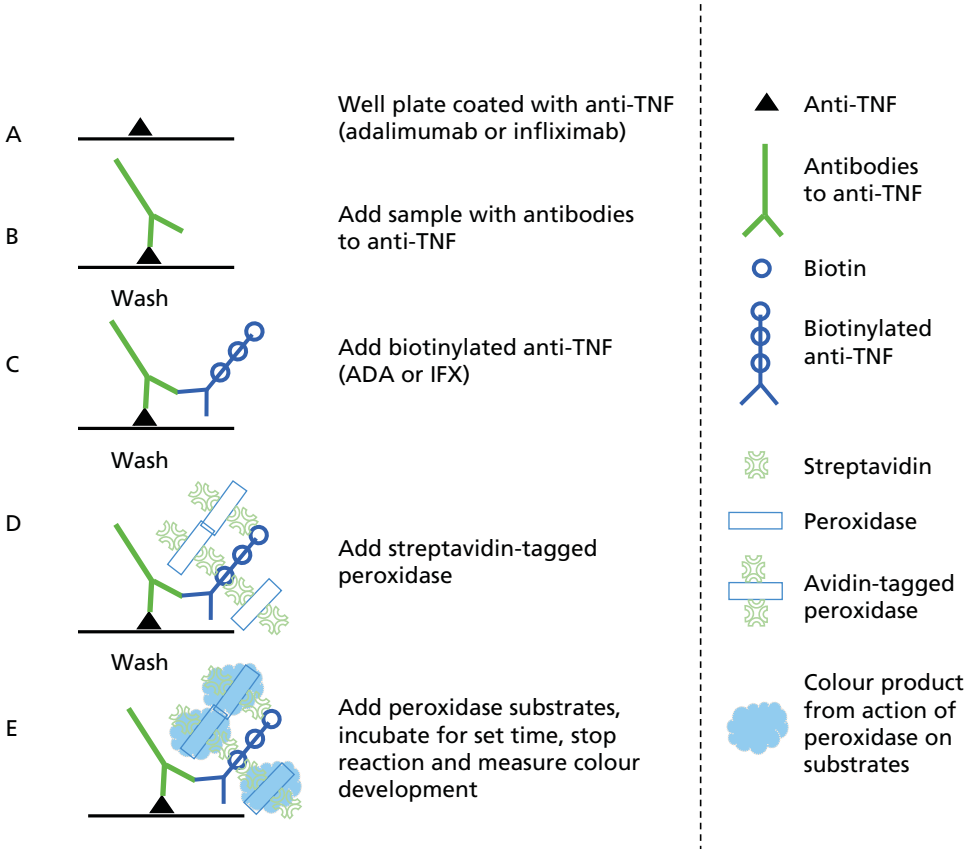


FIGURE 5 Diagrammatic representation of the LISA-TRACKER assay for antibodies to IFX or to ADA. This is a bridging assay in which antibodies to anti-TNF-α in the patient sample bridge the immobilised anti-TNF-α to the biotinylated anti-TNF-α. Procedural steps C and D are detection steps that function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

drugs.[51] These, and the other anti-drug antibodies, may hasten clearance of drug from the circulation as well as neutralising its binding capacity.

Tumour necrosis factor alpha and Promonitor assays differ from LISA-TRACKER assays in employing a single reagent for detecting well-bound anti-drug antibodies rather than two (biotinylated IFX or biotinylated ADA, plus avidin-conjugated peroxidase). *Table 2* summarises the information describing the mechanisms underlying these assays.[15]

TABLE 2 Summary of ELISAs to be considered in this review for detection of antibodies to IFX and ADA

Manufacturer and kit	Microtitre plate pre-coat	Detection reagent(s)
LISA-TRACKER bridge ELISA	Anti-TNF-α (i.e. IFX or ADA)	Biotinylated anti-TNF-α that binds to the paratope of the anti-drug antibodies in the sample
		Avidin-tagged peroxidase
TNF-α-Blocker ELISA IFX[a]	Confidential information has been removed	Confidential information has been removed
TNF-α-Blocker ELISA adalimumab[a]	Confidential information has been removed	Confidential information has been removed
Promonitor ELISA[a]	Confidential information has been removed[b]	Confidential information has been removed[c]

a Further details supplied were labelled commercial in confidence.
b Confidential information has been removed.
c Confidential information has been removed.

Brief overview of identified assay methods

There are no gold standard assays for anti-TNF-α agents or for antibodies to anti-TNF-α agents that might provide a robust basis for comparisons between the performances of different assays. According to US Medical Insurance assessments 'candidate' assays have been insufficiently investigated to establish any as a gold standard and, according to Steenholdt,[60] the evidence is incomplete on how these different assays may compare in practice.[14,61–64]

There appear to be four types of assay which differ fundamentally from each other. These are as follows.

1. ELISAs: solid-phase assays. These are available as commercial kits and several in-house methods are mentioned in the literature. Generally, the ELISAs quantitatively measure only 'free' anti-TNF-α and 'free' anti-drug antibodies, and it is acknowledged that the level of the unmeasured 'bound' anti-TNF-α 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-TNF-α and of anti-drug antibody present, in addition to the measured 'free' levels. In theory, this represents a potential deficiency in ELISAs, although whether or not this is serious in practice 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-TNF-α and total anti-drug antibody is supportable in practice.
2. RIAs: liquid-phase assays. These are provided as a total service rather than as purchasable kits. They measure total anti-TNF-α and total anti-drug antibody (probably as long as the anti-drug antibody light chain is λ class). These RIAs use ^{125}I-labelled human TNF-α and ^{125}I-labelled anti-TNF-α. 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}I-labelled TNF-α or ^{125}I-labelled anti-TNF-α further antibody (e.g. rabbit anti-human immunoglobulin λ-chain), which promotes the formation of immune complexes that are pelleted by centrifugation. The ^{125}I in the pellet is quantified in a gamma counter. Potential disadvantages include: (1) radiolabelled reagents do not keep indefinitely (^{125}I decays with a half-life of 59 days); (2) the laboratory needs to be equipped for handling hazardous (radioactive) materials; (3) some staff training may be necessary; and (4) the laboratory requires a gamma counter (preferably automated for high throughput). These factors obviously have cost implications for setting up RIAs.

3. Cell reporter assays: these assays utilise genetically engineered cells that respond to the presence of anti-TNF-α 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-TNF-α will lead to light emission and samples with antibodies to anti-TNF-α will quench light emission (for further information see *Appendix 2*).
4. Mobility shift assays: the mobility shift assay depends on detecting the shift in mobility of fluorescent probes when bound to either anti-TNF-α or anti-drug antibodies (for further information, see *Appendix 2*).

Timing and use of assays

The anti-TNF-α and anti-drug antibody assays are most frequently administered just before the next administration of the anti-TNF-α agent. This is said to allow measurement of a 'trough' level of anti-TNF-α and has been adopted to minimise effects from the presence of anti-TNF-α–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-TNF-α assays patients are classified as having therapeutic levels of anti-TNF-α or subtherapeutic 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]

Current usage of assays in the NHS

Current practice for monitoring TNF-α inhibitor antibody and drug levels in the UK is patchy because of 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.

Although 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 TNF-α inhibitors, has encouraged a cautious increase in uptake.

It is expected that therapeutic monitoring of TNF-α inhibitors might be useful in a number of clinical scenarios in the treatment of CD in the NHS, including for primary and LOR to anti-TNF-α therapy and in the optimisation of dosages for those who are already responding.

Chapter 2 Definition of decision problem

This report, undertaken for the NICE Diagnostics Assessment Programme, examines the clinical effectiveness and cost-effectiveness of ELISAs (LISA-TRACKER ELISA kits, TNF-α-Blocker ELISA kits and Promonitor ELISA kits) for measurement of patient blood levels of anti-TNF-α agents (IFX and ADA; also known as TNF-α inhibitors) and of antibodies to these agents (i.e. anti-drug antibody levels or anti-drug antibodies) in patients with CD whose disease responds to treatment with TNF-α inhibitor or who experience LOR during TNF-α inhibitor therapy. The report will help NICE to make recommendations about how well the assays work and whether or not 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-TNF-α 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. This review was registered on PROSPERO as CRD42014015278.

The decision questions for this project are shown in *Box 1*.

Overall aim of the assessment

The overall aim of this report was to present the evidence on the clinical effectiveness and cost-effectiveness of monitoring IFX and ADA and their antibodies in responders and patients with LOR when ELISA 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.

BOX 1 Decision questions

1. Does concurrent testing of TNF-α inhibitor levels and antibodies to TNF-α inhibitors represent a clinically effective and cost-effective use of NHS resources in patients with CD whose disease responds to treatment with a TNF-α inhibitor? Testing will be carried out:
   i. 3–4 months after start of treatment or
   ii. 3–4 months and every 12 months from start of treatment.
2. Does concurrent testing of TNF-α inhibitor levels and antibodies to TNF-α inhibitors represent a clinically effective and cost-effective use of NHS resources in patients with CD who experience LOR during maintenance treatment with a TNF-α inhibitor?
3. Does testing of TNF-α inhibitor levels followed by reflex testing of antibodies to TNF-α inhibitors if drug level is undetectable represent a clinically effective and cost-effective use of NHS resources in patients with CD whose disease responds to treatment with a TNF-α inhibitor? Testing will be carried out:
   i. 3–4 months after start of treatment or
   ii. 3–4 months and every 12 months from start of treatment.
4. Does testing of TNF-α inhibitor levels followed by reflex testing of antibodies to TNF-α inhibitors if drug level is undetectable represent a clinically effective and cost-effective use of NHS resources in patients with CD who experience LOR during maintenance treatment with a TNF-α inhibitor?

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 that compare an intervention test with a test method that 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 point 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 CD in therapeutic monitoring of TNF-α inhibitors (IFX and ADA) and their 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 tests or on a test that 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 CD following measurement of serum levels of anti-TNF-α drug and anti-drug antibodies.

To compare the algorithms used following therapeutic drug monitoring to the algorithms specified in the Trough level Adapted infliXImab Treatment (TAXIT) study for responders,[65] and in the reporting of LOR (algorithm adapted from the 2014 study by Scott and Lichtensteinin[66]).

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

To systematically review the literature comparing the clinical effectiveness of the intervention assays for anti-TNF-α agents and/or for anti-drug antibodies used in conjunction with a treatment algorithm in patients with CD treated with IFX or ADA with the clinical effectiveness of standard care (no tests performed or test-informed algorithm used) in patients with CD treated with IFX or ADA.

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, TNF-α-Blocker ELISA kits and Promonitor ELISA kits) (see *Objective A: review of comparative performance of tests*).

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

To analyse correlation studies that investigate the relationship between tests measuring anti-TNF-α and anti-drug antibody levels and clinical outcome in terms of response in patients with CD.

This objective was added because of the paucity of management studies that 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-TNF-α and anti-TNF-α antibody monitoring with LISA-TRACKER ELISA kits, TNF-α-Blocker ELISA kits and Promonitor ELISA kits in patients with CD compared with standard care (no anti-TNF-α monitoring).

To use a linked evidence approach when necessary (see objectives C1 and C2, 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.

Chapter 3 Clinical effectiveness review

Clinical effectiveness methods

Identification and selection of studies

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.*[43]) or population (e.g. Vande Casteele *et al.*[67]) 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 QoL 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.

The following bibliographic databases were searched from inception to the date of searching: MEDLINE; MEDLINE In-Process & Other Non-Indexed Citations; EMBASE; The Cochrane Library (including Cochrane Systematic Reviews, Database of Abstracts of Reviews of Effects, Cochrane Central Register of Controlled Trials, NHS Economic Evaluation Database and HTA databases); Science Citation Index and Conference Proceedings (Web of Science); Index to Theses; Digital Access to Research Theses-Europe; Dissertations & Theses; National Institute for Health Research (NIHR) HTA programme; and PROSPERO (International Prospective Register of Systematic Reviews). Searches were undertaken during the period October to November 2014 (see *Appendix 3* for exact search dates).

The following trial and patent databases were also searched: Current Controlled Trials, ClinicalTrials.gov, UK Clinical Research Network Portfolio Database, the World Health Organization's International Clinical Trials Registry Platform and Espacenet (European Patent Office).

Specific conference proceedings, selected with input from clinical experts and specialist committee members, were also checked from January 2010 to January 2015:

- European Crohn's and Colitis Organisation
- Digestive Diseases Week (meeting of the American Gastroenterology Association)
- British Society of Gastroenterology
- United European Gastroenterology Week
- 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 HTA publication – www.inahta.org/
- US Food and Drug Administration medical devices – www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/Databases/default.htm

- European Commission medical devices – http://ec.europa.eu/health/medical-devices/
- Theradiag – www.theradiag.com/en/
- Immundiagnostik AG – www.immundiagnostik.com/en
- Proteomika – www.proteomika.com/
- American College of Gastroenterology – http://gi.org/
- European Crohn's and Colitis Organisation – www.ecco-ibd.eu
- British Society of Gastroenterology – www.bsg.org.uk
- United European Gastroenterology – www.ueg.eu/
- The American Gastroenterology Association – 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 (Thomson Reuters, CA, USA). Included papers were checked for errata using PubMed.

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 a 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 expected to find only 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 (see *Objective A: review of comparative performance of test assays measuring anti-tumour necrosis factor alpha and/or anti-drug antibody levels*).

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.

Inclusion criteria

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

Objectives B and C1 Studies that satisfied the criteria outlined in *Table 3* were included.

Objective C2 Correlation studies were included if they provided at least one of the following:

1. a receiver operating characteristic (ROC) threshold analysis to determine an optimal drug-level threshold for predicting response (see *Results of threshold analysis studies*)
2. sufficient data to complete a 2 × 2 table of diagnostic accuracy of drug/anti-drug antibody level for prediction of response/LOR (*Table 4*) [see *Objective C2: studies relating test results to clinical state of patients (correlation studies)*]

Exclusion criteria

The criteria for exclusion of studies are presented in *Table 5*.

Objective C1 Tests that were not included under the heading of intervention but for which evidence was available on comparative diagnostic performance compared with an intervention test, and where clinical outcomes were also reported, were included for the purpose of performing linked evidence modelling (these included RIAs, cell reporter assays, liquid-phase mobility shift assays and in-house ELISAs).

TABLE 3 Criteria for study inclusion

Item	Criteria
Population	Patients with CD (adults and children) receiving IFX or ADA. Evidence on mixed patient groups containing CD and UC patients was included if patients with CD made up > 50% of the study population
Intervention	Use of LISA-TRACKER ELISA kits, TNF-α-Blocker ELISA kits and Promonitor ELISA kits to estimate plasma or sera levels of anti-TNF-α agents and/or of anti-drug antibodies in which test results are employed in conjunction with a treatment algorithm (*Box 2*). Other assay methods were considered for a linked evidence approach (see *Box 2*)
Comparator	Standard care (treatment decisions made on clinical judgement without measuring levels of anti-TNF-α and anti-drug antibodies)
Outcome	Any patient outcome (e.g. CDAI score-based response rate, any measure of change in severity of CD including physician's 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 QoL; and, secondarily, if two strategies compared are found clinically equivalent: time to result; number of inconclusive results; frequency of dose adjustment; and frequency of treatment switch)
Study design	All study designs were considered for inclusion
Health-care setting	Secondary and tertiary care

BOX 2 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).

BOX 2 Assay methods included as interventions in the review

Immundiagnostik TNF-α-Blocker ELISA kits [Immundiagnostik AG/Biohit Healthcare (Ellesmere Port, UK)]

- Immundiagnostik TNF-α-Blocker ADA, antibodies against infliximab (e.g. Remicade) ELISA (K9650).
- Immundiagnostik TNF-α-Blocker ADA, antibodies against adalimumab (e.g. Humira) ELISA (K9652).
- Immundiagnostik TNF-α-Blocker ADA, total antibodies against infliximab (e.g. Remicade) ELISA (K9654).
- Immundiagnostik TNF-α-Blocker ADA, total antibodies against adalimumab (e.g. Humira) ELISA (K9651).
- Immundiagnostik TNF-α-Blocker monitoring, infliximab drug level (e.g. Remicade) ELISA (K9655).
- Immundiagnostik TNF-α-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).

TABLE 4 Criteria for inclusion of studies that provide information for a 2 × 2 table

Item	Criteria
Population	Patients with CD (adults and children) receiving IFX or ADA. Evidence on mixed patient groups containing CD and UC patients was included if patients with CD made up > 50% of the study population
Intervention	Any assay to measure anti-TNF-α and/or anti-drug antibody levels
Outcome	2 × 2 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

TABLE 5 Criteria for exclusion of studies

Item	Criteria
Population	Studies with mixed patient groups containing < 50% CD patients
Intervention	Studies reporting an algorithm in which 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

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.

In addition, 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.

Review strategy

The general principles recommended in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement were used.[68] 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 retrieving the full publication and reaching consensus. 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.

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.

Quality assessment strategy

Objective A

For objective A, quality appraisal was completed using a modified Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) checklist.[69] In the patient selection domain three questions were included, using the standard version of the tool:

1. Was a consecutive or random sample of patients enrolled?
2. Was a case–control design avoided?
3. 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:

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

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

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

One additional question was added:

3. Were the number of failed results and measurement repeats reported?

For the reference standard, the two standard questions were used:

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

The best reference standard for test accuracy 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, with spiked samples test accuracy may not be reflective of test accuracy in clinical practice.

However, when 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, as a result of the lack of evidence, they are 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 or not 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.

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

Randomised controlled trials meeting the inclusion criteria were assessed using the Cochrane risk-of-bias tool.[70] The Downs and Black checklist[71] 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.

Methods of analysis/synthesis

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 that 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 positive reference standard results and negative reference standard results, or clinical outcomes when available) for therapeutic drug and detectable anti-drug for all index tests and comparators
2. characteristics of cases in which 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 positive reference standard sample results and negative reference standard sample results, when available) and Cohen's kappa. Two main secondary outcomes were also collected. First, the characteristics of cases in which there was disagreement and agreement between tests may provide information about the reason for and implications of the discordant results. Second, the shape of the Bland–Altman plots shows whether or not the difference between the two tests is dependent on absolute drug and anti-drug levels. Mean bias and the upper and lower limits of agreement were not particularly informative here, as we are interested in only one cut-off 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.[72] When there were sufficient studies, a meta-analysis of Cohen's kappa was considered.[72]

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 with the algorithm adapted from Scott and Lichtenstein[66] (for LOR) and to the algorithm adapted from Vande Casteele *et al.*[73] (for responders). 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.

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 (IFX or ADA) and patient group (patients with LOR or responders).

Study, treatment, population and outcome characteristics were summarised and compared qualitatively and, when possible, quantitatively in text, 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.[1] Evidence on outcomes reported by studies using other test methods (RIA, liquid-phase mobility shift assay and 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.

When relevant, Kaplan–Meier plots were available, individual patient data (IPD) were reconstructed using the method of Guyot *et al.*[74] Parametric models were fitted to reconstructed IPD using Stata version 11 (StataCorp LP, College Station, TX, USA).

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

For objective C2 we aimed to:

1. 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 a hierarchical meta-analysis of the data presented in the systematic reviews
2. 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 that 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, homogeneous mobility shift assay (HMSA)], the anti-TNF-α drug administered, dose regimen, patient inclusion and exclusion criteria, timing of testing, method for establishing clinical status, test cut-off point used and study design when these were reported. The populations of interest were (1) responders and responders who lost response; and (2) patients with LOR who continued with LOR or who regained a response.

Meta-analyses of single-test studies (i.e. measuring anti-TNF-α or anti-drug antibodies) were undertaken (1) to provide a pooled estimate for the probability of returning a specified test result after trough anti-TNF-α testing (useful for estimating reflex strategy test result probabilities); and (2) to provide pooled estimates for the probability of returning a specified test results by single test (i.e. anti-TNF-α or anti-drug antibodies) that can be compared for consistency with the corresponding probabilities from the identified patient-level studies.

RevMan version 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.[75,76] The prevalence of clinical status was meta-analysed using a random-effects model with 'MetaAnalyst' software (Tufts University, Medford, MA, USA).

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:

$$\text{Positive test} = (P \times \text{Sens}) + [(1 - P) \times (1 - \text{Spec})]. \quad (2)$$

And the probability of returning a negative test result is:

$$\text{Negative test} = [(1 - P) \times \text{Spec}] + [P \times (1 - \text{Sens})]. \quad (3)$$

Clinical effectiveness results

Search results

Figure 6 provides the PRISMA flow diagram for objectives A, B and C1 and 2. A total of 2428 records were identified through electronic searches. Six additional records were identified from other sources. The removal of duplicates left 1616 records to be screened, of which 1359 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 6* 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 ongoing trials in Clinical Trials.gov, Current Controlled Trials, the UK Clinical Research Network Portfolio and the World Health Organization's International Clinical Trials Registry Platform databases (carried out between 4 and 11 November 2014) retrieved seven relevant ongoing trials (see *Appendix 7*).

Of the included studies summarised in *Table 6*, studies comparing assay types address the performance of the different assays for a linked evidence approach (*Objective A: review of comparative performance of test assays measuring anti-tumour necrosis factor alpha and/or anti-drug antibody levels*). 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 prescribing patient management following test outcomes for drug and/or anti-drug antibody levels* and *Objective C1: clinical studies evaluating drug monitoring for the management of Crohn's disease patients (management studies)*]. Correlation studies with a ROC threshold analysis are summarised in *Results of threshold analysis studies* and correlation studies with sufficient 2 × 2 test accuracy data to contribute to the meta-analysis are reported in *C2: studies relating test results to clinical state of patients (correlation studies)*. Columns are not mutually exclusive; studies feature in more than one section.

Objective A: review of comparative performance of test assays measuring anti-tumour necrosis factor alpha and/or anti-drug antibody levels

Aim

To compare the performance of the different index tests (three specified ELISA kits) to one another, and to comparator tests that 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 or not therapeutic levels of drug and detectable levels of anti-drug 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).

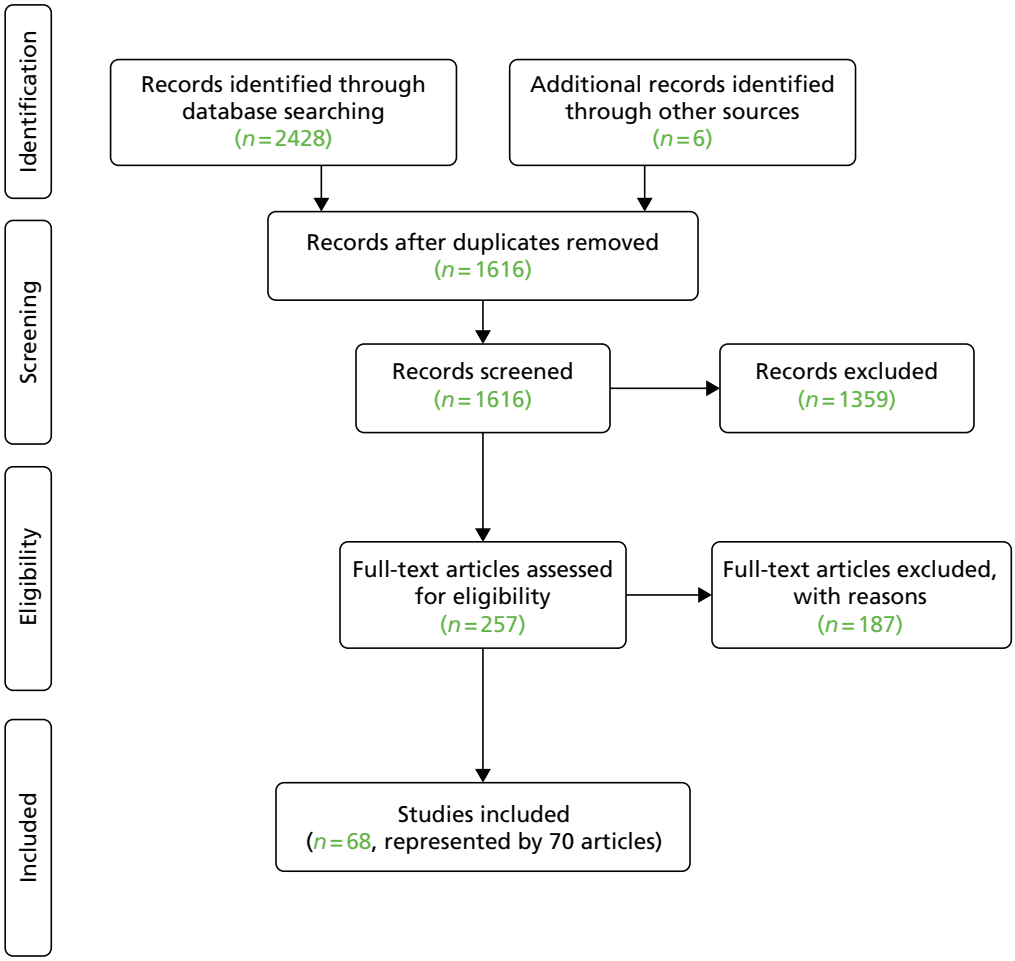


FIGURE 6 The PRISMA flow diagram describing the selection of included studies for the clinical effectiveness review.

TABLE 6 Overview of utility of included studies

Study (first author and year of publication)	Utility of included study and reference to relevant section in report			
	Assay comparison: *Objective A: review of comparative performance of test assays measuring anti-TNF-α and/or anti-drug antibody levels*	Management stipulated by algorithm and test outcome: *Objective C1: clinical studies evaluating drug monitoring for the management of CD patients (management studies)*	Correlation of drug/anti-drug antibodies and response – ROC: *Results of threshold analysis studies*	Correlation of drug/anti-drug antibodies and response – 2 × 2: *Objective C2: studies relating test results to clinical state of patients (correlation studies)*
Ainsworth *et al.*, 2008[47]	✗	✗	✗	✓
Baert *et al.*, 2014[77]	✗	✗	✓	✓
Ben-Bassat *et al.*, 2013;[78] abstract	✗	✗	✗	✓
Ben-Horin *et al.*, 2011[79]	✗	✗	✗	✓
Ben-Horin *et al.*, 2012[80]	✗	✗	✗	✓
Bodini *et al.*, 2014;[81] abstract	✓ (mapping only)[a]	✗	✗	✓

continued

TABLE 6 Overview of utility of included studies (*continued*)

Study (first author and year of publication)	Utility of included study and reference to relevant section in report			
	Assay comparison: ***Objective A: review of comparative performance of test assays measuring anti-TNF-α and/or anti-drug antibody levels***	**Management stipulated by algorithm and test outcome:** ***Objective C1: clinical studies evaluating drug monitoring for the management of CD patients (management studies)***	**Correlation of drug/anti-drug antibodies and response – ROC:** ***Results of threshold analysis studies***	**Correlation of drug/anti-drug antibodies and response – 2 × 2:** ***Objective C2: studies relating test results to clinical state of patients (correlation studies)***
Bortlik *et al.*, 2013[82]	✗	✗	✓	✓
Candon *et al.*, 2006[83]	✗	✗	✗	✓
Chiu *et al.*, 2013[84]	✗	✗	✓	✓
Cornillie *et al.*, 2014[85]	✗	✗	✓	✓
Corstjens *et al.*, 2013[86]	✓ (mapping only)	✗	✗	✗
Daperno *et al.*, 2013;[87] abstract	✓	✗	✗	✗
Dauer *et al.*, 2013;[88] abstract	✗	✗	✗	✓
Egea-Pujol *et al.*, 2013;[89] abstract	✓	✗	✗	✗
Eser *et al.*, 2013;[90] abstract	✓	✗	✗	✗
Eser *et al.*, 2013;[91] abstract	✓	✗	✗	✗
Farrell *et al.*, 2003[92]	✗	✗	✗	✓
Feagan *et al.*, 2012;[93] abstract	✗	✗	✓	✗
Frederiksen *et al.*, 2014[94]	✗	✗	✓	✓
Goldberg *et al.*, 2014;[95] abstract	✗	✗	✓	✗
Greathead *et al.*, 2014;[96] abstract	✓ (mapping only)	✗	✗	✗
Hanauer *et al.*, 2004[40]	✗	✗	✗	✓
Hauenstein *et al.*, 2012;[97] abstract	✓	✗	✗	✗
Hibi *et al.*, 2014[98]	✗	✗	✗	✓
Imaeda *et al.*, 2012[99]	✓ (mapping only)	✗	✗	✓
Imaeda *et al.*, 2014[100]	✓ (mapping only)	✗	✓	✓
Imaeda *et al.*, 2014[101]	✗	✗	✓	✗
Karmiris *et al.*, 2009[48]	✗	✗	✓	✗
Kong *et al.*, 2011;[102] abstract	✗	✗	✗	✓
Kopylov *et al.*, 2012[103]	✓ (mapping only)	✗	✗	✓
Lee *et al.*, 2012[63]	✗	✗	✗	✓ SR
Levesque *et al.*, 2014[104]	✗	✗	✓	✗

TABLE 6 Overview of utility of included studies (*continued*)

Study (first author and year of publication)	Utility of included study and reference to relevant section in report			
	Assay comparison: *Objective A: review of comparative performance of test assays measuring anti-TNF-α and/or anti-drug antibody levels*	Management stipulated by algorithm and test outcome: *Objective C1: clinical studies evaluating drug monitoring for the management of CD patients (management studies)*	Correlation of drug/anti-drug antibodies and response – ROC: *Results of threshold analysis studies*	Correlation of drug/anti-drug antibodies and response – 2 × 2: *Objective C2: studies relating test results to clinical state of patients (correlation studies)*
Marits *et al.*, 2014[105]	✗	✗	✓	✗
Marzo *et al.*, 2014;[106] abstract	✗	✗	✗	✓
Maser *et al.*, 2006[38]	✗	✗	✗	✓
Mazor *et al.*, 2013;[107] abstract	✗	✗	✓	✗
Mazor *et al.*, 2014[108]	✗	✗	✓	✓
McTigue *et al.*, 2013;[109] abstract	✓ (mapping only)	✗	✗	✗
Nagore *et al.*, 2015 (Dr Daniel Nagore, Progenika Biopharma, 2015, personal communication)	✓	✗	✗	✗
Nagore *et al.*, 2015;[110] abstract	✗	✗	✓	✓
Nanda *et al.*, 2013[111]	✗	✗	✗	✓ SR
Pallagi-Kunstar *et al.*, 2014[112]	✗	✗	✓	✗
Pariente *et al.*, 2012[59]	✗	✗	✗	✓
Paul *et al.*, 2012;[113] abstract	✗	✗	✓	✗
Paul *et al.*, 2013[58]	✗	✗	✓	✗
Paul *et al.*, 2014[114]	✗	✗	✗	✓ SR
Roblin *et al.*, 2014[115]	✗	✗	✓	✓
Ruiz-Arguello *et al.*, 2013[116]	✓	✗	✗	✗
Schatz *et al.*, 2013[117]	✓	✗	✗	✗
Semmler *et al.*, 2013;[118] abstract	✓ (mapping only)	✗	✗	✗
Singh *et al.*, 2014[119]	✗	✗	✓	✗
Steenholdt *et al.*, 2011[120]	✗	✗	✓	✓
Steenholdt *et al.*, 2013[121]	✓ (mapping only)	✗	✗	✗
Steenholdt *et al.*, 2013[52]	✗	✗	✗	✓
Steenholdt *et al.*, 2014[122]	✓	✓	✗	✗
Steenholdt *et al.*, 2014[123]	✓	✓	✗	✓
Steenholdt *et al.*, 2015[124]	✗	✓	✗	✗

continued

TABLE 6 Overview of utility of included studies (*continued*)

Study (first author and year of publication)	Utility of included study and reference to relevant section in report			
	Assay comparison: *Objective A: review of comparative performance of test assays measuring anti-TNF-α and/or anti-drug antibody levels*	**Management stipulated by algorithm and test outcome: *Objective C1: clinical studies evaluating drug monitoring for the management of CD patients (management studies)***	**Correlation of drug/anti-drug antibodies and response – ROC: *Results of threshold analysis studies***	**Correlation of drug/anti-drug antibodies and response – 2 × 2: *Objective C2: studies relating test results to clinical state of patients (correlation studies)***
Ungar *et al.*, 2014;[125] abstract	✓ (mapping only)	✗	✗	✗
Vande Casteele *et al.*, 2013[126]	✓	✗	✓	✓
Vande Casteele *et al.*, 2012[67]	✓	✗	✗	✗
Vande Casteele *et al.*, 2014;[127] abstract	✓ (mapping only)	✗	✗	✗
Vande Casteele *et al.*, 2015[73]	✗	✓	✗	✗
Vaughn *et al.*, 2014[128]	✗	✓	✗	✗
Wang *et al.*, 2010;[129] abstract	✓	✗	✗	✗
Wang *et al.*, 2011;[130] abstract	✓	✗	✗	✗
Wang *et al.*, 2012[131]	✓	✗	✗	✗
Ward *et al.*, 2013;[132] abstract	✗	✗	✓	✗
West *et al.*, 2008[133]	✗	✗	✗	✓
Yanai *et al.*, 2012;[134] abstract	✗	✗	✗	✓
Yarur *et al.*, 2013;[135] abstract	✗	✗	✓	✗
Total number of included references in each section	26 (11 mapping only)	5	24	31 and 3 SRs

✗, study did not provide data; ✓, study provided data; SR, systematic review.

a 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.

Rationale

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

1. drug positive and anti-drug antibody positive
2. drug positive and anti-drug antibody negative
3. drug negative and anti-drug antibody positive
4. drug negative and anti-drug antibody negative.

However, no trials were found that 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 or not any comparator tests (drug and anti-drug antibody tests) used in patients with CD 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 underestimation because of imperfect reference standard 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. When 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 or not 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.[72]

Results of assay type comparison studies

The search identified 25 relevant studies (reported in 26 references) that compared two or more assays to measure anti-TNF-α and/or anti-drug antibody levels in patients with CD (*Figure 7*). Of these, 10 were full texts (reported in 11 references)[67,86,99,100,103,116,121–123,126,131] and the remainder were conference abstracts, including one unpublished abstract provided by Proteomika (Dr Daniel Nagore, Progenika Biopharma, 2015, personal communication).[81,87,89–91,96,97,109,117,118,125,127,129,130] 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 (see *Appendix 8*).[81,86,96,99,100,103,109,118,121,125,127] Of the remaining 14 studies (15 references), which undertook relevant comparisons [including one unpublished abstract provided by Proteomika (Dr Daniel Nagore, Progenika Biopharma, 2015, personal communication)][67,87,89–91,97,116,117,122,123,126,129–131] (*Figure 8*), only five (six references) [including one unpublished abstract provided by Proteomika (Dr Daniel Nagore, Progenika Biopharma, 2015, personal communication)][67,87,117,122,123] reported concordance as numerical data or as Cohen's kappa (see *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 RIA,[123] the Leuven in-house ELISA,[73] the PROMETHEUS ELISA[128] and the PROMETHEUS HMSA.[128] All of the test comparisons identified in the search are detailed in *Figure 7*. The index tests are shown in light green, and those tests with some literature linking use of the test to clinical outcomes are marked in dark green.

Only those studies that 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 the use of the test to clinical outcomes, as described in *Objective C1: clinical studies evaluating drug monitoring for the management of Crohn's disease patients (management studies)*. Briefly, use of the RIA was linked to outcomes in a test–treat trial[123] and use of the PROMETHEUS ELISA and HMSA were linked to outcomes in a retrospective observational cohort.[128] The latter trial 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 a RCT.[73]

Quality appraisal

Only studies available as full texts[67,122,123,126,131] included in *Figure 8* were quality assessed. The results of the quality appraisal using a tailored QUADAS-2 tool are summarised in *Table 7* and *Appendix 10*. There was 'high' concern regarding patient selection across all papers, with a lack of clarity about the source of

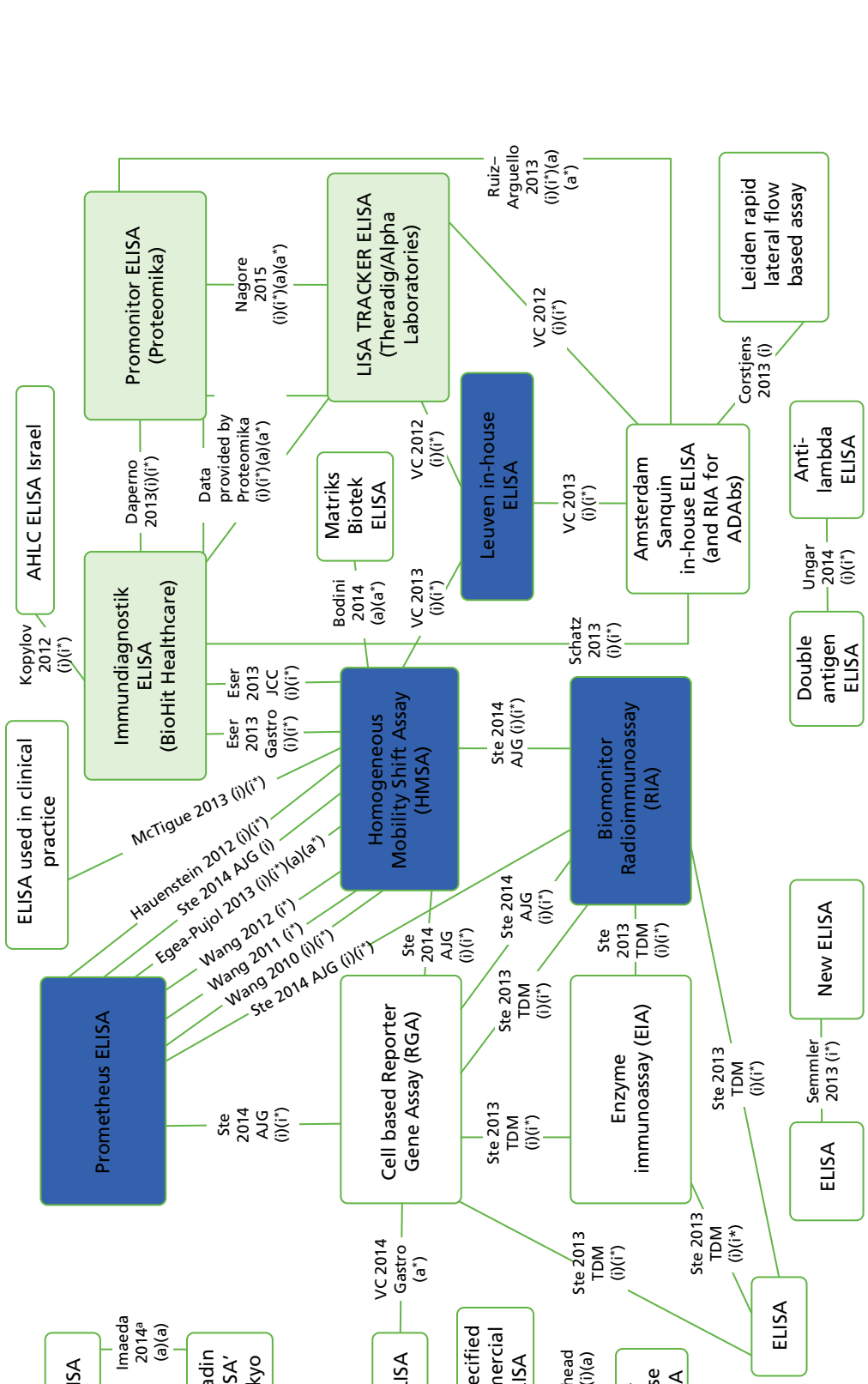


FIGURE 7 Summary of all of the tests for which there were comparisons in the identified literature. The index tests are shaded green and comparator tests are shaded blue. The comparisons are denoted as (i) for IFX, (a) for ADA, (i*) for anti-drug antibodies to IFX and (a*) for anti-drug antibodies to ADA. Studies include Bodini 2014;[81] Corstjens 2013;[86] Daperno 2013;[87] Egea-Pujol 2013;[89] Eser 2013 in *Gastroentrology*;[90] Eser 2013 in *J Crohns Colitis*;[91] Greathead 2014;[96] Hauenstein 2012;[97 a] Imaeda 2014;[100] Imaeda 2014;[99] Kopylov 2012;[103] McTigue 2013;[109] Nagore 2015 (Dr Daniel Nagore, Progenika Biopharma; 2015, personal communication); Ruiz-Arguello 2013;[116] Schatz 2013;[117] Semmler 2013;[118] Steenholdt 2014;[122,123] Steenholdt 2013;[121] Ungar 2014;[125] Vande Casteele 2012;[67] Vande Casteele 2013;[126] Vande Casteele 2014;[127] Wang 2010;[129] Wang 2011;[130] and Wang 2012.[131] ADAbs, anti-drug antibodies; AHLC, antihuman lambda chain antibody; AJG, *American Journal of Gastroenterology*; Gastro, *Gastroenterology*; JCC, *Journal of Crohn's and Colitis*; Ste, Steenholdt; TDM, *Therapeutic Drug Monitoring*; VC, Vande Casteele.

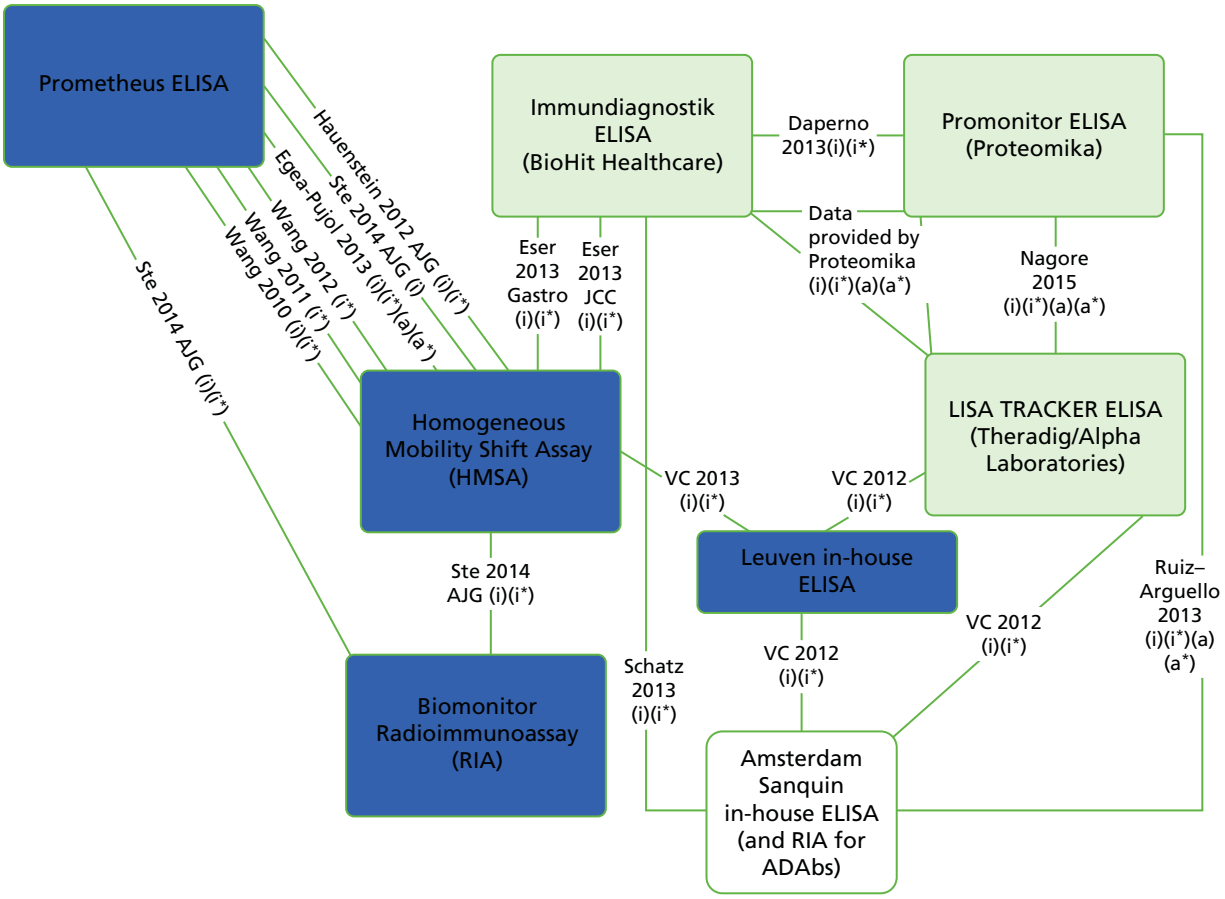


FIGURE 8 Comparisons that linked the index tests and comparator tests to each other. The index tests are shaded green and comparator tests shaded blue. The comparisons are denoted (i) for IFX, (a) for ADA, (i*) for anti-drug antibodies to IFX and (a*) for anti-drug antibodies to ADA. Studies include Daperno 2013;[87] Egea-Pujol 2013;[89] Eser 2013 in *Gastroentrology*;[90] Eser 2013 in *J Crohns Colitis*;[91] Hauenstein 2012;[97] Nagore 2015 (Dr Daniel Nagore, Progenika Biopharma, 2015, personal communication); Ruiz-Arguello 2013;[116] Schatz 2013;[117] Steenholdt 2014;[122,123] Vande Casteele 2012;[67] Vande Casteele 2013;[126] Wang 2010;[129] Wang 2011;[130] and Wang 2012.[131] AJG, *American Journal of Gastroenterology*; Gastro, *Gastroenterology*; JCC, *Journal of Crohn's and Colitis*; Ste, Steenholdt; VC, Vande Casteele.[131]

TABLE 7 Results of QUADAS-2[69] quality appraisal of included papers for objective A

	Concerns regarding risk of bias				Concerns regarding applicability		
Study (first author and year of publication)	Patient selection	Index test	Reference standard	Flow and timing	Patient selection	Index test	Reference standard
Vande Casteele *et al.*, 2012[67]	High	Low	High	High	High	Low	Low
Vande Casteele *et al.*, 2013[126]	High	Low	High	High	High	Low	Low
Steenholdt *et al.*, 2014[122,123]	High	Low	High	High	Low	High	Low
Wang *et al.*, 2012[131]	High	High	Unclear	High	High	High	Low

patients, and whether or not a consecutive series of patients was used. Some patients were selected from a biobank on the basis of index test results,[126] 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.[122] 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 UC as well as CD[126] and another included patients in unspecified numbers with unspecified conditions from departments of rheumatology and gastroenterology.[67]

There was 'low' concern overall about the implementation of the index tests but 'high' concern about applicability when 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.[131]

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[122,126] or the reference test was only conducted dependent on the results of the index test, thus introducing incorporation bias,[131] and one study did not include all patients in the analysis.[67]

Comparisons between the index tests

Results are presented here for all included studies, as outlined in *Figure 8*. This includes four full papers (five references),[67,122,123,126,131] as outlined in the quality assessment in *Quality appraisal*, and 10 abstracts [including one unpublished abstract provided by Proteomika (Dr Daniel Nagore, Progenika Biopharma, 2015, personal communication)].[87,89–91,97,116,117,129,130]

Adalimumab (Confidential information has been removed.)

One unpublished abstract that was provided by Proteomika [unpublished abstract provided by Proteomika (Dr Daniel Nagore, Progenika Biopharma, 2015, personal communication)] compared Promonitor assays with LISA-TRACKER assays for ADA. 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. For ADA, drug levels were different between the different assays: 6.0 [standard error of mean (SEM) 0.55] for Promonitor assays and 4.9 (SEM 0.39) for LISA-TRACKER assays. Pearson's R^2 was 0.83 and the authors concluded from the spiked samples that LISA-TRACKER assays underestimated ADA levels. In addition, 10% of samples were above the upper limit of quantification for LISA-TRACKER assays, and not for the Promonitor ELISA.

In summary, LISA-TRACKER assays may underestimate ADA 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 ADA

The same study that reported relationships between the index tests for ADA also reported some information on antibodies to ADA [unpublished abstract provided by Proteomika (Dr Daniel Nagore, Progenika Biopharma, 2015, personal communication)]. They reported a Cohen's kappa of 0.8 between Promonitor and LISA-TRACKER assays for antibodies to ADA, but it is unclear how many samples were included in this comparison. (Confidential information has been removed.)

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

Infliximab (Confidential information has been removed.)

There was one abstract comparing Promonitor to LISA-TRACKER assays for IFX [unpublished abstract provided by Proteomika (Dr Daniel Nagore, Progenika Biopharma, 2015, personal communication)]. In this abstract, 69 samples from an unspecified number of patients with IBD and an unspecified number of spiked samples were used. IFX drug levels were different between the different assays: 2.2 (SEM 0.24) for Promonitor and 3.4 (SEM 0.36) for LISA-TRACKER assays, for which the Pearson R^2 was 0.98 and the authors concluded from the spiked samples that LISA-TRACKER assays overestimated IFX levels. In addition, 23% of samples were above the upper limit of quantification for LISA-TRACKER assays and not for the Promonitor ELISA.

In one abstract, Daperno *et al.*[87] compared Immundiagnostik with Promonitor for IFX drug levels. In this study, Daperno *et al.*[87] enrolled a consecutive series of 66 patients (39 CD and 27 UC) undergoing regular IFX dosing

by intravenous therapy. It is unclear if additional samples were included. Bland–Altman plots of IFX drug levels showed mean bias of –1.8 µg/ml, indicating that Immundiagnostik estimates were on average lower than those of Promonitor by 1.8 µg/ml, and upper and lower limits of agreement of –10.8 and 7.1 µg/ml, respectively, indicating that 95% of IFX drug levels measured by Promonitor were between 10.8 µg/ml lower and 7.1 µg/ml higher than the same sample scores using the Immundiagnostik test (*Figure 9*).

In summary, LISA-TRACKER assays 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. (Confidential information has been removed.)

Antibodies to infliximab (Confidential information has been removed.)

One abstract described a comparison of Promonitor and LISA-TRACKER assays for antibodies to IFX [unpublished abstract provided by Proteomika (Dr Daniel Nagore, Progenika Biopharma, 2015, personal communication)]. 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. In the case of anti-drug measurements, 75% of samples had to be retested with LISA-TRACKER assays as the original measurements were above the upper limits of the measurement range.

One abstract, by Daperno *et al.*,[87] compared Immundiagnostik with Promonitor for measurement of antibodies to IFX. Daperno *et al.*[87] enrolled a consecutive series of 66 patients (39 CD and 27 UC) undergoing regular IFX dosing by intravenous therapy. 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 that were identical may provide some indication that the two tests for antibodies to IFX should not be considered equivalent.

In summary, one study found perfect agreement between Promonitor and LISA-TRACKER assays for antibodies to IFX and another study found few (6/63, i.e. < 10%) 'identical' results between Immundiagnostik and Promonitor.

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.

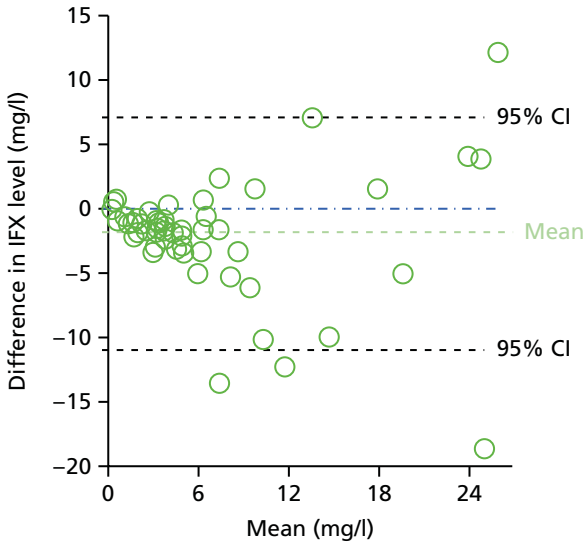


FIGURE 9 Reconstructed Bland–Altman plot comparing Promonitor IFX kits and Immundiagnostik TNF-α-Blocker ELISA kits. Based on data from Daperno *et al.*[87]

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

There is one study by Vande Casteele *et al.*[67] linking the LISA-TRACKER assay to the Leuven in-house ELISA for IFX and antibodies to IFX. This same study also compares LISA-TRACKER with 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 assays is also included here for interest.

Vande Casteele *et al.*[67] used 62 plasma samples from departments of gastroenterology and rheumatology. Of these, 36 were clinical samples from patients; the remaining 26 samples consisted of 24 spiked samples: 10 samples spiked with IFX, 10 spiked with antibodies to IFX, one spiked with ADA, three spiked with antibodies to ADA and two healthy control samples. 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 IFX, the LISA-TRACKER assay gave positive results for 11 samples that were negative when using either Amsterdam or Sanquin in-house ELISAs. Five of these were false-positive spiked samples that did not contain IFX but did contain antibodies to IFX (two samples) and antibodies to ADA (three samples). The remaining six 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 ADA was a true negative for IFX for both LISA-TRACKER and Leuven assays, but a false positive for the Amsterdam in-house ELISA. The Bland–Altman plots show that, although 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 slower than those using either Leuven or Amsterdam assays (*Figure 10*). This means that the levels of concordance between the LISA-TRACKER assay and the other assays will be dependent upon the particular threshold used. The Bland–Altman plot showed no pattern between the Leuven and the Amsterdam IFX tests, and from visual inspection the bias was near zero, with upper and lower limits of agreement between –10 and 10 mg/ml. The performance in detecting antibodies to IFX is less clear, with discordant results reported but not the type of sample in which these results were found. The Amsterdam in-house RIA detected antibodies to IFX in five samples when they were not detected by either of the other two assays, and both the Amsterdam and the LISA-TRACKER assays tests detected antibodies in three samples, which tested negative using the Leuven in-house ELISA. The thresholds used for drug and anti-drug levels, respectively, were 0.1 mg/l and 10 µg/l for LISA-TRACKER assays, 0.3 mg/l and 1 mg/l for Leuven and 0.002 mg/l and 12 arbitrary units/ml (1 arbitrary unit/ml equals approximately 10 µg/l) for Amsterdam. This higher threshold for the Leuven antibody ELISA may explain why fewer cases were detected.

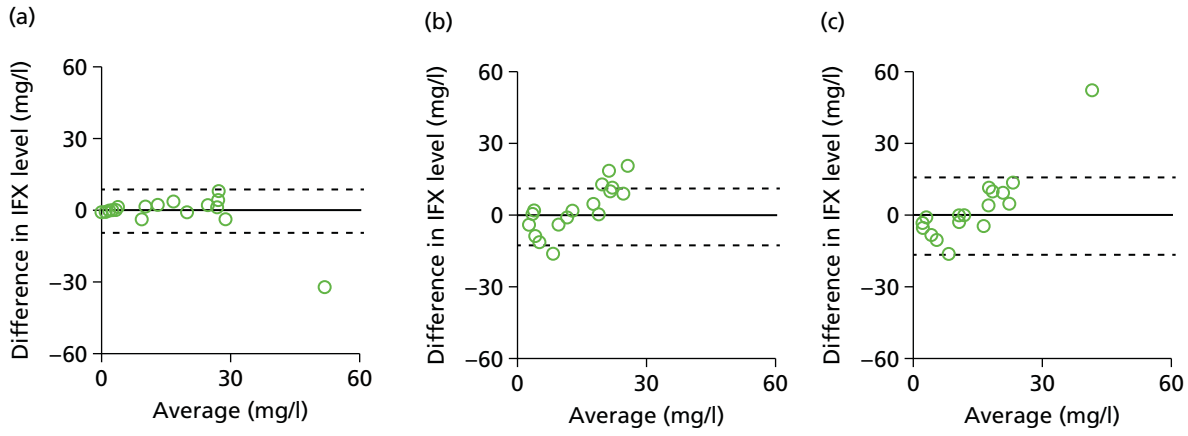


FIGURE 10 Reconstructed Bland–Altman plots of IFX levels (mg/l) comparing (a) Amsterdam in-house IFX ELISA and Leuven in-house IFX ELISA; (b) Amsterdam in-house IFX ELISA and LISA-TRACKER assays premium IFX kit; and (c) Leuven in-house IFX ELISA and LISA-TRACKER assays premium IFX kit. Based on data from Vande Casteele *et al.*[67]

In summary, in one study, using a range of clinical and spiked samples, there is some evidence that LISA-TRACKER assays may give false-positive results for IFX in the presence of antibodies to IFX or ADA, whereas 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 assays and then the Leuven in-house ELISA, but whether or not these antibodies detected are true positives or false positives is unclear.

Promonitor One letter, by Ruiz-Arguello *et al.*,[116] compared the Promonitor ELISA with the Amsterdam Sanquin ELISA for IFX and ADA, and the Amsterdam Sanquin RIA for both anti-drug antibodies. In addition, Vande Casteele *et al.*[67] 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 IFX, 30 samples over the range 0.001–5 µg/ml for ADA and 30 samples over the range 1–5000 arbitrary units/ml for both anti-drug antibodies to IFX and ADA.[116] The study defines set cut-off points and describes both assays as having no false-positive results for drug levels. However, no details are given of the sensitivity of each assay at those set cut-off 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: 10–30 ng/ml for IFX and 2–20 ng/ml for ADA, respectively. Bland–Altman plots of each assay in comparison with the known spiked concentrations gave a mean bias of –0.467 ng/ml [standard deviation (SD) 1.027 ng/ml] and 0.066 ng/ml (SD 0.196 ng/ml) for IFX and –1.140 ng/ml (SD 2.713 ng/ml) and –0.159 ng/ml (SD 0.488 ng/ml) for the Amsterdam Sanquin and Promonitor tests for ADA, respectively. The plots are not provided, but the authors describe a systematic overestimation of the drug levels by the Amsterdam Sanquin ELISA that increases with increased drug levels, which would explain the greater CIs for the mean bias estimate for the Sanquin ELISA. The authors describe this overestimation as occurring at drug levels of > 2 µg/ml. For antibodies to IFX and ADA, only correlation coefficients and analytical sensitivity were reported. Analytical sensitivity of the Promonitor assay was higher than that of Amsterdam Sanquin RIA, at 4 ng/ml and 20 arbitrary units/ml for IFX and 2 ng/ml and 30 arbitrary units/ml for ADA, respectively. Therefore, there is some evidence using spiked assays that the Amsterdam Sanquin assay may overestimate drug levels at higher concentrations, whereas the Promonitor assay may not. In a letter of response, however, Rispens and van der Kleij[136] 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 *LISA-TRACKER assays*. However, to recap in brief, for drug levels the Bland–Altman plot showed no pattern; therefore, 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 –10 and 10 mg/ml. For anti-drug antibodies to IFX, the Amsterdam RIA detected more cases than the Leuven assay, with the actual veracity of these unclear.

In summary, although 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 ADA or antibodies to ADA. In two abstracts, Eser *et al.*[90,91] compared the Immundiagnostik ELISA with the PROMETHEUS HMSA, and Schatz *et al.*[117] compared the Immundiagnostik ELISA with the Amsterdam Sanquin in-house tests (for IFX and antibodies to IFX), which are in turn compared with the Leuven in-house ELISAs by Vande Casteele *et al.*[67]

The two abstracts by Eser *et al.*[90,91] comparing Immundiagnostik ELISA with 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 IFX at the mid-infusion point, whereas ELISA returned inconclusive results because of interference from IFX. 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.*[117] linking Immundiagnostik ELISA to the Amsterdam in-house tests (ELISA for drug levels, RIA for antibody levels) compared performance of the two tests for IFX and anti-drug antibodies to IFX. They used serum samples from 202 paediatric patients, of whom 125 had been exposed to IFX and 77 were IFX naive. Samples were considered positive for IFX if they were above the limit of detectability, which was < 0.8 µg/ml for the Immundiagnostik ELISA and < 0.002 µg/ml for Amsterdam in-house ELISA. Overall agreement using Cohen's kappa was 0.792. Considering only the IFX-exposed patients, 25 were below the lower limit of detectability for both tests, leaving 87 who tested positive for both, 11 who tested positive only using the Amsterdam ELISA (measurements ranged from 0.1 to 2.3 µg/ml, so some of these will be below the lower limit of detectability using the Immundiagnostik test) and two whose results were not reported. For anti-drug antibodies to IFX, 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 (i.e. *LISA-TRACKER assays* and *Promonitor*). 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 –10 and 10 mg/ml; for anti-drug antibodies to IFX the Amsterdam RIA detected more cases than the Leuven assay, with actual veracity of these unclear.

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

Relationship between different comparator tests

One study, by Steenholdt *et al.*,[122] compared the performance of the Biomonitor RIA used in the test–treat trial with PROMETHEUS HMSA and PROMETHEUS ELISA for IFX and antibodies to IFX. Vande Casteele *et al.*[126] compared the Leuven in-house ELISA with PROMETHEUS HMSA. One full study, by Wang *et al.*,[131] and four further abstracts[89,97,129,130] compared the performance of HMSA with PROMETHEUS ELISA for IFX and antibodies to IFX, Egea-Pujol *et al.*[89] also made the same comparison for ADA and antibodies to ADA. However, three of these abstracts (Hauenstein *et al.*,[97] Egea-Pujol *et al.*[89] and Wang *et al.*[130]) did not provide data on concordance, Cohen's kappa or numbers of false-positive, true-positive, false-negative or true-negative test results, and will not be described further here.

The studies by Wang *et al.*[129,131] compared the performance of HMSA and of PROMETHEUS ELISA. Wang *et al.*[129] described 20 patients with IBD who had relapsed from treatment with IFX. ELISA detected IFX in 15 out of 20 patients and anti-drug antibodies to IFX in 15 out of 20 patients. HMSA detected IFX in 15 out of 20 patients and anti-drug antibodies to IFX in 18 out of 20 patients. It is not clear if the samples that tested positive on both tests were from the same 15 patients. The focus of Wang *et al.*[131] was to validate the performance of HMSA rather than to compare it with ELISA. Of 100 samples from healthy control participants, three were false-positive for antibodies to IFX for HMSA. This was to be expected as the cut-off point was determined from the same samples as mean plus two SDs. Repeat measurements of these three resulted in them being below the cut-off point, presumably regression to the mean. ELISA results for the 100 healthy control participants were not reported. Out of 100 patients with IBD selected as positive for antibodies for IFX on ELISA, five 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 that tested positive using HMSA, it is difficult to draw any conclusions at all. The only comparative data given constitute a plot of correlation that does not appear to show high correlation. The studies therefore did not provide useful concordance data for evaluation.

Vande Casteele *et al.*[126] compared the Leuven in-house ELISA to PROMETHEUS HMSA. Although 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 IFX in the presence of IFX; however, quantifying this is difficult, as reporting focused on other research questions. It is described in the discussion as the HMSA having detected the median 9 weeks earlier. However, in the absence of IFX and with the HMSA cut-off point for antibodies to IFX set at 7.95 U/ml, the Leuven in-house ELISA detected four more cases with antibodies to IFX. The authors report that PROMETHEUS has since lowered the threshold to 3.13 U/ml.

Steenholdt *et al.*[122] took 66 frozen patient samples (from patients with CD with LOR to IFX) from the test–treat trial[123] and reanalysed them using PROMETHEUS HMSA and PROMETHEUS ELISA. In all of these patients, the results of RIA using the samples before freezing were available and determined the treatment pathway at the time of the study. Threshold for positivity for IFX was unclear, but the lower limit of quantification was clearly defined as ≥ 0.15 µg/ml, ≥ 1 µg/ml and ≥ 1.4 µg/ml for RIA, HMSA and ELISA, respectively.[122] However, in the original paper reporting the RCT,[123] although the same lower limit of quantification thresholds were given, additional cut-off points of ≥ 0.5 µg/ml and ≥ 3 µg/ml were given for defining therapeutic drug levels for RIA and HMSA, respectively. However, no additional cut-off point was given for ELISA. Thresholds for anti-drug antibodies to IFX were ≥ 10 arbitrary units/ml, ≥ 3.13 arbitrary units/ml and ≥ 1.69 µg/ml for RIA, HMSA and ELISA, respectively. Using RIA, 54 out of 66 (82%) tested positive for IFX, in comparison with 58 out of 66 (88%) for HMSA and 50 out of 66 (76%) using ELISA. The concordance between the three tests is shown in *Figure 11*. In eight patients IFX was undetectable using all three tests and in 50 patients IFX was detectable using all three tests. In four patients IFX was detected by RIA and HMSA but not by ELISA, and in a further four patients IFX was detectable only by HMSA.

Using RIA, 18 out of 66 (27%) patients tested positive for anti-drug antibodies to IFX, in comparison with 22 out of 66 (33%) using HMSA and 6 out of 66 (9%) using ELISA. The concordance between the three tests is shown in *Figure 12*. In 43 patients, anti-drug antibodies to IFX were undetectable using all three tests, and in six patients anti-drug antibodies to IFX were detectable using all three tests. In 11 patients

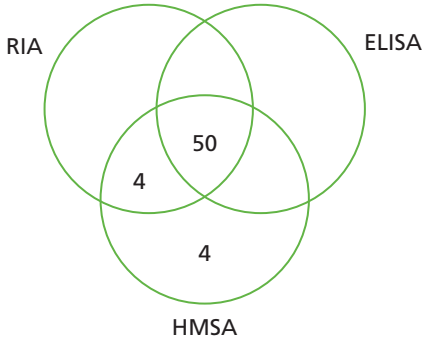


FIGURE 11 Concordance between RIA, HMSA and ELISA for detecting IFX in 66 patients with CD with LOR to IFX. In eight patients IFX was undetectable using all three tests.

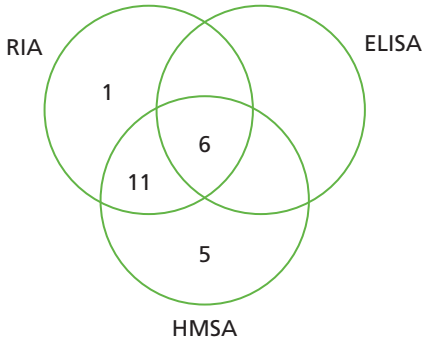


FIGURE 12 Concordance between RIA, HMSA and ELISA for detecting anti-drug antibodies to IFX in 66 patients with CD with LOR to IFX. In 43 patients anti-drug antibodies were not detectable in any of the tests.

anti-drug antibodies to IFX were detected by RIA and HMSA but not by ELISA, in a further five patients anti-drug antibodies to IFX were detectable only by HMSA and in one patient anti-drug antibodies to IFX were detectable only by RIA.

In summary, for IFX, RIA and HMSA agreed for 62 out of 66 patients, with the remaining four patients testing positive on HMSA but not RIA; HMSA and ELISA agreed in 58 out of 66 patients, with eight patients testing positive on HMSA but not ELISA; and RIA and ELISA agreed on 62 out of 66 patients, with the remaining four patients testing positive on RIA but not ELISA. The Bland–Altman plots comparing HMSA with RIA and ELISA with 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 to be independent of absolute drug concentrations.

In the case of anti-drug antibodies to IFX, RIA and HMSA agreed for 60 out of 66 patients, with five patients testing positive on HMSA and not on RIA, and one testing positive on RIA and not on HMSA. HMSA and ELISA agreed in 50 out of 66 patients, with 16 patients testing positive on HMSA but not on ELISA; RIA and ELISA agreed in 54 out of 66 patients, with the remaining 12 patients testing positive on RIA but not on ELISA.

Therefore, there is an indication that RIA and HMSA detect more patients with IFX than does PROMETHEUS ELISA, and this effect is more pronounced with anti-drug antibodies to IFX. We do not know the true measurements for these patients.

Summary

Figure 14 summarises the studies that quantify the link between the index tests and the comparator assays using concordance data. In comparing the three index tests with one another, we have data from two abstracts [including one unpublished abstract provided by Proteomika (Dr Daniel Nagore, Progenika Biopharma, 2015, personal communication)].[87]

For anti-drug antibodies, one abstract [unpublished abstract provided by Proteomika (Dr Daniel Nagore, Progenika Biopharma, 2015, personal communication)] describes complete agreement between Promonitor and LISA-TRACKER assays for anti-drug antibodies to IFX and a Cohen's kappa of 0.8 for anti-drug antibodies to ADA, but does not report how many samples were included. It also describes the upper limits of the measurement range for LISA-TRACKER assays as low. Daperno *et al.*[87] compared Immundiagnostik with Promonitor for anti-drug antibodies to IFX and found that the two tests showed identical results in only 6 out of 63 cases in a consecutive series of 66 patients (39 with CD and 27 with 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.

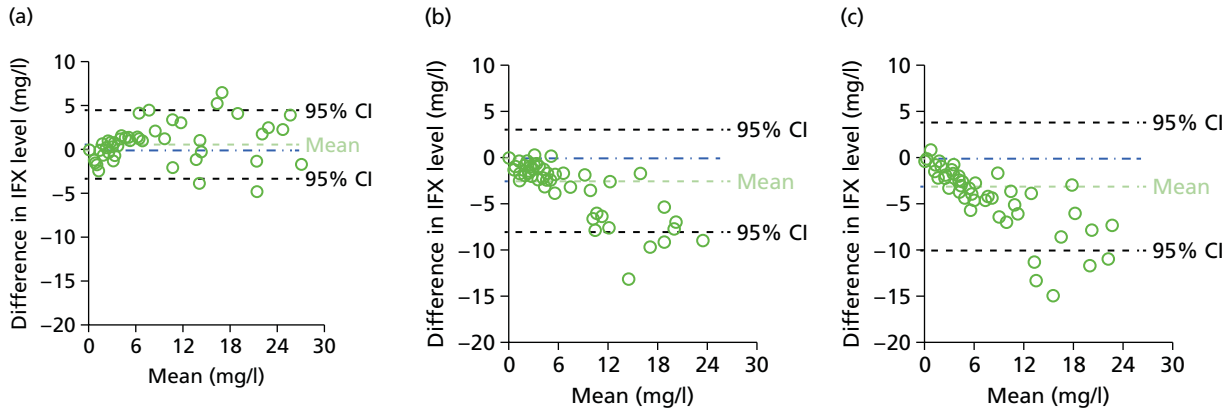


FIGURE 13 Reconstructed Bland–Altman plots comparing PROMETHEUS ELISA, HMSA and RIA. (a) ELISA vs. HMSA; (b) RIA vs. HMSA; and (c) RIA vs. ELISA. Based on data from Steenholdt *et al.*[122]

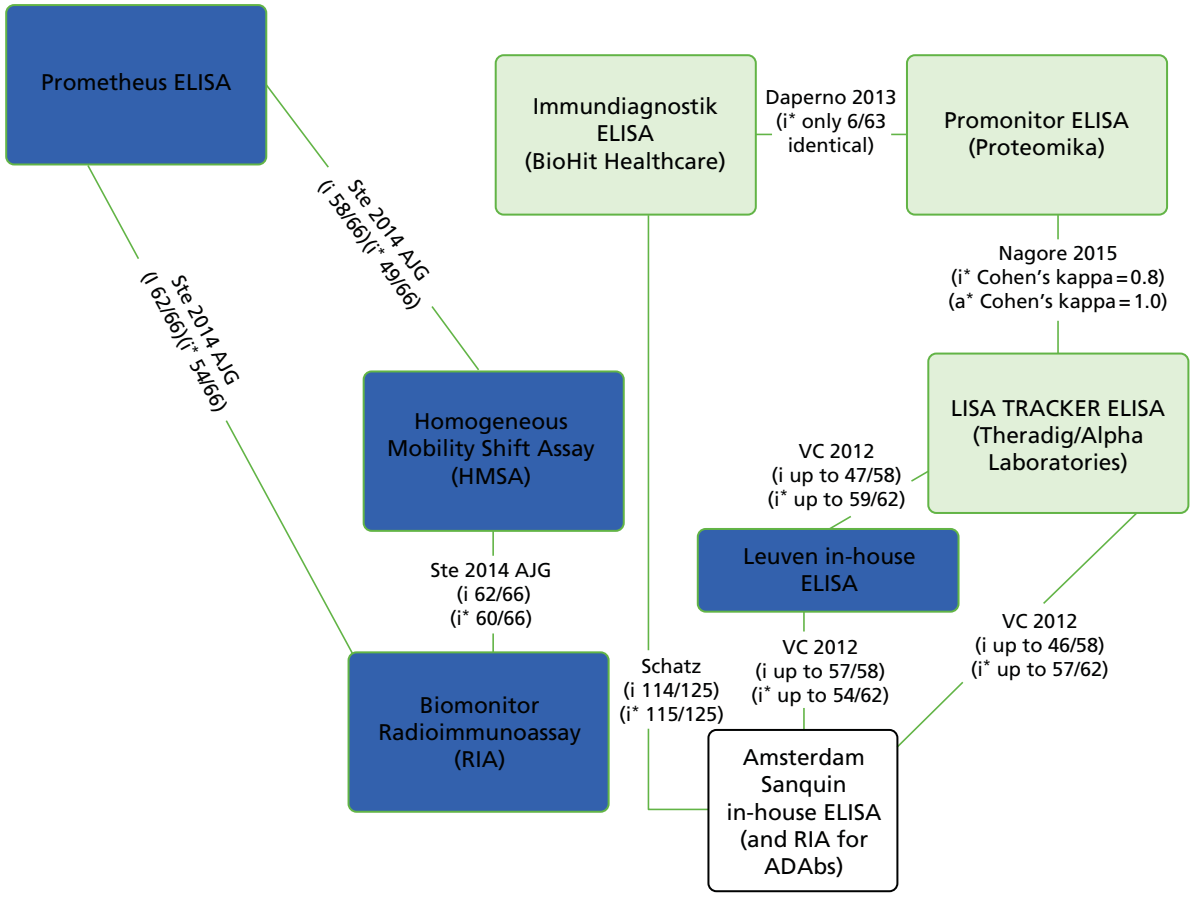


FIGURE 14 Results of comparisons which linked the index tests and comparator tests to each other of studies reporting concordance data. The index tests are shaded green and comparator tests are shaded blue. Results are listed as either Cohen's kappa or concordance levels displayed as a fraction for IFX (i), ADA (a), anti-drug antibodies to IFX (i*) and anti-drug antibodies to ADA (a*). Only studies which provide concordance at set threshold or Cohen's kappa for the comparisons between tests are included Daperno 2013;[87] Nagore 2015 (Dr Daniel Nagore, Progenika Biopharma, 2015, personal communication); Schatz 2013;[117] Steenholdt 2014;[122,123] and Vande Casteele 2012.[67] ADAbs, anti-drug antibodies; AJG, *American Journal of Gastroenterology*; Ste, Steenholdt; VC, Vande Casteele.

For each index test we investigated all links to the comparator tests. There was one link to the Amsterdam Sanquin tests (ELISA for IFX, RIA for anti-drug antibodies to IFX),[117] which showed agreement for 114 out of 125 samples for IFX and 115 out of 125 samples for anti-drug antibodies to IFX. However, there was only one study then linking the Amsterdam tests to the Leuven in-house ELISA,[67] which reported disagreement in at least 8 out of 62 samples, with a lack of clarity regarding the remainder of results. Similarly, there was only one study linking LISA-TRACKER assays to the Leuven in-house ELISA[67] and, although it reported disagreement for IFX in at least 11 out of 58 samples and for anti-drug antibodies to IFX in at least 3 out of 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 with each other, there was one study[122] that described reanalysing the same samples previously used in a test–treat trial.[123] There was agreement between PROMETHEUS ELISA and Biomonitor RIA for IFX in 62 out of 66 samples and for anti-drug antibodies to IFX in 54 out of 66 samples; agreement between PROMETHEUS ELISA and HMSA in 58 out of 66 samples for IFX and in 49 out of 66 samples for anti-drug antibodies to IFX; and, finally, agreement between RIA and HMSA in 62 out of 66 samples for IFX and in 60 out of 66 samples for anti-drug antibodies to IFX. 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).

Results of threshold analysis studies

The search identified 24 studies[48,58,77,82,84,85,93–95,100,101,104,105,107,108,110,112,113,115,119,120,126,132,135] that reported a ROC threshold analysis to determine optimal cut-off levels predictive of clinical response for IFX,[58,77,82,85,93,101,104,105,110,112,113,119,120,126] ADA[48,84,94,100,107,108,115,132,135] or both.[95] *Table 8* 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 > 0.5 indicates an informative test whereas 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 LOR are poorly defined and studies using different definitions will measure different outcomes, which will have implications on the reported thresholds.

Infliximab

Studies measuring IFX 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 point with maximum sensitivity and specificity, and maximising the 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),[95] whereas another favoured high specificity (1.00) at the expense of good sensitivity (0.33).[119] The reported IFX cut-off point ranged between 0.61 µg/ml[101] and 4.1 µg/ml,[105] and was reported to be as high as 7 µg/ml for the test with 100% specificity requiring a minimum drug trough level of 7 µg/ml at treatment week 14 to be predictive of good response at week 54.[119] One study reported a trough difference before and after dose optimisation as predictive of clinical outcome[58] and another reported the trough level that predicts response after reinitiation of IFX treatment.[77] Although the trough levels for the HMSA and RIA were not too dissimilar, one study reported trough levels for predicting anti-drug antibodies to IFX using HMSA that were exceptionally high (13 µg/ml).[126] There was great variation in the clinical marker used to assess clinical response, which included mainly subjective physician's assessment and disease activity scores as well as laboratory markers, such as CRP and FC, and objective assessments of mucosal healing. Six studies used a combination of different markers for the assessment.[77,82,95,101,104,105]

Adalimumab

Studies of ADA used mainly ELISAs, with only one study using HMSA[135] and one RIA.[94] One study reported only in the form of an abstract did not specify the test type used.[107] The reported thresholds for clinical markers such as response and clinical remission ranged from 3 µg/ml[95] to 6.85 µg/ml.[94] 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 ADA levels in one study of only ≥ 0.33 µg/ml.[48] One study reported the different threshold values for a test with maximum sensitivity of 14.5 µg/ml, maximum specificity of 0.35 µg/ml and sensitivity equal to specificity of 6.85 µg/ml.[94]

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 ADA concentration associated with clinical remission (CDAI score of < 150). This study therefore questioned the clinical utility of measuring ADA concentrations.[84]

Summary

The range of cut-off points illustrates that no validated threshold has been established to date. Cut-off points 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

TABLE 8 Cut-off points for drug levels from ROC analyses to predict clinical response

Study (first author and year of publication)	Cut-off point (µg/ml)	Performance measures				AUC (95% CI)	Clinical marker	Drug	Assay
		Sensitivity	Specificity	Positive predictive value	Negative predictive value				
Bortlik *et al.*, 2013[82]	3	0.70	0.62	0.41	0.84	0.70 (0.57 to 0.83)	Sustained response (no treatment failure or drug intolerance, no surgery, immunosuppressant introduction, steroids or IFX increase)	IFX	Commercial ELISA
Cornillie *et al.*, 2014[85]	3.5	0.64	0.78	0.56	0.83	0.75	Sustained response (CDAI score change)	IFX	Non-commercial ELISA
Goldberg *et al.*, 2014;[95] abstract	3	0.90	0.37	NR	NR	0.75	Disease activity (Physician's Global Assessment and CRP levels)	IFX	Commercial ELISA
Imaeda *et al.*, 2014[101]	0.6	0.73	0.62	NR	NR	0.67 (0.60 to 0.81)	CRP ≤ 0.3 mg/dl	IFX	Non-commercial ELISA
	1.0	0.67	0.71	NR	NR	0.72 (0.50 to 0.73)	Serum albumin (≥ 4.0 mg/dl)		
	1.1	0.72	0.56	NR	NR	0.63 (0.55 to 0.65)	FC (≤ 300 µg/g)		
	4.0	0.71	0.70	NR	NR	0.63 (0.56 to 0.70)	MH (Rutgeerts scoring system, 0 or 1)		
Marits *et al.*, 2014[105]	4.1	0.87	0.44	NR	NR	0.74 (SE 0.037)	Remission (HBI score < 5 and CRP level of < 3 mg/l)	IFX	Non-commercial ELISA
Nagore *et al.*, 2015[110]	0.8	0.86	0.75	NR	NR	0.86 (0.76 to 0.96)	Active disease	IFX	Commercial ELISA (Promonitor)
Pallagi-Kunstar *et al.*, 2014[112]	3.01	NR	NR	NR	NR	NR	Detecting anti-drug antibodies	IFX	Commercial ELISA
Paul *et al.*, 2012;[113] abstract	2	0.76	0.82	NR	NR	0.60	Remission (CDAI score of < 150)	IFX	Commercial ELISA (LISA-TRACKER)
Paul *et al.*, 2013[58]	0.5 (trough after optimisation minus trough before optimisation)	0.88	0.76	0.78	0.86	0.91 (0.83 to 1.0)	MH (FC < 250 µg/g)	IFX	Commercial ELISA (LISA-TRACKER Premium)

continued

TABLE 8 Cut-off points for drug levels from ROC analyses to predict clinical response (*continued*)

Study (first author and year of publication)	Cut-off point (µg/ml)	Performance measures				AUC (95% CI)	Clinical marker	Drug	Assay
		Sensitivity	Specificity	Positive predictive value	Negative predictive value				
Singh *et al.*, 2014[119]	4	0.53	0.75	0.76	0.52	0.64 (0.51 to 0.75)	Week 14 IFX levels as predictor of week 54 clinical remission according to CDAI	IFX	Commercial ELISA (up to July 2012) or HMSA (from July 2012)
	7	0.33	1.00	1.00	0.50	0.67 (0.58 to 0.75)			
Baert *et al.*, 2014[77]	2 (after re-exposure to IFX)	NR	NR	NR	NR	0.76 (0.62 to 0.90)	Long-term response [clinical assessment (HBI) and CRP levels (< 3 mg/l)]	IFX	HMSA
Levesque *et al.*, 2014[104]	3	NR	NR	NR	NR	NR	Disease activity at week 8 increase in CDAI score (≥ 70 points and a CRP level > 5 µg/l)	IFX	HMSA
Vande Casteele *et al.*, 2013[126]	13 (TL week 6)	0.72	0.81	NR	NR	0.87 (SE 0.06)	Anti-drug antibody formation	IFX	HMSA
Steenholdt *et al.*, 2011[120]	0.5	0.86	0.85	NR	NR	0.93 (0.85 to 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 *et al.*, 2012;[93] abstract	3	NR	NR	NR	NR	0.74	Disease activity	IFX	HPLC-based fluid phase assay
Chiu *et al.*, 2013[84]	No ADA concentration identified associated with clinical remission at any time point so clinical utility of measuring ADA concentrations was difficult to assess	NR	NR	NR	NR	Week 4, 0.51; week 24, 0.58; week 56, 0.57	Clinical remission (CDAI score of < 150)	ADA	Non-commercial ELISA

Study (first author and year of publication)	Cut-off point (µg/ml)	Performance measures				AUC (95% CI)	Clinical marker	Drug	Assay
		Sensitivity	Specificity	Positive predictive value	Negative predictive value				
Goldberg *et al.*, 2014;[95] abstract	3	0.83	0.63	NR	NR	0.8	Disease activity (Physician's Global Assessment and CRP levels)	ADA	Commercial ELISA
Imaeda *et al.*, 2014[100]	5.9	0.67	0.92	NR	NR	0.83 (0.80 to 0.95)	CRP level of ≤ 0.3 mg/dl	ADA	Non-commercial ELISA
Karmiris *et al.*, 2009[48]	0.33	0.95	NR	0.81	NR	NR	Sustained clinical benefit (patient reporting lasting control of disease with possible dose escalation)	ADA	Non-commercial ELISA
Mazor *et al.*, 2014[108]	5.85	0.68	0.71	NR	NR	0.75 (0.66 to 0.84)	Remission according to two physicians' assessments	ADA	Non-commercial ELISA
Roblin *et al.*, 2014[115]	4.85	0.81	0.67	0.84	0.57	0.73	Clinical remission (CDAI score of < 150)	ADA	Commercial ELISA (LISA-TRACKER Premium)
	4.9	0.66	0.85	0.88	0.51	0.77	MH (disappearance of all ulcerations on endoscopy)		
Ward *et al.*, 2013;[132] abstract	4.9	0.83	0.65	NR	NR	0.75	Remission	ADA	Commercial ELISA (LISA-TRACKER Premium)
Yarur *et al.*, 2013;[135] abstract	5	NR	NR	NR	NR	0.71	Elevation of CRP levels	ADA	HMSA
Frederiksen *et al.*, 2014[94]	14.5	1.00	0.12	0.41	1.00	0.77 (0.62 to 0.93)	LOR (Physician's Global Assessment)	ADA	RIA
	0.35	0.50	0.96	0.89	0.76				
	6.85	0.69	0.69	0.58	0.78				
Mazor *et al.*, 2013;[107] abstract	5	NR	NR	NR	NR	0.77 (0.67 to 0.86)	Clinical response and normal CRP	ADA	NR

HPLC, high-performance liquid chromatography; MH, mucosal healing; NR not reported; SE, standard error; TL, trough level.

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 LOR. An additional variable that impacts on the threshold of anti-TNF-α drug levels (which is insufficiently depicted in *Table 8* because of poor reporting in the studies) is the time of testing and the time of clinical assessment.

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

Aim

To provide a narrative description of algorithms used in studies that report clinical outcomes for patients whose treatment options were directed by a test-informed algorithm; and to compare these with related algorithms identified in the literature during scoping as relevant to the NHS.[66,73]

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 patients with CD.[44,56,57,61,64,137–139] Here we describe the test-based algorithms that have actually been implemented in studies with patients with CD and briefly compare these with the most similar 'precursor' algorithms identified during scoping as relevant to the NHS. None of the algorithms described in this section was used in conjunction with one of the index tests.

Algorithms from management studies

No management studies of CD or IBD patients treated with ADA were found. Three IFX studies fulfilled inclusion criteria for this objective: (1) a RCT in patients with CD with LOR to IFX;[123] (2) the TAXIT RCT[73] of IBD patients responding to IFX; and (3) a retrospective observational study of IBD patients responding to IFX.[128]

Table 9 summarises the algorithm used by Steenholdt *et al.*[123] 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. IFX negative and anti-drug antibody positive
2. IFX negative and anti-drug antibody negative
3. IFX positive and anti-drug antibody negative
4. IFX positive and anti-drug antibody positive.

An important feature is the proposal of different causes for secondary treatment failure in groups 1–3; 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 LOR who have therapeutic levels of IFX 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. As most patients fall into this group, these less prescriptive aspects add to uncertainty about treatment cost of the algorithm-based strategy and whether or not 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 IFX and anti-drug antibody) are reviewed with suspicion and require retesting in case of error.

The TAXIT algorithm for patients responding to IFX is based on the hypothesis that an IFX trough level between 3 and 7 µg/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.

TABLE 9 Summary of the concurrent testing-based algorithm used by Steenholdt *et al.*[123]

Cut-off point	Detectable anti-IFX antibodies	Undetectable anti-IFX antibodies
Subtherapeutic IFX < 0.5 µg/ml	Group 1 Insufficient IFX bioavailability because of the induced immunogenicity of IFX ↓ Change to different TNF-α-inhibitor: ADA 80 mg s.c. at inclusion followed by 40 mg s.c. every other week; dose intensification allowed	Group 2 Insufficient IFX bioavailability because of non-immune-mediated pharmacokinetics of IFX ↓ Intensify IFX treatment: IFX 5 mg/kg intravenously every 4 weeks
Therapeutic IFX ≥ 0.5 µg/ml	Group 4 Consider: • pharmacodynamics • non-functional anti-IFX antibodies • false-positive test ↓ Repeat IFX and anti-IFX antibody analyses and handle accordingly If unchanged results, then act as group 3	Group 3 Pharmacodynamics: inhibition of TNF-α is ineffective because of non-TNF-α-driven disease ↓ TNF-α-inhibitors not effective, so are 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

s.c., subcutaneously.

Figure 15 summarises the TAXIT study algorithm as described in the recently published paper.[73] Patients are categorised into four groups according to their trough IFX level: (1) undetectable, (2) low, (3) optimum (3–7 µg/ml) or (4) high (> 7 µg/ml). Group 1 is reflex tested for anti-IFX antibodies and further divided into two subgroups on the basis of anti-drug antibody test results: (1) in those with high anti-drug antibody levels, IFX therapy is stopped; (2) in those with anti-drug antibodies at a lower level (< 8 µg/ml), the dose of IFX is increased. In patients in group 2 (detectable but low trough levels of IFX; < 3 µg/ml) dosing interval is first reduced and then, if necessary, the dose is increased, in attempt to bring IFX trough concentrations within the 'optimum' range (3–7 µg/ml). In group 3, whose trough IFX levels are already in the optimum range, dose adjustment is not necessary. In patients in group 4 (high trough levels of IFX) dose interval is increased, followed, if required, by a 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.

Vaughn *et al.*[128] describe trough monitoring of IFX 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, whereas for those identified later the commercial HMSA method was used. Initially, dose adjustments aimed to bring trough IFX into the detectable range, but later the target range was changed to 5–10 µg/ml. The authors quote a typical dose adjustment for those with an undetectable trough IFX level to be an increase in dose to 7 mg/kg with a 6-week interval to the next infusion, followed by a return to 8-week infusion intervals. The authors state that for trough levels < 5 µg/ml, the dose of IFX was increased by '50 or 100 mg'. For patients with a trough level of IFX > 10 µg/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 typographical error) the infusion interval was increased. No dose adjustment was made for those in target range. This algorithm

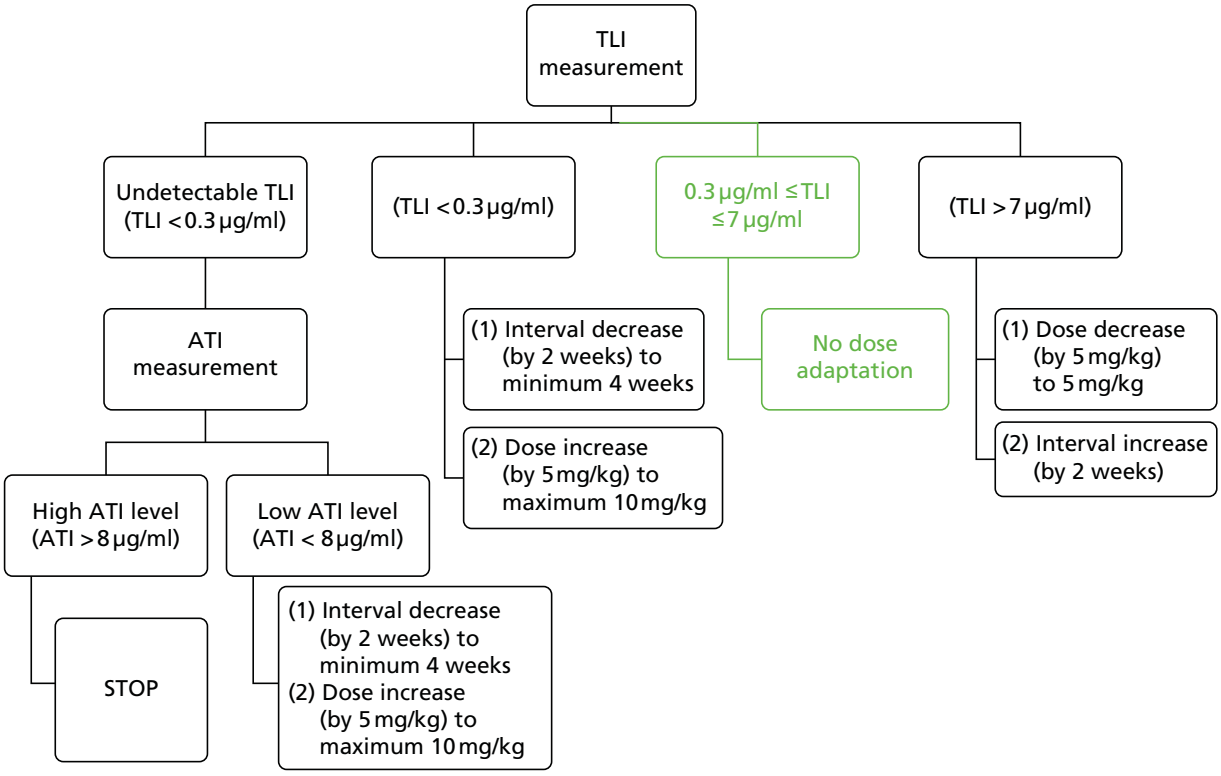


FIGURE 15 The TAXIT study algorithm presented in Vande Casteele *et al.*[73] Reprinted from *Gastroenterology*, Vol. 148, Vande Casteele N, Ferrante M, Van Assche G, Ballet V, Compernolle G, Van Steen K, *et al.*, Trough concentrations of IFX guide dosing for patients with inflammatory bowel disease, pp. 1320–9.e3, Copyright 2015 with permission from Elsevier. ATI, antibodies to IFX; TLI, trough level of IFX.

would be somewhat difficult to implement without assuming that the authors' preferred policy is to use HMSA testing with a target range of 5–10 µg/ml and to manipulate dosage and dose intervals at the clinicians' discretion so as to bring the patient into the target range.

Comparison of algorithms that 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 LOR for use with an unspecified anti-TNF-α (based on Scott and Lichtenstein,[66] who specified IFX) and one for responders to IFX based on the public domain description of the TAXIT trial. The algorithm proposed by Scott and Lichtenstein[66] requires concurrent testing (*Figure 16*).

This algorithm is similar to that of Steenholdt *et al.*[123] (see *Table 9*) in categorising patients with LOR 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 subtherapeutic as in Steenholdt *et al.*[123] The suggested treatments for 'anti-drug antibody positive l drug trough level low', 'anti-drug antibody negative l drug trough level therapeutic', and 'anti-drug antibody negative l drug trough level low' groups are the same as those proposed by Steenholdt *et al.*,[123] namely treatment with an alternative anti-TNF-α drug, treatment with a non-TNF-α drug with a different mechanism of action and escalation of drug exposure, respectively. For the 'anti-drug antibody positive l drug trough level therapeutic' group, Scott and Lichtenstein[66] make no recommendations, but Steenholdt *et al.*[123] 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.*[123] 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' and undetectable levels of anti-drug antibodies are termed

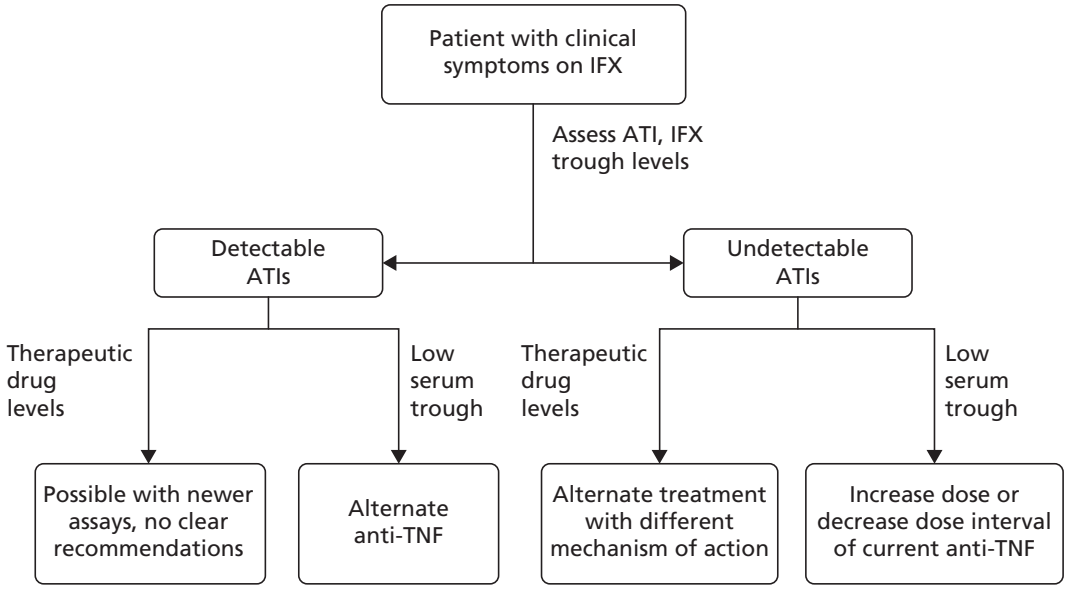


FIGURE 16 The Scott and Lichtenstein[66] algorithm for patients with LOR. ATI, antibodies to IFX. Reproduced from *Current Treatment Options in Gastroenterology*, Therapeutic drug monitoring of anti-TNF therapy in inflammatory bowel disease, vol. 12, 2014, pp. 59–75, Scott FI, Lichtenstein GR[66] (© Springer Science + Business Media, LLC 2014), with permission of Springer.

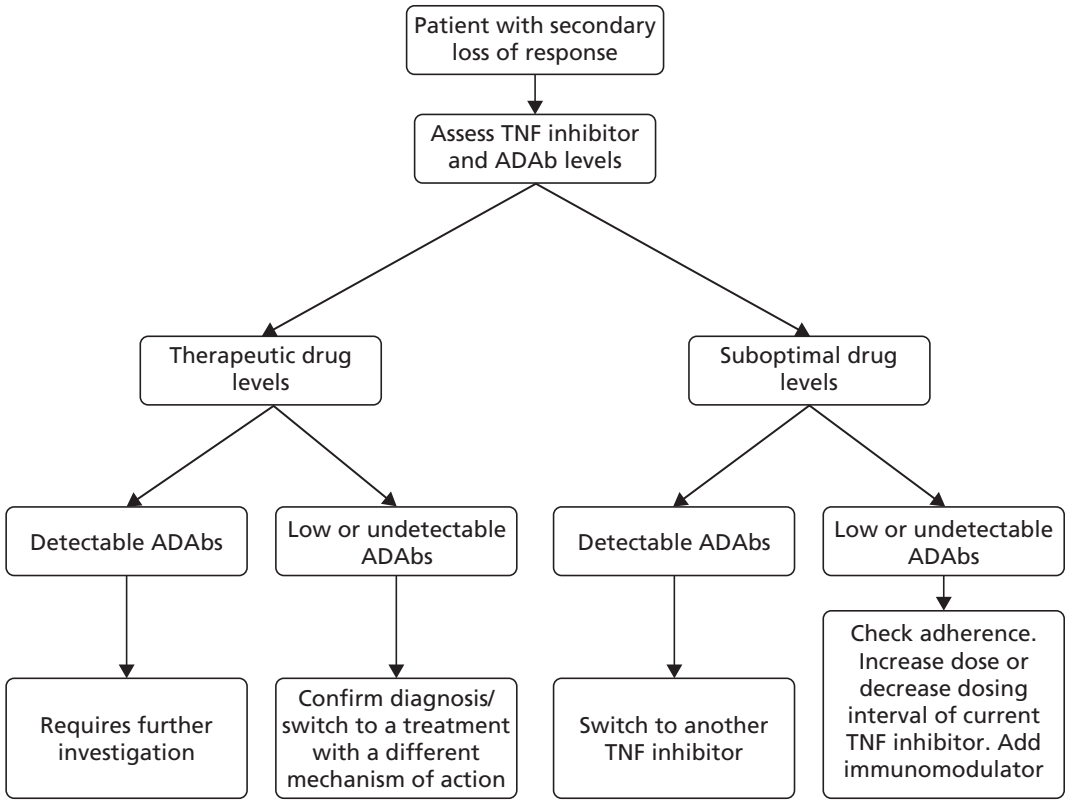


FIGURE 17 The precursor algorithm for LOR identified in scoping (based on Scott and Lichtenstein 2014[66]). ADAb, anti-drug antibody.

'low or undetectable'. In patients with 'suboptimal drug' and 'low or undetectable' anti-drug antibodies it is recommended that adherence to treatment should be checked and that an immunosuppressant may be added. This version of the algorithm suggests that, in the case of the patients with 'therapeutic drug' levels and 'low or undetectable anti-drug antibodies', the 'diagnosis' of CD is confirmed.

The scope precursor version of the TAXIT trial algorithm is almost identical to the public domain version shown in *Figure 18*. The differences are (a) no specific trough drug levels are specified; and (b) the addition of an immunosuppressant is recommended for the group with undetectable trough drug levels and low anti-drug antibody levels.

Summary

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

Vaughn *et al.*[128] recommended trough IFX testing for IBD patients to bring trough IFX 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 IFX responders hypothesised the therapeutic target range of 3–7 µg/ml based on analyses using the HMSA method;[140] the trial used an in-house ELISA. The algorithm prescribes dose adjustments for patients with trough IFX levels of < 3 µg/ml and > 7 µg/ml to bring the trough drug level to within the target range. Patients with trough IFX levels of < 0.3 µg/ml were reflex tested for anti-drug antibody and dichotomised as above or below 8 µg/ml; the algorithm recommended cessation of IFX for those with trough IFX levels > 8 µg/ml. The algorithm is sufficiently detailed to be replicable.

The Steenholdt *et al.*[123] algorithm for patients with LOR to IFX employs concurrent testing for IFX and anti-drug antibody and generates four categories of patient: (1) IFX negative and anti-drug antibody positive; (2) IFX negative and anti-drug antibody negative; (3) IFX positive and anti-drug antibody negative; and (4) IFX positive and anti-drug antibody positive. The trial used RIAs, and cut-off points for dichotomising test results were based on a previous study.[120] The algorithm specifies treatments for each category of patient that are based on hypothesised mechanisms underpinning the LOR. Treatment for patients with a positive test for drug and a negative test for anti-drug antibodies was not sufficiently prescriptive to be easily replicated and would probably vary between clinician(s).

The precursor/scoping algorithms represent minor differences of those proposed by Steenholdt *et al.*[123] and the TAXIT trial[73] investigators.

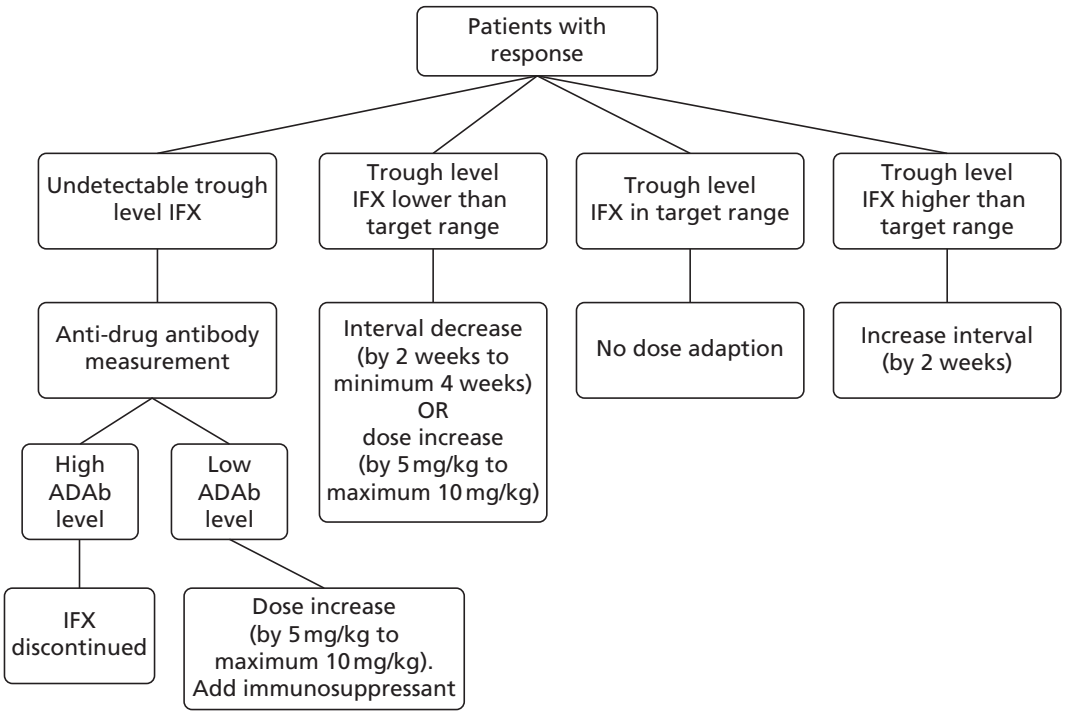


FIGURE 18 The precursor algorithm based on the TAXIT trial algorithm for IFX responders. ADAb, anti-drug antibody.

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 8*. This table demonstrates that cut-off levels are study specific and are not readily generalisable.

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

Aim

The aim of this section is to assess the evidence from studies that report the clinical impact of implementing a test-informed treatment algorithm for anti-TNF-α recipients with CD.

Results

After screening 2428 studies, we identified three which matched our inclusion criteria for management studies. All three investigated patients treated with IFX. There were no management studies for patients treated with ADA and none of the three studies used one of the index tests. The three studies that fulfilled our inclusion criteria (i.e. Steenholdt *et al.*,[123] Vaughn *et al.*[128] and Vande Casteele *et al.*[73]) address several aspects of the decision questions. Other studies were identified (e.g. Afif *et al.*,[56] Pariente *et al.*,[59] Roblin *et al.*[57] and Paul *et al.*[58]) that investigated the clinical utility of therapeutic drug monitoring of anti-TNF-α to predict response to a change in treatment mainly because of LOR to anti-TNF-α. 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-informed algorithm strategy for patient outcomes and compare that with a strategy that might be similar to standard care. For completeness, 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.

Assessment of the risk of bias in the management studies

The risk-of-bias assessment, using the Cochrane risk-of-bias tool,[70] for the two included RCTs[73,123] is summarised in *Table 10* and *Figure 19*. Steenholdt *et al.*[123] described a treatment algorithm in patients with LOR and, more recently, provided updated longer-term data (20 weeks' follow-up).[124] The quality assessment of the retrospective observational pilot study by Vaughn *et al.*[128] was assessed using the Downs and Black checklist.[71] Further details on the quality assessment of these three studies are provided in *Appendix 10* and *Table 10*.

TABLE 10 Risk of bias by study: summary of reviewers' judgements on each risk-of-bias item

Bias item	Study (first author and year of publication)	
	Steenholdt *et al.* 2014,[123] 2015[124]	Vande Casteele *et al.*, 2015[73]
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 of bias; ?, unclear risk of bias; ITT, intention to treat; PP, per protocol.

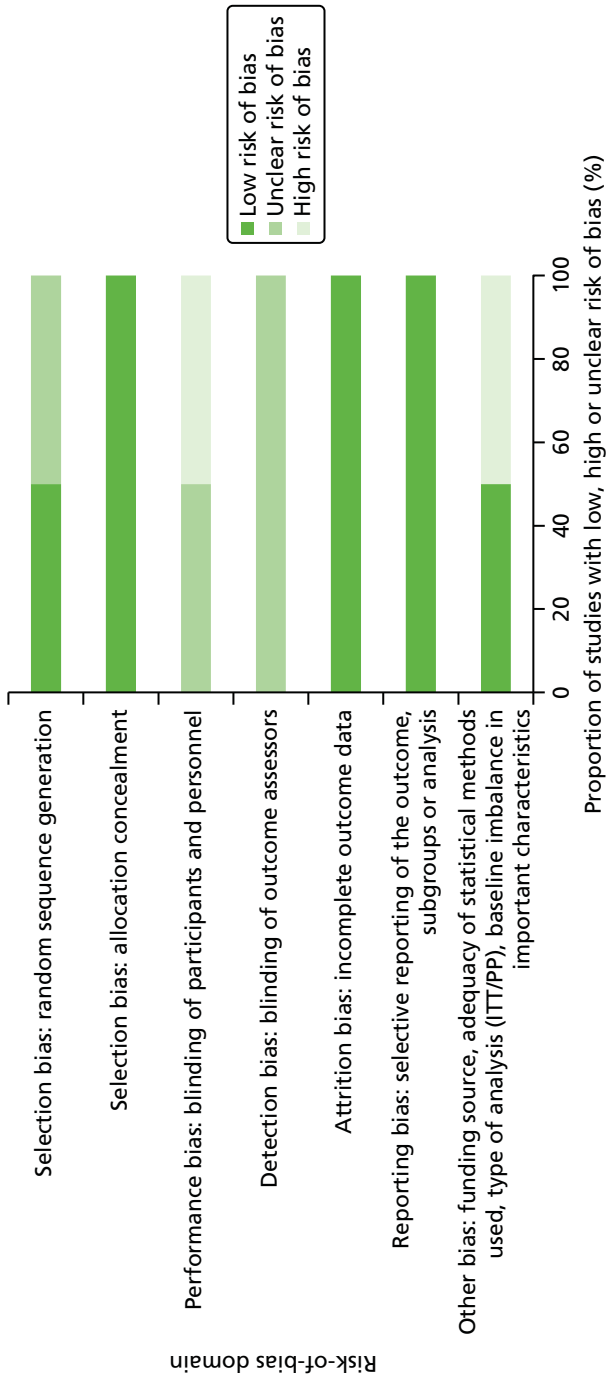


FIGURE 19 Risk-of-bias graph across two included RCTs: reviewers' judgements about each risk-of-bias item. ITT, intention to treat; PP, per protocol.

Randomised controlled trials

One RCT reported an adequate method for random sequence generation[73] and one[123,124] was judged at unclear risk of bias because 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.*[123,124] 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 IFX and IFX antibodies in the treatment of those patients who were randomised to the algorithm group. The TAXIT study[73] was considered at unclear risk of performance bias because there was insufficient information about blinding to IFX trough and antibodies to IFX 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[73] was considered adequate in terms of other potential bias (e.g. funding source, statistical methods used, analysis and baseline characteristics), there was concern about potential high risk of bias in Steenholdt *et al.*,[123,124] as 42% of patients were not treated in accordance with the algorithm, resulting in patients crossing over to the 'comparator-like' treatment. Overall, Steenholdt *et al.*[123,124] was rated at high risk of bias and Vande Casteele *et al.*[73] was rated as unclear risk of bias according to the *Cochrane Handbook for Systematic Reviews of Interventions* on summarising risk of bias.[141]

Non-randomised study

The retrospective observational pilot study by Vaughn *et al.*[128] 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 favourably as (1) there was no list of principal confounders in each group; (2) it was not possible to determine if those participants selected, invited or agreeing to take part in the study were representative of a target population; and (3) we were unable to determine if the staff, places and facilities where the patients were treated were 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 favourably than the other items, with concerns raised about whether or not 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.

Summary of the management studies

Overview of the included management studies

Two of the management studies included a substantial minority of UC patients along with CD patients. Steenholdt *et al.*[123] described management with a treatment algorithm in patients with CD with LOR and compared this with standard dose intensification treatment in a RCT design. Vaughn *et al.*[128] investigated the impact of proactive drug concentration monitoring of IFX in retrospectively identified cohorts of IBD patients who were in remission on IFX (responders); proactive drug monitoring was compared with standard dose intensification treatment for relapse. The TAXIT study[73] was a RCT comparing clinical management with management using dose adjustment based on trough drug levels in IBD patients previously brought to target trough level of IFX using a dose adjustment algorithm.

During the review process, Steenholdt *et al.* were contacted and provided further information and clarification (Casper Steenholdt, Herlev University Hospital, 2015, personal communication). In addition, the authors drew our attention to an extension study of the original RCT.[124]

The three studies were heterogeneous with regard to populations, treatment algorithms and methods of testing for IFX and antibodies to IFX. These difference precluded meaningful pooling of study outcomes. *Table 11* summarises the major features and findings from the three studies, and is accompanied by detailed descriptions of each of the studies.

TABLE 11 Summary of the main features of the management studies

Feature	Management study (first author and year of publication)		
	Steenholdt *et al.*, 2014[123]	Vande Casteele *et al.*, 2015[73]	Vaughn *et al.*, 2014[128]
Patient population	LOR to IFX	Responders to IFX	Responders to IFX
Type of IBD	CD	68% CD; 31% UC	71.4% CD; 27% UC
Study design	RCT	RCT	Retrospective pilot study
Setting	Six Danish centres	University hospital	Tertiary health-care centre
Follow-up	12 weeks; 20 weeks	1 year	≈4 years
Aim of the study	Assess cost-effectiveness of AL-based treatment vs. dose intensification (II) treatment	To compare clinical and biological remission in ClinBD vs. ConBD at 1 year after randomisation	Investigate the usefulness of proactive drug monitoring to bring TLI target vs. no-proactive drug monitoring
Comparisons	AL vs. II	ClinBD vs. ConBD	Proactive drug monitoring vs. no-proactive drug monitoring
Algorithm by drug and antibody levels	Yes, both	IFX mainly (antibodies for those with undetectable IFX)	IFX mainly (some reflex testing of antibodies)
Test used	RIA	Leuven in-house ELISA	ELISA and HMSA
Time of analysis	At IFX treatment failure	Before each infusion time	Unclear
Drug and antibodies: cut-off point/target	RIA: drug ≥ 0.5 µg/l; antibodies detectability	Target IFX: 0.3–0.7 µg/ml; antibodies to IFX (if IFX negative): > 8 µg/ml	Target IFX range: initially detectable, later 5–10 µg/ml
Limit of quantification	RIA: IFX 0.15 µg/ml; antibodies: 10 arbitrary units/ml	IFX 0.3 µg/ml; antibodies to IFX 1.0 µg/ml	Variable during study
Definition of clinical response	≥ 70-point CDAI score reduction from baseline (luminal disease); ≥ 50% reduction in active fistulas of from baseline (fistulising disease)	Symptom free or clear clinical improvement and decrease of disease activity but with clinical symptoms	Unclear/physicians' judgement
Definition of clinical remission	CDAI score of ≤ 150 and complete closure of all fistulas despite gentle pressure	HBI score ≤ 4 for CD and partial Mayo score of ≤ 2, with no individual subscore > 1 for UC. Biological remission = CRP ≤ 5 mg/l	Lack of symptoms attributable to underlying IBD (by treating physicians' documentation)
Definition of clinical progression	Withdrawal for lack of effect of treatment	Not given	Not given
Definition of relapse	N/A	Need of IFX dose escalation or addition of steroids, or switch to another anti-TNF-α (on physician's assessment)	Not given
Major findings			
Clinical response by subgroup	II/AL at 12 weeks	N/A	N/A
Group 1[a] ($n = 14$; AL $n = 5$, II $n = 9$)	ITT: 4 (44%)/2 (40%); PP: 4 (44%)/2 (40%)	N/A	N/A
Group 2[a] ($n = 3$; AL $n = 1$, II $n = 2$)	ITT: 1 (50%)/0 (0%); PP: 1 (50%)/0 (0%)	N/A	N/A

TABLE 11 Summary of the main features of the management studies (*continued*)

Feature	Management study (first author and year of publication)		
	Steenholdt *et al.*, 2014[123]	Vande Casteele *et al.*, 2015[73]	Vaughn *et al.*, 2014[128]
Group 3[a] (*n* = 48; AL *n* = 26, II *n* = 22)	ITT: 12 (55%)/16 (62%); PP: 12 (55%)/7 (54%)	N/A	N/A
Group 4[a] (*n* = 4; AL *n* = 1, II *n* = 3)	ITT: 2 (67%)/0 (0%); PP: 2 (67%)/0 (0%)	N/A	N/A
Clinical response 20 weeks (all)	ITT: 56%/76%; PP: 56%/74%	N/A	N/A
Clinical remission	II/AL at 20 weeks: ITT, 39%/55%; PP, 39%/58%	Clinical + biological (1 year): ConBD, 68.8%; ClinBD, 65.9%; *p* = 0.880 CD only: ConBD, 63%; ClinBD, 55%; *p* = 0.353	Not given
Probability of remaining on IFX	N/A	Not reported	Proactive drug monitoring vs. no-proactive drug monitoring: hazard ratio of 0.3 (95% CI 0.1 to 0.6; *p* = 0.0006). At 5 years in treatment: proactive drug monitoring 86% vs. 52% no-proactive drug monitoring

AL, algorithm; ClinBD, clinically based dosing; ConBD, concentration-based dosing; II, IFX intensification; ITT, intention to treat; N/A, not applicable; PP, per protocol; TLI, trough level of IFX.

a Group identities of patients with lost response to IFX were based on concurrent test results as follows: group 1, IFX negative, anti-drug antibodies positive; group 2, IFX negative, anti-drug antibodies negative; group 3, IFX positive, anti-drug antibodies negative; and group 4, IFX positive, anti-drug antibodies positive.

Steenholdt et al.[123]

Study design This was a single-blind RCT of 69 adults with CD who were previously responsive to maintenance therapy with 'regular' infusions of IFX (at 5 mg/kg) (i.e. patients on maintenance IFX with LOR). Participants were randomised to either an IFX-intensified arm (*n* = 36) or an algorithm arm (*n* = 33). In the former, the dose frequency of 5 mg/kg IFX was increased every 4 weeks. In the latter, participants received treatment in accordance with a defined algorithm based on serum concentrations of IFX and of antibodies to IFX. It is unclear if the randomisation method was the most appropriate because 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 IFX 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 IFX failure was reported and were analysed by RIA (samples were stored for retrospective analysis by alternative assays methods including ELISA and HMSA). The RIA cut-off points used were therapeutic IFX $\geq$ 0.5 µg/l and subtherapeutic IFX < 0.5 µg/l, and the cut-off point for anti-drug antibody was the limit of quantification (10 arbitrary units/ml). These samples were taken immediately before IFX 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 to receive a defined treatment, as shown in *Table 9*.

The authors suggests the following underlying mechanisms for LOR:

- group 1 – insufficient bioavailability of IFX because of immunogenicity of IFX
- group 2 – insufficient bioavailability of IFX because of non-immune-mediated pharmacokinetics
- group 3 – inhibition of TNF-α ineffective because of non-TNF-α-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 with 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 19–81 years) and the majority were female (61%). The mean duration of disease was slightly greater in the IFX-intensified arm than in the algorithm arm [10 years (range 1–35 years) vs. 7 years (range 1–27 years)]. Around one-quarter (26%) of the participants gave a history of smoking and 30% had undergone previous surgery, whereas around 20% of patients had received anti-TNF-α therapy previously. Mean treatment duration at anti-TNF-α failure was 657 days (range 97–3313 days). Mean CRP level was 9 mg/ml (range 2–22 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'.[123] The study objective of estimating 'disease control' in each arm was undertaken using the proportion of patients with the primary outcome of clinical response. Other secondary outcomes included (1) the proportion with remission, defined as an absolute CDAI score of < 150 and complete closure of all fistulas despite gentle pressure [at baseline the mean CDAI score was 296 (range 221–526) and 301 (range 230–487) in algorithm and control arms, respectively; three and four patients in each arm had fistulising disease]; and (2) the proportion with a CDAI score of 100 response (a reduction of CDAI score from baseline of ≥ 100). Mean decrease in CDAI and PDAI scores and mean increase in IBDQ scores were also reported, together with changes in laboratory measures (white blood cell count, haemoglobin level and albumin level).

Intention-to-treat and per-protocol populations and handling of withdrawals Outcome analyses were reported for intention-to-treat (ITT) and for per-protocol (PP) populations. All 36 patients in the dose intensification arm received allocated treatment; there were eight withdrawals for lack of effect or severe infusion reaction.

In the intervention arm, patients in group 1 received ADA therapy during the study in accordance with local guidelines at the participating centres. Some used ADA in the dosing registered by European Medicines Agency (80 mg at week 0 followed by 40 mg every other week), whereas 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 optimisation of ADA was allowed (Dr C Steenholdt, Herlev University Hospital, Denmark, 25 January 2015, personal communication).

In the algorithm arm, 14 out of 33 patients did not receive treatment allocated in accordance with the local algorithm, leaving a PP population of 19. Most of these 14 non-PP patients continued to receive IFX. Of these, IFX was continued in 12 patients (nine patients in group 3 and one in group 4). The IFX regimen administered was as follows (with all patients receiving 5 mg/kg): IFX q8 regimen (two infusions during the trial, i.e. weeks 0 and 8), $n = 5$; IFX q4 regimen (four infusions during the trial, i.e. weeks 0, 4, 8 and 12), $n = 2$; IFX q4 regimen but not throughout the entire trial (three infusions during the trial), $n = 1$; IFX q4 regimen but not throughout the entire trial (two infusions during the trial), $n = 2$; and IFX q4 regimen but not throughout the entire trial (one infusion during the trial), $n = 2$.

The remaining two patients were switched to ADA because of misinterpretation of test results. Both patients were in group 3. The applied ADA regimen was ADA induction (160 mg–80 mg–40 mg) followed

by 40 mg every other week for one patient and ADA induction (80 mg–40 mg) followed by 40 mg every other week for the other patient (Dr Casper Steenholdt, Herlev University Hospital, Denmark, 25 January 2015, personal communication).

There were two withdrawals in the algorithm arm and eight in the dose-intensified arms. 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 because it is unclear if the eight patients who withdrew from dose intensification contributed to medical costs carried forward (but for treatment which they did not receive) or if post-withdrawal drug costs were zero.

Test results according to intention-to-treat and per-protocol populations In the algorithm arm, concurrent testing categorised the 33 patients to the four groups as follows: 26 to group 3, five to group 1 and one to each of 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 12*.

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

These test results imply that LOR 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 TNF-α is ineffective due to non-TNF-α-driven disease. TNFα inhibitors not

TABLE 12 Proportion of patients according to concurrent testing (ITT population)

Grouping in algorithm	Arm		
	Algorithm ($N = 33$), *n* (%)	IFX intensified ($N = 36$), *n* (%)	All ($N = 69$), *n* (%)
Group 1: subtherapeutic IFX and anti-drug antibody positive	5 (15)	9 (25)	14 (20)
Group 2: subtherapeutic IFX and anti-drug antibody undetectable	1 (3)	2 (6)	3 (4)
Group 3: therapeutic IFX and anti-drug antibody undetectable	26 (79)	22 (61)	48 (70)
Group 4: therapeutic IFX and anti-drug antibody positive	1 (3)	3 (8)	4 (6)

TABLE 13 Proportion of patients in each algorithm group (PP population)

Grouping in algorithm	Arm		
	Algorithm ($N = 19$), *n* (%)	IFX intensified ($N = 36$), *n* (%)	All ($N = 55$), *n* (%)
Group 1: subtherapeutic IFX and anti-drug antibody positive	5 (26)	9 (25)	14 (26)
Group 2: subtherapeutic IFX and anti-drug antibody undetectable	1 (5)	2 (6)	3 (5)
Group 3: therapeutic IFX and anti-drug antibody undetectable	13 (68)	22 (61)	35 (64)
Group 4: therapeutic IFX and anti-drug antibody positive	0 (0)	3 (8)	3 (5)

effective and is discontinued'.[123] 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 in the dose intensification arm (19/36; 53%) and in the algorithm arm (9/19; 47%) (RR 0.898, 95% CI 0.510 to 1.580; $p = 0.781$).

Table 14 summarises the rates of clinical response in ITT and PP populations according to test-defined subgroups. Group 3 (i.e. therapeutic IFX levels with undetectable anti-IFX antibodies) contributed the majority of the patients (ITT, 66.7%; PP, 63.6%) and also most of the clinical responses (ITT, 75.6%; PP, 67.9%), and thereby greatly influences the overall comparison between arms.

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-TNF-α mechanism; however, explanations other than intensified IFX 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.

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

For the ITT population, the coprimary outcome measure of mean cost was lower in the algorithm arm (€6038, SD €4146) than in the dose intensification arm (€9178, SD €2058) (mean difference –€3141,

TABLE 14 Clinical response according to test-defined subgroups

Subgroup	Population	Arm, *n*/*N* (%)		Algorithm vs. IFX-intensified arm, RR (95% CI); *p*-value
		IFX intensified	Algorithm	
Group 1: subtherapeutic IFX and anti-drug antibody positive	ITT	4/9 (44)	2/5 (40)	0.900 (0.246 to 3.297); 1.00
Insufficient IFX bioavailability because of the induced immunogenicity of IFX	PP	4/9 (44)	2/5 (40)	0.900 (0.246 to 3.297); 1.00
Group 2: subtherapeutic IFX and anti-drug antibody undetectable	ITT	1/2 (50)	0/1 (0)	NC
Insufficient IFX bioavailability because of non-immune-mediated pharmacokinetics	PP	1/2 (50)	0/1 (0)	NC
Group 3: therapeutic IFX and anti-drug antibody undetectable	ITT	12/22 (55)	16/26 (62)	1.128 (0.693 to 1.837); 0.770
Inhibition of TNF-α ineffective because of non-TNF-α-driven disease	PP	12/22 (55)	7/13 (54)	0.987 (0.525 to 1.856); 1.00
Group 4: therapeutic IFX and anti-drug antibody positive	ITT	2/3 (67)	0/1 (0)	NC
Pharmacodynamics or non-functional anti-IFX antibodies or false-positive test	PP	2/3 (67)	0/0 (0)	NC
All four subgroups	ITT	19/36 (53)	18/33 (58)	1.09 (0.713 to 1.673); 0.810
	PP	19/36 (53)	9/19 (47)	0.898 (0.510 to 1.580); 0.781

NC, not calculated.

95% CI –€4617 to –€1373; $p < 0.001$). For the PP population, mean costs were €4062 (SD €2763) in the algorithm arm compared with €9178 (SD €2058) in the dose intensification arm (mean difference –€5116, 95% CI –€6482 to –€3561; $p < 0.001$). *Table 15* summarises the mean cost in ITT and PP populations according to test-defined subgroups.

As for response, the coprimary outcome of mean cost was predominantly contributed by 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 a mean cost per patient of €7726.

Remission According to ITT and PP analysis, clinical remission was achieved by more patients in the IFX-intensified arm 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 vs. dose intensification)].

Extension study Clinical outcome findings to 20 weeks and mean cost to 52 weeks have been published.[124] Of 69 patients included in the trial, 45 (17 in the algorithm arm; 28 in the IFX-intensified arm) completed 12 weeks of PP treatment and 29 patients (16 in the algorithm arm; 13 in the IFX-intensified arm) completed 20 weeks of PP 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 16* summarises the results.

TABLE 15 Mean cost according to test-defined subgroups

Subgroup	Population	Arm (€), mean (SD)		Algorithm vs. IFX intensification (€), mean difference (95% CI); *p*-value
		IFX intensified	Algorithm	
Group 1: subtherapeutic IFX and anti-drug antibody positive	ITT	8299 (1796)	6837 (990)	–1462 (–2819 to 712); 0.090
Insufficient IFX bioavailability because of the induced immunogenicity of IFX	PP	8299 (1796)	6837 (990)	–1462 (–2819 to 712); 0.090
Group 2: subtherapeutic IFX and anti-drug antibody undetectable	ITT	8666 (1111)	9814 (N/A)	1148 (N/A); N/A
Insufficient IFX bioavailability because of non-immune-mediated pharmacokinetics	PP	8666 (1111)	9814 (N/A)	1148 (N/A); N/A
Group 3: therapeutic IFX and anti-drug antibody undetectable	ITT	9898 (1901)	5728 (4606)	–4169 (–5968 to –1788); 0.001
Inhibition of TNF-α ineffective because of non-TNF-α-driven disease	PP	9898 (1901)	2552 (1639)	–7349 (–8557 to –6032); < 0.001
Group 4: therapeutic IFX and anti-drug antibody positive	ITT	6883 (2309)	6003 (N/A)	–880 (N/A); N/A
Pharmacodynamics or non-functional anti-IFX antibodies or false-positive test	PP	6883 (2309)	N/A (N/A)	N/A; N/A
All four subgroups	ITT	9178 (2058)	6038 (4146)	–3141 (–4617 to –1373); < 0.001
	PP	9178 (2058)	4062 (2763)	–5116 (–6482 to –3561); < 0.001

N/A, not applicable.

TABLE 16 Clinical response and remission at 20 weeks (Steenholdt *et al*. 2015[124])

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

None of the differences between dose-intensified and algorithm groups reached statistical significance. According to ITT analyses in this and the original study, of the 33 patients in the algorithm arm, 18 patients showed response at 12 weeks, and this number increased to 25 by week 20. Of the 36 patients in the dose-intensified arm, 19 patients showed response at week 12, increasing to 20 at week 20. Among dose-intensified patients, 14 out of 36 were in remission at both 12 and 20 weeks, whereas in the algorithm arm the proportion increased from 10 out of 33 at week 12 to 18 out of 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 2012 US dollars (US$) rather than euros, as in the earlier report. This made it problematic to compare cost over the first 12 weeks with those subsequently accumulated to week 20 or 52. According to ITT analysis, mean costs related to CD at 20 weeks were US$11,940 and US$17,236 in the algorithm and dose-intensified arms, respectively (mean difference –US$5296, 95% CI – US$8453 to – US$1566; $p = 0.005$). At 52 weeks (ITT analysis), the corresponding values were US$22,066 and US$29,072 (mean difference – US$7006, 95% CI – US$12,848 to – US$874; $p = 0.022$).

Summary and conclusions One published study described the implementation of an algorithm in patients with CD with LOR. The authors concluded that the treatment of LOR to IFX using an algorithm based on concurrent IFX plus anti-drug antibody measurements significantly reduces average treatment costs per patient compared with routine IFX 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 IFX-intensified arm; follow-up was short; and a large proportion of patients in the algorithm arm did not receive the algorithm-recommended treatment (42%), raising the question of whether or not the efficacy of the algorithm has, in fact, been tested.

In addition, little information was provided on the components contributing to the coprimary outcome of mean cost. Test costs were 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, unfortunately, was the one in which treatments were least well described and largely depended on clinicians' judgement and reflection, which is unlikely to be replicable between clinicians (note that further details of treatments were provided in the extension study[124]).

Vande Casteele et al.*: Trough level Adapted infliXImab Treatment study*[73]

Study design This RCT[73] included 251 patients with IBD (173 with CD; 78 with UC) with stable response to IFX therapy who were randomised (1 : 1) to two different treatment strategies: (1) clinically based dosing ($n = 123$); or (2) IFX trough concentration-based dosing ($n = 128$) that targeted an IFX trough level of 3–7 µg/ml. Prior to randomisation, a consecutive cohort of 275 IBD patients (186 with CD; 89 with UC) were screened and subjected to an optimisation phase using an algorithm for dose adjustment to identify patients whose trough IFX levels could be successfully brought to the target range. All randomised patients entered with trough IFX levels within the target range. For patients randomised to the clinically based dosing arm, subsequent IFX dosing was in accordance with clinical symptoms and CRP levels (recorded at each infusion) and followed standard clinical criteria. For those randomised to the trough concentration-based dosing arm, IFX dosing continued in accordance with the algorithm. Patients were followed for 52 weeks post randomisation.

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

Specified primary and secondary outcomes The primary outcome was the rate of clinical plus biological remission 1 year after randomisation (clinical remission required a HBI score of ≤ 4 for CD and partial Mayo score of ≤ 2 for UC; biological remission required a CRP concentration of ≤ 5 mg/l). Early terminations were considered failures for the primary end point [criteria for termination included safety and failure of IFX therapy defined as persisting clinical symptoms (HBI score of > 4 or partial Mayo score of > 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 IFX levels in target range, anti-drug antibody positivity, European Quality of Life-5 Dimensions (EQ-5D) QoL score 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 IFX.[73]

Optimisation phase During optimisation patients were first categorised into one of four categories on the basis of trough IFX levels: (1) trough IFX levels > 7 µg/ml; (2) trough IFX levels in the target range (3–7 µg/ml); (3) trough IFX levels < 3 µg/ml; or (4) trough IFX levels undetectable and anti-drug antibodies < 8 µg/ml (patients with undetectable trough IFX levels but anti-drug antibodies > 8 µg/ml were excluded at screening). In each category, dose was adjusted in accordance with the algorithm shown in *Figure 15*.

Of 72 patients with trough IFX levels < 3 µg/ml (categories 3 and 4), dose escalation brought 69 to a target trough level of IFX. In total, 115 category 2 patients were in range and were randomised, and in 67 out of 72 category 1 patients dose was de-escalated to achieve the target range. A total of 251 patients were randomised: 128 to trough IFX level-monitored dosing and 123 to clinically based dosing.

Infliximab and antibody measurement The trough IFX and anti-drug antibody levels were measured using an in-house developed ELISA (Leuven in-house ELISA). The trough IFX level was measured using direct ELISA and anti-drug antibody levels using bridging ELISA. The lowest quantification value for IFX and anti-drug antibody limit for the test was 0.3 and 1.0 µg/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 ($n = 263$): the mean age of patients was 41.0 years (range 30–48.5 years) and, 77.2% of patients were in remission, mean CRP level was 1.7 mg/l and the mean IFX trough level was 4.6 µg/ml (2.5–7.7 µg/ml). Around 5% of patients were receiving immunosuppressant.

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 (range 6.3–19.9 years), median duration of disease at first IFX exposure was 5.8 years (range 1.7–13.5 years), median time since first IFX was 4.6 years (range 2.1–7.5 years), 82.5% of patients were in remission (CD, 79.8%; UC, 88.5%) and mean CRP and mean IFX trough concentration were 1.4 mg/l (range 0.6–4.2 mg/l) and 4.9 µg/ml (range 3.9–8.5 µg/ml), respectively.

Results: optimisation phase The results for patients with CD are summarised in *Table 17*.

Of 178 patients with CD entering optimisation, 131 were in clinical remission (HBI ≤ 4). After optimisation, 138 out of 173 were in remission (ITT: RR 1.053, 95% CI 0.936 to 1.186).

Of 44 patients with CD in the dose escalation group entering optimisation, 43 achieved target trough IFX levels. Of these, 28 were in clinical remission at entry, rising to 38 in clinical remission after optimisation [reported PP, odds ratio (OR) 4.1, 95% CI 1.3 to 12.5; RR 1.297, 95% CI 1.008 to 1.669; $p = 0.020$]. These patients also showed a significant decrease in mean CRP concentration at the end of optimisation (from 4.3 mg/l to 3.2 mg/l; $p < 0.001$). The corresponding results for UC patients did not reach statistical significance ($p = 1.0$ and $p = 0.16$, respectively). Among patients with CD who underwent dose reduction during optimisation (PP, $n = 51$), the proportion in remission decreased from 80.4% to 69.4% (PP: 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 IFX levels ($p = 0.3$ and $p = 1.0$, respectively, for clinical remission and $p = 0.56$ and $p = 0.86$, respectively, for CRP levels).

In the dose escalation group, an average of 2.1 optimisations were required to reach target trough IFX levels, and at the end of optimisation the median infusion interval was 6 weeks (range 4–8 weeks). In the dose reduction group, a mean of 1.4 optimisations was required and the median infusion interval was 8 weeks (range 6–12 weeks).

Results: maintenance phase 52 week primary outcome Almost 90% of patients completed the maintenance phase. The reasons for not completing were, in the clinically based and concentration-based dosing arms respectively, discontinuation because of active disease (4 and 4), serious adverse event (1 and 1),

TABLE 17 Remission rates for patients with CD; comparison of after optimisation vs. before optimisation

	Clinical remission, *n*/*N* (%)		
Patient group	After optimisation	Before optimisation	Statistic (after vs. before) (95% CI)
All patients with CD[a]	138/173 (79.8)	131/178 (73.6)	RR 1.053 (0.936 to 1.186)
Patients with CD dose escalated[b]	38/43 (88.4)	28/43 (65.1)	OR 4.071 (1.324 to 12.524); RR 1.297 (1.008 to 1.669)
Patients with CD dose reduced[b]	35/51 (69.4)	41/51 (80.4)	OR 0.534 (0.215 to 1.325); RR 0.854 (0.678 to 1.074)

OR, odds ratio.
a ITT analysis.
b PP analysis, numbers of patients estimated from reported percentages.

lost to follow-up (2 and 1), pregnancy (3 and 1), inability to maintain the target trough level (0 and 1) and other reasons (2 and 1).

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

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

Results: maintenance phase secondary outcomes There was little difference between groups in the probability of maintaining durable remission (26% and 27% in the concentration-based dosing arm and the clinically based dosing arm, respectively; $p = 0.88$).

A higher proportion of patients in the concentration-based dosing arm than in the clinically based dosing arm (74% vs. 57%) had IFX trough concentrations between 3 and 7 µg/ml ($p < 0.001$), whereas the risk of patients in the clinically based arm having undetectable trough levels of IFX 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 was 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 because of acute appendicitis and another because of ileostomy complications. There were 12 discontinuations in the clinically based dosing arm and 13 in the concentration-based dosing arm.

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 2.4, 95% CI 1.2 to 5.1; $p = 0.018$). Relapse was defined as 'the need for IFX 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'.[73] Among those relapsing and requiring rescue therapy, relatively more (9/21, 43%) in the clinically based arm than in the concentration-based arm (2/9, 22%) had trough IFX levels of < 3 µg/ml.

Relapse free-survival time was longer in the concentration-based dosing arm than in the clinically based dosing arm (*Figure 20*).

Authors' conclusions The authors concluded that optimisation of IFX dose to achieve the target trough levels of 3–7 µg/ml is more efficient and is cost-effective relative to clinically based adjustment. Therefore, the authors recommended using dose-to-target optimisation of IFX to achieve the target trough IFX 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.

Vaughn et al.[128]

Study design and conduct The aim was to investigate the usefulness of proactive therapeutic concentration monitoring and titration of IFX to a target concentration.[128] This was a retrospective observational pilot study of patients with IBD in clinical remission receiving IFX 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 groups, clinical remission was defined as 'lack of symptoms attributable to underlying IBD based on the treating gastroenterologist's documentation'.[128]

The IFX and antibodies to IFX concentrations were measured initially using solid-phase ELISA (PROMETHEUS Laboratories) and later with the HMSA (PROMETHEUS laboratories). The latter test could detect IFX

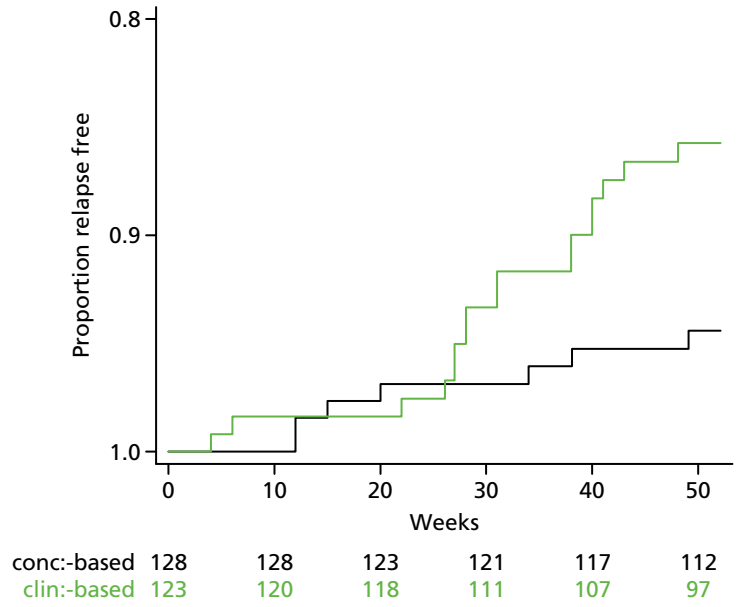


FIGURE 20 Kaplan–Meier analysis of time to relapse during maintenance phase. IPD reconstructed using the method of Guyot *et al.*[74] Conc:-based, concentration-based strategy; clin:-based, clinically based strategy.

concentrations as low as 1 µg/ml, compared with 1.4 µg/ml for ELISA. In the proactive drug monitoring group, serum trough IFX levels were used to guide dose modifications to achieve target drug levels. Initially the target was detectable IFX; later the target was changed to an IFX concentration between 5 and 10 µg/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 IFX infusion was increased to 7.5 mg/kg. The next infusion was given after 6 weeks, and after which IFX was given every 8 weeks.
- For patients with detectable trough drug levels < 5 µg/ml, the dose of IFX was increased by 50 or 100 mg.
- In patients with trough drug levels of > 10 µg/ml on at least two occasions, the dose was reduced. However, in those patients who were already receiving 5 mg/kg IFX, instead of dose modification, the treatment interval was increased.
- In patients who had trough levels in the range 5–10 µg/ml, no changes were made.

(Trough concentration was defined as 'IFX concentration measured at any time up to 7 days before the next infusion'.[128])

Reactive testing was done in both groups whenever there was LOR or if there was a concern that the patient was experiencing side effects as a result of antibody formation. For patients in the control group with LOR, the dose of IFX was increased at the treating physician's discretion in accordance with a standard of care guideline (typically to 10 mg/kg, but the dose did not reach > 10 mg/kg every 4 weeks).

Patient populations There were 48 IBD patients in the proactive drug monitoring group and 78 in the control group. 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 performed only reactively in the control group, but both reactively and proactively in the proactive drug monitoring group to determine any dose changes judged necessary to reach target trough concentration; furthermore, when the dose was escalated in the control group, IFX exposure was probably doubled (e.g. to 10 mg/kg), but dose escalations in the proactive drug monitoring group were of much smaller magnitude (e.g. by 50–100 mg; for a 70-kg 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) occurred only in the proactive drug monitoring group.

Two patients were 'IBD unclassified', 90 (69%) were diagnosed as having CD and 34 (29%) were diagnosed as having UC. Almost 70% were male. Median age at IFX initiation was 34.9 years in the proactive drug monitoring arm and 35 years in the control arm [interquartile range (IQR) 26.2–49.7 years]. The median age at diagnosis was 23.5 and 25 years in the proactive drug monitoring and control arms, respectively; 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; and 52 patients (41%) received combination therapy (44% and 40% of the proactive drug monitoring and control group, respectively). The median duration of IFX before proactive drug monitoring was 43 weeks (IQR 32–72 weeks).

The main reported outcomes comparing the proactive drug monitoring and control groups were the time remaining on IFX (Kaplan–Meier analysis) and the reasons for stopping IFX. 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 IFX 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 21* shows the reconstructed Kaplan–Meier comparison between groups. Beyond 5 years there are very few patients at risk; the median duration of IFX treatment before proactive drug monitoring implementation was reported to be 43 weeks (IQR 32–72 weeks). In multiple Cox regression analysis, the probability of patients remaining on IFX therapy was found to be significantly related only to proactive drug monitoring of IFX.

The authors also reported on the subgroup of patients that started maintenance IFX after 1 January 2009. As the implementation of proactive drug monitoring was reported to be from 2009, this subgroup would appear to be the more relevant population. *Figure 22* shows the reconstructed Kaplan–Meier comparison between the proactive drug monitoring and control subgroups. The reported hazard ratio was 0.3 (95% CI 0.1 to 0.7; $p = 0.003$). The reconstructed hazard ratio was similar, at 0.24 (95% CI 0.12 to 0.51); it is possible that the authors stratified their analysis by baseline variables (e.g. monotherapy or combination therapy, previous surgery, etc.). Parametric modelling based on reconstructed IPD is provided in *Appendix 11*.

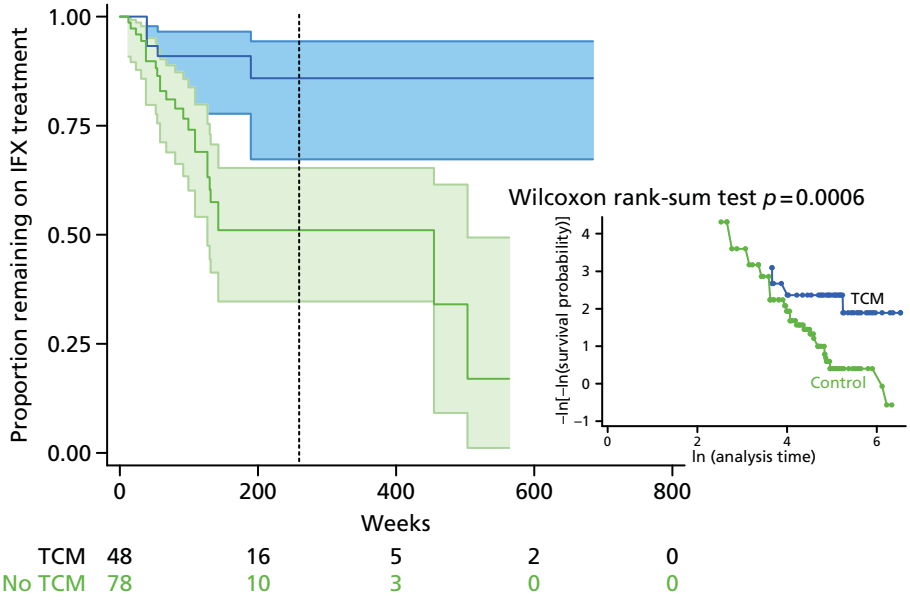


FIGURE 21 Kaplan–Meier analysis of time to stopping IFX treatment. IPD reconstructed using the method of Guyot *et al.*[74] The vertical line represents 5 years. Inset shows proportional hazards test plot. TCM, trough concentration monitoring.

Outcomes: reasons for stopping infliximab The reasons for stopping IFX are summarised in *Table 18*; 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.

Outcomes: proactive drug monitoring group trough levels and dose changes For the proactive drug monitoring group, the authors reported data concerning proactive tests undertaken; all reactive tests were omitted. Of initial proactive tests ($n = 48$), 37 used ELISA and 11 used HMSA methodology. Of subsequent proactive tests ($n = 40$), seven used ELISA and 33 used HMSA.

Median trough IFX at initial testing was 5 µg/ml (IQR 2.8–9.9 µg/ml), whereas the median subsequent trough level was 7.6 µg/ml (IQR 4.3–12.3 µg/ml).

Dosing adjustment after the initial proactive test was implemented in 17 patients (35%) as follows: dose escalation in 12 patients (71%), dose reduction in three patients (18%) and termination of therapy in two patients (12%). Dosing adjustment in subsequent proactive tests involved 10 patients (25%); these adjustments were described as eight patients (80%) who received dose escalation and two (20%) a dose decrease.

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

A trough level of IFX ≥ 5 µg/ml (lower end of the later target range) was reportedly achieved in 75% of patients (36/48). Of those in who reached this level of IFX, none developed antibodies to IFX or an immune reaction. In one patient, IFX was stopped after colectomy, which was undertaken for flat low-grade dysplasia.

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

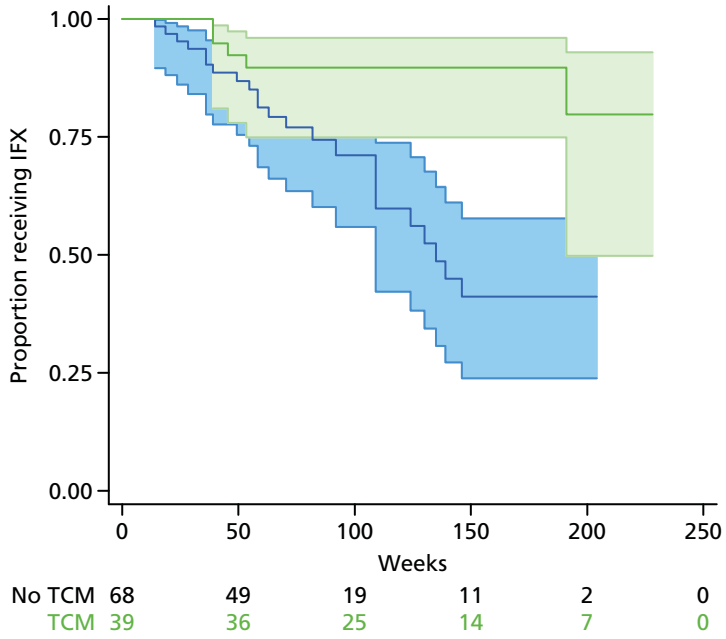


FIGURE 22 Kaplan–Meier analysis of time to stopping IFX treatment, patients starting maintenance January 2009. IPD reconstructed using the method of Guyot *et al.*[74] TCM, trough concentration monitoring.

TABLE 18 Reasons for stopping IFX therapy

Reason for stopping	Group (*n*)	
	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 IFX[a]	1	2

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

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 patients with CD only, and the other two examined both CD and UC patients. Two were RCTs and the other was a retrospective observational study. None employed designated intervention tests (LISA-TRACKER assays, TNF-α-Blocker or Promonitor ELISA).

The Steenholdt *et al.* RCT[123] used concurrent RIA testing prior to implementation of a treatment algorithm for patients with CD with LOR to IFX; the comparator group received IFX 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 IFX but no anti-drug antibodies) the algorithm recommended cessation of IFX therapy. That cessation of IFX therapy was not associated with reduced disease control suggests IFX may not be useful for most patients with CD with LOR; however, the criteria for LOR may have been imprecise so that patients appeared to regain response 12 weeks later. Weaknesses of the study include short duration, small size, a large number of withdrawals, the fact 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 IFX in the algorithm arm. Further studies are required to test the reliability of findings.

The TAXIT RCT (Vande Casteele *et al.*[73]) used a test algorithm for IFX responders to optimise the trough IFX level to a set target range. Tests employed in-house ELISAs. Trough optimisation with dose adjustments did not change the proportion of patients with CD in clinical remission (RR for after vs. before 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 patients with CD (54.9% vs. 62.6%; $p = 0.353$). Time to relapse for CD plus UC patients was superior with test monitoring than with 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.*[128] compared proactive trough concentration monitoring with dosing based on clinical judgement (test results did not influence treatment). Among

clinically managed patients, dose escalations were likely to involve a doubling of IFX 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 IFX treatment (hazard ratio 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.

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[73,123,128] were found in which patients treated with IFX were investigated, but no corresponding studies were found for ADA. 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.*[123]) 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; therefore, 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 LOR to IFX may be more costly than the alternative strategy proposed in the authors' 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 IFX.[73] Patients in both randomised groups were optimised to a target trough level of IFX; therefore, a comparator group in which from outset 'treatment decisions made on clinical judgement without measuring levels of TNF[-]α inhibitor and antibodies to TNF[-]α inhibitors'[73] did not exist. The pilot study of Vaughn *et al.*[128] 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.

Evidence taken forward to the economic evaluation

Data from the three management studies[73,123,128] have been taken forward for economic evaluation. The two RCTs, one for responders and the other for IFX recipients with LOR, have informed model structure and provided information for the base-case economic analysis. The study by Vaughn *et al.*,[128] which reports substantial clinical advantage for a test algorithm strategy in terms of retention in IFX treatment, has been used in economic evaluation sensitivity analysis.

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

Search results

The search identified three systematic reviews with meta-analytic pooling of results from multiple studies[32,63,111,114] and 31 primary studies[38,40,47,52,59,77–85,88,92,94,98–100,102,103,106,108,110,115,120,123,126,133,134] that reported the relationship between test outcomes and clinical status of patients in sufficient detail to allow us to extract 2 × 2 data of diagnostic performance when using a drug and/or anti-drug antibody test to diagnose/predict response or LOR. The systematic reviews are summarised in *Published meta-analyses of studies relating test results to clinical state of patients* and the primary studies are analysed in *Analysis of correlation studies of anti-tumour necrosis factor alpha/anti-drug antibodies level and response*.

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

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.

Rationale

Anti-TNF-α drug and anti-drug antibody levels can be used to aid the management of patients with CD on anti-TNF-α drugs if test results are used to predict response or LOR and prompt appropriate action. How good the tests are will therefore depend not only on the choice of treatment (change) following the test results (prescribed by the algorithms discussed in *Objective B: description of algorithms prescribing patient management following test outcomes for drug and/or anti-drug antibody levels*) but also on the diagnostic performance of the test to predict response or lack of response correctly. We therefore reviewed systematic reviews that carried out 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 the tests are measured against is clinical assessment, which is far from perfect.

Results

The literature search yielded several reviews that addressed the relationship between test results and the clinical state of patients with IBD.[36,44,49,61–63,111,114,139] Of these, four were systematic reviews that meta-analysed pooled results from multiple studies.[49,63,111,114] One meta-analysis encompassed several inflammatory conditions in addition to IBD and is not considered further here.[49] The three remaining meta-analyses considered anti-drug antibodies, and one also examined drug trough level tests.[114] Although many of the primary studies included in the meta-analyses 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 RR of a negative test result in LOR relative to state no LOR or, conversely, RR of LOR in patients with negative test relative to those with a positive test). Viewing the tests as diagnostic/predictive permits hierarchical (bivariate) meta-analysis that incorporates covariance between sensitivity and specificity estimates. The RR statistic does not formally allow for covariance between estimated associations. Each of the meta-analyses is considered in turn.

Nanda *et al.*[111] The authors estimated the pooled RR of LOR to IFX 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 patients who were antibody positive than in patients who were antibody negative generates a RR of > 1.0). Eleven studies were included,[40,47,55,59,79,83,92,99,103,120,142] 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 (*Figure 23a*; RR 3.16, 95% CI 2.00 to 4.98; $I^2 = 70.1\%$) indicated that the risk of LOR was about threefold higher in those with a positive anti-drug antibodies test than in those with a negative test.

When viewed as a predictive/diagnostic test[143] 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).[144] In this case, a positive test for anti-drug antibodies is viewed as predictive/diagnostic of LOR. *Figure 23b* indicates marked heterogeneity among the studies and the trade-off between sensitivity and specificity in the different studies. *Figure 23c* and *d* summarises the summary receiver operating characteristic (sROC) meta-analysis results. The meta-analysis test accuracy results are summarised in *Table 19*. The large RCT-based study by Hanauer *et al.*[40] was identified as both influential and 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 (see *Figure 23c*). 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.

The implication of these test accuracy results was explored in terms of predictive values as suggested in the *Cochrane Handbook for Systematic Reviews of Interventions*.[141] As predictive values are influenced by prevalence of the target condition, we determined a pooled random-effects estimate of prevalence (LOR) among the studies (34.7%, 95% CI 25.1% to 44.4%). The point estimates for positive predictive values

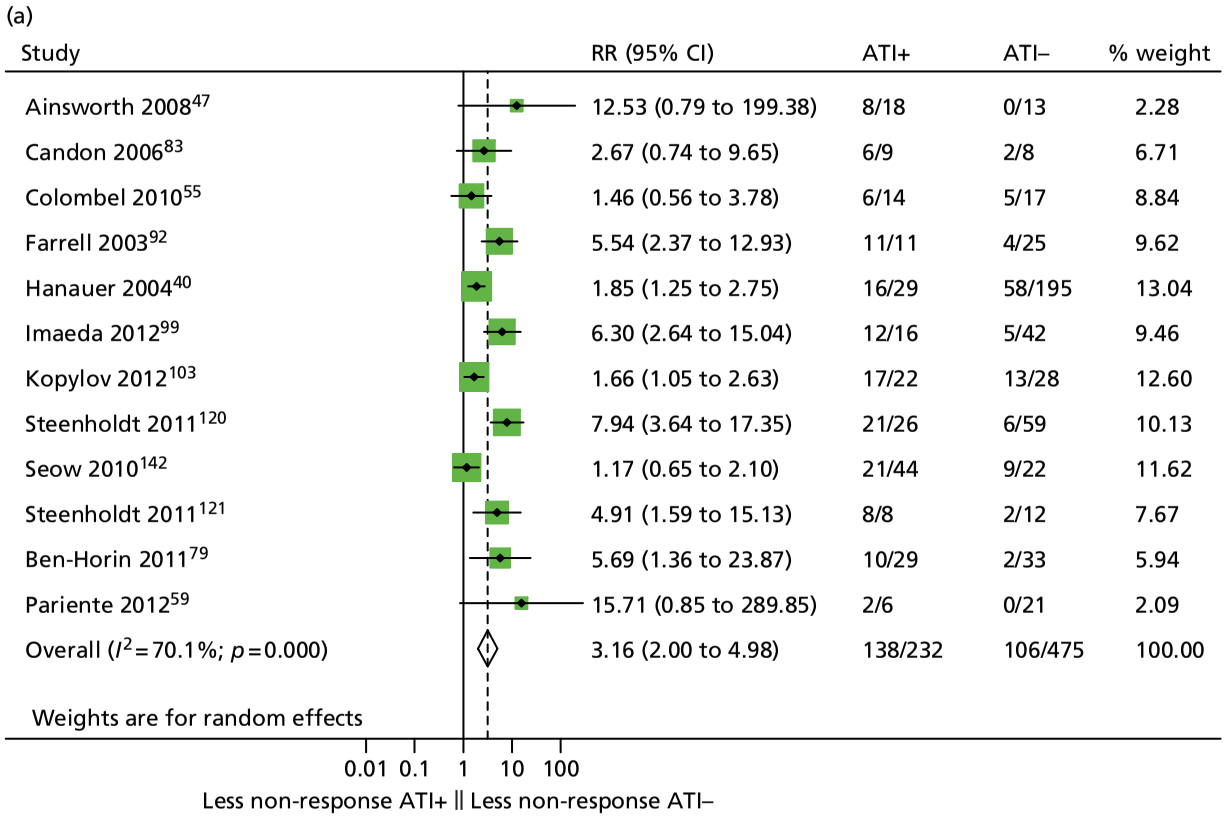
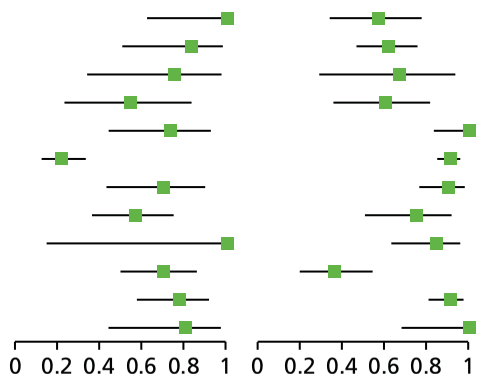
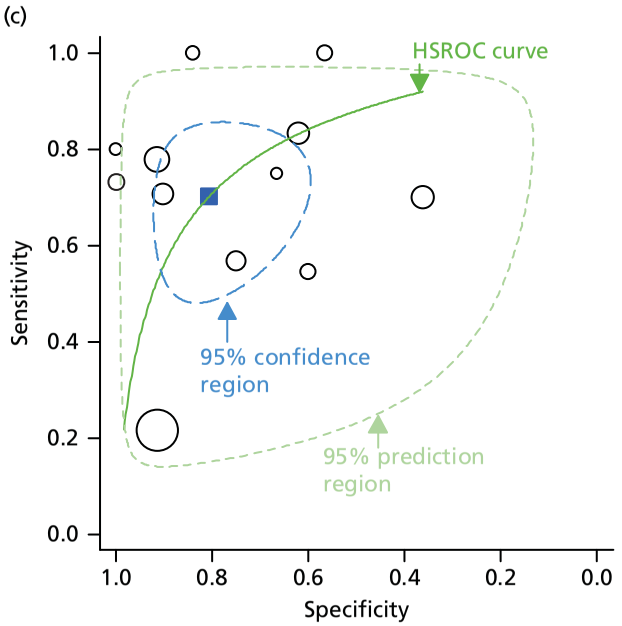


FIGURE 23 Meta-analysis of data based on that from Nanda *et al.*[111] (a) Random-effects meta-analysis of the RR of LOR to IFX (anti-drug antibodies positive vs. anti-drug antibodies negative); (b) sensitivity and specificity forest plot of the included studies; (c) summary ROC bivariate meta-analysis of sensitivity–specificity pairs (hollow symbols) with pooled point estimate square solid symbol; and (d) summary ROC bivariate meta-analysis of sensitivity–specificity pairs with pooled point estimate square solid symbol, excluding an influential outlier study. ATI, antibodies to IFX; FN, false negative; FP, false positive; HSROC, hierarchical summary receiver operating characteristic; TN, true negative; TP, true positive. (*continued*)

(b)

Study	TP	FP	FN	TN	Condition	Sensitivity (95% CI)	Specificity (95% CI)
Ainsworth 2008[47]	8	10	0	13	CD	1.00 (0.63 to 1.00)	0.57 (0.34 to 0.77)
Ben-Horin 2011[79]	10	19	2	31	IBD	0.83 (0.52 to 0.98)	0.62 (0.47 to 0.75)
Candon 2006[83]	6	3	2	6	CD	0.75 (0.35 to 0.97)	0.67 (0.30 to 0.93)
Colombel 2010[55]	6	8	5	12	CD	0.55 (0.23 to 0.83)	0.60 (0.36 to 0.81)
Farrell 2003[92]	11	0	4	21	CD	0.73 (0.45 to 0.92)	1.00 (0.84 to 1.00)
Hanauer 2004[40]	16	13	58	137	CD	0.22 (0.13 to 0.33)	0.91 (0.86 to 0.95)
Imaeba 2012[99]	12	4	5	37	CD	0.71 (0.44 to 0.90)	0.90 (0.77 to 0.97)
Kopylov 2012[103]	17	5	13	15	CD	0.57 (0.37 to 0.75)	0.75 (0.51 to 0.91)
Pariente 2012[59]	2	4	0	21	IBD	1.00 (0.16 to 1.00)	0.84 (0.64 to 0.95)
Seow 2010[142]	21	23	9	13	UC	0.70 (0.51 to 0.85)	0.36 (0.21 to 0.54)
Steenholdt 2011[120]	21	5	6	53	CD	0.78 (0.58 to 0.91)	0.91 (0.81 to 0.97)
Steenholdt 2011[120]	8	0	2	10	UC	0.80 (0.44 to 0.97)	1.00 (0.69 to 1.00)





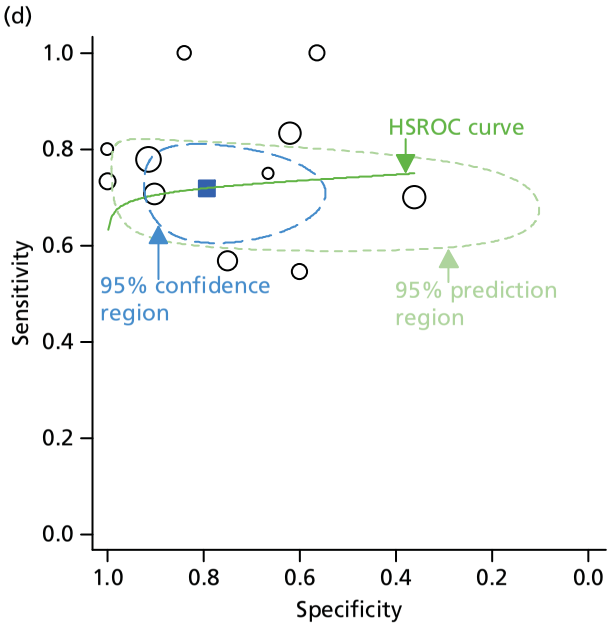


FIGURE 23 Meta-analysis of data based on that from Nanda *et al.*[111] (a) Random-effects meta-analysis of the RR of LOR to IFX (anti-drug antibodies positive vs. anti-drug antibodies negative); (b) sensitivity and specificity forest plot of the included studies; (c) summary ROC bivariate meta-analysis of sensitivity–specificity pairs (hollow symbols) with pooled point estimate square solid symbol; and (d) summary ROC bivariate meta-analysis of sensitivity–specificity pairs with pooled point estimate square solid symbol, excluding an influential outlier study. ATI, antibodies to IFX; FN, false negative; FP, false positive; HSROC, hierarchical summary receiver operating characteristic; TN, true negative; TP, true positive.

TABLE 19 Test accuracy parameters generated by hierarchical MA[144]

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

(PPVs) and negative predictive values (NPVs) 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.

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

Lee *et al.*[63] The authors estimated the pooled RR of remission in patients with a positive test for anti-drug antibodies to IFX relative to those with a negative test for anti-drug antibodies (a RR of < 1.0 indicates that anti-drug antibodies are associated with lower risk of remission, consistent with the hypothesis that anti-drug antibodies reduce response to IFX therapy) based on nine studies.[38,40,47,55,83,92,142,145,146] Comparative numbers of events and patients were reported. The fixed- and random-effects RRs 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 IFX is considered as predictor of, or diagnostic of, a lack of remission, then meta-analysis yielded low joint sensitivity specificity values of 0.42 and 0.69, respectively (*Figure 25*).

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.

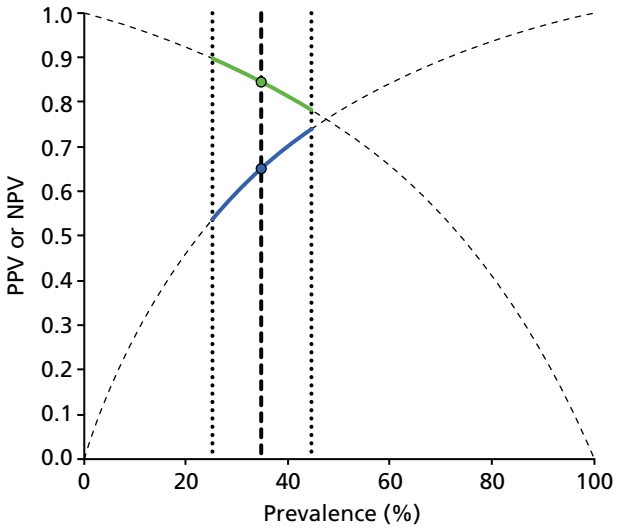


FIGURE 24 Positive predictive values and NPVs, according to prevalence of LOR at the sROC, model estimates of sensitivity and specificity, as prevalence increases PPV increases and NPV decreases. Data points are PPV and NPV at sROC sensitivity and specificity and pooled prevalence. Dashed vertical lines are pooled prevalence and 95% CI. Thick curves are PPV and NPV for hierarchical model sensitivity and specificity at the pooled prevalence and 95% CI.

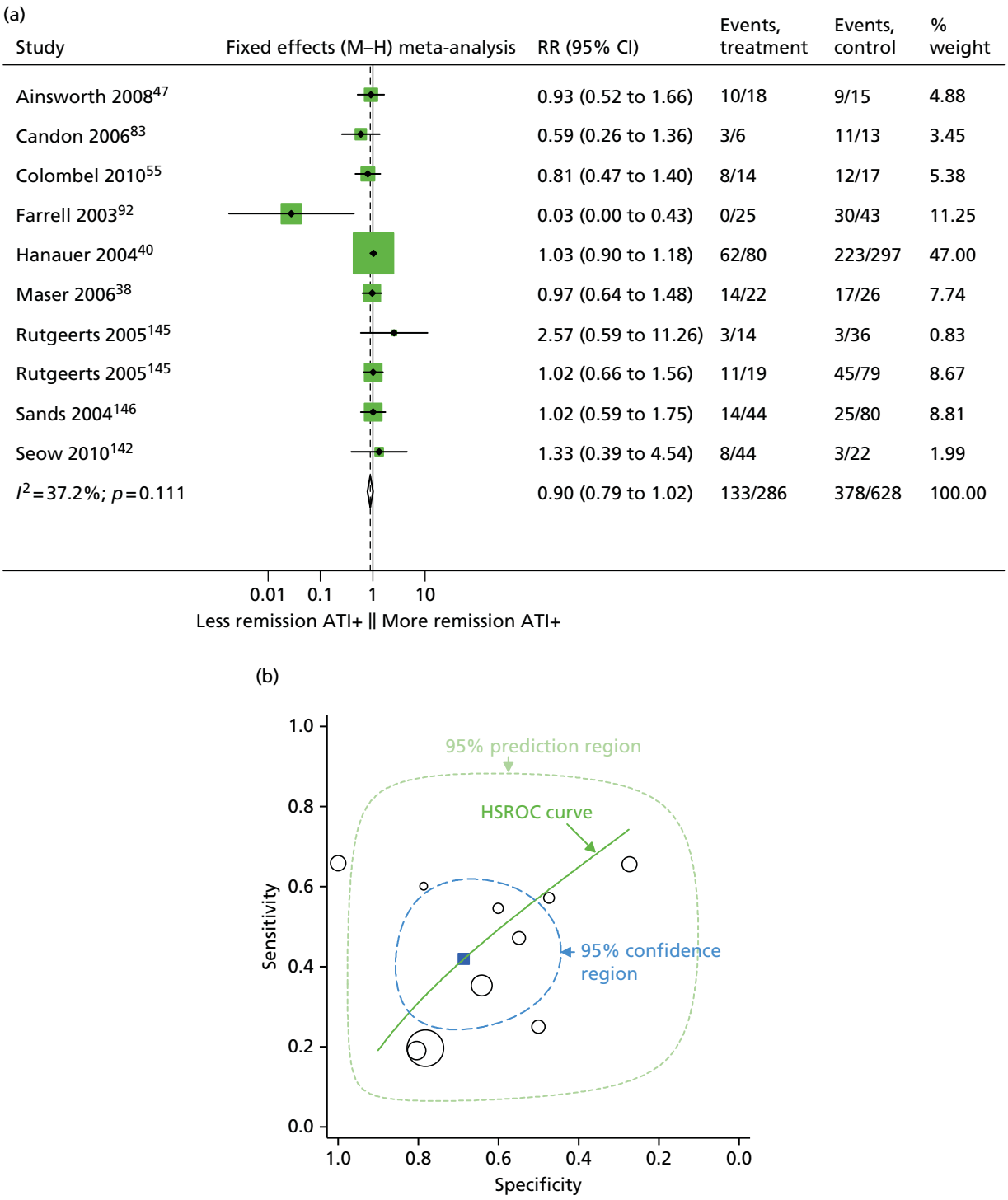


FIGURE 25 Meta-analysis of data based on that from Lee *et al.*[63] (a) Fixed-effects meta-analysis of the RR of remission (presence of antibodies to IFX vs. absence of antibodies to IFX); and (b) sROC bivariate meta-analysis of sensitivity specificity pairs (hollow symbols) with pooled point estimate square solid symbol. ATI, antibodies to IFX; HSROC, hierarchical summary receiver operating characteristic; M–H, Mantel–Haenszel.

Lee *et al.*[63] 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[38,40,43,45,53–55,142,146–148] and the authors generated a fixed-effects RR (antibodies present with suppressants vs. 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. The results of fixed- and random-effects meta-analyses are illustrated in *Figure 26*.

Paul *et al.*[114] The authors estimated the pooled ratio for the odds of a lack of response in association with a negative test for ADA (i.e. subtherapeutic) compared with the odds of a lack of response in those with a positive test for ADA (an OR of > 1.0 indicates that subtherapeutic ADA 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 three studies of adults with CD (pooled OR 7.5, 95% CI 3.58 to 13.90) and two studies of children with CD (pooled OR 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 subtherapeutic 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 *et al.*[111] review using the test for antibodies to IFX 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 ADA compared with the odds of lack of response in those with a positive test for anti-drug antibodies (an OR of > 1.0 indicates that the presence of anti-drug antibodies is 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 ADA) as a predictor for the lack of clinical response. This value is similar to the DOR for antibodies to IFX as a predictor for the lack of response [using the pooled sensitivity specificity pair (0.72 and 0.79) derived from the studies in the Nanda *et al.*[111] review, which provides a DOR of 9.67 [(0.72/0.28)/(0.21/0.79)].

Summary

When viewed as predictive tests for lack of response and or lack of remission, the published meta-analyses 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.

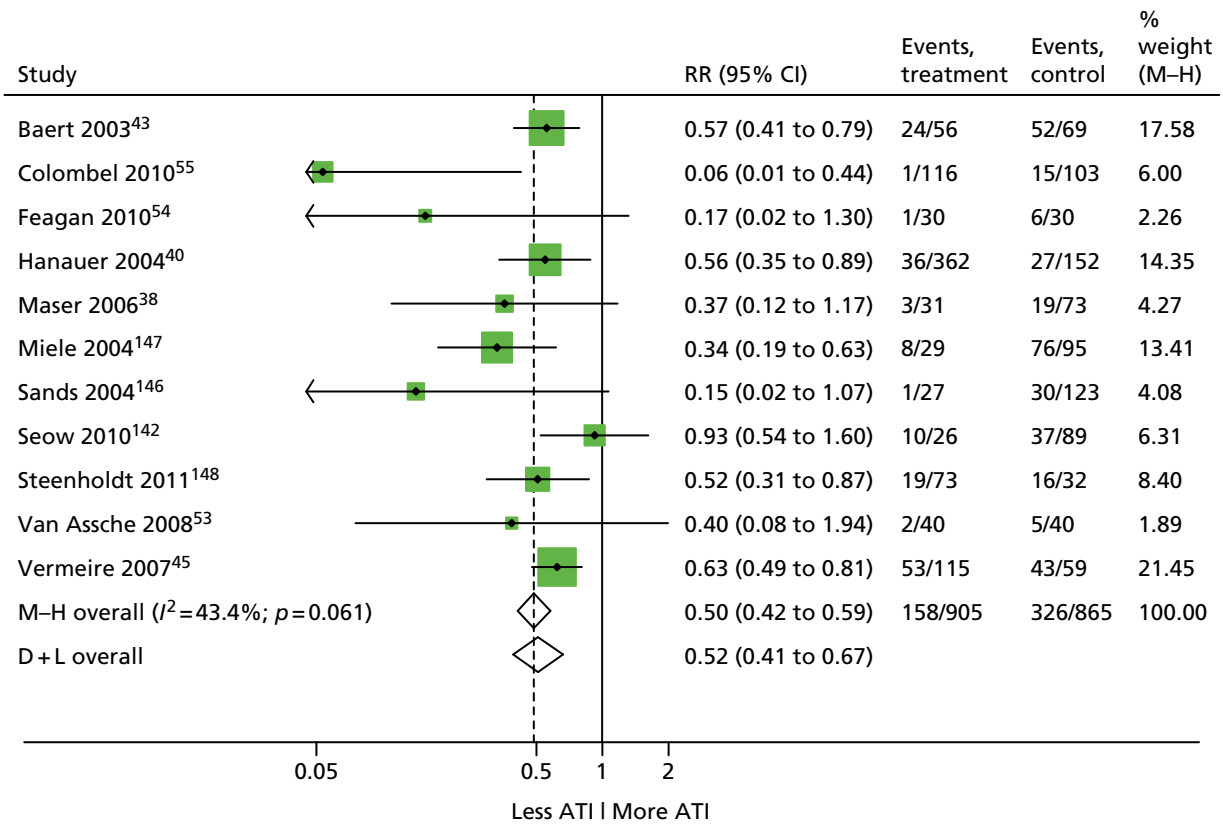


FIGURE 26 The RR of anti-drug antibodies with immunosuppressants vs. without suppressants. Data based on Lee *et al.*[63] ATI, antibodies to IFX; D + L, DerSimonian–Laird; M–H, Mantel–Haenszel.

Analysis of correlation studies of anti-tumour necrosis factor alpha/anti-drug antibodies level and response

Aim

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

Rationale

No published and tested multivariable prognostic models were found that incorporated test results with other variables to predict clinical status. The majority of the identified studies of tests for anti-TNF-α and anti-drug antibody levels were classified as correlation studies (see *Table 3*). These reported correlations or associations between test results and other patient-dependent variables. Only one published management study[123] 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:

1. In concurrent testing, levels of anti-TNF-α drug and anti-drug antibodies are measured at the same time. When tests are dichotomised using cut-off points, four patient categories are generated:

    i. anti-drug antibodies positive and anti-TNF-α negative
    ii. anti-drug antibodies positive and anti-TNF-α positive
    iii. anti-drug antibodies negative and anti-TNF-α negative
    iv. anti-drug antibodies negative and anti-TNF-α positive.

2. In reflex testing, tests for anti-TNF-α levels precede subsequent testing for anti-drug antibodies, and anti-drug antibodies tests are carried out only in those with subtherapeutic levels of anti-TNF-α. When tests are dichotomised using cut-off points, three patient categories are generated:

    i. anti-TNF-α positive
    ii. anti-TNF-α negative and anti-drug antibodies positive
    iii. anti-TNF-α negative and anti-drug antibodies negative.

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 also be obtained from studies in which test results are reported by group rather than by individual; when several such studies are available, the option of meta-analysis of multiple studies offers greater power. Studies that undertook both anti-TNF-α and anti-drug antibody tests but did not provide test results for each patient (but only by group) were not useful for obtaining 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 to provide a comparison (by a 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.

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-TNF-α levels and measures of serum CRP concentration, or of FC, 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 2 × 2 diagram similar to that shown in *Table 20*.

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 RR of a positive test result in state A relative to state B {[a/(a + d)]/[b/(b + c)]} or, conversely, RR of state A in patients with a positive test relative to those with a negative test {[a/(a + b)]/[d/(d + c)]}.

The 31 studies[38,40,47,52,59,77–85,88,92,94,98–100,102,103,106,108,110,115,120,123,126,133,134] taken forward for meta-analysis (see *Table 53*, *Appendix 12*) were heterogeneous in terms of populations, treatments, tests used, completeness of reporting, test cut-off points 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 when data from medical records about clinical state were available and serum samples had been collected and stored for future assay. The most common threats to validity of study findings in this collection of studies are 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.[99,100,123] 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 positive and antibody positive; drug positive and antibody negative; drug negative and antibody positive; and drug negative and antibody positive).[99,100,123] The results summarised in *Tables 21–23* indicate the probability of LOR according each possible test result category for the three studies.

Reflex testing: test result probabilities The test results for studies that 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: (1) anti-TNF-α positive; (2) anti-TNF-α negative and anti-drug antibodies positive; and (3) anti-TNF-α negative and anti-drug antibodies negative. The results summarised in *Tables 24–26* indicate the probability of LOR according to each possible test result category.

Meta-analytic test result probabilities: trough infliximab levels Meta-analysis results for single-test studies using trough IFX levels as a test for LOR in responders and failure to regain response in patients with LOR are summarised in *Appendix 12*, *Infliximab trough level tests for loss of response or lack of regaining response*. For responders the probability of returning a positive test result (i.e. IFX 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

TABLE 20 Illustration of 2 × 2 table data from correlation studies

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

FN, false negative; FP, false positive; TN, true negative; TP, true positive.

TABLE 21 Concurrent testing for responders receiving ADA

Study: Imaeda *et al.*, 2014[100]	ADAbs positive	ADAbs negative	Total	Population and anti-TNF-α therapy; tests
Anti-TNF-α negative	LOR = 8	LOR = 2	LOR = 10	Responders on ADA maintenance; ELISA; prevalence of LOR = 37.5%
	RESP = 0	RESP = 2	RESP = 2	
Anti-TNF-α positive	LOR = 2	LOR = 3	LOR = 5	
	RESP = 4	RESP = 19	RESP = 23	
Total	LOR = 10	LOR = 5	LOR = 15	
	RESP = 4	RESP = 21	RESP = 25	

ADAb, anti-drug antibody; RESP, responders.
Note
The probabilities of a patient returning each of the four possible test result combinations were ADAbs positive and anti-TNF-α negative = 0.200; ADAbs positive and anti-TNF-α positive = 0.150; ADAbs negative and anti-TNF-α negative = 0.10; and ADAbs negative and anti-TNF-α positive = 0.550.
The probabilities of losing response according to category of test result were 1.00, 0.333, 0.500 and 0.136, respectively.

TABLE 22 Concurrent testing for responders receiving IFX

Study: Imaeda *et al.*, 2012[99]	ADAbs positive	ADAbs negative	Total	Population and anti-TNF-α therapy; tests
Anti-TNF-α negative	LOR = 9	LOR = 0	LOR = 9	Responders on IFX maintenance; ELISA; prevalence of LOR = 29.3%
	RESP = 1	RESP = 7	RESP = 8	
Anti-TNF-α positive	LOR = 3	LOR = 5	LOR = 8	
	RESP = 3	RESP = 30	RESP = 33	
Total	LOR = 12	LOR = 5	LOR = 17	
	RESP = 4	RESP = 37	RESP = 41	

ADAb, anti-drug antibody; RESP, responders.
Note
The probabilities of a patient returning each of the four possible test result combinations were ADAbs positive and anti-TNF-α negative = 0.172; ADAbs positive and anti-TNF-α positive = 0.103; ADAbs negative and anti-TNF-α negative = 0.121; and ADAbs negative and anti-TNF-α positive = 0.603.
The probabilities of losing response according to category of test result were 0.900, 0.500, 0.000 and 0.143, respectively.

TABLE 23 Concurrent testing for patients with LOR receiving IFX

Study: Steenholdt *et al.*, 2014[123]	ADAbs positive	ADAbs negative	Total	Population and anti-TNF-α therapy; tests
Anti-TNF-α negative	NOR = 8	NOR = 2	NOR = 10	Failure on IFX, continued failure or gain of response at 12 weeks; RIA; prevalence of NOR = 44.9%
	RESP = 6	RESP = 1	RESP = 7	
Anti-TNF-α positive	NOR = 1	NOR = 20	NOR = 21	
	RESP = 3	RESP = 28	RESP = 31	
Total	NOR = 9	NOR = 22	NOR = 31	
	RESP = 9	RESP = 29	RESP = 38	

ADAb, anti-drug antibody; NOR, no regain of response; RESP, responders.
Note
The probabilities of a patient returning each of the four possible test result combinations were ADAbs positive and anti-TNF-α negative = 0.203; ADAbs positive and anti-TNF-α positive = 0.058; ADAbs positive and anti-TNF-α negative = 0.0.043; and ADAbs negative and anti-TNF-α positive = 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.

TABLE 24 Reflex testing for responders receiving ADA

Study: Imaeda *et al.*, 2014[100]	ADAbs positive	ADAbs negative	Total	Population and anti-TNF-α therapy; tests
Anti-TNF-α negative	LOR = 8	LOR = 2	LOR = 10	Responders on ADA maintenance; ELISA; prevalence of LOR = 37.5%
	RESP = 0	RESP = 2	RESP = 2	
Anti-TNF-α positive	LOR = 5	–	LOR = 5	
	RESP = 23	–	RESP = 23	
Total	LOR = 13	LOR = 2	LOR = 15	
	RESP = 23	RESP = 2	RESP = 25	

ADAb, anti-drug antibody; RESP, responders.
Note
The probabilities of a patient returning each of the three possible test result combinations were anti-TNF-α positive = 0.700; anti-TNF-α negative and ADAbs positive = 0.200; and anti-TNF-α negative and ADAbs negative = 0.100.
The probabilities of losing response according to category of test result were 0.179, 1.00 and 0.500, respectively.

TABLE 25 Reflex testing for responders receiving IFX

Study: Imaeda *et al.*, 2012[99]	ADAbs positive	ADAbs negative	Total	Population and anti-TNF-α therapy; tests
Anti-TNF-α negative	LOR = 9	LOR = 0	LOR = 9	Responders on IFX maintenance; ELISA; prevalence of LOR = 29.3%
	RESP = 1	RESP = 7	RESP = 8	
Anti-TNF-α positive	LOR = 8	–	LOR = 8	
	RESP = 33	–	RESP = 33	
Total	LOR = 17	LOR = 0	LOR = 17	
	RESP = 34	RESP = 7	RESP = 41	

ADAb, anti-drug antibody; RESP, responders.
Note
The probabilities of a patient returning each of the three possible test result combinations were anti-TNF-α positive = 0.707; anti-TNF-α negative and ADAbs positive = 0.172; and anti-TNF-α negative and ADAbs negative = 0121.
The probabilities of losing response according to category of test result were 0.195, 0.00 and 0.900, respectively.

TABLE 26 Reflex testing for patients with LOR receiving IFX

Study: Steenholdt *et al.*, 2014[123]	ADAbs positive	ADAbs negative	Total	Population and anti-TNF-α therapy; tests
Anti-TNF-α negative	NOR = 8	NOR = 2	NOR = 10	Failure on IFX, continued failure or gain of response at 12 weeks; RIA; prevalence of NOR = 44.9%
	RESP = 6	RESP = 1	RESP = 7	
Anti-TNF-α positive	RESP = 31	–	NOR = 21	
	NOR = 21	–	RESP = 31	
Total	NOR = 29	NOR = 2	NOR = 31	
	RESP = 37	RESP = 1	RESP = 38	

ADAb, anti-drug antibody; NOR, no regain of response; RESP, responders.
Note
The probability of a patient returning each of the three possible test result combinations was: anti-TNF-α positive = 0.754; anti-TNF-α negative and ADAbs positive = 0.203; and anti-TNF-α negative and ADAbs negative = 0.044.
The probabilities of not gaining response according to category of test result were 0.404, 0.667 and 0.571, respectively.

0.271 when prevalence was set to that of the single available patient-level study of Imaeda *et al.*,[100] 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 LOR 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 IFX as a test for LOR in responders and failure to regain response in patients with LOR are summarised in *Appendix 12*, *Antibodies to infliximab tests for loss of response or lack of regaining response*. The probability of returning a positive test result (i.e. anti-IFX antibodies undetectable) was 0.345 at the pooled prevalence (the range based on 95% CI for prevalence was 0.324 to 0.365); and for a negative test result the probability was 0.655 (the range based on 95% CI for prevalence was 0.635 to 0.686); these do 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.*,[100] which returned a similar positive test probability of 0.276 (95% CI 0.167 to 0.409).

Seven heterogeneous studies[47,59,77,79,83,123,126] were available for patients with LOR (see *Table 53, Appendix 12*). The probability of returning a positive test result (i.e. anti-IFX 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.*,[123] which returned a much lower positive test probability of 0.261 (95% CI 0.163 to 0.381).

Meta-analytic test result probabilities: trough ADA levels Meta-analysis results for single-test studies using trough ADA levels as a test for LOR in responders and failure to regain response in patients with LOR are summarised in *Appendix 12*, *Adalimumab trough level test for loss of response or lack of regaining response*. The probability of returning a positive test result (i.e. ADA 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.*,[99] which returned a lower positive test probability of 0.300 (95% CI 0.166 to 0.465).

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

Meta-analytic test result probabilities: anti-ADA antibody levels Meta-analysis results for single-test studies using trough anti-ADA antibody levels as a test for LOR in responders are summarised in *Appendix 12*, *Antibodies to adalimumab as test for loss of response or lack of regaining response*.

The probability of returning a positive test result (i.e. anti-ADA 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.*,[99] which returned a higher positive test probability of 0.350 (95% CI 0.206 to 0.517).

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 IFX-treated responders, one for IFX-treated patients with LOR and one for responders treated with ADA). These studies allowed estimation of the proportion of patients who would enter each of the treatment categories following from concurrent or reflex testing strategies.

Representativeness of available evidence As 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, partly because of 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 three patient-level studies 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 antibody level tests as predictors of clinical status was moderate (see *Appendix 12*). PPVs and NPVs across clinical prevalence ranges indicated that 20–30% of positive and negative test results were incorrect at plausible prevalence settings for clinical status (see *Appendix 12*, *Predictive values for drug and anti-drug antibodies tests for LOR or failure to regain response*).

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.*[99] of patients treated with IFX and that of Steenholdt *et al.*[123] of patients with LOR to maintenance IFX. As the Steenholdt *et al.*[123] study coupled testing results with prospective implementation of a treatment algorithm, it was used in the base-case economic analysis. Data from the Imaeda *et al.* study[99] were used in a sensitivity analysis in the cost-effectiveness comparison of testing strategies with 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.

Summary of clinical effectiveness findings

Assays based on different principles have been developed to measure anti-TNF-α agents and antibodies to anti-TNF-αs 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 patients with CD responding to anti-TNF-α agents. The TAXIT trial described outcomes when a test algorithm strategy based on trough IFX levels was implemented for IBD patients responding to IFX, but a standard care comparator population was not available because all randomised TAXIT study patients received test-directed optimisation of IFX dosing. A single retrospective case series of IBD patients responding to IFX reported better retention in IFX treatment for those whose dose changes were based on prospective testing than for 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 with an intensified dose strategy in patients with CD who had lost response to IFX. 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) and was small (69 patients); 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-TNF-α agents and/or antibodies to anti-TNF-α drugs provides a clinical advantage over standard anti-TNF-α strategies used for responders or for patients with LOR. Ongoing trials may deliver more relevant data to inform a decision.

The main points of the clinical effectiveness can be summarised:

- ELISAs 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 the 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-off points 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 from the ones presented to us in the NICE scope for this work, reflecting the influence of the variation in clinical judgement.
- 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 because of reduced administration of IFX in patients who do not require IFX (drug positive and anti-drug antibody negative) according to one RCT.
- Drug optimisation during the 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 IFX treatment according to one observational study.
- Problems with the RCTs included:
    - mixed patient populations
    - short follow-up
    - small patient numbers
    - no evidence on ADA
    - 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 patients 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[73,123,128] both informed 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 correlated clinical status only with either test results for anti-TNF-α or results for antibodies to anti-TNF-α). 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.

Chapter 4 Cost-effectiveness review and health economic modelling

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, TNF-α-Blocker ELISA kits and Promonitor ELISA kits for measuring levels of TNF-α inhibitors and of anti-drug antibodies in detail.

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, TNF-α-Blocker ELISA kits and Promonitor ELISA kits for measuring levels of TNF-α inhibitors and of anti-drug antibodies in detail.

Methods

Search strategy

A comprehensive search of the literature for published economic evaluations (including any existing models), cost studies and QoL (utility) studies was performed. The systematic search included searching the following electronic databases during December 2014 (from 12 to 17 December 2014):

- MEDLINE (via Ovid) (1946 to Week 3 November 2014)
- MEDLINE In-Process Citations and Daily Update (via Ovid) (11 December 2014)
- EMBASE (via Ovid) (1947 to 15 December 2014)
- NHS Economic Evaluation Database (The Cochrane Library)
- Science Citation Index (Web of Knowledge) (1970–present)
- Cost-effectiveness Analysis Registry
- EconPapers (Research Papers in Economics)
- School of Health and Related Research Health Utilities Database.

The search included terms for CD, anti-TNF-α drugs and the different assay kits, combined with economic and 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*.

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 CD
- intervention – anti-TNF-α drugs (ADA and IFX) and antibody drug testing (LISA-TRACKER ELISA kits, TNF-α-Blocker ELISA kits and Promonitor ELISA kits) for any dosage or treatment regimen
- comparator – standard care treatment: anti-TNF-α drugs (ADA and IFX) 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, Short Form questionnaire-6 Dimensions score or quality-adjusted life-years (QALYs)].

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-TNF-α drugs have also been used, such as UC, rheumatoid arthritis, psoriasis and tuberculosis.

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 reasons.

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 1 considered all eligible studies (fully published economic evaluations including any economic models) and stage 2 considered studies assessed for usefulness for populating the economic model. Data extracted during stage 1 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.

Quality assessment

The quality of full economic evaluation studies that were identified was 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: title and abstract, introduction, methods, results, discussion and other. Under these dimensions, a series of questions check whether or not 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.*[149] (see *Appendix 15*). The Philips checklist contains two main dimensions: structure of the model and data used to parameterise the model. Under these dimensions several questions assess whether or not the criteria has been clearly reported.

Data synthesis

Information extracted from the included studies were summarised and tabulated. Findings from individual studies were compared narratively.

Results

In developing the economic model, we have consulted the previous technology appraisal guideline and HTA report by Dretzke *et al.*[5] even though this work did not include any assay kits for measuring levels of TNF-α 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.

Summary of the *Health Technology Assessment* report by Dretzke *et al.*[5]

The main aim of this HTA report was to assess the cost-effectiveness of anti-TNF-α drugs in the management of adult patients with moderate to severe CD in the UK NHS. The authors described induction therapy as the use of anti-TNF-α therapy with the aim of achieving remission (a repeated reinduction treatment was considered, rather than a one-off induction therapy) and maintenance therapy as the use of anti-TNF-α therapy to maintain remission in patients who have responded (and continue to respond) to anti-TNF-α therapy when in relapse. Response by the authors was defined as remission within 8 weeks.

The authors developed a Markov model from a NHS and Personal Social Services (PSS) perspective to estimate the incremental cost per QALY gained for both ADA and IFX (anti-TNF-α therapy) compared with standard care. Mortality was not included in the model, as the authors found no difference in the mortality rates that 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 Silverstein *et al.*[150] 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 the study by Gregor *et al.*,[20] which used the time-trade off measure to estimate the health-related QoL in CD. A utility value for surgery was not available in the 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 the sum of the anti-TNF-α costs and type-specific health-state costs. The costs of anti-TNF-α therapy, both induction and maintenance, were derived from the BNF (2007/8),[151,152] and administration costs were also included for IFX. Type-specific health-state costs included costs for surgery, which were modelled as the cost of inpatient IBD interventions, and post-surgery remission costs, which 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 to 2006*[153]).

Incremental cost-effectiveness ratios (ICERs) and cost-effectiveness acceptability curves were presented. One-way sensitivity analyses and probabilistic sensitivity analyses (PSAs) using 10,000 simulations were conducted to characterise uncertainty in the model.

For induction therapy for severe CD, both ADA and IFX dominated standard care (i.e. they were cheaper and more effective). For maintenance therapy for severe CD, neither drug was cost-effective (well above NICE thresholds). For moderate CD, for maintenance therapy for both drugs and induction therapy for IFX, these were not cost-effective (well above NICE thresholds); however, for induction therapy, ADA dominated standard care.

Sensitivity analysis showed that, in patients with severe disease, IFX induction treatment was cost-effective relative to maintenance treatment and standard care in > 99% of cases at all points up to £100,000 per QALY. Likewise, ADA 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 (1-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.*[150] data for transition probabilities, which inherently had their own problem (i.e. surgery rates were higher and relapse rates much lower than in routine practice).

Search results for objective D

The literature search identified 2466 records through electronic database searches and other sources. After removing duplicates, 1527 records were screened for inclusion. On the basis of a title and abstract sift only, 1518 records were excluded. The remaining nine records were subjected to full-text screening. A further five articles[5,26,154–156] were excluded at the full-text stage, as these studies did not use assay kits to measure levels of TNF-α inhibitors and anti-drug antibodies. The literature search identified four studies[73,123,124,157] of the cost-effectiveness of different assay kits for measuring levels of TNF-α inhibitors and of anti-drug antibodies (*Figure 27*).

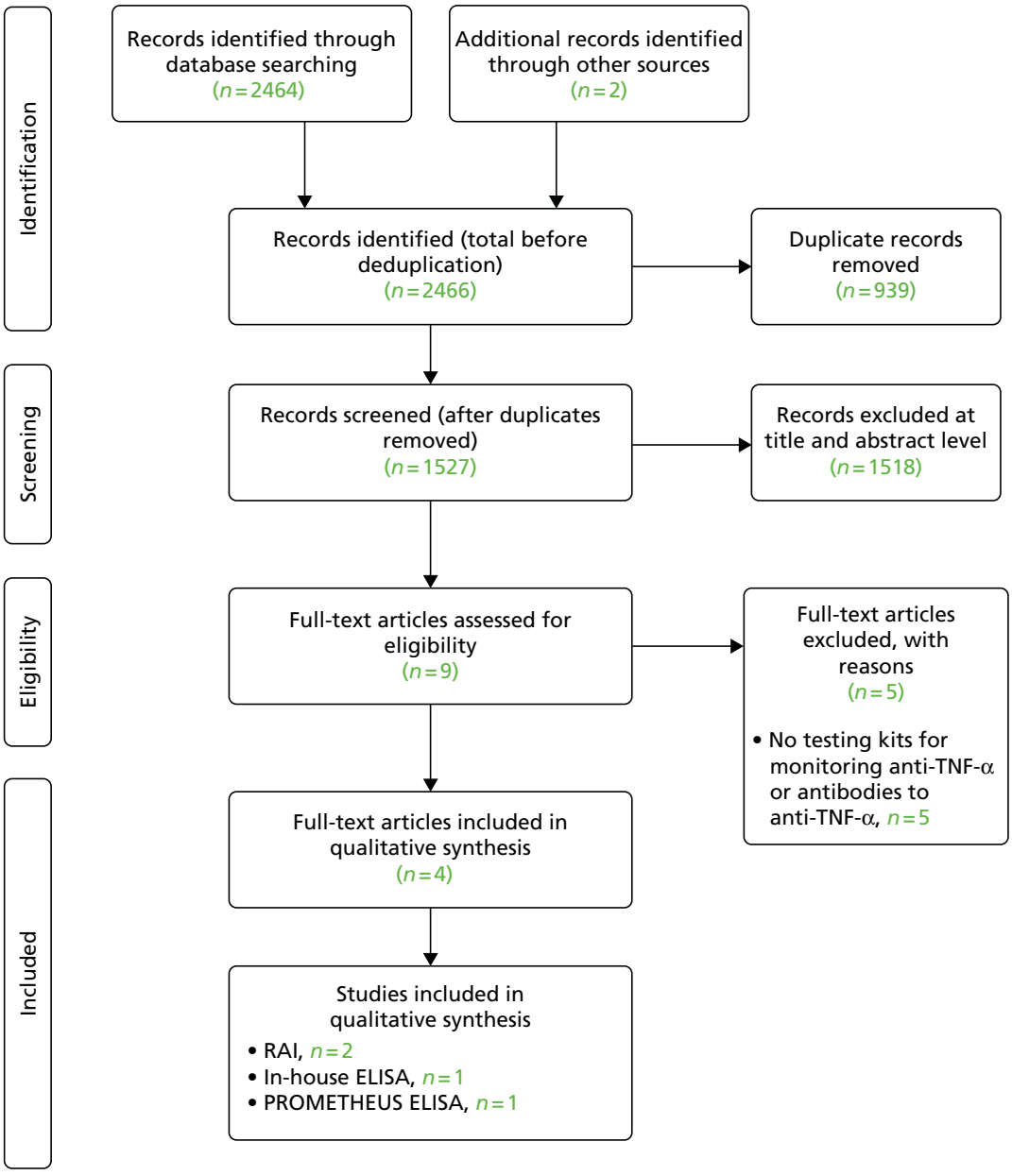


FIGURE 27 The PRISMA flow diagram of cost-effectiveness studies.

Overview of included studies

The literature search identified four studies[73,123,124,157] that met our inclusion criteria (studies looking at the cost-effectiveness of different assay kits for measuring levels of TNF-α inhibitors and of anti-drug antibodies) and were reviewed. In the following sections we present an overview of the included studies by population (responders and those showing LOR) of interest.

Vande Casteele et al.[73]

Vande Casteele *et al.*[73] aimed to determine whether or not concentration-based IFX dosing was more cost-effective than clinically based IFX dosing. These authors conducted a RCT and assigned people with moderate to severe CD or UC to receive concentration-based or clinically based IFX dosing. Included patients were those who were treated with maintenance IFX 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)'.[73] Patients eligible for the study were dose optimised until IFX trough concentrations between 3 and 7 µg/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 patients had a trough concentration between 3 and 7 µg/ml. Briefly, depending on IFX trough concentration, patients received an increase dose of IFX treatment, no dose adaptation or a decrease in IFX 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 1 year. The EQ-5D was used to calculate 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 ICER based on the outcome of cost per QALY gained. Uncertainty in incremental QALYs and costs was determined by non-parametric bootstrapping consisting of 1000 iterations and plotted onto a cost-effectiveness plane. The base-case results demonstrated that concentration-based dosing was slightly less effective (0.8227 vs. 0.8421) and less costly (€20,700 vs. €21,000) than clinically based dosing, but overall differences were small.

Steenholdt et al.[123]

Steenholdt *et al.*[123] assessed the cost-effectiveness of receiving treatment based on serum concentrations of IFX and IFX antibodies at the time of IFX treatment failure in accordance with the algorithm (for further details of the algorithm, see *Chapter 3*, *Objective B: description of algorithms prescribing patient management following test outcomes for drug and/or anti-drug antibody levels*) compared with receiving IFX at an increased dose frequency of 5 mg/kg every 4 weeks. The study included patients who experienced failure of IFX treatment while on maintenance treatment. Failure of IFX treatment was defined in the study as recurrence of active disease with a CDAI score of $\geq$ 220 and/or a minimum of one draining fistula. Serum IFX and IFX antibodies were analysed using RIA. Samples were stored and further analysed using ELISA and HMSA after study completion. The study was a single-blind RCT 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. Clinical effectiveness was based on clinical response rates, which is regaining response or continuing to lose response to IFX therapy. Resource use and costs were based on IFX doses and all inpatient and outpatient contacts in hospitals, which also included diagnostic and treatment procedures that were recorded in the National Patient Registry database. Costs were reported in Danish krone and converted to euros in 2012 prices. The base-case results were expressed as cost per ITT and PP population. Costs were compared using arithmetic means and were assessed by non-parametric bootstrapping. One-way sensitivity analyses of key primary and secondary end points were conducted. The base-case results showed that costs were significantly lower in the algorithm group than in the IFX intensification group in both the ITT and PP population.

Steenholdt et al.[124]

In follow-up to their study published in 2014,[123] Steenholdt *et al.*[124] extended the time horizon to 1 year to assess the long-term costs and clinical outcomes of treatment of CD in patients with LOR to IFX maintenance therapy using a proposed algorithm compared with intensified IFX treatment. Serum IFX and IFX antibodies

were analysed using RIA, and were further analysed using ELISA and HMSA after study completion. IFX levels were classified as therapeutic or subtherapeutic (≥ 0.5 µg/ml and < 0.5 µg/ml, respectively); IFX antibodies were classified as detectable or undetectable. Costs were assessed at the 20-week scheduled trial visit and again at 1 year. Clinical outcomes were assessed after 20 weeks. Costs were reported in Danish krone and converted to US dollars in 2012 prices. The base-case results were expressed as cost per ITT population, cost PP population, cost PP population completion at end of trial week 12 and cost PP population completion at end of follow-up week 20. Sensitivity analyses on inclusion of estimated costs for administering biologic agents, use of actual IFX dosing and a reduction in the price of biologic agents of 3.5% and 7% were conducted to determine the robustness of the base-case results. At the 20-week follow-up, the costs were significantly lower in the algorithm group than in the IFX intensification group, and this differential was maintained throughout the 1-year study period. The base-case results, in terms of ITT for patients randomised to the algorithm group, showed costs of approximately US$11,900 for one patient at the 20-week follow-up, compared with US$22,100 at the 1-year follow-up. Among patients randomised to the IFX intensification group, the corresponding costs were US$17,200 and US$29,100, respectively. In terms of PP, among those randomised to the algorithm and the IFX groups, costs at the 20-week follow-up were approximately US$8700 and US$17,200, respectively, whereas at the 1-year follow-up the costs were approximately US$15,700 and US$29,100, respectively. The results from the sensitivity analyses were similar to the base-case results.

Velayos et al.[157]

Velayos *et al.*[157] 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 CD who become unresponsive to therapy with IFX. These authors used the algorithm proposed by Afif *et al.*[56] to form the basis of the testing-based strategy, whereas the empiric dose escalation strategy was informed by the consensus statement from the World Congress of Gastroenterology.[157] The study was conducted from the perspective of the third-party payer and a time horizon of 1 year, with a 4-week cycle length. Outcomes were reported as QALYs. QALYs gained were derived based on utility values obtained from the study undertaken by Gregor *et al.*[20] Briefly, utility scores for 180 individuals with CD were obtained using various elicitation methods (standard gamble, time trade-off or visual analogue scale). Gregor *et al.*[20] suggested that the standard gamble technique reflected the true value for health states related to patients with CD. Resource use and costs included the cost of interventions – IFX, ADA, certolizumab pegol, natalizumab and surgery – and the cost of diagnostics – anti-IFX antibody/serum IFX measurement, computerised tomography enterography and colonoscopy. Costs were expressed in US dollars, but the price year was not reported. The base-case results were expressed as an 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 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 fivefold more than in the base case. In addition, 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 than in the testing-based group. Furthermore, increasing the cost for testing 25-fold resulted in the testing-based strategy being more expensive than the empiric strategy. Results from the PSA showed that the testing-based strategy has approximately 69% probability of being cost-effective compared with empiric dose escalation at a willingness-to-pay of US$50,000 per QALY.

Comparison of the included studies

All four studies included in this review have been summarised in *Table 27*. Three studies were based on RCTs[73,123,124] and only one study[157] presented an economic model. Of the RCTs, two[123,124] were conducted in Denmark and one[73] was conducted in Belgium. All four studies[73,123,124,157] conducted cost-effectiveness analyses: Vande Casteele *et al.*[73] compared concentration-based with clinician-based dosing; Steenholdt *et al.*[123,124] compared IFX treatment failure using a treatment algorithm compared with IFX dose increasing; and Velayos *et al.*[157] compared a testing-based strategy with an empiric dose escalation strategy. All studies[73,123,124,157] clearly stated the type of assay used to analyse serum levels and antibodies to anti-TNF-αs. Two studies[123,124] used RIA in the base case, one study[73] used an assay developed in house and the remaining study[157] used a PROMETHEUS ELISA.

TABLE 27 Summary characteristics of the economic studies comparing ELISA kits

Study (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-TNF-αs	Outcome measure(s)	Model type	Health states	Utility values	Results (base-case and sensitivity analysis)
Responder									
Vande Casteele *et al.*, 2015, Belgium[73]	To determine whether or not continued concentration-based dosing is superior to clinically based dosing of IFX for maintaining remission in patients with moderate to severe CD and UC	RCT 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	EQ-5D – individual utility values not reported	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 3, where concentration-based dosing was less costly and less effective
LOR									
Steenholdt *et al.*, 2014, Denmark[123]	To determine whether or not individualised therapy is more cost-effective than dose intensification in patients with CD who lose response to anti-TNF-α treatment	RCT 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 and antibodies to anti-TNF-αs	RIA	Cost per ITT and cost PP population	Not applicable	Not applicable	Not applicable	Costs lower in the algorithm group than in the IFX intensification group in both the ITT population (mean difference per patient –€3141) and the PP population (mean difference per patient –€5116). ICERs not reported. Analysing serum samples using ELISA and HMSAs resulted in similar classification for the proposed algorithm in 72–78% of patients

continued

TABLE 27 Summary characteristics of the economic studies comparing ELISA kits (*continued*)

Study (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-TNF-αs	Outcome measure(s)	Model type	Health states	Utility values	Results (base-case and sensitivity analysis)
Steenholdt *et al.*, 2015, Denmark[124]	To assess the cost-effectiveness of individualised therapy as a long-term method compared with dose intensification in patients with CD-failing IFX	RCT with a cost-effectiveness analysis, perspective not reported, six Danish hospitals	Receive treatment based on serum concentrations of IFX and antibodies to IFX at the time of IFX treatment failure in accordance with the algorithm	RIA	Cost per ITT and cost PP population	Not applicable	Not applicable	Not applicable	Incremental costs in favour of the algorithm group. Results from the sensitivity analyses showed similar findings to the base-case results
Velayos *et al.*, 2013, USA[157]	To determine whether or not 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

The patient populations for three studies[73,123,124] included eligible patients with moderate to severe CD, whereas the study by Vande Casteele *et al.*[73] included patients with UC. The study perspective was not reported in two studies,[123,124] whereas the other two studies[73,157] conducted the analysis from a third-party payer perspective. The time horizon varied from 12 weeks to 1 year. Steenholdt *et al.*[123] based their analysis on a 12-week horizon, whereas the other three studies[73,124,157] used a 1-year time horizon to estimate the cost-effectiveness of the different strategies.

In two studies,[73,157] outcomes were reported as cost per QALYs gained. Vande Casteele *et al.*[73] used the EQ-5D measure to estimate QALYs, whereas Velayos *et al.*[157] did not explicitly report how the QALYs were estimated, except to say that they were obtained from a secondary source.[20] The two studies by Steenholdt *et al.*[123,124] reported outcomes in terms of cost per ITT and cost PP population.

Three studies[123,124,157] provided quite a comprehensive breakdown of resource use and costs, whereas the study by Vande Casteele *et al.*[73] did not elaborate on resource use, apart from the drug costs. Three studies[73,123,124] reported costs in 2012 prices, whereas Velayos *et al.*[157] did not report the price year explicitly; however, we assumed that costs are most likely to be in 2012 prices, as the study was published in 2013.

No studies conducted discounting for either costs or benefits as the time horizon for these studies was ≤ 1 year.

The results and conclusions reported differed between studies, Vande Casteele *et al.*[73] demonstrated that concentration-based dosing was slightly less effective and less costly than clinically based dosing, but overall differences were small. Steenholdt *et al.*[123] 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 1 year.[124] Velayos *et al.*[157] showed that the testing strategy was cheaper and more effective than the empiric strategy.

All four studies[73,123,124,157] conducted sensitivity analyses to deal with uncertainty around key parameters. The sensitivity analyses ranged from the most simplistic one-way sensitivity analyses[123,124] to the more sophisticated probabilistic analyses.[157]

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.[158] Using a 25-point CHEERS checklist, one article[73] 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[73,157] reported the viewpoint of the economic analysis. All studies described the comparators fully and reported the time horizon. However, because of the short time horizon, no studies conducted discounting of costs and benefits. In addition, the choice of health outcomes was well reported by all four studies;[73,123,124,157] however, only one study[73] reported how these health states were valued. Resource use and costs were well reported in three studies[123,124,157] apart from that by Vande Casteele *et al.*,[73] who described only the drug costs. The majority of the studies[73,123,124] conducted an economic analysis alongside a RCT, whereas one study[157] 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.[123,124,157]

From the studies identified, one[157] conducted a model-based economic analysis to determine whether or not 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 Philips's checklist.[149] In general, Velayos *et al.*[157] conformed to best practice for reporting model-based economic evaluations in terms of clearly stating the decision problem, adequately outlining the objectives, clearly stating the viewpoint of the analysis and describing the model structure, which represented the clinical pathway that patients with CD may follow. Time horizon and cycle length were stated and justified. In terms of the data required to

populate the model, Velayos *et al.*[157] adequately provided references, but they were unclear on the choices made between data sources and the quality of information used in the model. In addition, it was unclear whether or not any expert opinion had been used when choosing baseline information for the model. The other limitations identified were 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.

Discussion and conclusion

The evidence available on the cost-effectiveness of LISA-TRACKER ELISA kits, TNF-α-Blocker ELISA kits and Promonitor ELISA kits for measuring levels of TNF-α inhibitors and of anti-drug antibodies appears to be limited. We identified four cost-effectiveness analyses,[73,123,124,157] 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 to severe CD and were considered responders to IFX maintenance treatment. Studies ($n = 2$) mainly used RIA kits to analyse serum levels and antibodies to anti-TNF-αs. 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 *et al.*[157] 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 it varied between studies. In addition, the definition of patients with moderate to severe CD varied across studies. Second, owing to the small sample sizes, the studies 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-TNF-α levels and antibodies to anti-TNF-αs. Fourth, the method used to choose between data sources and the quality of information used in the model was unclear. Of the two studies[73,157] that reported their outcomes in terms of cost per QALY, only one[157] 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 for the 4-week cycle length and the lack of transparency on how transition probabilities were obtained and derived in the modelling study by Velayos *et al.*[157] and, in the case of the study conducted by Vande Casteele *et al.*,[73] the lack of detail on the resource use and costs.

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 *Developing the model structure*, we outline the development of economic models to determine the cost-effectiveness of various assays to inform on the treatment algorithm for patients who are considered responders and patients with LOR.

Considerations of using the former *Health Technology Assessment* model by Dretzke *et al.*[5] to inform the current model structure

The previous HTA model[5] used natural history data, which are now outdated. The current model for the standard care arm is restricted to starting with IFX (through lack of data for ADA) but otherwise adopts the general approach used in the HTA model but 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, as the latter requires considerable added complexity because 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 IFX.

Health economic methods

Objective

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

Standard care for patients during maintenance of disease (responders) is shown in *Figure 28*.

Standard care of patients with CD may vary across hospitals in the UK. Based on expert clinical input, we assumed that patients categorised as responders will continue to receive IFX maintenance therapy every 8 weeks until they lose response. Patients who lose response will receive an increased dose of IFX. Patients will either respond to this increased dose or continue to exhibit LOR, in which case they will receive another agent in addition to their current treatment. Patients who receive another agent may regain response or continue to exhibit LOR, in which case their anti-TNF-α treatment will be changed. Patients who do not respond a new a anti-TNF-α treatment will be considered for surgery. We have assumed that patients who respond to treatment will remain on that treatment until they lose response. We assume that patients who are in the post-surgery health state might receive various treatments (anti-TNF-α, a combination of anti-TNF-α and immunosuppressant or no treatment). Patients who experience LOR 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 IFX and follow the same treatment regime until they require repeat surgery.

Developing the model structure

We developed a Markov model using TreeAge Pro 2013 software program (TreeAge Software, Williamstown, MA, USA). The model was developed with clinical input, and represents the clinical pathway patients would undergo while being treated for moderate to severe CD. The illustrative model structures for responders and for those who lose response are shown in *Figures 29* and *30*, respectively. More detailed decision trees on the

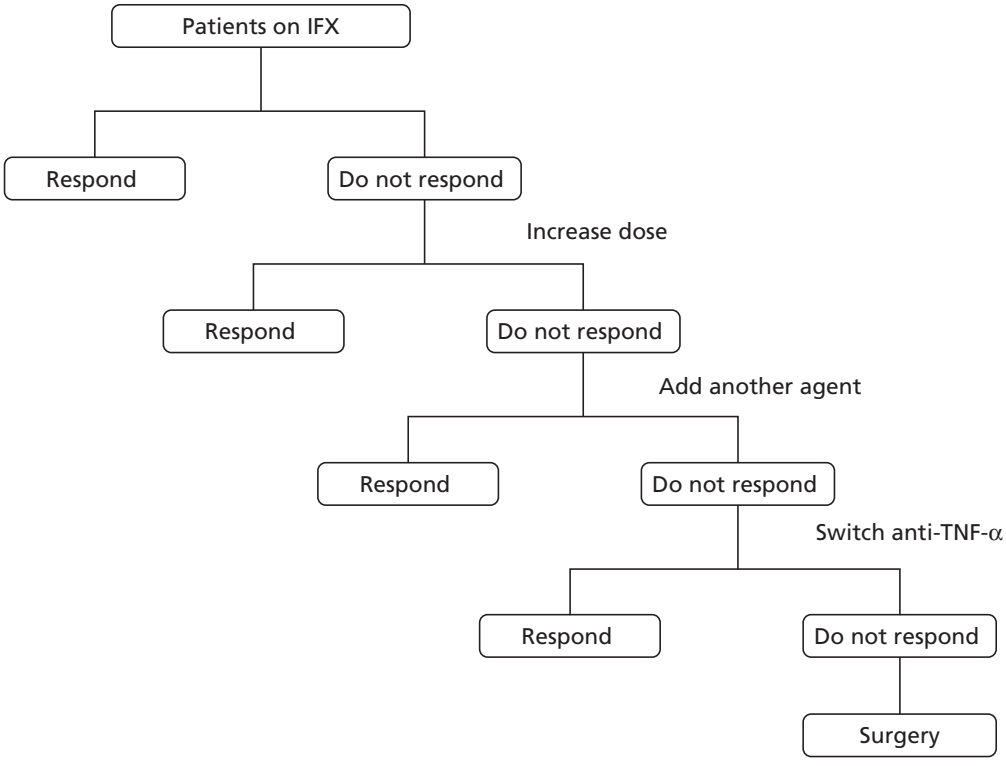


FIGURE 28 Standard care pathway for patients on maintenance therapy.

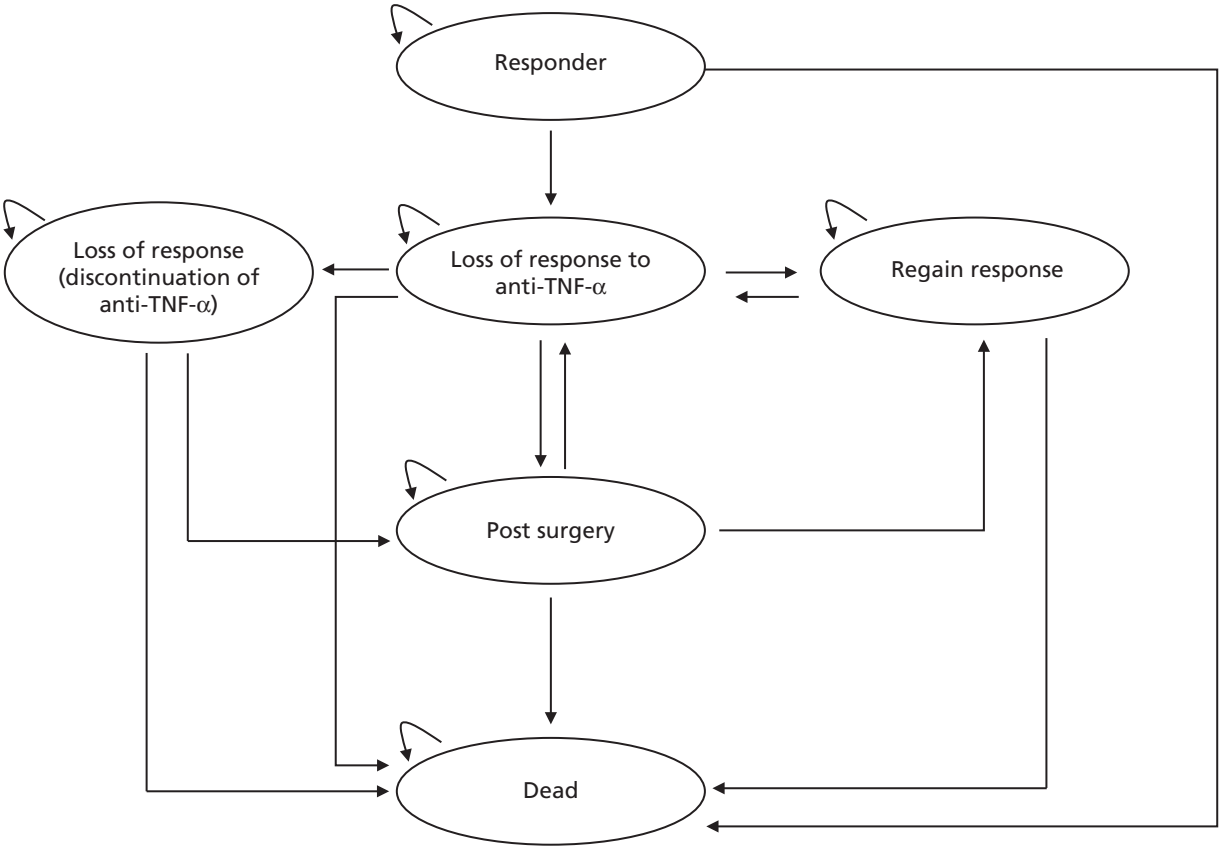


FIGURE 29 Illustrative structure for responders.

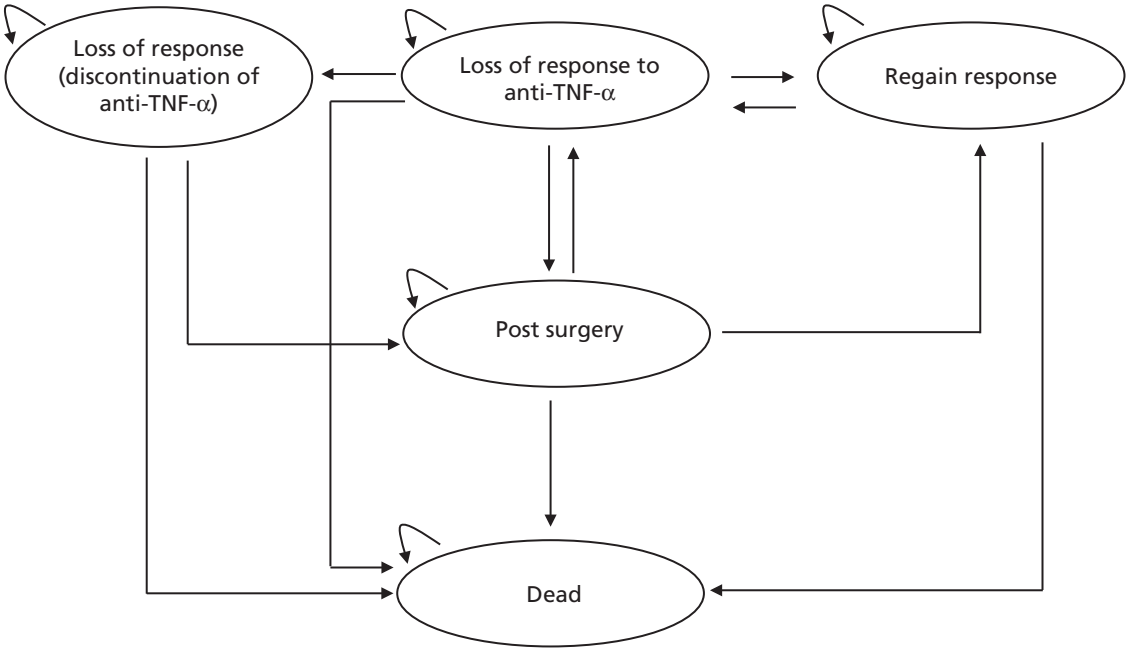


FIGURE 30 Illustrative structure for LOR.

patient pathways can be found in *Appendix 16*. In the models, we compared concurrent and reflex testing conducted every 3 months with standard care for responders and those who experience LOR:

- standard care
- concurrent testing – testing for TNF-α inhibitor levels and antibodies to TNF-α inhibitors
- reflex testing – testing for TNF-α inhibitor levels followed by testing of antibodies to TNF-α inhibitors depending on the drug test level.

The 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 have provided evidence of significant difference between testing and standard regimens in CD mortality.[5] Hence, we assumed a 10-year time horizon with 4-week cycle lengths to be appropriate to capture all benefits of testing and treatment.

Table 28 shows the health states required for the responder and LOR models.

In the following sections we discuss the testing strategies (concurrent and reflex testing) to be compared in both models (responders and patients with LOR).

Concurrent testing

In the concurrent testing strategy, patients undergo tests for serum anti-TNF-α levels and antibodies to anti-TNF-α simultaneously, and once the test results are available follow the proposed algorithm. Patients are classified, on the basis of their test results, into one of four groups: drug absent and antibodies present, drug absent and antibodies absent, drug present and antibodies present, or drug present and antibodies absent. Alternatively, patients may be categorised according to levels of drug regardless of antibody levels (e.g. as in the TAXIT trial[73]). Details of test results and proposed algorithms from Steenholdt *et al.*[123] for patients with LOR and from Vande Casteele *et al.*[73] for responders are presented in *Chapter 3*, *Objective B: description of algorithms prescribing patient management following test outcomes for drug and/or anti-drug antibody levels*.

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,[73] the only clinical study of an implemented and defined algorithm for responders:

1. if drug is absent and antibodies are present in a concentration > 8 mg/ml, patients receive a switch in TNF-α inhibitor

TABLE 28 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
LOR	Recurrence of active symptoms while on treatment with maintenance regimen, after having responded to treatment
LOR (no anti-TNF-α)[a]	Recurrence of active symptoms having discontinued anti-TNF-α treatment with maintenance regimen, 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 Patients who have discontinued anti-TNF-α treatment, but are receiving best supportive care.

2. if drug is absent and antibodies are present in a concentration < 8 mg/ml, patients receive an increased dosage of current treatment (i.e. IFX dose to 10 mg/kg every 8 weeks)
3. if the drug is present (there is no need to measure antibodies), and depending on the trough levels, patients 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 of these algorithm treatments, patients may remain responders, lose response (move to the LOR health state) or die.

Loss of response

After LOR to anti-TNF-α, testing and algorithm treatments are based on those used by Steenholdt *et al.*[123] in patients who lost response to anti-TNF-α (IFX); this is the only clinical study of implementation of an algorithm for patients with lost response:

1. Drug absent and antibodies present – patients would receive a switch in TNF-α inhibitor.
2. Drug absent and no antibodies – patients would receive an increase dosage of current treatment.
3. Drug present and antibodies present – we have assumed that patients will either have symptoms not requiring surgery and discontinue anti-TNF-α treatment or have active symptoms that require surgery. Patients in the former group would discontinue maintenance treatment and move to the LOR health state (discontinuation of anti-TNF-α) and receive best supportive care. Patients 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 patients with drug and antibodies present is identical to the pathway for patients with drug present without antibodies.

As a result of the treatment algorithm, patients may remain with LOR, regain response or could die.

Loss of response health state (discontinuation of anti-tumour necrosis factor alpha)

Patients who occupy this health state are those who have discontinued anti-TNF-α maintenance treatment and are receiving best supportive care. As in the LOR health state (see *Concurrent testing*), we have assumed that patients who remain in this health state have symptoms of CD that do not require surgery. Patients who develop active symptoms that require surgery move to the post-surgery health state or could die.

Regain 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 the TAXIT study[73] algorithm in *Chapter 3*, *Objective B: description of algorithms prescribing patient management following test outcomes for drug and/or anti-drug antibody levels*). As a result of the treatment algorithm, patients can remain in the regain response health state, lose response (move to the LOR health state) or could die.

Post-surgery (remission) health state

For patients who move to the post-surgery health state, treatment options are an anti-TNF-α, an immunosuppressant, a combination of an anti-TNF-α and an immunosuppressant or no treatment. Patients who are receiving an anti-TNF-α or a combination of anti-TNF-α and an immunosuppressant can regain response or lose response. For patients who regain response or who lose response, we have assumed that the pathway is similar to patients in the regain response health state (see *Concurrent testing*, *Regain response health state*) or the LOR health state (see *Concurrent testing*, *Loss of response*), respectively. Patients who are receiving immunosuppressants or no treatment could remain in the post-surgery health state until further surgery is required or die.

Reflex testing

In the reflex testing strategy, patients would receive a test to analyse serum anti-TNF-α levels. As a result of testing, two test outcomes are likely: drug absent or drug present. Based on the drug result, patients would undergo further testing for the presence or absence of antibodies. In this section we outline the health states and the pathways for patients undergoing reflex testing for both responder and LOR 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[73] trial for responders and the Steenholdt *et al.*[123] algorithm for patients with LOR using concurrent testing. Further details of test results and proposed algorithms are presented in *Chapter 3*, *Objective B: description of algorithms prescribing patient management following test outcomes for drug and/or anti-drug antibody levels*.

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 – patients with antibodies present would receive a switch in TNF-α inhibitor. Patients with no antibodies would receive an increase dosage of current treatment (i.e. IFX dose to 10 mg/kg every 8 weeks).
2. If drug is absent and there are no antibodies – patients would receive an increase dosage of current treatment (i.e. IFX dose to 5 mg/kg every 4 weeks).
3. If the drug is present and depending on the trough levels – patients would have 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 in the dosing intervals (if trough level is above the target range).

As a result of the treatment algorithm, patients could remain responders, lose response (move to the LOR health state) or could die.

Loss of response

1. Drug absent and antibodies present: patients would receive a switch in TNF-α inhibitor.
2. Drug absent and no antibodies: patients would receive an increased dosage of current treatment.
3. Drug present and antibodies present: we have assumed that some patients will have symptoms not requiring surgery and discontinue anti-TNF-α treatment or have active symptoms that require surgery. Patients in the former would discontinue maintenance treatment and move to the LOR health state (discontinuation of anti-TNF-α) and receive best supportive care. Patients who develop active symptoms that require surgery move to the post-surgery health state or could die.

As a result of the treatment algorithm, patients could remain in the LOR state, regain response or die.

Loss of response health state (discontinuation of anti-tumour necrosis factor alpha)

Patients who occupy this health state are those who have discontinued anti-TNF-α maintenance treatment and who are receiving best supportive care. As in the LOR health state (see *Reflex testing, Loss of response*), we have assumed that patients who remain in this health state have symptoms of CD that do not require surgery. Patients who develop active symptoms that require surgery move to the post-surgery health state or could die.

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-TNF-α. We have assumed that they would follow the same treatment algorithm for patients categorised as responders (see TAXIT study[73] algorithm in *Chapter 3*, *Objective B: description of algorithms prescribing patient management following test outcomes for drug and/or anti-drug antibody levels*). As a result of the treatment algorithm, patients can remain in the regain response health state, lose response (move to the LOR health state) or could die.

Post-surgery (remission) health state

For patients who move to the post-surgery health state, the treatment options are to receive an anti-TNF-α, immunosuppressant, a combination of anti-TNF-α and an immunosuppressant or no treatment. Patients who are receiving an anti-TNF-α or a combination of anti-TNF-α and an immunosuppressant can regain or lose response and follow the same pathways as outlined in *Reflex testing*, *Regain response health state* and *Reflex testing*, *Loss of response*. For patients who are receiving immunosuppressants or no treatment, the modelled options are to remain in the post-surgery health state until further surgery is required or to die.

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 CD.
2. Patients were assumed to have received intravenous infusions of 5 mg/kg IFX at weeks 0, 2 and 6. Here we assumed that patients weighed > 70 kg.
3. Patients who regained response have the same utility as those who are considered to be responders.
4. We have assumed that patients with CD are not at increased risk of dying from the disease, and that there is no difference in mortality between testing and standard care. However, in the case of patients who have undergone surgery, the model assumes an increased risk of 0.0015 of dying as a result of the procedure.
5. Treatment effects for patients receiving dose escalation (from 5 mg/kg to 10 mg/kg IFX) and a decreased interval (from 8 weeks to 6 weeks) are the same.
6. Patients who are categorised as responders 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 Juillerat *et al.*[159]
8. Patients who remain in the LOR health state (discontinuation of anti-TNF-α) have symptoms of CD that in time may require surgery. Patients will receive best supportive care until the development of active symptoms necessitating surgery.

Data required for the model

The model was populated with clinical information from the current clinical 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.

Proportions

The proportions of patients required to populate various model decision tree branches were obtained from secondary sources [e.g. management studies described in *Chapter 3, Objective C1: clinical studies evaluating drug monitoring for the management of Crohn's disease patients (management studies)*] and, when such data were lacking, from clinical input. Proportions that were estimated included partitioning of patients by presence or absence of IFX and of antibodies to IFX in responders and in those with LOR; partitioning of responders according to defined IFX trough levels; and partitioning by treatment options following surgery.

Table 29 summarises the partitioning of IFX responders based on the study of Imaeda *et al.*,[99] discussed in *Chapter 3*, *Analysis of correlation studies of tumour necrosis factor alpha/anti-drug antibodies level and response*, that used concurrent monitoring for the absence or presence of IFX and antibodies to IFX.

The proportions of IFX responders with various trough levels of IFX were based on Vande Casteele *et al.*[73] (discussed in *Chapter 3*, *Vande Casteele et al.: Trough level Adapted infliXImab Treatment study*[73]). These authors screened a cohort of patients with IBD who were receiving maintenance IFX treatment, and further

TABLE 29 Proportions derived based on concurrent testing of patients responding to IFX

Result	Proportion	Source
IFX absent and antibodies to IFX present	0.17241	Imaeda *et al.*[99]
IFX absent and antibodies to IFX absent	0.12069	
IFX present	0.7069	

categorised patients by drug concentration based on test result. Drug levels < 3 µg/ml were considered below the target range, levels between 3 and 7 µg/ml were considered within range and those > 7 µg/ml were above target range. *Table 30* shows the proportions of responders with different trough drug levels and the proportions of responders derived from this study.

The partitioning of IFX patients with LOR according to concurrent test monitoring of IFX and antibodies to IFX was based on information obtained from Steenholdt *et al.*[123] *Table 31* summarises these proportions.

Patients who have undergone surgery may receive post-operative treatment to maintain remission. These options include an anti-TNF-α, an immunosuppressant, a combination of an anti-TNF-α and an immunosuppressant or no treatment. *Table 32* shows these proportions based on the study of van der Have *et al.*[160]

TABLE 30 Proportions according to IFX trough levels of patients responding to IFX

Trough level	Threshold (µg/ml)	Proportion	Source
1	< 3	0.2310	Vande Casteele *et al.*[73]
2	3–7	0.4821	
3	> 7	0.2869	

TABLE 31 Proportions based on concurrent testing of patients with LOR to IFX

Result	Proportion	Source
IFX absent and antibodies to IFX present	0.1515	Steenholdt *et al.*[123]
IFX absent and antibodies to IFX absent	0.0303	
IFX present and antibodies to IFX present	0.0303	
IFX present and antibodies to IFX absent	0.7879	

TABLE 32 Treatment following surgery

Result	Proportion	Source
Anti-TNF-α	0.1250	van der Have *et al.*[160]
Immunosuppressant	0.5000	
Combination of anti-TNF-α and immunosuppressant	0.1250	
No treatment	0.2500	

Table 33 summarises the proportions of IFX responders based on the study by Imaeda *et al.*,[99] in which reflex testing was used to test for the absence or presence of IFX. In patients in whom IFX was present, we used the proportions according to IFX trough levels based on the Vande Casteele *et al.*[73] study, as shown in *Table 30*.

The partitioning of IFX patients with LOR according to reflex test monitoring of IFX was based on information obtained from Steenholdt *et al.*[123] *Table 34* summarises these proportions.

Time-to-event transition probabilities

Table 35 summarises the transition probabilities for time-to-event outcomes used in the models.

Transition probabilities from time-to-event studies

The transition probabilities provided in *Table 35* 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*.

Resource use and costs

The resource use and costs included were those directly incurred by the NHS. The costs of reagents for monitoring trough concentration of anti-TNFs and of antibody-measuring kits, treatment for CD and 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 36*. The majority of the cost information used in the analyses was obtained from secondary sources.

The costs of monitoring kits for IFX and for antibodies to IFX were obtained from Theradiag/Alpha Laboratories. In *Appendix 18*, we present a breakdown of the resource use and costs associated with monitoring kits for IFX and antibodies to IFX. In the models, we used a cost of £39.58 per person for concurrent testing for IFX and antibodies to IFX. In the case of reflex testing, we used a cost of £43.48 for patients in whom testing for IFX was followed by testing for antibodies because the results of the former were negative. For patients in whom a test for IFX was positive, no subsequent antibodies monitoring test was undertaken and, hence, we used a cost of £21.74.

The costs of maintenance treatment were obtained from the BNF (2013/14).[166] The costs of treatment associated with the induction phase (weeks 0–6) were not included. IFX treatment costs comprised its acquisition and administration costs. In the base case, we assumed that patients receiving maintenance therapy have received infusions of IFX 5 mg/kg every 8 weeks and that patients weighed, on average, 70 kg. For IFX maintenance, we derived a cost of £1966.41 (assuming four 100-mg vials at £419.62 plus administration costs of £287.93 per infusion) every 8 weeks. For patients switching to ADA, we derived a cost of £704.28 (2 × £352.14, assuming 40 mg of ADA is required every 2 weeks) per 4-week cycle. We assumed that patients would self-administer ADA; hence, no administration costs were included.

The estimated costs of management (outpatient visits to consultants and further investigations) associated with occupying all health states except the dead state were obtained from *NHS Reference Costs 2013 to 2014*[167] and in consultation with a clinical expert. These health-state costs include outpatient visits, colonoscopy and magnetic resonance imaging. In *Table 36*, we present the unit costs per year associated with each health state.

TABLE 33 Proportions derived based on reflex testing of patients responding to IFX

Result	Proportion	Source
IFX absent	0.2931	Imaeda *et al.*[99]
IFX present	0.7069	

TABLE 34 Proportions based on reflex testing of patients with for LOR to IFX

Result	Proportion	Source
IFX absent and antibodies present	0.2029	Steenholdt *et al.*[123]
IFX and antibodies absent	0.0435	
IFX present	0.7536	

TABLE 35 Summary of parametric models used for estimating transition probabilities for time-to-event outcomes

Transition	Transition probabilities (95% CI)	Source	Comments/assumptions
Standard care			
1. IFX maintenance to LOR	0.008075 (0.007179 to 0.009084)	Juillerat *et al.*[159]	Observed data to 10 years, patients with CD only
2. IFX maintenance to LOR after IFX escalation	0.017415 (0.014443 to 0.020991)	Ma *et al.*[161]	Time to LOR after dose escalation. Observed data to > 6 years
3. ADA after IFX, failure to LOR	0.058553 (0.052622 to 0.065129)	Sandborn *et al.*;[162] Karmiris *et al.*[48]	RCT of ADA for CD (Sandborn *et al.*[162]) and prospective study of 168 patients (Karmiris *et al.*[48])
All			
4. Time to surgery	0.002591 (0.002279 to 0.002945)	Nguyen *et al.*[163]	Large study with 7 years of data; surgery incidence similar to small UK study
5. Time to recurrent surgery	0.003122 (0.002398 to 0.004065)	Nguyen *et al.*[163]	As 4 above
6. Time to post-surgical relapse on no therapy	0.049792 (0.042951 to 0.057538)	Gordon *et al.*[164]	Limited data
7. Time to post-surgical relapse on immunosuppressant	0.029714 (0.024975 to 0.035022)	Gordon *et al.*[164]	Limited data
8. Time to post-surgical relapse on anti-TNF-α	0.020784 (0.0143 to 0.030162)	Baert *et al.*[77]	Limited data, assumes applicability of study population
9. Time to post-surgical relapse on anti-TNF-α and immunosuppressant	As 8 above	Lack of data	Assumed as anti-TNF-α alone
Intervention arm: test algorithm strategy			
10. IFX maintenance to LOR (dose-escalation group)	As 1 above	Juillerat *et al.*[159]	No evidence for advantage relative to standard care
11. IFX maintenance to LOR (dose-unchanged group)	As 1 above	Juillerat *et al.*[159]	No evidence for difference according to trough group
12. IFX maintenance to LOR (dose-decreased group)	As 1 above	Juillerat *et al.*[159]	No evidence for difference according to trough group
13. Regained response on ADA to LOR (group 1, IFX negative/ antibodies to IFX positive)	As 3 above	As 3 above	As 3 above
14. Regained response on intensified IFX to LOR (group 2, IFX negative/antibodies to IFX negative)	As 2 above	As 2 above	As 2 above
15. Regained response on un-prescribed treatment for LOR (group 3 or 4 IFX positive/antibodies to IFX positive or negative) to LOR	0.086173 (0.04727 to 0.140943)	Rutgeerts *et al.*[165]	Constant hazard for loss of regained response based on placebo group from the 1999 Rutgeerts *et al.*[165] RCT

TABLE 36 Resource use and costs and utilities used in the models

Variable	Base-case value	Range for sensitivity analysis	Distribution	Reference(s)
Resource use and costs				
Monitoring IFX	£21.74		Fixed	NICE (Sarah Bond, NICE, 2014, personal communication)
Monitoring antibodies to IFX (reflex testing)	£41.98		Fixed	
Monitoring IFX and antibodies to IFX (concurrent testing)	£38.83		Fixed	
Maintenance IFX[a]	£1966.41		Fixed	BNF 2013/14[166]
Maintenance ADA[b]	£352.14		Fixed	BNF 2013/14[166]
Azathioprine[c]	£8.40		Fixed	BNF 2013/14[166] and expert opinion
Mercaptopurine[d]	£100.94		Fixed	
Prednisolone[e]	£14.25		Fixed	
Nutritional therapy (Modulen)[f]	£15.06		Fixed	
Laparoscopic ileocolic resection[g]	£6908		Fixed	*NHS Reference Costs 2013 to 2014*[167] and expert opinion
Responder[h]	£725.69		Fixed	
LOR[h]	£1241.38		Fixed	
Regain response[h]	£725.69		Fixed	
Post surgery[h]	£790.69		Fixed	
Utility values				
Responder	0.77	0.70–0.84	Beta(117.04, 34.96)	Velayos *et al.*[157]
LOR	0.62	0.59–0.66	Beta(465, 750)	Derived from Gregor *et al.*[20]
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.*[168]
Post surgery	0.86	0.82–0.90	Beta(301, 49)	Velayos *et al.*[157]
Dead	0		Fixed	By definition
Other				
Mortality (age-specific death rates)	Life tables		Fixed	Office for National Statistics, 2014[169]
Mortality associated with surgical procedure	0.0015		Fixed	Velayos *et al.*[157]
Discount rate per annum (costs and QALYs)	3.5%		Fixed	

a Patients receiving 5 mg/kg IFX during maintenance therapy every 8 weeks. See *Appendix 18*, *Table 63* for details.
b Patients receiving of ADA during maintenance therapy every 40 mg/kg every 2 weeks. See *Appendix 18*, *Table 63* for details.
c Cost based on a 50-mg (56-tablet) pack and recommended dosage of 2.5 mg/kg per day.
d Cost based on a 50-mg (25-tablet) pack and recommended dosage of 1.25 mg/kg per day.
e Cost based on a 20-mg/100-ml single dose and recommended dosage of 30 mg in week 1, then 5 mg each week for the next 3 weeks.
f Cost based on 400 g of Modulen®, IBD (Nestlé, York, UK).
g Patients undergoing a laparoscopic ileocolic procedure. Detail resources used are provided in *Appendix 18*, *Table 63*.
h Unit cost (per year) associated with occupying this health state. Please see *Appendix 18*, *Table 64* for further details on resource use.

Costs obtained from published sources were adjusted to 2013/14 prices using the Hospital and Community Health Service Pay and Price Index[170] and future costs were discounted at a rate of 3.5% per annum, as recommended by NICE.

Outcomes

The outcome measure used in our analyses was the number of QALYs gained. To calculate the estimated QALYs associated with the health states described in the model, we obtained utility weights from published literature[157] reported in our review of cost-effectiveness, and combined these utility values with data on life expectancy from the Office for National Statistics.[169] Utility values reported in Velayos *et al.*[157] were obtained from the study undertaken by Gregor *et al.*,[20] who compared various elicitation techniques (standard gamble, time trade-off and visual analogue scale) in 180 consecutive CD patients. These authors suggested that the standard gamble technique reflected the true value for health states related to patients with CD, and these values may be the most appropriate for an economic analysis. *Table 36* shows the utility weights used in the model. In each cycle of the model, patients will incur a utility pay-off depending on the health state being occupied. In the model, we applied a utility weight of 0.77 for individuals categorised as responders 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.

Analysis

The model was constructed to assess the cost-effectiveness of concurrent testing, reflex testing and no testing of blood levels of anti-TNF-α agents and of antibodies to these agents in patients with severe CD. The model estimated the mean costs and effects associated with each testing strategy, and was simulated over a 10-year time horizon with 4-weekly cycle lengths. The starting point for the responder population was a hypothetical cohort of patients aged 30 years whose disease responds to a maintenance course of TNF-α inhibitor therapy. This age was chosen because the onset for CD is likely to occur from the late teens to age 30 years.[171] We define a maintenance course as 5 mg/kg intravenous IFX every 8 weeks. The analysis was undertaken from a NHS perspective in an outpatient care setting, and outcomes were reported as ICERs, expressed in terms of cost per cost per QALY gained.

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 LOR models
2. estimate the mean costs and effects associated with each strategy using a 1-year time horizon with 4-week cycle lengths
3. in the responders model – three possible modes of one-off testing:

    i. one-off testing at 3 months followed by yearly retesting
    ii. one-off testing at 3 months and one retest for those who regained response
    iii. one-off testing at 3 months and no retesting for responders/regained response.

4. in the LOR model – 3-monthly testing for patients with LOR; no testing for patients who have regained response
5. no regain of response following best supportive care (responders)
6. no regain of response following best supportive care (LOR).

Probabilistic sensitivity analyses

Probabilistic sensitivity analyses 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 PSA, 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 36*. We have calculated probabilities that each strategy is the most cost-effective, at a willingness to pay of £20,000/QALY.

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 adults aged 30 years with severe CD, the results of concurrent testing, reflex testing and no testing (standard practice), in terms of QALYs gained, are presented in *Table 37*. At the 10-year time horizon, in the standard practice cohort, reflex testing resulted in a mean gain of 6.2761 QALYs, with a corresponding mean cost of £138,700. The concurrent testing cohort gained 6.2637 QALYs, with a mean cost of £139,800. The no-testing cohort gained 6.5084 QALYs, with a mean cost of £150,500. These results show that the reflex testing strategy was less costly and produced more QALYs than the concurrent testing strategy, hence dominating the concurrent testing. The no-testing strategy was the most costly and effective strategy with an ICER of approximately £50,800 per QALY.

Table 38 presents the results of the analyses based on an outcome of cost per QALY in the LOR model with testing (concurrent and reflex) undertaken every 3 months. The results show that at the 10-year time horizon the concurrent testing strategy resulted in 6.1807 QALYs, with a corresponding mean cost of approximately £129,400. Reflex testing produced marginally more QALYs at an incremental cost of approximately £94,700 per QALY. The no-testing strategy has a mean cost of approximately £215,800 and costs approximately £84,800 more than reflex testing, with a total effectiveness of 6.4961 QALYs. This result indicates that, in this LOR 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 with reflex testing costs £284,100 in a cohort of patients with LOR.)

Results of sensitivity analyses

We undertook a number of one-way sensitivity analyses to determine the impact on the results of changing key model input parameters (*Table 39*).

First, in the responder model, we changed the testing strategy from 3 months to annual testing. The results showed that concurrent testing was the cheapest strategy, with a mean cost of approximately £114,000 and generating 6.2201 QALYs. In the reflex testing arm, this strategy was marginally more expensive and provided more QALYs, with an ICER of approximately £12,500 per QALY. As expected,

TABLE 37 Base-case results for the analysis cost per QALY (2013/14 prices)

Strategy	Mean cost per strategy (£)	Difference in costs (£)	Effectiveness (QALYs)	Incremental QALYs	ICER (£)
Reflex testing	138,700	–	6.2761	–	–
Concurrent testing	139,800	1100	6.2637	–0.0124	Dominated
No testing	150,500	11,800	6.5084	0.2323	50,800

TABLE 38 Base-case results for the analysis cost per QALY (2013/14 prices) (LOR model)

Strategy	Mean cost per strategy (£)	Difference in costs (£)	Effectiveness (QALYs)	Incremental QALYs	ICER (£)
Concurrent testing	129,400	–	6.1807	–	–
Reflex testing	131,000	1600	6.1976	0.0169	94,700
No testing	215,800	84,800	6.4961	0.2985	284,100

TABLE 39 Univariate sensitivity analyses

Parameter varied	Mean cost per strategy (£)	Difference in costs (£)	Effectiveness (QALYs)	Incremental QALYs	ICER (£)
Base case					
Reflex testing	138,700	–	6.2761	–	–
Concurrent testing	139,800	1100	6.2637	–0.0124	Dominated
No testing	150,500	11,800	6.5084	0.2323	50,800
Annual testing in responder model					
Concurrent testing	114,000	–	6.2201	–	–
Reflex testing	114,100	100	6.2281	0.0080	12,500
No testing	150,500	36,400	6.5084	0.2803	129,900
Annual testing in LOR model					
Concurrent testing	106,900	–	6.1406	–	–
Reflex testing	108,100	1200	6.1532	0.0126	95,200
No testing	215,800	107,700	6.4961	0.3429	314,100
1-year time horizon in responder model					
No testing	14,900	–	0.7686	–	–
Concurrent testing	18,500	3600	0.7549	–0.0137	Dominated
Reflex testing	19,200	4300	0.7543	–0.0143	Dominated
1-year time horizon in LOR model					
Concurrent testing	12,000	–	0.6870	–	–
Reflex testing	12,500	500	0.6915	0.0045	111,100
No testing	23,500	11,000	0.7560	0.0645	170,500
One-off testing at 3 months followed by yearly retesting					
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
One-off testing at 3 months and one retest for those who regained 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
One-off testing at 3 months and no retesting for responders/regained response					
Concurrent testing	102,000	–	6.2255	–	–
Reflex testing	102,900	900	6.2390	0.0135	66,700
No testing	150,500	47,600	6.5084	0.2694	176,700
In the LOR model: 3-monthly testing for patients with LOR; no testing for patients who have regained response					
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

continued

TABLE 39 Univariate sensitivity analyses (*continued*)

Parameter varied	Mean cost per strategy (£)	Difference in costs (£)	Effectiveness (QALYs)	Incremental QALYs	ICER (£)
No regain of response following best supportive care (responders)					
Reflex testing	87,900	–	5.7853	–	–
Concurrent testing	89,900	2000	5.7838	–0.0015	Dominated
No testing	150,500	62,600	6.5084	0.7231	86,600
No regain of response following best supportive care (LOR)					
Concurrent testing	54,000	–	5.4649	–	–
Reflex testing	57,700	3700	5.4992	0.0343	107,900
No testing	215,700	158,000	6.4961	0.9969	158,500

the mean cost and effectiveness of the no-testing strategy remained unchanged. A no-testing strategy compared with a reflex testing strategy had a reported ICER of £129,900 per QALY.

Second, changing the 3-month testing to annual testing in the LOR 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 1 year with 3-month cycles, we found that the no-testing strategy dominated both testing strategies. In the LOR model, the no-testing strategy was the most expensive and most effective strategy, with a mean cost of approximately £23,500 and corresponding QALYs of 0.7560.

Finally, changing the testing regime in the responder model to one-off testing at 3 months followed by yearly testing (for those responding to treatment), at 3 months followed by one retest for those who regained response, and at 3 months and no retesting for responders or those who regained response showed that no testing was more expensive than testing and was more effective. Similar results were shown obtained for the one-off testing in the LOR model, and assuming that patients could not regain response following best supportive care.

In further sensitivity analyses, we varied key model input parameters to determine which inputs influence the ICER. *Figures 31* and *32* show the percentage change in the cost per QALY as a result of increasing or decreasing these inputs by 10% of the base-case value. The results showed that the model is stable to most of these changes, but is sensitive to a 10% increase in the utility value for patients who regain response in both reflex and concurrent testing.

Results of probabilistic sensitivity analysis and cost-effectiveness acceptability curves

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 no testing. Scatterplots of the 10,000 runs of the Monte Carlo simulations 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 in *Figure 34*. Cost-effectiveness acceptability curves give the probability that a strategy is cost-effective at various values of willingness to pay for a QALY. The willingness-to-pay threshold used by NICE is between £20,000 and £30,000 per QALY. From the information and assumptions used in the model, the results in *Figure 34* show that, at £20,000 per QALY, the no-testing strategy is 92% likely to be cost-effective compared with concurrent and reflex testing.

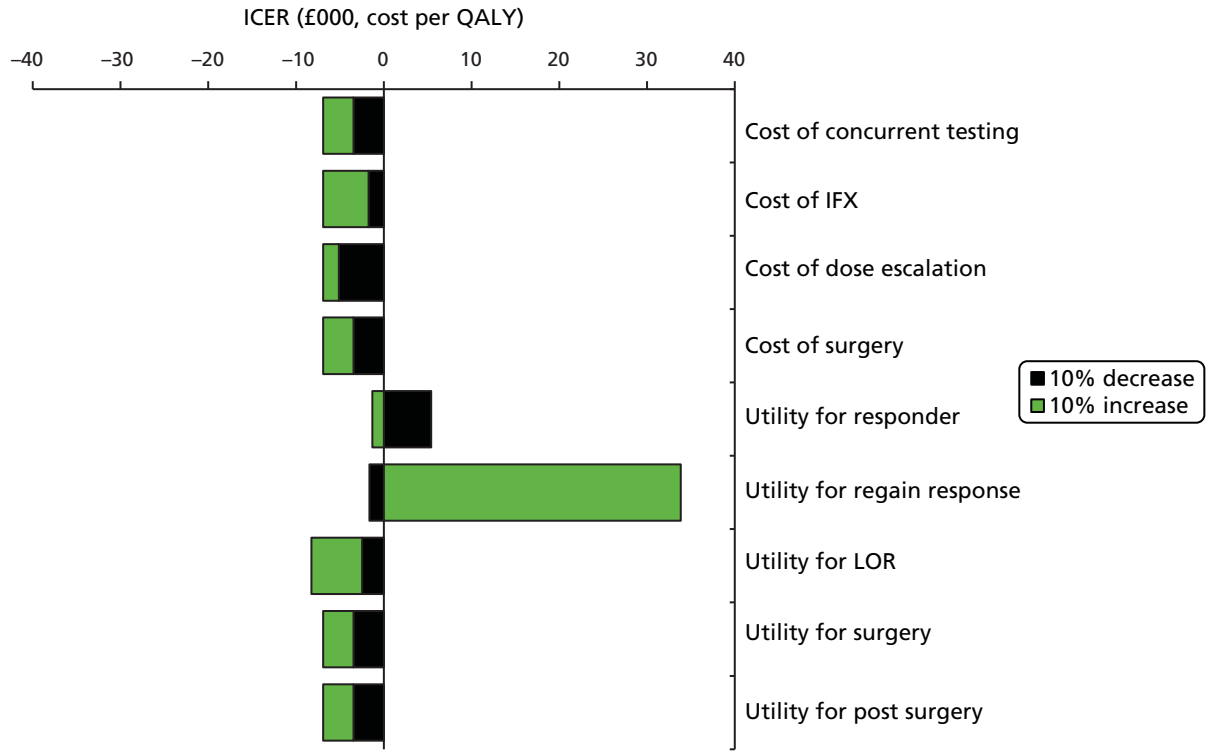


FIGURE 31 Tornado diagram comparing no testing with reflex testing.

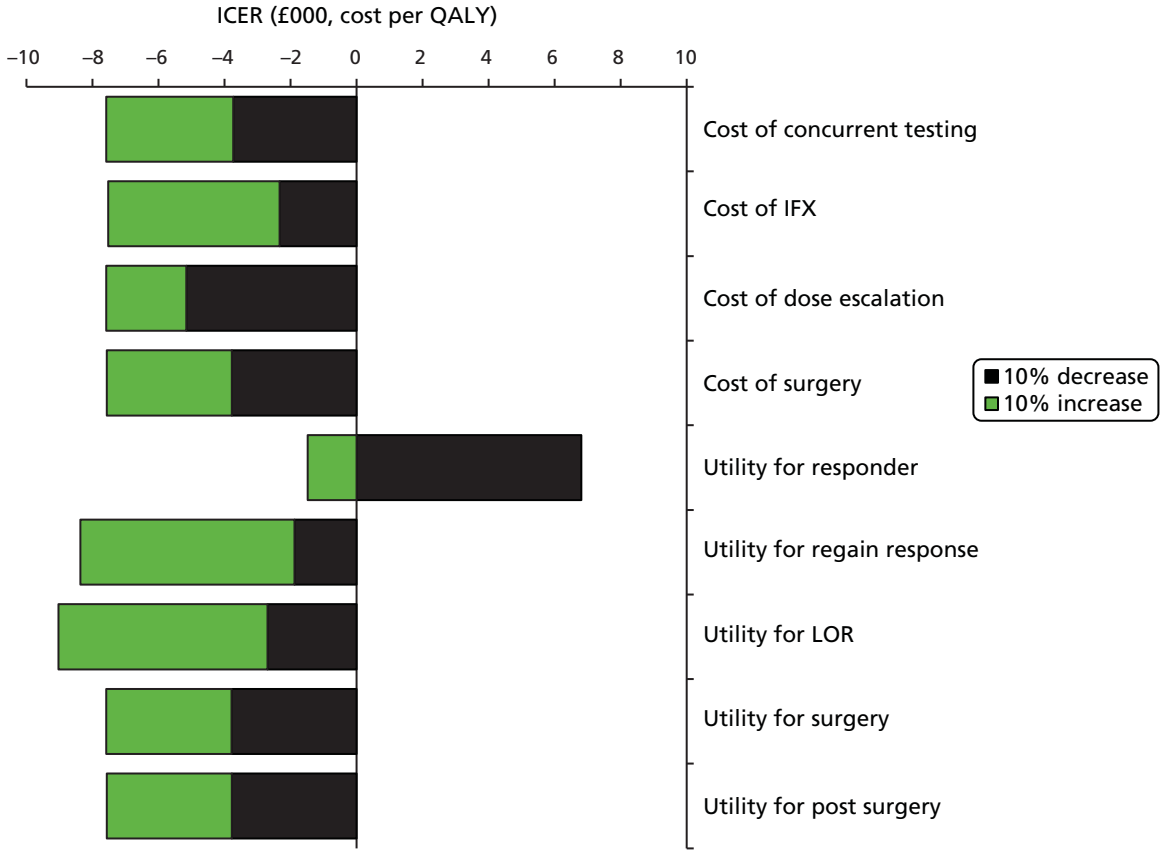


FIGURE 32 Tornado diagram comparing no testing with concurrent testing.

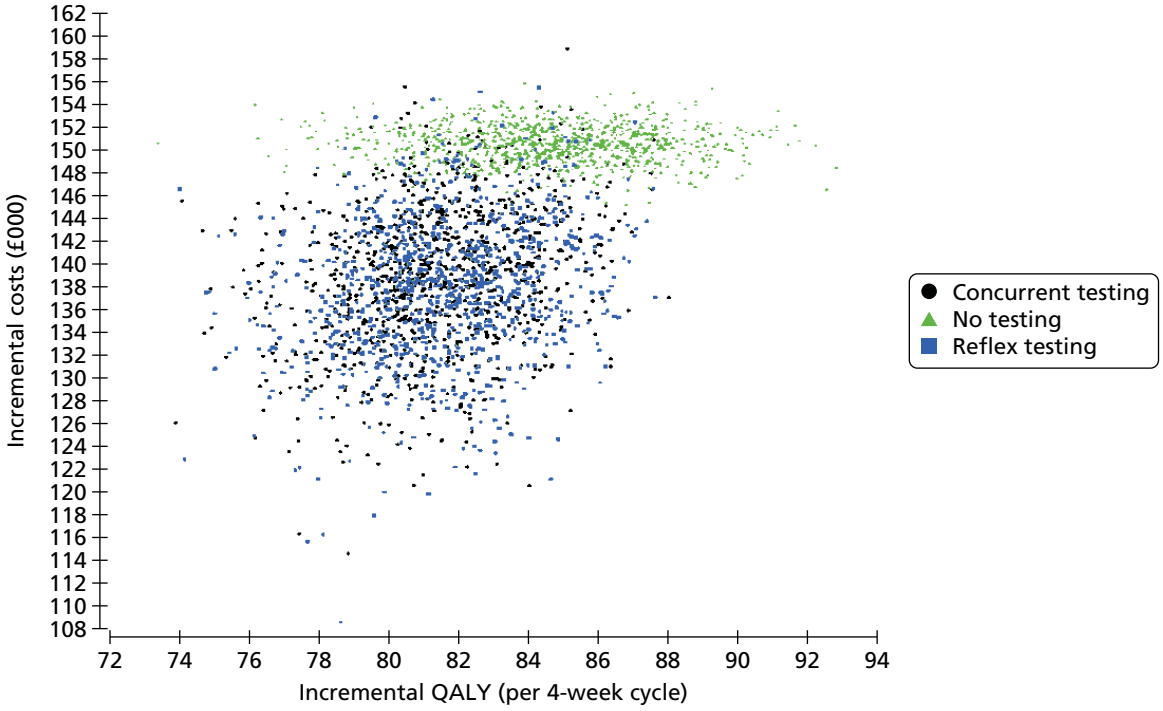


FIGURE 33 Probabilistic sensitivity analysis results for concurrent and reflex testing and no testing. Scatterplot using distributions around model input parameters.

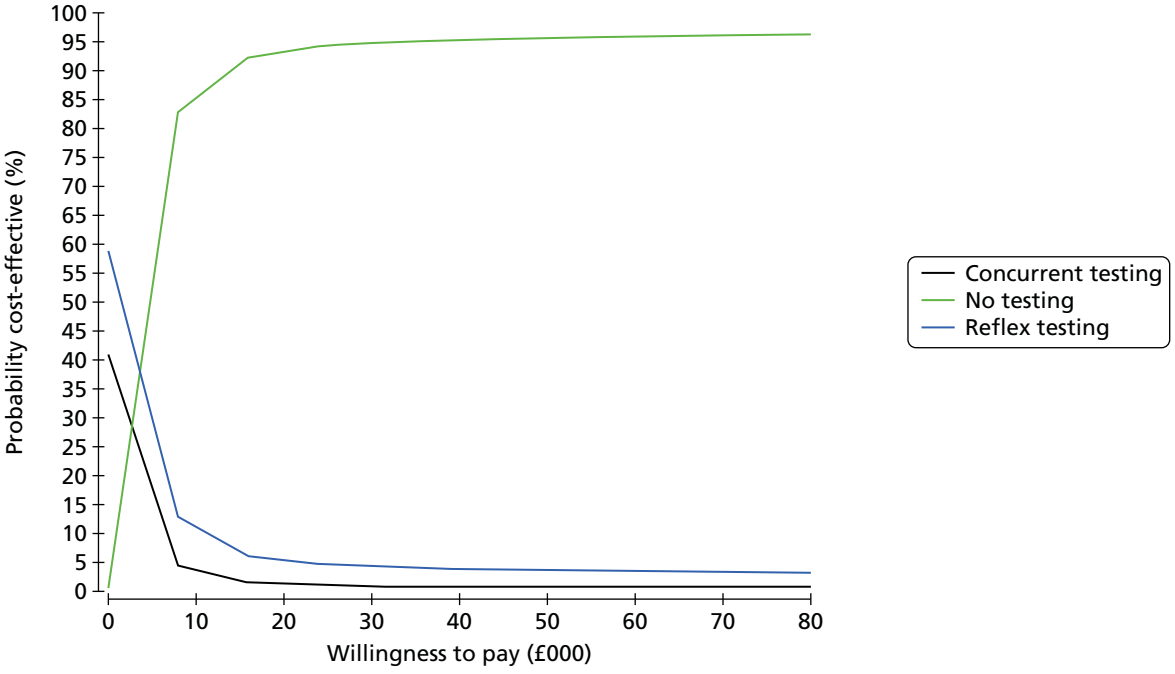


FIGURE 34 Cost-effectiveness acceptability curve using distributions around outcomes.

Summary of cost-effectiveness

In summary, a de novo Markov model was built in TreeAge Pro 2013 to evaluate the cost-effectiveness of test algorithm-based treatment strategies compared with standard care. Two test strategies were assessed: concurrent testing of drugs and of antibodies to the drugs, and sequential or reflex testing (i.e. a drug test first, and then an anti-drug antibody test depending on the results 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 10-year time horizon and adopted NHS and PSS perspectives. Costs were adjusted to 2013/14 prices and annually discounted at 3.5%. The starting point was a hypothetical cohort of patients aged 30 years. Outcomes are reported as ICERs, expressed in terms of cost per 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 4-week cycle lengths, altered transition probabilities for LOR, altering the proportions of patients in the different testing results categories and an arbitrary 10% change in the main input parameters. PSA 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 patients with CD and the treatment pathways dictated by the algorithms. Populating the model with information from the two management studies was problematic because the studies were of small size and short duration, and reported outcomes that were not directly relevant to an economic model; in addition, one study lacked an appropriate standard care arm for economic modelling and neither reported 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, we conclude that QALY gains are likely to very similar in both arms (concurrent/reflex) whereas the cost of the testing strategy (concurrent/reflex) appears to be more than twice the cost of standard care.

Chapter 5 Discussion

It has been proposed that measuring levels of anti-TNF-α drug and antibodies raised against the drug during an immune response can aid the management of patients with CD who are on maintenance therapy. This implies that patients have responded to induction therapy of anti-TNF-α 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. In this assessment we investigated to what extent drug and anti-drug antibody measurements obtained using three different types 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, TNF-α-Blocker ELISA kits and Promonitor ELISA kits.

Decision problem and objectives

Our overall objective was to undertake a clinical effectiveness and cost-effectiveness analysis of testing anti-TNF-α levels and antibodies to anti-TNF-α in patients with CD who either are responding to anti-TNF-α 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-TNF-α. We aimed to systematically review the evidence on the clinical effectiveness of monitoring anti-TNF-α drugs and their antibodies in responders and patients with LOR when ELISA 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 ELISAs and develop a cost-effectiveness model.

Summary of methods and findings

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 focus of the studies 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 administering dichotomised testing for drug and anti-drug antibody levels. These are drug present and antibodies absent; drug absent and antibodies absent; drug present and antibodies present; and drug absent and 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-TNF-α 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 2428 records, of which 62 studies were included; an additional six studies were identified through other sources, making a total of 68 included studies. Of these studies, three were management studies measuring levels of IFX using RIA in patients with LOR, a commercial ELISA and HMSA in responders and an in-house ELISA in responders. 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 IFX 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-TNF-α therapy for CD. Furthermore, the cut-off points of drug and anti-drug antibody levels used to determine therapeutic levels were not comparable in these studies. The sample collection times and analysis times were different, as were the definitions used for clinical response, remission, progression and relapse. Steenholdt *et al.*[123,124] was the only RCT that compared drug monitoring and treatment change according to an algorithm with standard care (dose intensification) in patients with LOR. The primary outcome of this study was cost, and the authors concluded that combined measurement of drug and anti-drug antibodies reduces average treatment costs per patients compared with routine IFX 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 RCT (TAXIT)[73] indicated no benefit in clinical remission from test-directed dose optimisation (RR after compared with before optimisation 1.053, 95% CI 0.936 to 1.186), and no difference at 1 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.*[128] reported superior retention in IFX treatment, implying clinical benefit from test monitoring, but 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 a gold standard; this limits the comparative data that are useful for a linked evidence approach using concordance data and/or Cohen's kappa. The only direct link that was found was a study[67] comparing the performance of LISA-TRACKER assays with that of the Leuven in-house ELISA used in the TAXIT trial investigating the clinical effectiveness of dose adjustment in responders to IFX.[73] It reported disagreement for IFX level measurements in at least 11 out of 58 samples and for anti-drug antibodies to IFX in at least 3 out of 62 samples; the results for the remainder were unclear.[67] Overall, there were no concordance data linking any of the index tests to any of the comparator tests at a clinically meaningful threshold. From these 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; PPV and NPV estimates derived from meta-analyses indicated that between 20% and 30% of positive and negative test results are likely to be inaccurate.

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 TNF-α inhibitors and of anti-drug antibodies in patients with severe CD. Of these, one study[157] used a decision-analytic model to assess the cost-effectiveness of using test-based strategy compared with dose escalation in patients who have lost responsiveness to IFX. 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, the title of one study[73] failed to state that an economic evaluation was conducted. Second, resource use and costs reported in Vande Casteele *et al.*[73] were not comprehensive, including only costs related to drug treatment. In the case of the study[157] that conducted a model-based economic evaluation, the 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. However, the study had limitations: first, that there was lack of clarity regarding the methods used to extrapolate short-term results into final outcomes; and, second, it was unclear if the model was developed with any clinical input. Finally, these authors did not undertake half-cycle correction or 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 compared with standard care. Two test strategies were assessed: (1) concurrent testing of drugs and antibodies to the drugs, and (2) sequential or reflex testing (i.e. a drug test first, and then an anti-drug antibody test depending on the results 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 4-week cycle and a 10-year time horizon, and adopted NHS and PSS perspectives. Costs were adjusted to 2013/14 prices and annually discounted at 3.5%. The starting point was a hypothetical cohort of patients aged 30 years. Outcomes are reported as ICERs, expressed in terms of cost per 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 4-week cycle lengths, altered transition probabilities for LOR and altering the proportions of patients in the different testing results categories. PSA 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 QALY showed that reflex testing dominated the concurrent testing strategy at the 10-year time horizon. Standard practice was more costly and produced more QALYs than the other strategies. No testing resulted in 6.5084 QALYs, with a corresponding mean cost of £150,500. Reflex testing resulted in 6.2761 QALYs, with a mean cost of £138,700. Concurrent testing generated 6.2637 QALYs, with a mean cost of £139,800.

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.

The no-testing strategy dominated both of the other testing strategies when making changes to the model time horizon in the responder model.

Varying key model input parameters by an arbitrary 10% showed that the no-testing strategy continued to dominate the testing strategies in most cases.

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 probable equivalence of assay methods specified as interventions

compared with 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 studies relating test results to clinical state of patients) to include correlation studies. Correlation studies that 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 by individual were used to provide a pooled estimate for the probability of returning a specified test result after trough anti-TNF-α testing (useful for estimating reflex strategy test result probabilities). This information was used when 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 on the reasonable assumption that the commercial ELISA kits with RIA and the Leuven in-house ELISA are equivalent. 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 these data to undertake a sensitivity analysis. Unfortunately, the study by Steenholdt *et al.*[122] 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 that we have presented may not be reflective of the cost-effectiveness of LISA-TRACKER ELISA kits, TNF-α-Blocker ELISA kits and Promonitor ELISA kits.

Furthermore, most of the available evidence about tests and algorithms for prescribing treatment in accordance with 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. Although a number of algorithms have been suggested, only two have been tested in RCTs. Steenholdt *et al.*[123] tested a prescriptive algorithm for the clinical management of patients with LOR following combined drug and anti-drug antibody testing. This is representative of our concurrent testing strategy. However, no RCT presenting and evaluating 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 and antibody present; drug absent and antibody absent; drug present). As the drug-positive and anti-drug antibodies-positive group is treated identically to the drug-positive and anti-drug antibodies-negative group in the RCT by Steenholdt *et al.*[123] (using combined drug and anti-drug antibodies testing), the model structure of the strategies of reflex testing and concurrent testing are identical.

It seems unlikely that reflex testing could be clinically feasible, as the delay in treatment because of conditional testing of antibodies on absence of drug could be up to 4 weeks. The evidence for responders comes from Vande Casteele *et al.*[73] 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 or 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 currently relevant ones are those that have been implemented prospectively with patients and compared with 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 reasons (generally > 50%) for initiating testing are LOR, partial response or a flare.[56,95,172–174] Other reasons for testing include routine monitoring and adverse events. Although this finding is reasonably constant across studies, the consideration of test results in clinical decision-making in the absence of a prescriptive treatment algorithm varied widely. Although one study reported that drug levels but not the presence of antibodies influenced treatment decisions,[95] another study reported that more patients with positive anti-drug antibody tests than with negative ones received a treatment change.[175] There is also evidence that test results affected treatment decisions in only 73% of patients tested[56] and that an appropriate treatment change of switching anti-TNF-α agents in patients with a positive anti-drug antibody test and dose increase in patients with subtherapeutic IFX levels occurred in only 57% and 21% of patients, respectively.[173]

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.*[123] revealed, however, that the algorithm for 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-TNF-α 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 patients with CD 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-TNF-α treatment can be successful only if multiple factors rather than just a single test result are considered in the management of patients.[30]

This presented a considerable challenge for the modelling. Although 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 be prescriptive, restricting treatment options to one or two possible treatments with little consideration 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-TNF-α 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 could still be receiving this treatment option. Approval of vedolizumab for the treatment of CD in the UK would therefore change the model outcome.

When considering using ELISAs or other assays to predict clinical outcomes (e.g. response) of patients with CD on anti-TNF-α 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 if prescribed by an algorithm considering only test outcomes. Unfortunately, because of a 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 reference standard for the assessment of response in CD 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 or not anti-drug antibodies are measured, and the type of clinical marker used to evaluate response (CRP level, serum albumin, FC or endoscopic scoring of mucosal healing). Although we are aware that the various assays do not measure the precise levels of drug and anti-drug antibody because of the difficulties of interference and drug–antibody complexes, the uncertainties around the test threshold and definitions of response question the value of knowing precise measurements of drug and anti-drug antibody levels in patients.

The included studies recruited only adult patients with CD, and the applicability of outcomes to children therefore remains unknown. However, in an abstract, Turon *et al.*[176] reported that measurements of IFX 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 clinical effectiveness of monitoring patients on ADA. Although both ADA and IFX are anti-TNF-α agents, they are different molecules and 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 IFX as equivalent to outcomes for patients on ADA. For these reasons the economic modelling was limited to IFX-treated patients.

The impact of immunosuppressants on patient outcome was not formally assessed in this review, as it was outside the scope. However, the evidence suggests that immunosuppressants generally improve patient outcome. The role that immunosuppressants might play in the monitoring of drug levels and anti-drug antibody levels is an area for future research. Evidence is also emerging that FC can be used to monitor patients with CD, as it is a good marker for IBD activity and predicts relapse in time before symptoms return, thereby providing the clinician time to optimise therapy.[177] Future research is needed into how this marker and anti-TNF-α monitoring might complement each other in the management of CD patients.

Finally, although this review was restricted to the population of patients with CD, a substantial number of studies included in the review recruited patients with IBD. We remain uncertain about the impact of patient mix on the reported outcomes. Even though IFX and ADA have recently been approved by NICE for the treatment of UC (positive NICE technology appraisal published in December 2014[178]), outcomes from this assessment should not be readily transferred to UC patients; for example, Bar-yoseph *et al.*[179] reported that IFX is more immunogenic and reaches lower trough levels in patients with UC than in those with CD. This seems to suggest that there are may be differences in the response to IFX between the two patient groups that may be of importance for the cost-effectiveness of monitoring anti-TNF-α agents in UC patients.

Overall, owing to the paucity of evidence, and especially the lack of 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 with outcome and test data from the management studies was problematic because of the small population size and short duration of these studies, and difficulties in allocating outcomes to categories of patients returning different defined test results. Inputs for the economic model need to be drawn from disparate studies so our 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 because of 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 are available about duration of anti-TNF-α 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 because of 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 patients who are reintroduced to anti-TNF-α treatment, for example the substantial proportion of patients who fail on IFX and are then switched to receive a variety of non-anti-TNF-α therapies that may or may not improve their clinical status, but who in the longer term are given an anti-TNF-α again.

All the studies used for modelling included mixed patient populations, a substantial proportion of whom were already being treated with immunosuppressants and had previously been exposed to steroids. Steroids are used intermittently for flare, and information on frequency and duration of use is missing. The goal of treatment with immunosuppressants added to anti-TNF-α therapy is to restore a better response to anti-TNF-α. The clinical effects of these agents are subsumed within the analysis of time to anti-TNF-α cessation. Information about the timing and frequency of addition of immunosuppressants and the duration of their use is inadequate. Therefore, it was difficult to model the addition of immunosuppressants and use of steroids for patients on anti-TNF-α, and clinical expert opinion was sought regarding the proportion of time over 10 years patients treated with anti-TNF-α agents would spend using steroid and immunosuppressants. These data were assumed to apply for both testing and standard care arms of the model and were used for costing purposes.

We were not able to include adverse events occurring as a result of drug treatments in the models. In addition, we have not included any health states or costs for patients who may have complications following surgery. As a result, this may underestimate the costs.

Chapter 6 Conclusions

The systematic review evidence gives some indication that the use of testing for IFX 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 for IFX have been presented. These indicate that very similar QALY gains are likely in both arms. No testing appears to be the most cost-effective option. This finding was robust when investigated in various sensitivity analyses. Conclusions for ADA cannot be drawn from the current evidence.

Recommendations for further research

We are aware that more comparative evidence is becoming available with the publication of the Trial Assigning Individualized Options for Treatment (Rx) (ClinicalTrials.gov identification number NCT01442025).[180,181] This trial will provide further insight into the effectiveness of sustaining therapeutic IFX trough levels by measuring drug levels followed by dose increase if criteria are met in CD patients. Furthermore, the Personalised anti-TNF therapy in Crohn's disease (UK Clinical Research Network identification number 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:

1. How does measuring anti-TNF-α drug and their antibodies by ELISA kits vary from using RIA and other methods?
2. What are the clinically significant differences in the performance of ELISAs, RIA and HMSA?
3. What are the best criteria for estimating response, non-response and LOR, and at what time should an assessment take place?
4. What is the most widely acceptable algorithm in the UK?
5. What are the barriers of following an algorithm for clinical management according to anti-TNF-α and anti-drug antibody levels?
6. What is the best time point of measuring drug and antibody, and how frequently should measurements be taken in responders and in patients with LOR?
7. What is the validated drug threshold that predicts clinical outcome?
8. What is the clinical effectiveness and cost-effectiveness of monitoring patients with CD on ADA and for paediatric patients with CD?
9. What is the relevance of cotreatment with immunosuppressants in the monitoring of anti-TNF-α agents and their antibodies?
10. Is there a benefit of measuring total drug/antibodies over free drug/antibody only?

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.

Contributions of authors

Karoline Freeman (Research Fellow) co-ordinated the review, and wrote the introduction and discussion.

Martin Connock (Senior Research Fellow) conducted the data analysis.

Martin Connock, **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 papers against the inclusion criteria, appraising the quality of papers and abstracting data from papers for synthesis.

Peter Auguste (Research Assistant) and **Hema Mistry** (Assistant Professor) undertook the health economic work.

Rachel Court (Information Specialist) developed the search strategy and undertook searches.

Ramesh Arasaradnam (Associate Professor of Gastroenterology) provided clinical comment and guidance.

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.

All authors were involved in writing draft versions of the report.

Data sharing statement

All available data can be obtained from the corresponding author.

References

1. 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**:343–50. http://dx.doi.org/10.1017/S0266462313000287
2. NICE. *Crohn's Disease: Management in Adults, Children and Young People*. Clinical guideline 152. London: NICE; 2012. URL: www.nice.org.uk/guidance/cg152 (accessed 6 October 2014).
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 Technol Assess* 2003;**7**(3). http://dx.doi.org/10.3310/hta7030
4. NHS Choices. *Crohn's Disease*. NHS; 2013. URL: www.nhs.uk/Conditions/Crohns-disease/Pages/Introduction.aspx (accessed 6 October 2014).
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). http://dx.doi.org/10.3310/hta15060
6. NICE. *Infliximab (Review) and Adalimumab for the Treatment of Crohn's Disease*. Technology appraisal 187. London: NICE; 2010. URL: www.nice.org.uk/guidance/TA187 (accessed 6 October 2014).
7. Hanauer SB, Sandborn W. Management of Crohn's disease in adults. *Am J Gastroenterol* 2001;**96**:635–43. http://dx.doi.org/10.1111/j.1572-0241.2001.03671.x
8. Jenkins HR. Inflammatory bowel disease. *Arch Dis Child* 2001;**85**:435–7. http://dx.doi.org/10.1136/adc.85.5.435
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):11–17. http://dx.doi.org/10.1046/j.1365-2036.17.s2.17.x
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**:863–73. http://dx.doi.org/10.1053/j.gastro.2006.12.003
11. Vermeire S, Van Assche G, Rutgeerts P. Laboratory markers in IBD: useful, magic, or unnecessary toys? *Gut* 2006;**55**:426–31. http://dx.doi.org/10.1136/gut.2005.069476
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**:512–30. http://dx.doi.org/10.1053/gast.2002.31072
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**:27–32. http://dx.doi.org/10.1097/00004836-199501000-00008
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**:355–66. http://dx.doi.org/10.1016/j.crohns.2010.04.004

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*. London: NICE; 2014. URL: 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 (accessed 21 April 2015).

16. Lichtenstein GR, Hanauer SB, Sandborn WJ. Management of Crohn's disease in adults. *Am J Gastroenterol* 2009;**104**:465–83. http://dx.doi.org/10.1038/ajg.2008.168

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**:28–62. http://dx.doi.org/10.1016/j.crohns.2009.12.002

18. Joint Formulary Committee. *British National Formulary* (online). London: BMJ Group and Pharmaceutical Press. URL: www.medicinescomplete.com (accessed 6 October 2014).

19. Joint Formulary Committee. *British National Formulary for Children* (online) London: BMJ Group and Pharmaceutical Press. URL: www.medicinescomplete.com (accessed 6 October 2014).

20. 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**:265–76. http://dx.doi.org/10.1002/ibd.3780030405

21. Cohen RD. The quality of life in patients with Crohn's disease. *Aliment Pharmacol Ther* 2002;**16**:1603–9. http://dx.doi.org/10.1046/j.1365-2036.2002.01323.x

22. IBD Standards Group. *Standards for the Healthcare of People who have Inflammatory Bowel Disease (IBD): 2013 Update*. IBD Standards; 2013. URL: www.ibdstandards.org.uk/uploaded_files/IBDstandards.pdf (accessed 25 February 2015).

23. 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**:1471–8. http://dx.doi.org/10.1136/gut.2004.041616

24. 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**:1047–52. http://dx.doi.org/10.1097/00042737-200510000-00007

25. 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**:1357–63. http://dx.doi.org/10.1111/j.1365-2036.2010.04482.x

26. 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. *J Crohns Colitis* 2011;**5**:301–16. http://dx.doi.org/10.1016/j.crohns.2011.02.005

27. 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**:1541–9. http://dx.doi.org/10.1016/S0140-6736(02)08512-4

28. Sands BE, Blank MA, Patel K, van Deventer SJ, ACCENT II Study. Long-term treatment of rectovaginal fistulas in Crohn's disease: response to infliximab in the ACCENT II Study. *Clin Gastroenterol Hepatol* 2004;**2**:912–20. http://dx.doi.org/10.1016/S1542-3565(04)00414-8

29. 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**:52–65. http://dx.doi.org/10.1053/j.gastro.2006.11.041

30. Ben-Horin S, Chowers Y. Tailoring anti-TNF therapy in IBD: drug levels and disease activity. *Nat Rev Gastroenterol Hepatol* 2014;**11**:243–55. http://dx.doi.org/10.1038/nrgastro.2013.253

31. 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**:1029–35. http://dx.doi.org/10.1056/NEJM199710093371502

32. 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**:323–33. http://dx.doi.org/10.1053/j.gastro.2005.11.030

33. de Boer N, Lowenberg M, Hoentjen F. Management of Crohn's disease in poor responders to adalimumab. *Clin Exp Gastroenterol* 2014;**7**:83–92.

34. Bendtzen K. Anti-TNF-alpha biotherapies: perspectives for evidence-based personalized medicine. *Immunotherapy* 2012;**4**:1167–79. http://dx.doi.org/10.2217/imt.12.114

35. Gisbert JP, Panes J. Loss of response and requirement of infliximab dose intensification in Crohn's disease: a review. *Am J Gastroenterol* 2009;**104**:760–7. http://dx.doi.org/10.1038/ajg.2008.88

36. 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**:674–84. http://dx.doi.org/10.1038/ajg.2011.60

37. 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;**146**:S-196. http://dx.doi.org/10.1016/S0016-5085(14)60690-9

38. 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**:1248–54. http://dx.doi.org/10.1016/j.cgh.2006.06.025

39. Cassinotti A, Travis S. Why don't we just measure infliximab drug levels in IBD? *Pract Gastroenterol* 2010;**34**:11–20.

40. 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**:542–53. http://dx.doi.org/10.1016/S1542-3565(04)00238-1

41. 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. http://dx.doi.org/10.1016/S1873-9946(14)60087-8

42. Barry D, Bloomfeld RS. The prevalence of human antichimeric antibodies in patients on infliximab increases with age. *Gastroenterology* 2012;**142**:S-387. http://dx.doi.org/10.1016/s0016-5085(12)61466-8

43. Baert F, Noman M, Vermeire S, Van Assche G, D' Haens G, Carbonez A, *et al.* Influence of immunogenicity on the long-term efficacy of infliximab in Crohn's disease. *N Engl J Med* 2003;**348**:601–8. http://dx.doi.org/10.1056/NEJMoa020888

44. 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**:447–59. http://dx.doi.org/10.1111/apt.12407

45. 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**:1226–31. http://dx.doi.org/10.1136/gut.2006.099978

46. 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. http://dx.doi.org/10.1016/S1873-9946(12)60273-6

47. 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**:944–8. http://dx.doi.org/10.1111/j.1572-0241.2007.01638.x

48. 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**:1628–40. http://dx.doi.org/10.1053/j.gastro.2009.07.062

49. 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**:1947–55. http://dx.doi.org/10.1136/annrheumdis-2012-202220

50. 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**:669–73. http://dx.doi.org/10.1097/FTD.0000000000000074

51. 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**:311–14. http://dx.doi.org/10.1136/annrheumdis-2014-206237

52. 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**:1172–83. http://dx.doi.org/10.1111/apt.12330

53. 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**:1861–8. http://dx.doi.org/10.1053/j.gastro.2008.03.004

54. 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;**138**:S167–8. http://dx.doi.org/10.1016/S0016-5085(10)60767-6

55. 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**:1383–95. http://dx.doi.org/10.1056/NEJMoa0904492

56. Afif W, Loftus EV, Faubion WA Jr, 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**:1133–9. http://dx.doi.org/10.1038/ajg.2010.9

57. 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**:1250–6. http://dx.doi.org/10.1038/ajg.2014.146

58. 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**:2568–76. http://dx.doi.org/10.1097/MIB.0b013e3182a77b41

59. 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**:1199–206. http://dx.doi.org/10.1002/ibd.21839

60. 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**:B4616.

61. Ben-Horin S, Chowers Y. Review article: loss of response to anti-TNF treatments in Crohn's disease. *Aliment Pharmacol Ther* 2011;**33**:987–95. http://dx.doi.org/10.1111/j.1365-2036.2011.04612.x

62. 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**:1264–75. http://dx.doi.org/10.1002/ibd.20899

63. 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**:1078–85. http://dx.doi.org/10.1097/MEG.0b013e32835558cf

64. 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**:971–86. http://dx.doi.org/10.1111/j.1365-2036.2012.05057.x

65. 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 Europ Gastroenterol J* 2013;**1**:A1.

66. Scott FI, Lichtenstein GR. Therapeutic drug monitoring of anti-TNF therapy in inflammatory bowel disease. *Curr Treat Options Gastroenterol* 2014;**12**:59–75. http://dx.doi.org/10.1007/s11938-013-0004-5

67. 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**:765–71. http://dx.doi.org/10.1111/apt.12030

68. 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. http://dx.doi.org/10.1136/bmj.b2535

69. 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**:529–36. http://dx.doi.org/10.7326/0003-4819-155-8-201110180-00009

70. 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. http://dx.doi.org/10.1136/bmj.d5928

71. 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**:377–84. http://dx.doi.org/10.1136/jech.52.6.377

72. Chang SM, Matchar DB, Smetana GW, Umscheid CA, Gray R, Torchia MM. *Methods Guide for Medical Test Reviews*. Agency for Healthcare Research and Quality (AHRQ) publication no. 12-EC017. Rockville, MD: AHRQ; 2012. URL: www.effectivehealthcare.ahrq.gov/ehc/products/246/558/Methods-Guide-for-Medical-Test-Reviews_Full-Guide_20120530.pdf (accessed 11 March 2015).

73. 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* 2015;**148**:1320–9.e3. http://dx.doi.org/10.1053/j.gastro.2015.02.031

74. 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**:9. http://dx.doi.org/10.1186/1471-2288-12-9

75. Harbord R, Whiting P. metandi: meta-analysis of diagnostic accuracy using hierarchical logistic regression. *Stata J* 2009;**9**:211–29.

76. Harris RJ, Bradburn MJ, Deeks JJ, Harbord R, Altman DG, Sterne JAC. Metan: fixed- and random-effects meta-analysis. *Stata J* 2008;**8**:3–28.

77. 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**:1474–81.e2. http://dx.doi.org/10.1016/j.cgh.2014.01.033

78. 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;**144**:S-775. http://dx.doi.org/10.1016/S0016-5085(13)62866-8

79. 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**:41–8. http://dx.doi.org/10.1136/gut.2009.201533

80. 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**:714–22. http://dx.doi.org/10.1111/j.1365-2036.2012.04997.x

81. 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. http://dx.doi.org/10.1016/S1873-9946(14)60625-5

82. 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**:736–43. http://dx.doi.org/10.1016/j.crohns.2012.10.019

83. 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**:11–19. http://dx.doi.org/10.1016/j.clim.2005.07.010

84. 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**:1112–22. http://dx.doi.org/10.1097/MIB.0b013e3182813242

85. 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**:1721–7. http://dx.doi.org/10.1136/gutjnl-2012-304094

86. 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**:7367–75. http://dx.doi.org/10.1007/s00216-013-7154-0

87. 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**:S213–14. http://dx.doi.org/10.1016/S1873-9946(13)60529-2

88. Dauer RM, Yarur AJ, Abreu MT. Infliximab re-induction outcomes after a failure to treatment. *Gastroenterol* 2013;**144**:S-430. http://dx.doi.org/10.1016/S0016-5085(13)61583-8

89. 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.

90. 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;**144**:S-779. http://dx.doi.org/10.1016/S0016-5085(13)62880-2

91. 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**:S231–2. http://dx.doi.org/10.1016/S1873-9946(13)60572-3

92. 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**:917–24. http://dx.doi.org/10.1053/gast.2003.50145

93. 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;**142**:S-114. http://dx.doi.org/10.1016/S0016-5085(12)60430-2

94. 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**:1714–21. http://dx.doi.org/10.1097/MIB.0000000000000138

95. 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. http://dx.doi.org/10.1016/S1873-9946(14)60498-0

96. Greathead L, Kelleher P, Steel A. Development and validation of ELISA to measure serum anti TNFa levels. *J Crohns Colitis* 2014;**8**:S97–8. http://dx.doi.org/10.1016/S1873-9946(14)60208-7

97. 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;**142**:S-538. http://dx.doi.org/10.1016/S0016-5085(12)62067-8

98. 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**:254–62. http://dx.doi.org/10.1007/s00535-013-0807-0

99. 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**:136–43. http://dx.doi.org/10.1007/s00535-011-0474-y

100. 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**:100–9. http://dx.doi.org/10.1007/s00535-013-0803-4

101. 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**:674–82. http://dx.doi.org/10.1007/s00535-013-0829-7

102. 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.

103. 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**:1628–33. http://dx.doi.org/10.1002/ibd.21919

104. 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**:1126–35. http://dx.doi.org/10.1111/apt.12733

105. 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**:881–9. http://dx.doi.org/10.1016/j.crohns.2014.01.009

106. 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. http://dx.doi.org/10.1016/S1590-8658(14)60224-3

107. 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;**144**:S-778. http://dx.doi.org/10.1016/S0016-5085(13)62874-7

108. 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**:620–8. http://dx.doi.org/10.1111/apt.12869

109. 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.

110. 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. *Gut* 2015;**64**:A99. http://dx.doi.org/10.1136/gutjnl-2015-309861.202

111. 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**:40–7. http://dx.doi.org/10.1038/ajg.2012.363

112. 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**:5031–5. http://dx.doi.org/10.3748/wjg.v20.i17.5031

113. 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;**142**:S354. http://dx.doi.org/10.1016/S0016-5085(12)61332-8

114. 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**:1288–95. http://dx.doi.org/10.1097/MIB.0000000000000037

115. 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**:80–4.e2. http://dx.doi.org/10.1016/j.cgh.2013.07.010

116. 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**:e287–9. http://dx.doi.org/10.1515/cclm-2013-0461

117. 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, London, 8–11 May 2013. *J Pediatr Gastroenterol Nutr* 2013;**56**:19.

118. 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**:eA27–8.

119. 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**:1708–13. http://dx.doi.org/10.1097/MIB.0000000000000137

120. 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**:310–18. http://dx.doi.org/10.3109/00365521.2010.536254

121. 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**:530–8. http://dx.doi.org/10.1097/FTD.0b013e31828d23c3

122. 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**:1055–64. http://dx.doi.org/10.1038/ajg.2014.106

123. 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**:919–27. http://dx.doi.org/10.1136/gutjnl-2013-305279

124. Steenholdt C, Brynskov J, Thomsen OØ, Munck LK, Fallingborg J, Christensen LA, *et al.* Individualized therapy is a long-term cost-effective method compared to dose intensification in Crohn's disease patients failing infliximab. *Dig Dis Sci* 2015;**60**:2762–70. http://dx.doi.org/10.1007/s10620-015-3581-4

125. 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;**146**:S-245. http://dx.doi.org/10.1016/S0016-5085(14)60862-3

126. 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**:962–71. http://dx.doi.org/10.1038/ajg.2013.12

127. 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;**146**:S-242. http://dx.doi.org/10.1016/S0016-5085(14)60852-0

128. 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**:1996–2003. http://dx.doi.org/10.1097/MIB.0000000000000156

129. 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;**138**:S684–5. http://dx.doi.org/10.1016/s0016-5085(10)63147-2

130. 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**:S475–6.

131. 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**:177–88. http://dx.doi.org/10.1016/j.jim.2012.06.002

132. 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.

133. 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**:1122–6. http://dx.doi.org/10.1111/j.1365-2036.2008.03828.x

134. 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;**142**:S788–9. http://dx.doi.org/10.1016/S0016-5085(12)63061-3

135. Yarur AJ, Deshpande AR, Sussman DA, Hauenstein S, Lockton S, Barkin JS, *et al.* TU1147 serum adalimumab levels and antibodies correlate with endoscopic intestinal inflammation and inflammatory markers in patients with inflammatory bowel disease. *Gastroenterology* 2013;**144**:S774–5. http://dx.doi.org/10.1016/S0016-5085(13)62863-2

136. 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**:e291–2. http://dx.doi.org/10.1515/cclm-2013-0570

137. 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 (SAT0479). *Ann Rheum Dis* 2013;**71**:634–5. http://dx.doi.org/10.1136/annrheumdis-2012-eular.3425

138. Lichtenstein GR. Comprehensive review: antitumor necrosis factor agents in inflammatory bowel disease and factors implicated in treatment response. *Therap Adv Gastroenterol* 2013;**6**:269–93. http://dx.doi.org/10.1177/1756283X13479826

139. Yanai H, Hanauer SB. Assessing response and loss of response to biological therapies in IBD. *Am J Gastroenterol* 2011;**106**:685–98. http://dx.doi.org/10.1038/ajg.2011.103

140. 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;**64**:1539–45. http://dx.doi.org/10.1136/gutjnl-2014-307883

141. Higgins JPT, Green S, editors. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0*. The Cochrane Collaboration; 2011. URL: www.cochrane-handbook.org (accessed 15 January 2015).

142. Seow CH, Newman A, Irwin SP, Steinhart AH, Silverberg MS, Greenberg GR. Trough serum infliximab: a predictive factor of clinical outcome for infliximab treatment in acute ulcerative colitis. *Gut* 2010;**59**:49–54. http://dx.doi.org/10.1136/gut.2009.183095

143. Pepe NS. *The Statistical Evaluation of Medical Tests for Classification and Prediction*. New York, NY: Oxford University Press; 2003.

144. Harbord RM, Deeks JJ, Egger M, Whiting P, Sterne JA. A unification of models for meta-analysis of diagnostic accuracy studies. *Biostatistics* 2007;**8**:239–51. http://dx.doi.org/10.1093/biostatistics/kxl004

145. Rutgeerts P, Sandborn WJ, Feagan BG, Reinisch W, Olson A, Johanns J, *et al.* Infliximab for induction and maintenance therapy for ulcerative colitis. *N Engl J Med* 2005;**353**:2462–76. http://dx.doi.org/10.1056/NEJMoa050516

146. Sands BE, Anderson FH, Bernstein CN, Chey WY, Feagan BG, Fedorak RN, *et al.* Infliximab maintenance therapy for fistulizing Crohn's disease. *N Engl J Med* 2004;**350**:876–85. http://dx.doi.org/10.1056/NEJMoa030815

147. Miele E, Markowitz JE, Mamula P, Baldassano RN. Human antichimeric antibody in children and young adults with inflammatory bowel disease receiving infliximab. *J Pediatr Gastroenterol Nutr* 2004;**38**:502–8. http://dx.doi.org/10.1097/00005176-200405000-00008

148. Steenholdt C, Svenson M, Bendtzen K, Thomsen O, Brynskov J, Ainsworth MA. Can measurements of anti-infliximab antibodies predict acute severe infusion reactions to infliximab? *Gastroenterology* 2011;**140**:S774.

149. 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). http://dx.doi.org/10.3310/hta8360

150. 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**:49–57. http://dx.doi.org/10.1016/S0016-5085(99)70549-4

151. Joint Formulary Committee. *British National Formulary*. 53 ed. London: BMJ Group and Pharmaceutical Press; 2007.

152. Joint Formulary Committee. *British National Formulary*. 55 ed. London: BMJ Group and Pharmaceutical Press; 2008.

153. Department of Health. *NHS Reference Costs 2005 to 2006*. London: Department of Health; 2006.

154. 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**:77–85. http://dx.doi.org/10.1016/j.crohns.2011.07.007

155. 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**:265–74. http://dx.doi.org/10.1111/j.1365-2036.2009.04033.x

156. 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**:1509–20. http://dx.doi.org/10.1111/j.1365-2036.2007.03548.x

157. 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**:654–66. http://dx.doi.org/10.1016/j.cgh.2012.12.035

158. 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**:117–22. http://dx.doi.org/10.1017/S0266462313000160

159. 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**:60–70. http://dx.doi.org/10.1097/MIB.0000000000000225

160. 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**:93–106. http://dx.doi.org/10.1016/j.crohns.2013.04.007

161. 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**:1454–63. http://dx.doi.org/10.1016/j.crohns.2014.05.007

162. 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**:829–38. http://dx.doi.org/10.7326/0003-4819-146-12-200706190-00159

163. 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**:90–7. http://dx.doi.org/10.1053/j.gastro.2011.03.050

164. 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. http://dx.doi.org/10.1002/14651858.cd010233.pub2

165. 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**:761–9. http://dx.doi.org/10.1016/S0016-5085(99)70332-X

166. Joint Formulary Committee. *British National Formulary*. 66 ed. London: BMJ Group and Pharmaceutical Press; 2013.

167. Department of Health. *NHS Reference Costs 2013 to 2014*. London: Department of Health; 2014. URL: www.gov.uk/government/publications/nhs-reference-costs-2013-to-2014 (accessed 24 March 2015).

168. 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**:815–24. http://dx.doi.org/10.1586/14737167.2014.957682

169. Office for National Statistics. *National Life Tables, United Kingdom, 1980–82 to 2011–13*. 2014. URL: www.ons.gov.uk/ons/search/index.html?newquery=National+Life+Tables%2C+United+Kingdom%2C+1980-82+to+2011-13 (accessed 31 March 2015).

170. Curtis L. *Unit Costs of Health and Social Care 2014*. Canterbury: Personal Social Services Research Unit, University of Kent; 2014. URL: www.pssru.ac.uk/project-pages/unit-costs/2014/ (accessed 24 March 2015).

171. Saito S, Shimizu U, Zhang N, Yokoyama J, Watanabe M, Terajima K, *et al.* A health economic analysis of combination therapy with infliximab plus elemental diet for moderately to severely active Crohn's disease. *Health (NY)* 2014;**6**:107–14. http://dx.doi.org/10.4236/health.2014.61017

172. 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.

173. 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.

174. 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;**146**:S-242. http://dx.doi.org/10.1016/S0016-5085(14)60853-2

175. 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;**144**:S-423. http://dx.doi.org/10.1016/S0016-5085(13)61559-0

176. 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;**144**:S531. http://dx.doi.org/10.1016/S0016-5085(13)61975-7

177. Turvill J. Mapping of Crohn's disease outcomes to faecal calprotectin levels in patients maintained on biologic therapy. *Frontline Gastroenterol* 2014;**5**:167–75. http://dx.doi.org/10.1136/flgastro-2014-100441

178. NICE. *Infliximab, Adalimumab and Golimumab for Treating Moderately to Severely Active Ulcerative Colitis after the Failure of Conventional Therapy. NICE technology appraisal guidance [TA329]*. London: NICE; 2015.

179. 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;**144**:S-780. http://dx.doi.org/10.1016/S0016-5085(13)62883-8

180. D'Haens GR, Vermiere S, Lambrecht G, Baert FJ, Bassuyt P, Nachury M, *et al.* 692 Drug-level based dosing versus symptom-based dose adaptation in patients with crohn's disease: a prospective, randomized multicenter study (TAILORIX). *Gastroenterology* 2016;**150**:S143. http://dx.doi.org/10.1016/S0016-5085(16)30583-2

181. D'Haens G, Vermeire S, Lambrecht G, Baert F, Bossuyt P, Nachury M, *et al. OP029 Drug-Concentration Versus Symptom-Driven Dose Adaptation of Infliximab in Patients with Active Crohn's Disease: A Prospective, Randomised, Multicentre Trial (Tailorix)*. European Crohn's and Colitis Organisation; 2016. URL: www.ecco-ibd.eu/index.php/publications/congress-abstract-s/abstracts-2016/item/op029-drug-concentration-versus-symptom-driven-dose-adaptation-of-infliximab-in-patients-with-active-crohnix2019is-disease-a-prospective-randomised-multicentre-trial-tailorix.html (accessed 29 September 2016).

182. Nagore D, Ruiz Del Agua A, Pascual J, Llinares-Tello F, Herreros B, Martinez A. TU1325 therapeutic cut-off of infliximab in patients with inflammatory bowel diseases. *Gastroenterology* 2015;**148**:S-860. http://dx.doi.org/10.1016/S0016-5085(15)32913-9

183. 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**:1675–83. http://dx.doi.org/10.1016/j.crohns.2014.07.010

184. 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**:1200–6. http://dx.doi.org/10.1136/gut.2009.202101

185. 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**:e623–9. http://dx.doi.org/10.1016/j.crohns.2013.04.020

Appendix 1 Details of manufacturers' enzyme-linked immunosorbent assay kits

Sections of this appendix have been reproduced with permission from NICE.[15]

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 enzyme-linked immunosorbent assay kits (Theradiag/Alpha Laboratories)

LISA-TRACKER assay kits are ELISAs for the quantitative determination of TNF-α inhibitor levels and antibodies against TNF-α inhibitor. There are six LISA-TRACKER ELISA kits relevant to this assessment (*Table 40*): two of these kits measure the levels of free anti-drug antibodies; two kits measure the levels of free TNF-α inhibitor; and two kits measure the levels of both free anti-drug antibodies and TNF-α inhibitor.

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 TNF-α inhibitor or levels of anti-drug antibody in the patient's serum.

Detecting levels of tumour necrosis factor alpha inhibitor

Patient samples, the standards and controls are added to the pre-coated microtitre plate. The TNF-α inhibitor (ADA or IFX) 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 and binds to the TNF-α inhibitor attached to the coated plate. Any unbound reagent is removed by a second wash step before peroxidase-labelled streptavidin is added to the

TABLE 40 LISA-TRACKER ELISA kits

Name (code)	Detects	Microtitre plate pre-coat	Secondary reagent	Incubation times
LISA-TRACKER Adalimumab (LTA002)	Free ADA	TNF-α	Biotinylated anti-human IgG antibody	1 hour; 1 hour; 30 minutes; and 15 minutes
LISA-TRACKER Infliximab (LTI002)	Free IFX	TNF-α		
LISA-TRACKER anti-Adalimumab (LTA003)	Free anti-ADA antibodies	ADA	Biotinylated ADA	1 hour; 1 hour; 30 minutes; and 15 minutes
LISA-TRACKER anti-Infliximab (LTI003)	Free anti-IFX antibodies	IFX	Biotinylated IFX	
LISA-TRACKER Duo Adalimumab (LTA005)	As above; the Duo Adalimumab kit consists of a LISA-TRACKER Adalimumab kit and a LISA-TRACKER anti-Adalimumab kit			
LISA-TRACKER Duo Infliximab (LTI005)	As above; the Duo Infliximab kit consists of a LISA-TRACKER Infliximab kit and a LISA-TRACKER anti-Infliximab kit			

Note
There are two additional LISA-TRACKER ELISA kits that are available in some European countries, but not in the UK. The LISA-TRACKER Premium Adalimumab and the LISA-TRACKER Premium Infliximab assays both measure three parameters: TNF-α inhibitor, TNF-α levels and anti-drug antibody levels.

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 TNF-α 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 41*.

Detecting levels of antibodies to tumour necrosis factor alpha inhibitor

Patient samples, the standards and the controls are added to the pre-coated microtitre plate. The free anti-IFX antibodies or free anti-ADA 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 TNF-α inhibitor in each patient sample. The limits of detection and assay ranges are presented in *Table 41*.

Tumour necrosis factor alpha-Blocker enzyme-linked immunosorbent assay kits (Immundiagnostik AG)

There are six Immundiagnostik ELISA kits relevant to this assessment, which are distributed in the UK by BioHit Healthcare Ltd (*Table 42*): two of these kits measure the levels of free anti-drug antibodies; two kits measure the levels of total anti-drug antibodies (free antibodies and antibodies already bound to the drug); and two kits measure the levels of free TNF-α inhibitor.

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 IFX or ADA (K9655 and K9657) follow a standard ELISA procedure for detecting levels of TNF-α inhibitor, as described in *Chapter 1, Intervention technologies*, 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-ADA antibodies or free anti-IFX antibodies (K9650 and K9652) follow a standard ELISA procedure for detecting levels of antibodies to TNF-α inhibitor, as described in *Chapter 1, Intervention technologies*, except that the secondary reagent is directly labelled with peroxidase and, therefore, there is no biotin–streptavidin binding step. Furthermore, standards are not used in the anti-drug antibody ELISAs; therefore, the results are interpreted semiquantitatively using a cut-off control. Details on the interpretation of results, limits of detection and assay measurement ranges are presented in *Table 43*.

TABLE 41 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	0.1–8 µg/ml
LISA-TRACKER Infliximab (LTI002)		0.1 µg/ml	0.1–8 µg/ml
LISA-TRACKER anti-Adalimumab (LTA003)	Quantitative. Generation of standard curve and determination of ADAb level in ng/ml	10 ng/ml	10–160 ng/ml
LISA-TRACKER anti-Infliximab (LTI003)		10 ng/ml	10–200 ng/ml

TABLE 42 Immundiagnostik ELISA kits

Name (code)	Detects	Microtitre plate pre-coat	Secondary reagent	Incubation times
Immundiagnostik TNF-α-Blocker monitoring, infliximab drug level (e.g. Remicade) ELISA (K9655)	Free IFX	Monoclonal anti-IFX antibody	Peroxidase-labelled antibody	1 hour; 1 hour; and 10–20 minutes
Immundiagnostik TNF-α-Blocker monitoring, adalimumab drug level (e.g. Humira) ELISA (K9657)	Free ADA	Monoclonal anti-ADA antibody	Peroxidase-labelled antibody	
Immundiagnostik TNF-α-Blocker ADA, antibodies against infliximab (e.g. Remicade) ELISA (K9650)	Free anti-IFX antibodies	IFX $F(ab)_2$ fragments	Peroxidase-labelled IFX	2 × 15 minutes; 16–20 hours; 1 hour; and 10–20 minutes
Immundiagnostik TNF-α-Blocker ADA, antibodies against adalimumab (e.g. Humira) ELISA (K9652)	Free anti-ADA antibodies	ADA $F(ab)_2$ fragments	Peroxidase-labelled ADA	16–20 hours; 1 hour; and 10–20 minutes
Immundiagnostik TNF-α-Blocker ADA, TOTAL antibodies against infliximab (e.g. Remicade®) ELISA (K9654)	Total anti-IFX antibodies	Streptavidin	N/A	20 minutes; 1 hour; 1.5 hours; and 10–20 minutes
Immundiagnostik TNF-α-Blocker ADA, TOTAL antibodies against adalimumab (e.g. Humira) ELISA (K9651)	Total anti-ADA antibodies	Streptavidin	N/A	

N/A, not applicable.

TABLE 43 Interpretation of results, limits of detection and assay ranges for the Immundiagnostik ELISAs

Name (code)	Results interpretation	Limit of blank	Assay range
Immundiagnostik TNF-α-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	0.4–45 µg/ml
Immundiagnostik TNF-α-Blocker monitoring, adalimumab drug level (e.g. Humira) ELISA (K9657)		2.3 ng/ml	0.4–45 µg/ml
Immundiagnostik TNF-α-Blocker ADA, antibodies against infliximab (e.g. Remicade) ELISA (K9650)	Semiquantitative. Evaluated by a cut-off control (10 arbitrary units/ml) to give a positive or negative result	5.787 arbitrary units/ml	N/A
Immundiagnostik TNF-α-Blocker ADA, antibodies against adalimumab (e.g. Humira) ELISA (K9652)		N/A	N/A
Immundiagnostik TNF-α-Blocker ADA, TOTAL antibodies against infliximab (e.g. Remicade) ELISA (K9654)	Semiquantitative. Evaluated by a cut-off control (10 arbitrary units/ml) to give a positive or negative result	2.653 arbitrary units/ml	N/A
Immundiagnostik TNF-α-Blocker ADA, TOTAL antibodies against adalimumab (e.g. Humira) ELISA (K9651)		2.765 arbitrary units/ml	N/A

N/A, not applicable.

The TOTAL anti-drug antibody ELISA kits (K9654 and K9651) enables the measurement of anti-drug antibodies in the presence of TNF-α inhibitor. During sample preparation, immune complexes between anti-drug antibodies and ADA or IFX are dissociated using an acidic buffer. Biotinylated and peroxidase-labelled ADA or IFX 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.

Promonitor enzyme-linked immunosorbent assay kits (Proteomika)

There are four Promonitor ELISA kits relevant to this assessment (*Table 44*): two of these kits measure the levels of free anti-drug antibodies; and two kits measure the levels of free TNF-α inhibitor.

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 TNF-α inhibitor, as described in *Chapter 1, Intervention technologies*, 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 TNF-α inhibitor, as described in *Chapter 1, Intervention technologies*, 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 45*.

TABLE 44 Promonitor ELISA kits

Name (code)	Detects	Microtitre plate pre-coat	Secondary reagent	Incubation times
Promonitor ADL ELISA (5080230000)	Free ADA	Anti-ADA human monoclonal antibody	Peroxidase-labelled anti-ADA monoclonal antibody	1 hour; 1 hour; and 25–35 minutes
Promonitor IFX ELISA (5060230000)	Free IFX	Anti-TNF-α human monoclonal antibody bound to human recombinant TNF-α	Peroxidase-labelled anti-IFX monoclonal antibody	1 hour; 1 hour; and 10–20 minutes
Promonitor anti-ADL ELISA (5090230000)	Free anti-ADA antibodies	ADA	Peroxidase-labelled ADA	1 hour; 1 hour; and 25–35 minutes
Promonitor anti-IFX ELISA (5070230000)	Free anti-IFX antibodies	IFX	Peroxidase-labelled IFX	1 hour; 1 hour; and 25–35 minutes

TABLE 45 Limits of quantification and assay ranges for Promonitor ELISAs

Name	Results interpretation	Limit of quantification	Assay range
Promonitor ADL ELISA	Semiquantitative. Evaluated using a cut-off value (0.024 µg/ml for ADA and 0.035 µg/ml for IFX) to give a positive or negative result	2.9 ng/ml	0.024–12 µg/ml
Promonitor IFX ELISA	Quantitative. Generation of standard curve and determination of drug level in µg/ml	1.7 ng/ml	0.035–14.4 µg/ml
Promonitor anti-ADL ELISA	Semiquantitative. Evaluated using a cut-off value (10 arbitrary units/ml for anti-ADA antibodies and 5 arbitrary units/ml for anti-IFX antibodies) to give a positive or negative result	3.7 arbitrary units/ml	3.5–2000 arbitrary units/ml
Promonitor ADLanti-IFX ELISA	Quantitative. Generation of standard curve and determination of anti-drug antibody level in arbitrary units/ml	2 arbitrary units/ml	2–1440 arbitrary units/ml

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 enzyme 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 TNF-α signalling pathway so that when the cells are incubated in the presence of TNF-α 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-TNF-α is present, then the TNF-α response is partially quenched and the quenching estimated. If anti-drug antibodies are present, quenching by anti-TNF-α 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 a luminometer (although these are not necessarily routinely available) and equipment for culture of growth-arrested genetically engineered cells under controlled conditions (oxygen, carbon dioxide and 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 high-performance liquid chromatography (SE-HPLC). The mobility shift assay is a liquid-phase assay based on SE-HPLC, which separates free probe (small size) from probe in an immune complex (large size). The probe for the assay of antibodies to an anti-TNF-α drug (e.g. antibodies to IFX) is a fluorescent dye-labelled anti-TNF-α drug (e.g. IFX); the probe for the assay of anti-TNF-α drug is fluorescent dye-labelled TNF-α.

The anti-drug antibody assays (e.g. assays for antibodies to IFX) use fluorescent dye-labelled anti-TNF-α (D*) as the probe; in the presence of antibodies to the anti-TNF-α drug (e.g. antibodies to IFX) 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 high-performance liquid chromatography (HPLC) column at different mobilities is measured with a fluorimeter (*Figure 35*).

The anti-TNF-α drug assay (e.g. assay for IFX) uses fluorescent dye-labelled TNF-α (TNF-α*) as the probe; in the presence of anti-TNF-α drug (e.g. IFX) some TNF-α* form immune complexes with the anti-TNF-α and these have greater mobility on the SE-HPLC than the free TNF-α*. The amount of TNF-α* shifted to greater mobility is proportional to the amount of anti-TNF-α drug 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-TNF-α and anti-drug antibody so that all anti-TNF-α and anti-drug antibody are 'free'; after neutralisation the sample is incubated with fluorescent dye-labelled anti-TNF-α drug (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 and 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

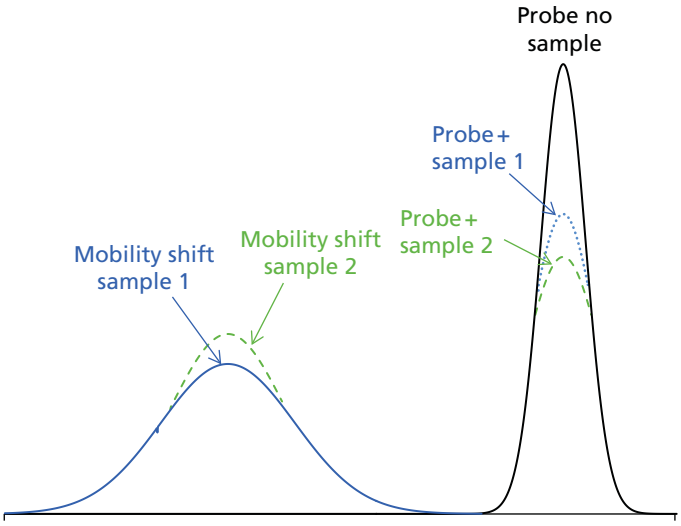


FIGURE 35 Illustration of chromatograms obtained after size exclusion of probe-labelled samples using HPLC. The vertical axis represents the fluorescence signal.

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 available only as a service and may not be practicable for use in UK patients. Setting up mobility shift assays in a hospital laboratory and constructing requisite reagents would be a major and expensive undertaking.

Appendix 3 Search strategies

Clinical effectiveness: database searches

MEDLINE(R) (via Ovid)

Exact database searched: 1946 to October week 2 2014.

Date searched: 22 October 2014.

Search strategy

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	24,181
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/	129,174
15	enzyme* link* immunoassay*.mp.	2873
16	enzyme* link* immuno* assay*.mp.	158,537
17	ELISA*.mp.	113,426
18	11 or 12 or 13 or 14 or 15 or 16 or 17	205,224
19	*Radioimmunoassay/	7091
20	(radioimmuno* or radio immuno* or radio-immuno*).mp.	101,819
21	RIA.tw.	17,353
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	124,775

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/	14,444
34	Crohn Disease/	31,596
35	crohn*.tw.	32,370
36	inflammator* bowel* disease*.tw.	26,840
37	IBD.tw.	11,936
38	33 or 34 or 35 or 36 or 37	58,401
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/	3,983,380
45	43 not 44	349

MEDLINE(R) In-Process & Other Non-Indexed Citations (via Ovid)

Exact database searched: 21 October 2014.

Date searched: 22 October 2014.

Search strategy

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	10,101
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 and EMBASE

Exact database searched: 1947 to 2014 week 42.

Date searched: 22 October 2014.

Search strategy

1	adalimumab.tw.	7379
2	*adalimumab/	3997
3	ADA.tw.	10,848
4	infliximab.tw.	13,600
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	35,630
14	lisa* tracker*.tw.	11
15	(immundiagnostik* or immunodiagnostik* or immunediagnostik*).tw.	74
16	(proteomika* or promonitor*).tw.	27
17	*enzyme linked immunosorbent assay/	14,622
18	enzyme* link* immunoassay*.tw.	3275
19	enzyme* link* immuno* assay*.tw.	71,923
20	ELISA*.tw.	166,866
21	14 or 15 or 16 or 17 or 18 or 19 or 20	207,373
22	*radioimmunoassay/	17,240
23	(radioimmuno* or radio immuno* or radio-immuno*).tw.	74,895
24	RIA.tw.	20,769
25	reporter* gene* assay*.tw.	4396
26	RGA.tw.	400
27	semi* fluid* phase* enzyme* immuno*.tw.	1
28	EIA.tw.	10,836
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	113,752
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/	34,280
37	crohn*.tw.	50,039
38	inflammator* bowel* disease*.tw.	41,418
39	IBD.tw.	23,266
40	36 or 37 or 38 or 39	82,551
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/	3,490,973
47	45 not 46	917

The Cochrane Library (via Wiley Online Library)

Date searched: 22 October 2014.

Search strategy

#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).
- 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)

Date searched: 22 October 2014.

Search strategy

#40	#39 OR #38 OR #37 OR #36	806
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#39	#35 AND #32	324
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#38	#35 AND #31 AND #9	26
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#37	#35 AND #16 AND #9	128
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#36	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"))	539
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#35	#34 OR #33	80,743
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#34	TS=(((inflammator* near/1 bowel*) near/1 disease*) or IBD)	53,142
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#33	TS=crohn*	50,398
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	

#32	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*)))	1366
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#31	#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	79,288
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#30	TS=(Prometheus* or "Anser IFX" or "Anser ADA")	713
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#29	TS=((Matriks* near/1 Biotek*) or Shikari*)	10
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#28	TS=(Biomonitor* or iLite)	8841
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#27	TS=HMSA	107
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#26	TS=((homogenous* or homogeneous*) near/1 (mobilit* near/1 (shift* near/1 assay*)))	11
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#25	TS=EIA	8832
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#24	TS=((semi* near/1 fluid*) near/3 (enzyme* near/1 immuno*))	1
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#23	TS=((semi* near/1 fluid*) near/2 (enzyme* near/1 immuno*))	0
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#22	TS=(semi* near/1 fluid* near/1 phase* near/1 enzyme* near/1 immuno*)	0
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#21	TS=(((semi* near/1 fluid*) near/1 phase*) near/1 (enzyme* near/1 immuno*))	0
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#20	TS=RGA	1230
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#19	TS=(reporter* near/1 gene* near/1 assay*)	4518
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#18	TS=RIA	12,773
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#17	TS=(radioimmuno* or (radio near/1 immuno*) or radio-immuno*)	46,937
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#16	#15 OR #14 OR #13 OR #12 OR #11 OR #10	146,389
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#15	TS=ELISA*	113,120
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#14	TS=((enzyme* near/1 link*) near/1 (immuno* near/1 assay))	60,666
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	

#13	TS=((enzyme* near/1 link*) near/1 immunoassay*)	2850
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#12	TS=(proteomika* or promonitor*)	1
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#11	TS=(immundiagnostik* or immunodiagnostik* or immunediagnostik*)	9
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#10	TS=(lisa* near/1 tracker*)	0
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#9	#8 OR #7 OR #6 OR #5 OR #4 OR #3 OR #2 OR #1	32,262
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#8	TS=ADAb	35
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#7	TS=((anti* near/1 drug*) near/1 antibod*)	2534
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#6	TS=((anti* near/1 tumo$r*) near/1 (necrosis* near/1 factor*))	4072
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#5	TS=((anti-TNF* or antiTNF* or TNF*) near/2 inhibitor*)	4065
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#4	TS=IFX	373
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#3	TS=infliximab	13,729
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#2	TS=ADA	8006
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
#1	TS=adalimumab	4973
	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	

Index to Theses

Date searched: 22 October 2014.

Search strategy

((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))

Fourteen documents 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")))

Four documents retrieved.

Digital Access to Research Theses – Europe
Date searched: 28 October 2014.

Search strategy
(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)

One hundred and thirteen documents retrieved.

Dissertations and Theses
Date searched: 29 October 2014.

Search strategy
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)))

Twenty-one documents retrieved.

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"))

Fifteen documents retrieved.

National Institute for Health Research Health Technology Assessment programme
Date searched: 29 October 2014.

Search strategy
adalimumab

Sixteen documents retrieved.

infliximab

Twenty-three documents retrieved.

TNF

Seventeen documents retrieved.

PROSPERO
Date searched: 29 October 2014.

Search strategy
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

Twenty-nine records retrieved.

ClinicalTrials.gov

Date searched: 4 November 2014.

Search strategy

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

Fourteen studies retrieved.

Current Controlled Trials

Date searched: 4 November 2014.

Search strategy

(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*)

Thirty studies retrieved.

UK Clinical Research Network Portfolio Database

Date searched: 4 November 2014.

Search strategy

Specialty: Gastroenterology

Research Summary: adalimumab infliximab TNF AntiTNF Anti-TNF ADAb

'Any' selected (combines terms with Boolean OR)

Four studies retrieved.

World Health Organization's International Clinical Trials Registry Platform

Date searched: 10 November 2014.

Advanced search

In Title: adalimumab OR infliximab OR AntiTNF* OR Anti-TNF* OR TNF inhibitor* OR TNF-α 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*

Thirty-nine trials found.

Espacenet (European Patent Office)

Date searched: 10 November 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 two 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

Cooperative Patent Classification: 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

Date searched: 22 January 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

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. Five potentially relevant abstracts saved.

Digestive Diseases Week (meeting of the American Gastroenterology Association)

www.ddw.org

Abstracts in *Gastroenterology*

2009–2014 Indexed in EMBASE. Checked and the search of EMBASE has picked them up.

Note: Promonitor have sent two abstracts submitted to Digestive Diseases Week May 2015.

British Society of Gastroenterology

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: www.bsg.org.uk/education/meeting/index.html

Searches

infliximab

adalimumab

TNF

Sifted online. Two potentially relevant abstracts saved.

United European Gastroenterology Week

Searched 2013 and 2014 in *United European Gastroenterology* 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	36,891
#2	Search crohn*	42,212
#1	Search "United European Gastroenterol J"[Journal]	149

Previous years not available.

American College of Gastroenterology

Meeting abstracts in *American Journal of Gastroenterology.*

2010–13 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: 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 2 February 2015.

European Crohn's and Colitis Organisation

URL: www.ecco-ibd.eu

Browsed consensus statements for Crohn's Disease, Publications and Research Projects.

No additional 'management' studies identified.

The American Gastroenterology Association

URL: www.gastro.org

Browsed: 'Technical Reviews'.

Browsed: Research > Research Resource Library > Immunology, Microbiology and IBD.

No additional 'management' studies identified.

British Society of Gastroenterology

URL: www.bsg.org.uk

Browsed: 'Research' and 'Clinical' sections.

No additional 'management' studies identified.

United European Gastroenterology

URL: www.ueg.eu

Browsed: 'Research' section.

No additional 'management' studies identified.

American College of Gastroenterology

URL: http://gi.org

Browsed: 'Research and Awards' and 'Clinical Guidelines' sections.

No additional 'management' studies identified.

International Network of Agencies for Health Technology Assessment publication

URL: www.inahta.org

Searched within publications for:

infliximab

adalimumab

TNF

No additional 'management' studies identified.

US Food and Drug Administration medical devices

URL: 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

URL: http://ec.europa.eu/growth/sectors/medical-devices

Searches

infliximab

adalimumab

TNF

No additional 'management' studies identified.

Theradiag

URL: www.theradiag.com/en

Browsed Theranostics > LISA-TRACKER

Saved and sifted list of publications for LISA-TRACKER assays. No additional 'management' studies identified.

Immundiagnostik

URL: 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

URL: 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

MEDLINE(R) (via Ovid)

Exact database searched: 1946 to October week 3 2014.

Date searched: 12 December 2014.

Search strategy

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	24,434
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/	129,940
23	enzyme* link* immunoassay*.mp.	2879
24	enzyme* link* immuno* assay*.mp.	159,574
25	ELISA*.mp.	114,330

26	19 or 20 or 21 or 22 or 23 or 24 or 25	206,726
27	*Radioimmunoassay/	7654
28	(radioimmuno* or radio immuno* or radio-immuno*).mp.	102,645
29	RIA.tw.	17,539
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	125,716
40	Inflammatory Bowel Diseases/	14,609
41	Crohn Disease/	31,828
42	crohn*.tw.	32,634
43	inflammator* bowel* disease*.tw.	27,171
44	IBD.tw.	12,128
45	40 or 41 or 42 or 43 or 44	58,950
46	18 and 45	3875
47	26 and 45	1771
48	39 and 45	278
49	exp Economics/	513,380
50	exp "Costs and Cost Analysis"/	190,833
51	Health Status/	63,445
52	exp "Quality of Life"/	126,611
53	exp Quality-Adjusted Life Years/	7642
54	(pharmacoeconomic* or pharmaco-economic* or economic* or cost*).tw.	461,021
55	(health state* or health status).tw.	40,275
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.	138,384
57	(markov or time trade off or TTO or standard gamble or hrql or hrqol or disabilit* or disutilit*).tw.	129,972
58	(quality adj2 life).tw.	148,233
59	(decision adj2 model).tw.	3980
60	(visual analog* scale* or discrete choice experiment* or health* year* equivalen* or (willing* adj2 pay)).tw.	31,394
61	("resource use' or resource utili?ation).tw.	9307
62	(well-being or wellbeing).tw.	44,692
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	1,298,647
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

MEDLINE(R) In-Process & Other Non-Indexed Citations (via Ovid)

Exact database searched: 11 December 2014.

Date searched: 16 December 2014.

Search strategy

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	10,654
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.	54,972
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.	15,909
39	(markov or time trade off or TTO or standard gamble or hrql or hrqol or disabilit* or disutilit*).tw.	13,731
40	(quality adj2 life).tw.	17,497
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	101,172
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

EMBASE Classic and EMBASE (via Ovid)

Exact database searched: 1947 to 15 December 2014.

Date searched: 16 December 2014.

Search strategy

1	adalimumab.tw.	7509
2	*adalimumab/	4043
3	ADA.tw.	10,949
4	infliximab.tw.	13,814
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	36,094
14	lisa* tracker*.tw.	11
15	(immundiagnostik* or immunodiagnostik* or immunediagnostik*).tw.	76
16	(proteomika* or promonitor*).tw.	27
17	*enzyme linked immunosorbent assay/	14,705
18	enzyme* link* immunoassay*.tw.	3301
19	enzyme* link* immuno* assay*.tw.	72,608
20	ELISA*.tw.	169,424
21	14 or 15 or 16 or 17 or 18 or 19 or 20	210,314
22	*radioimmunoassay/	17,241
23	(radioimmuno* or radio immuno* or radio-immuno*).tw.	75,063
24	RIA.tw.	20,852
25	reporter* gene* assay*.tw.	4446
26	RGA.tw.	401
27	semi* fluid* phase* enzyme* immuno*.tw.	1
28	EIA.tw.	10,934
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	114,144
35	*crohn disease/	34,603
36	crohn*.tw.	50,590
37	inflammator* bowel* disease*.tw.	42,049
38	35 or 36 or 37	79,897
39	13 and 38	6882
40	21 and 38	2411
41	34 and 38	394
42	exp *health economics/	200,481
43	exp health status/	150,318
44	exp "quality of life"/	283,712
45	exp quality adjusted life year/	13,007
46	(pharmacoeconomic* or pharmaco-economic* or economic* or cost*).tw.	656,408
47	(health state* or health status).tw.	51,749
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.	195,997
49	(markov or time trade off or TTO or standard gamble or hrql or hrqol or disabilit* or disutilit*).tw.	189,075
50	(quality adj2 life).tw.	233,390
51	(decision adj2 model).tw.	5912
52	(visual analog* scale* or discrete choice experiment* or health* year* equivalen*).tw.	42,481

53	("resource use" or resource utili?ation).tw.	15,005
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	1,506,135
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 (via The Cochrane Library)

Date searched: 17 December 2014.

Search strategy

#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 #3	9
#43	#40 or #41 or #42	373

All results (373)

- Economic evaluations (30)

Science Citation Index 1970 – present (via Web of Knowledge)

Date searched: 17 December 2014.

Search strategy

#42	(#41) AND LANGUAGE: (English) Indexes=SCI-EXPANDED Timespan=All years	784
#41	#40 AND #39 Indexes=SCI-EXPANDED Timespan=All years	820
#40	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	1,328,585
#39	#38 OR #37 OR #36 Indexes=SCI-EXPANDED Timespan=All years	8339
#38	#34 AND #31 Indexes=SCI-EXPANDED Timespan=All years	246
#37	#34 AND #16 Indexes=SCI-EXPANDED Timespan=All years	1971

#36	#34 AND #9	6311
	Indexes=SCI-EXPANDED Timespan=All years	
#35	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"))	560
	Indexes=SCI-EXPANDED Timespan=All years	
#34	#33 OR #32	80,169
	Indexes=SCI-EXPANDED Timespan=All years	
#33	TS=(((inflammator* near/1 bowel*) near/1 disease*) or IBD)	52,825
	Indexes=SCI-EXPANDED Timespan=All years	
#32	TS=crohn*	50,019
	Indexes=SCI-EXPANDED Timespan=All years	
#31	#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	77,531
	Indexes=SCI-EXPANDED Timespan=All years	
#30	TS=(Prometheus* or "Anser IFX" or "Anser ADA")	588
	Indexes=SCI-EXPANDED Timespan=All years	
#29	TS=((Matriks* near/1 Biotek*) or Shikari*)	11
	Indexes=SCI-EXPANDED Timespan=All years	
#28	TS=(Biomonitor* or iLite)	8544
	Indexes=SCI-EXPANDED Timespan=All years	
#27	TS=HMSA	102
	Indexes=SCI-EXPANDED Timespan=All years	
#26	TS=((homogenous* or homogeneous*) near/1 (mobilit* near/1 (shift* near/1 assay*)))	13
	Indexes=SCI-EXPANDED Timespan=All years	
#25	TS=EIA	8367
	Indexes=SCI-EXPANDED Timespan=All years	
#24	TS=((semi* near/1 fluid*) near/3 (enzyme* near/1 immuno*))	1
	Indexes=SCI-EXPANDED Timespan=All years	
#23	TS=((semi* near/1 fluid*) near/2 (enzyme* near/1 immuno*))	0
	Indexes=SCI-EXPANDED Timespan=All years	
#22	TS=(semi* near/1 fluid* near/1 phase* near/1 enzyme* near/1 immuno*)	0
	Indexes=SCI-EXPANDED Timespan=All years	
#21	TS=(((semi* near/1 fluid*) near/1 phase*) near/1 (enzyme* near/1 immuno*))	0
	Indexes=SCI-EXPANDED Timespan=All years	
#20	TS=RGA	962
	Indexes=SCI-EXPANDED Timespan=All years	
#19	TS=(reporter* near/1 gene* near/1 assay*)	4550
	Indexes=SCI-EXPANDED Timespan=All years	
#18	TS=RIA	12,369
	Indexes=SCI-EXPANDED Timespan=All years	

#17	TS=(radioimmuno* or (radio near/1 immuno*) or radio-immuno*)	46,687
	Indexes=SCI-EXPANDED Timespan=All years	
#16	#15 OR #14 OR #13 OR #12 OR #11 OR #10	145,530
	Indexes=SCI-EXPANDED Timespan=All years	
#15	TS=ELISA*	112,098
	Indexes=SCI-EXPANDED Timespan=All years	
#14	TS=((enzyme* near/1 link*) near/1 (immuno* near/1 assay))	60,765
	Indexes=SCI-EXPANDED Timespan=All years	
#13	TS=((enzyme* near/1 link*) near/1 immunoassay*)	2846
	Indexes=SCI-EXPANDED Timespan=All years	
#12	TS=(proteomika* or promonitor*)	1
	Indexes=SCI-EXPANDED Timespan=All years	
#11	TS=(immundiagnostik* or immunodiagnostik* or immunediagnostik*)	10
	Indexes=SCI-EXPANDED Timespan=All years	
#10	TS=(lisa* near/1 tracker*)	1
	Indexes=SCI-EXPANDED Timespan=All years	
#9	#8 OR #7 OR #6 OR #5 OR #4 OR #3 OR #2 OR #1	31,622
	Indexes=SCI-EXPANDED Timespan=All years	
#8	TS=ADAb	31
	Indexes=SCI-EXPANDED Timespan=All years	
#7	TS=((anti* near/1 drug*) near/1 antibod*)	2570
	Indexes=SCI-EXPANDED Timespan=All years	
#6	TS=((anti* near/1 tumo$r*) near/1 (necrosis* near/1 factor*))	4119
	Indexes=SCI-EXPANDED Timespan=All years	
#5	TS=((anti-TNF* or antiTNF* or TNF*) near/2 inhibitor*)	4113
	Indexes=SCI-EXPANDED Timespan=All years	
#4	TS=IFX	381
	Indexes=SCI-EXPANDED Timespan=All years	
#3	TS=infliximab	13,827
	Indexes=SCI-EXPANDED Timespan=All years	
#2	TS=ADA	7173
	Indexes=SCI-EXPANDED Timespan=All years	
#1	TS=adalimumab	5046
	Indexes=SCI-EXPANDED Timespan=All years	

Cost-effectiveness Analysis Registry

Date searched: 17 December 2014.

Search strategy

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 (research papers in economics)

Date searched: 17 December 2014.

Search strategy

crohn* OR inflammatory bowel disease* among working papers and articles and books & chapters and software and authors (25)

School of Health and Related Research Health Utilities Database

Date searched: 17 December 2014.

Search strategy

crohn* in Any field

OR

inflammatory bowel disease* in Any field

Total: 1.

Appendix 4 Information provided by Theradiag/ Alpha Laboratories, Proteomika and Immundiagnostik

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 (Number 18 October 2013)
    - background
    - method
    - finding
    - summary.

Full text of abstracts:

- Unsworth 2013 – Measurement of infliximab and anti-infliximab antibodies analytical aspects and clinical implications
- Swart 2013 – Acceptance and adjustment in a districts general cohort of IBD patients: finding and implications
- 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 Laboratories (1) in manufacturer's 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 of IFX
4. Steenholdt 2013 *Gut* – IBD economic
5. Velayos 2013 *Clin 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 *Curr 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 *J Crohn Colitis* – consensus paediatric CD.

Presentation

The submission consists of presentation hand-outs on 'Monitoring antiTNF α drugs in chronic inflammatory disease-impact on tailoring therapies'.

LISA-TRACKER assays information

Detailed information about LISA-TRACKER assays from the company websites in two different languages, that is English and French:

(a) LISA-TRACKER Duo Infliximab (French)
(b) LISA-TRACKER Duo Infliximab (English)
(c) LISA-TRACKER anti-Infliximab (French)
(d) LISA-TRACKER anti-Infliximab (English)
(e) LISA-TRACKER Infliximab (French)
(f) LISA-TRACKER Infliximab (English)
(g) LISA-TRACKER Duo Adalimumab (French)
(h) LISA-TRACKER Duo Adalimumab (English)
(i) LISA-TRACKER anti-Adalimumab (French)
(j) LISA-TRACKER anti-Adalimumab (English)
(k) LISA-TRACKER Adalimumab (French)
(l) LISA-TRACKER Adalimumab (English).

Information from Proteomika

This submission consists of the following.

Annex 1

Lists of promonitor peer-reviewed articles (indexed in PubMed) ($n = 8$):

1. Chen DY, Chen YM, Tsai WC, Tseng JC, Chen YH, Hsieh CW, *et al.* Significant associations of anti-drug antibody levels with serum drug trough levels and therapeutic response of adalimumab and etanercept treatment in rheumatoid arthritis. *Ann Rheum Dis* 2014;**74**:e16.
2. Llinares-Tello F, Rosas J, de la Torre I, Valor L, Barber X, Senabre JM. Comparative study of both versions of an immunoassay commercialised for therapeutic drug monitoring of adalimumab in rheumatoid arthritis. *Reumatol Clin* 2014;**10**:105–8.
3. Llinares-Tello F, Rosas-Gómez de Salazar J, Senabre-Gallego JM, Santos-Soler G, Santos-Ramírez C, Salas-Heredia E, *et al.* Practical application of acid dissociation in monitoring patients treated with adalimumab. *Rheumatol Int* 2014;**34**:1701–8.
4. Llinares-Tello F, de Salazar JR, Gallego JM, Soler GS, Ramírez CS, Heredia ES, *et al.* Analytical and clinical evaluation of a new immunoassay for therapeutic drug monitoring of infliximab and adalimumab. *Clin Chem Lab Med* 2012;**50**:1845–7.
5. Mazilu D, Opriş D, Gainaru C, Iliuta M, Apetrei N, Luca G, *et al.* Monitoring drug and anti-drug 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;**2014**:702701.
6. Pascual-Salcedo D, Plasencia C, Ramiro S, Nuño L, Bonilla G, Nagore D, *et al.* Influence of immunogenicity on the efficacy of long-term treatment with infliximab in rheumatoid arthritis. *Rheumatology* 2011;**50**:1445–52.
7. Plasencia C, Pascual-Salcedo D, Nuño L, Bonilla G, Villalba A, Peiteado D, *et al.* Influence of immunogenicity on the efficacy of long-term treatment of spondyloarthritis with infliximab. *Ann Rheum Dis* 2012;**71**:1955–60.
8. Ruiz-Argüello B, del Agua AR, Torres N, Monasterio A, Martínez 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**:e287–9.

Annex 2

Lists of promonitor abstracts presented at international congresses in 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*. Proceedings of the American Academy of Allergy, Asthma & Immunology, 2014 Annual Meeting, 28 February–4 March, San Diego, CA, abstract no. 639. ABS 5.2.0 DTD Abstracts AB185.
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*. Proceedings of the American Gastroenterological Association (AGA), 2013, abstract no. Tu1173.
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*. Proceedings of the American Gastroenterological Association (AGA), 2013, abstract no. Tu1168.
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)*. Proceedings of the European Crohn's and Colitis Organisation (ECCO), Poster presentations: Clinical: Therapy and observation, 2013, abstract no. P508.
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*. Proceedings of Digestive and Liver Disease 45S, 19th National Congress, 2013, abstract no. P.03.13.
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*. Proceedings of the European Crohn's and Colitis Organisation (ECCO), Poster presentations: Clinical: Therapy and observation, 2013, abstract no. P569.
7. Diana M, Iliuta M, Gainaru C, Luca G, Apetrei N, Gudu T, *et al. Correlation between Serum Rituximab Level and Clinical Response in Rheumatoid Arthritis Patients Treated with B Cell Depletion Therapy*. Proceedings of the European League Against Rheumatism (EULAR), 2014 Annual meeting, Paris, France, abstract no. FRI0026.
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* 2014;**73**(Suppl. 2):715–16, abstract no. SAT0340.
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* 2013;**72**(Suppl. 3);237, abstract no. THU0215.
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* 2013;**72**(Suppl. 3):429, abstract no. FRI0171.
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*. Proceedings of the American College of Rheumatology (ACR), 2013 ACR/ARHP Annual Meeting, abstract no. 1444.
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* 2014;**73**(Suppl. 2):934–5, abstract no. AB0388.
13. Jauregui-Amezaga A, Ordas I, Gallego M, Ramirez A, Pino S, Masamunt MC, *et al.* [*Impacto de la medición de niveles de anti-TNF y título de anticuerpos contra el fármaco en el manejo de la enfermedad inflamatoria intestinal*.] XVI Reunión Anual de la Asociación Española de Gastroenterología, 2013, abstract no. 93, poster no. 50.

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*. Proceedings of the European Crohn's and Colitis Organisation (ECCO), Poster presentations: Clinical: Therapy and observation, 2013, abstract no. P481.
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*. Proceedings of the European Crohn's and Colitis Organisation (ECCO), Poster presentations: Clinical: Therapy & observation, 2014, abstract no. P302.
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* 2013;**72**(Suppl. 3):A234, abstract no. THU0207.
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* 2014;**73**(Suppl. 2):237–8, abstract no. THU0166.
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*. Proceedings of the 8th International Congress on Autoimmunity, 9–13 May 2012, Granada, Spain, abstract no. 609.
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* 2010;**69**(Suppl. 3):55, abstract no. OP0017.
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. *Ann Rheum Dis* 2014;**73**:946–7, abstract no. AB0422.
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* 2011;**70**(Suppl. 3):412, abstract no. FRI0207.
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. *Am Coll Rheumatol* 2011; abstract no. 2636.
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. *Int Congress Autoimmun* 2012; abstract no. 1590.
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 no. 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* 2011;**70**(Suppl. 3):82, abstract no. OP0045.
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. *Am Coll Rheumatol* 2012; abstract no. 546.
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. *Am Coll Rheumatol* 2011; abstract no. 2211.
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. *Int Congress Autoimmun* 2012; abstract no. 1568.
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* 2013;**72**(Suppl. 3):233, abstract no. THU0206.

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* 2014;**73**(Suppl. 2), abstract no. AB0389.
31. Ruiz del Agua A, Pascual-Salcedo D, Balsa A, Ramos I, Novalbos L, Ramiro S, *et al.* Monitoring of anti-TNF biological treatments. *J Transl Med* 2010;**8**(Suppl. 1):P32.
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* 2014;**73**(Suppl. 2): abstract no. FRI0265.
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* 2013;**72**(Suppl. 3):623, abstract no. SAT0125.
34. Valor L, Hernández D, de la Torre I, Llinares F, Rosas J, Yagüe J, *et al.* Infliximab and adalimumab levels and anti-drug antibodies detection in patients with rheumatoid arthritis (RA): an interlaboratory comparison using a commercial ELISA assay. *Ann Rheum Dis* 2014;**73**(Suppl. 2): abstract no. AB0396.

Full paper

Rosas 2014 *Clinical and Experimental Rheumatology*

Clinical relevance of monitoring serum levels of adalimumab in patients with rheumatoid arthritis in daily practice.

Presentation

Topic (confidential information has been removed).

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 25 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 2014 ($n = 5$):

1. Ghia *ACR* 2014–2436 – Analytical and clinical evaluation of an immunoassay for estimating immunogenicity of infliximab and etanercept in Indian population.
2. 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.
3. Llinares-Tello *ACR* 2014–1519 – Implementation of an acid dissociation procedure for immunogenicity detection in patients treated with anti-TNF drugs.
4. Opris *ACR* 2014–1539 – Relation between number of previous anti TNF agents and clinical response in rheumatoid arthritis patients treated with rituximab.
5. 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).

Information from Immundiagnostik/BioHit

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 anti-human 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 anti-drug 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 paediatric 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 e-mail after discussion)

Data on assays regarding the limit of blank.

Manual

There is a manual that provides information on technology in two different versions (i.e. English version and German version).

i. TNF-α Blocker adalimumab, total antibodies against adalimumab (e.g. Humira®) (German and English version).
ii. TNF-α Blocker adalimumab, antibodies against adalimumab (e.g. Humira®) (German and English version).
iii. TNF-α Blocker adalimumab, total antibodies against infliximab (e.g. Remicade®) (German and English version).
iv. TNF-α Blocker adalimumab, antibodies against infliximab (e.g. Remicade®) (German and English version).
v. TNF-α Blocker monitoring adalimumab drug level (e.g. Humira®) (German and English version).
vi. TNF-α Blocker monitoring infliximab drug level (e.g. Remicade®) (German and English version).

Request for information.

Appendix 5 Data extraction sheets

Data extraction form anti-tumour necrosis factor alpha 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.)		
First author surname		
Year of publication		
Country		
Study design		
Publication (full/abstract)		
Study setting		
Number of centres (by arm)		
Duration of study		
Follow-up period		
Funding		
Competing interests		
Aim of the study		
Inclusion/exclusion criteria for patients		
Test comparison		
Tests	*Name*	*Details*
Intervention test		
Comparison test 1		
Comparison test 2		
Comparison test 3		
Details of any repeat measurements (to check reliability, performance across different laboratories)		
Drug type tested		
IFX		
Anti-IFX		
ADA		
Anti-ADA		
Selection and storage of patients/plasma samples		
Description of method of selection		
Description of method and duration of storage		
Number of clinical samples		
Number of calibrator samples (spiked) for anti-TNF		

Selection and storage of patients/plasma samples
Number of calibrator samples (spiked) for antibodies
Number of blank (control) samples
Total number of plasma samples
Results of comparison *Name of test*
Threshold for drug
Number positive for drug
Threshold for antibodies
Number positive for antibodies
Details of correlation/overlap between the tests
Other information
Results of comparison of drug levels *Name of tests to be compared*
Total number concordant/all tested
Number of positive cases concordant/all positive cases
Number of negative cases concordant/all negative cases
Correlation of drug measurement
Regression method
Linearity test/cusum test?
R^2 (95% CI)
Slope (95% CI)
Intercept (95% CI)
From Bland–Altman plot for drug measurement
Percentage bias (95% CI)
Upper limit of agreement
Lower limit of agreement
Details of outliers
Visually is there a pattern between the mean value and the difference?
Results of comparison for antibody levels *Name of tests to be compared*
Total number concordant/all tested
Number of positive cases concordant/all positive cases
Number of negative cases concordant/all negative cases
Correlation of antibody measurement
Regression method
Linearity test/cusum test?
R^2 (95% CI)
Slope (95% CI)
Intercept (95% CI)

From Bland–Altman plot for antibody measurement	
Percentage bias (95% CI)	
Upper limit of agreement	
Lower limit of agreement	
Details of outliers	
Visually is there a pattern between the mean value and the difference?	
Authors' conclusion	
Reviewer's conclusion	

Data extraction form for anti-tumour necrosis factor alpha drug monitoring: management studies

Name of first reviewer: Deepson S Shyangdan.

Name of second reviewer: Martin Connock.

Study details	
Study identification number	[123]
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 weeks 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 CD and a previous beneficial clinical response to standard IFX 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 a 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 screened	95		
n excluded (ineligible)	26		
Randomisation / blinding	Randomised to algorithm or IFX 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 non-participants	14 not treated according to algorithm protocol (*n* = 7 continued IFX no assessment; *n* = 5 continued IFX no inflammation; and *n* = 2 misinterpreted analyses)		
Item	*IFX-intensified arm*	*Algorithm arm*	*All*
n study sample at baseline randomised (if applicable)	N/A	N/A	N/A
Withdrawals	8 (*n* = 7 lack of effect; *n* = 1 severe infusion reaction)	2 (lack of effect)	10
Lost to follow-up/dropouts (sample attrition)	Unclear	Unclear	Unclear

Study flow (CONSORT diagram)

Available in paper

Treatment algorithm for patients randomised to algorithm arm

	Detectable anti-IFX antibodies	**Undetectable anti-IFX antibodies**
Subtherapeutic IFX < 0.5 µg/ml	Group 1 Insufficient IFX bioavailability due to induced immunogenicity of IFX Change to different TNF-α-inhibitor: ADA 80 mg s.c. at inclusion followed by 40 mg s.c. every other week: dose intensification allowed	Group 2 Insufficient IFX bioavailability due to non-immune-mediated pharmacokinetics of IFX Intensify IFX treatment: IFX 5 mg/kg i.v. every 4 weeks
Therapeutic IFX ≥ 0.5 µg/ml	Group 4 Consider: Pharmacodynamics Non-functional anti-IFX antibodies False-positive test Repeat IFX and anti-IFX antibody analyses and handle accordingly. If unchanged results, then act as group 3	Group 3 Pharmacodynamics: inhibition of TNF-α is ineffective as a result of non-TNF-α-driven disease TNF-α-inhibitors not effective discontinued. Review of clinical condition at discretion of the investigator: if relapse of CD, use drug(s) with other target, for example conventional immune suppressives, glucocorticoids, and/or other biological agents. Consider surgery if appropriate. If no relapse, treat underlying problem

i.v., intravenously; s.c., subcutaneously.

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 PP (100)	33 ITT (100); 19 PP (58)	69 ITT (100); 55 PP (80)
Patient group (responders/secondary LOR)	Secondary LOR	Secondary LOR	Secondary LOR
Age (years), mean (range)	37 (19–63)	36 (19–81)	37 (19–81)
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
CDAI score, mean (range)	301 (230–487)	296 (221–526)	299 (221–526)
n (%) patients in remission	All patients at inclusion had recurrence of active disease		
n (%) patients with active CD			
CD classification (Vienna/Montreal)	Not clear	Not clear	Not clear
Disease duration (years), mean (range)	10 (1–35)	7 (1–27)	9 (1–35)
Smoking, *n* (%)	12 (33)	6 (18)	18 (26)
Previous surgery, *n* (%)	10 (28)	10 (30)	20 (29)
Concomitant treatment (specify),* n *(%)			
Immunosuppressants	14 (39)	13 (39)	27 (39)
Systemic corticosteroids or budesonide	1 (3)	1 (3)	2
Treatment duration at anti-TNF failure (days)	635 (range 97–1913)	681 (range 126–3313)	657 (range 97–3313)
Previous anti-TNF therapy	6 (17)	8 (24)	14 (20)
CRP (mg/ml)	6 (range 1–28)	9 (range 3–21)	9 (range 2–22)
Calprotectin (μg/g)	NR	NR	NR
Treatment			
Item	*IFX-intensified arm*		*Algorithm arm*
Anti-TNF drug (name)	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 weeks		Not clear planned 12 weeks

Intervention test assay (please specify):	
Manufacturer	RIA (probably Biomonitor A/S, Copenhagen, Denmark) Post hoc paper ELISA and HMSA (Prometheus Laboratories, San Diego, CA, 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 IFX/ADA (therapeutic/ subtherapeutic) (µg/ml)	• RIA: therapeutic ≥ 0.5 µg/l; subtherapeutic < 0.5 µg/l Post hoc • ELISA: 1.4 µg/ml for IFX • HMSA: therapeutic ≥3 µg/l; subtherapeutic <3 µg/l
Limit of quantification of anti-TNF antibodies [U/ml (arbitrary units/ml)] for antibodies detectable/ non-detectable	• RIA: 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 weeks (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; white blood cells change; haemoglobin change; albumin change
Timing of assessments (including information 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 that 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 IBDQ; PDAI; number of draining fistulas

Item	*Algorithm arm*	*Intensification arm*	*Comparison*
(A) Rates of response (coprimary outcome). Note that all patients started with secondary LOR	ITT: 19/33 (58%)	ITT: 19/36 (53%)	ITT: RR 1.09, 95% CI 0.713 to 1.673; $p = 0.810$; difference = 5% (–19% to 28%)
	PP: 9/19 (47%)	PP: 19/36 (53%)	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%)	ITT: 17/36 (47%)	ITT: RR 1.027, 95% CI 0.627 to 1.681; $p = 1.0$
	PP: 8/19 (42%)	PP: 17/36 (47%)	PP: RR 0.892, 95% CI 0.475 to 1.675; $p = 0.781$
(C) Clinical remission	ITT: 10/33 (30%)	ITT: 14/36 (39%)	ITT: RR 0.779, 95% CI 0.403 to 1.507; $p = 0.613$
	PP: 4/19 (29%)	PP: 14/36 (39%)	PP: RR 0.541, 95% CI 0.207 to 1.417; $p = 0.234$
Clinical response by subgroups, n (%)	***Algorithm arm***	***Intensification arm***	***Comparison***
Group 1 ($n = 14$; algorithm arm: $n = 5$; IFX-intensified arm: $n = 9$)	ITT: 2/5 (40)	ITT: 4/9 (44)	ITT: RR 0.90, 95% CI 0.246 to 3.297; $p = 1.00$
(Subtherapeutic IFX + detectable anti-IFX Abs + insufficient IFX bioavailability due to induce immunogenicity of IFX)	PP: 2/5 (40)	PP: 4/9 (44)	PP: RR 0.90, 95% CI 0.246 to 3.297; $p = 1.00$
Group 2 ($n = 3$: algorithm arm, $n = 1$; IFX-intensified arm, $n = 2$)	ITT: 0/1 (0)	ITT: 1/2 (50)	ITT: not calculable
(Subtherapeutic IFX + undetectable anti-IFX Abs + insufficient IFX bioavailability due to non-immune-mediated pharmacokinetics)	PP: 0/1 (0)	PP: 1/2 (50)	PP: not calculable
Group 3 ($n = 48$: algorithm arm, $n = 26$; IFX-intensified arm, n=22)	ITT: 16/26 (62)	ITT: 12/22 (55)	ITT: RR 1.128, 95% CI 0.693 to 1.837; $p = 0.770$
(Therapeutic IFX + undetectable anti-IFX Abs + inhibition of TNF-alpha ineffective due to non-TNF drive disease)	PP: 7/13 (54)	PP: 12/22 (55)	PP: RR 0.987, 95% CI 0.525 to 1.856; $p = 1.00$
Group 4 in algorithm ($n = 4$: algorithm arm, $n = 1$; IFX-intensified arm, $n = 3$)	ITT: 0/1 (0)	ITT: 2/3 (67)	ITT: not calculable
(Therapeutic IFX + detectable anti-IFX Abs + pharmacodynamics or non-functional anti-IFX Abs or FP test)	PP: 0/0	PP: 2/3 (67)	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:	***Algorithm arm***	***Intensification arm***	***Comparison***
Response	NR	NR	NR
Relapse	NR	NR	NR
Remission	NR	NR	NR
Rates of hospitalisation, *n* (%)	NR	NR	NR
Rates of surgical intervention, *n* (%)	NR	NR	NR
Time to surgical intervention, yes/no	NR	NR	NR
Health-related QoL, yes/no	Yes	Yes	Yes
Length of follow-up reported, yes/no	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 (subtherapeutic IFX)	Group 1: 5	Group 1: 9	Group 1: 14
	Group 2: 1	Group 2: 2	Group 2: 3
	Group 3: 0	Group 3: 0	Group 3: 0
	Group 4: 0	Group 4: 0	Group 4: 0
Mean anti-TNF (mg/m²/week) (SD)	NR		
Number of patients dose increased	Unclear treatments for group 3 of algorithm arm; all patients were increased in the dose intensification arm		
Number of patients dose reduced	Unclear, group 3 of algorithm arm should have stopped IFX but many did not		
Health-related QoL			
Item	*Algorithm arm, mean (standard error)*	*IFX-intensified arm, mean (standard error)*	*Mean difference*
PDAI score decrease from baseline	ITT: 2.4 (0.8)	ITT: 1.5 (0.7)	0.9 (95% CI −1.4 to 3.2); $p = 0.421$
	PP: 1.4 (0.5)	PP: 1.5 (0.7)	−0.1 (95% CI −2.1 to 1.9); $p = 0.911$
IBDQ score increase from baseline	ITT: 8.8 (1.7)	ITT: 8.8 (1.9)	0 (95% CI −5.1 to 5.2); $p = 0.996$
	PP: 5.4 (2.0)	PP: 8.8 (1.9)	−3.4 (95% CI −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; that is, IFX-intensified arm in which IFX treatment was intensified and an algorithm arm in which patients would receive interventions based on the serum IFX and IFX antibody 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 the 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

CONSORT, Consolidated Standards of Reporting Trials; FP, false positive; LOQ, limit of quantification; N/A, not applicable; NR, not reported.

Name of first reviewer: Martin Connock.

Name of second reviewer: Paul Sutcliffe.

Study details	
Study identification number	[73]
First author surname	Vande Casteele
Year of publication	2015
Country	Belgium
Study design	RCT
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 or not dosing based on therapeutic drug monitoring increases the rate of remission and whether or not continued concentration-based dosing is superior to clinically based dosing of IFX 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 • Aged ≥ 18 years • On IFX at least 14 weeks • Clinically stable
Exclusion criteria	Non-standard higher dosing regimen for secondary LOR to IFX 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 screened	Optimisation preceded randomisation		275
n excluded (ineligible)	Unclear	Unclear	24
n included for optimisation			263
n randomised	123	128	251
n of non-participants at study entry (those refused, etc.)	6 of 263 withdrew consent during 'optimisation' before randomisation; 6 further either developed LOR 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)	128 (71.1)	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 LOR)	Responders	Responders	Responders
Age (years), median (IQR)	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
CD activity; CRP (mg/l)	1.3 (IQR 0.6–4.5)	1.5 (IQR 0.7–4.0)	1.4 (IQR 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 (6.3–19.9)
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 *(%)			
Immunosuppressants	7 (5.7)	6 (4.7)	13 (5.2)
Systemic corticosteroids or budesonide	NR	NR	NR
Treatment duration at anti-TNF failure (days)	N/A	N/A	N/A
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)	IFX	IFX
Anti-TNF dose	Various based on clinical decisions (CRP and symptoms)	Various based on trough IFX testing
Duration of treatment	Patients entered on IFX	
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-α detection: limit of detection	In-house ELISA; reference provided to previous study	
	Lower limit of detection 0.3 µg/ml IFX	
Antibody detection: limit of detection	In-house ELISA; reference provided to previous study	
	Lower limit of detection 1.0 µg/ml IFX	
Outcomes reported *Item*		
Primary outcome(s)	Proportion with clinical (HBI $\leq$ 4 for CD, Mayo $\leq$ 2 UC) and biological (CRP $\leq$ 5 m/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), EQ-5D; costs of treatment	
Timing of assessments (including information on parallel or sequential)	Probably at each infusion, or every 8 weeks	
Time to test result	NR	
Number of inconclusive results, *n* (%)	Authors used defined cut-off points, inconclusive results = 0%	
Frequency of dose adjustment, *n* (%)	Unclear; both groups received dose adjustments, if required, during optimisation phase so as to bring drug trough levels with target range (3–7 µg/ml). 115 no adjustment, 76 dose escalated, 72 dose reduction (CD plus UC). Post-randomisation dose adjustments unclear	
Frequency of treatment switch, *n* (%)		
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 week 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 weeks)	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%) (RR 2.4, 95% CI 1.2 to 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 patients with CD = HBI score of ≤ 4 corresponds to remission. Biological remission = CRP concentration of ≤ 5 mg/l	
Rates of hospitalisation, *n* (%)	Two out of 263 hospitalised: one for appendectomy and one for ileostomy complications; both patients were in the clinically based dosing group
Rates of surgical intervention, *n* (%)	NR/unclear
Time to surgical intervention, yes/no	NR
Health-related QoL, yes/no	Yes
Length of follow-up reported, yes/no	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),* n *(%)*	*Concentration-based dosing (N=128),* n *(%)*
Adverse event		
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 IFX/ADA (therapeutic/subtherapeutic) (μg/ml)	Trough level defined groups at start of optimisation: 1. IFX < 0.3 μg/ml ADAb < 8 μg/ml 2. IFX < 3 μg/ml 3. IFX 3–7 μg/ml 4. IFX > 7 μg/ml

Dose monitoring		
Item (please define if necessary)		
Limit of quantification of anti-TNF antibodies	1.0 µg/ml (see above)	
Algorithm specified for management, yes/no (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 per week) (SD)	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 QoL		
Item	*Concentration based*	*Clinically based*
EQ-5D completed	Unclear	Unclear
Author's conclusion		
Targeting patients' IFX TCs to 3–7 µg/ml results in a more efficient use of the drug. After dose optimisation, 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		

ATI, antibodies to IFX; CONSORT, Consolidated Standards of Reporting Trials; N/A, not applicable; NR, not reported.

Name of first reviewer: Deepson S Shyangdan.

Name of second reviewer: Martin Connock.

Study details	
Study identification number	[128]
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, USA)
Number of centres (by arm)	One

Study details	
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 TCM of IFX-treated patients in clinical remission on IFX using dose adjustment based on testing to bring IFX 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 to a control group (i.e. that did not receive proactive 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, USA). 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 centre 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 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 screened	88 identified from Prometheus Laboratories data	84 identified from infusion centre
n excluded (ineligible)	14 did not meet inclusion criteria; 22 did not reach clinical remission; four patients did not have level when in remission or level was not a trough	10 did not reach clinical remission
n enrolled/included (eligible)	48 included as 'proactive TCM of IFX'	74 + 4 from PROMETHEUS record = 78
n non-participants at study entry (those refused, etc.)	N/A	N/A
n study sample at baseline randomised (if applicable)	N/A	N/A
Withdrawals	N/A	N/A
Lost to follow-up/drop outs (sample attrition)	N/A	N/A

Participants (characteristics and numbers)

Item	***TCM group***	***Control (non-TCM) group***
Total number of participants at baseline (% CD)	48 (38/48; 79%)	78 (45/78; 67%)
n (%) followed up	48	78
n (%) included in analysis	48	78
Patient group (responders/secondary LOR)	Responders (in remission)	Responders (in remission)

Study flow (CONSORT diagram)		
Participants (characteristics and numbers)		
Age, median (range), years	35 (29–42.5) at start of IFX therapy	34.9 (26.2–49.7) at start of IFX therapy
Sex (women), *n* (%)	15 (31)	33 (42)
Diagnostic criteria for CD	NR	NR
Children, *n* (%)	None	None
CDAI score, mean (SD)	NR	NR
n (%) patients in remission	All patients in remission	All patients in remission
n (%) patients with active CD		
CD classification (Vienna/Montreal)	NR	NR
Disease duration (years)	Not clear	Not clear
Smoking, *n* (%) – tobacco status	Current: 5 (10)	Current: 7 (9)
	Former: 12 (25)	Former: 14 (18)
	Never: 31 (56)	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* (%)	IFX (100%)	IFX (100%)
CRP (mg/ml)	NR	
Calprotectin (µg/g)	NR	
Treatment		
Item		
Anti-TNF drug (name)	IFX	
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, USA) (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-α detection: ELISA, HMSA details	NR	
Antibody 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-IFX treatment cessation TCM vs. control group	
Timing of assessments (including information 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; and dose unchanged 31/48	
	Subsequent trough tests: dose escalated 8/40; dose decreased 2/40; and dose unchanged 30/40	
Frequency of treatment switch, *n* (%)	N/A	
Measure of disease activity (e.g. CDAI, others?)	Physicians' judgement of remission assessed on medical notes; cessation of treatment	
Rates of:		
Response, yes/no	Time-to-event analysis of time to IFX treatment cessation. Others: no	
Relapse, yes/no		
Remission, yes/no		
Describe definition of progression:		
Equivalent to IFX 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 gastroenterologist's documentation')		
Duration of:		
Response	Time-to-event analysis of time to IFX treatment cessation	
Relapse		
Remission		
Rates of hospitalisation, *n* (%)	NR	
Rates of surgical intervention, *n* (%)	NR	
Health-related QoL, yes/no	No	
Length of follow-up reported, yes/no	Various; follow-up to IFX cessation in Kaplan–Meier analysis	
Incidence of adverse effects of treatment (reasons for stopping IFX)		
Item	*TCM group*	*No-TCM group*
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

Incidence of adverse effects of treatment (reasons for stopping IFX)		
Item	*TCM group*	*No-TCM group*
Delayed	1	0
Other (unrelated to IFX)	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 IFX/ADA (therapeutic/ subtherapeutic) (µg/ml)	Initially undetectable IFX was defined as subtherapeutic, later the target range of 5–10 µg/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 [U/ml (arbitrary unit/ml)] for Ab detectable/ non-detectable	HMSA for IFX IFX 1 µg/ml; ELISA 1.4 µg/ml	
Algorithm specified for management, yes/no (specify)	Yes. Algorithm resulted in dose increases of typically 50–100 mg (for a 70-kg patient receiving 5 mg/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/week) (SD)	Unclear	
Number of patients dose increased	See above	
Number of patients dose reduced	See above	
Health-related QoL		
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 proactive and non-proactive 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 proactive 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 with no TCM provided time-to-event outcomes for retention in IFX 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		

ATI, antibodies to IFX; CONSORT, Consolidated Standards of Reporting Trials; N/A, not applicable; NR, not reported; TCM, trough concentration monitoring.

Appendix 6 Excluded studies with reason

Full text exclusions with reason

TABLE 46 Full-text exclusions from the review with reason for exclusion

Reference	Reason for exclusion
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**:1133–9	C insufficient data M no algorithm specified/acted on
Baert F, Noman M, Vermeire S, Van Assche G, D'Haens G, Carbonez A, *et al.* Influence of immunogenicity on the long-term efficacy of infliximab in Crohn's disease. *N Engl J Med* 2003;**348**:601–8	C insufficient data
Balzola F, Bernstein C, Ho GT, Lees C. Clinical utility of measuring infliximab and human antichimeric antibody concentrations in patients with inflammatory bowel disease: commentary. *Inflamm Bowel Dis Monitor* 2010;**11**:85–6	Commentary no original data
Balzola F, Cullen G, Ho GT, Russell RK. Clinical utility of newly developed immunoassays for serum concentrations of adalimumab and anti-adalimumab antibodies in patients with Crohn's disease. *Inflamm Bowel Dis Monitor* 2013;**14**:19	Commentary no original data
Ben-Horin S, Chowers Y. Review article: loss of response to anti-TNF treatments in Crohn's disease. *Aliment Pharmacol Ther* 2011;**33**:987–95	Review without meta-analysis
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**:674–84	SR without meta-analysis
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**:1264–75	Review without meta-analysis
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**:971–86	SR without meta-analysis
Colombel JF, Feagan BG, Sandborn WJ, Van Assche G, Robinson AM. Therapeutic drug monitoring of biologics for inflammatory bowel disease. *Inflamm Bowel Dis* 2012;**18**:349–58	Review without meta-analysis
Ebert EC, Das KM, Mehta V, Rezac C. Non-response to infliximab may be due to innate neutralizing anti-tumour necrosis factor-alpha antibodies. *Clin Exp Immunol* 2008;**154**:325–31	Measurement of antibodies to TNF-α, not to anti-TNF-α drugs
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**:1947–55	> 50% RA patients
Hamalainen A, Sipponen T, Kolho KL. Serum infliximab concentrations in pediatric inflammatory bowel disease. *Scand J Gastroenterol* 2013;**48**:35–41	C insufficient data
Hibi T, Sakuraba A, Watanabe M, Motoya S, Ito H, Motegi K, *et al.* Retrieval of serum infliximab level by shortening the maintenance infusion interval is correlated with clinical efficacy in Crohn's disease. *Inflamm Bowel Dis* 2012;**18**:1480–7	C insufficient data
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**:447–59	SR without meta-analysis
Lazebnik LB, Sagynbaeva VE. [Level of adalimumab and its antibody titers define the effectiveness of the biological (anticytokine) therapy in Crohn's disease.] *Eksp Klin Gastroenterol* 2013;**7**:18–22	Non-English
Lichtenstein GR. Comprehensive review: antitumor necrosis factor agents in inflammatory bowel disease and factors implicated in treatment response. *Therap Adva Gastroenterol* 2013;**6**:269–93	SR without meta-analysis

continued

TABLE 46 Full-text exclusions from the review with reason for exclusion (*continued*)

Reference	Reason for exclusion
Malickova K, Duricova D, Bortlik M, Machkova N, Janatkova I, Lukas M. [Serum infliximab trough levels and induction of antibodies to infliximab during the biological treatment of patients with inflammatory bowel diseases.] *Alergie* 2011;**13**:216–22	Non-English
Rivero Marcotegui A, Ibanez Bosch R, Zuniga Vera A, Arin Letamendia A, Burusco Paternain MJ. [Clinical usefulness in measuring infliximab and human anti-chimeric antibodies.] *Rev Labor Clin* 2014;**7**:68–72	> 50% RA patients
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**:1250–6	C insufficient data M no algorithm specified/ acted on
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**:761–9	C insufficient data
Sono K, Yamada A, Yoshimatsu Y, Takada N, Suzuki Y. Factors associated with the loss of response to infliximab in patients with Crohn's disease. *Cytokine* 2012;**59**:410–16	C insufficient data
Steenholdt C, Svenson M, Bendtzen K, Thomsen OO, Brynskov J, Ainsworth MA. Severe infusion reactions to infliximab: aetiology, immunogenicity and risk factors in patients with inflammatory bowel disease. *Aliment Pharmacol Ther* 2011;**34**:51–8	C insufficient data
Ungar B, Chowers Y, Yavzori M, Picard O, Fudim E, Har-Noy O, *et al.* The temporal evolution of anti-drug antibodies in patients with inflammatory bowel disease treated with infliximab. *Gut* 2014;**63**:1258–64	C insufficient data
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. *Gastroenterol* 2008;**134**:1861–8	C insufficient data
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**:1226–31	C insufficient data
Yamada A, Sono K, Hosoe N, Takada N, Suzuki Y. 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* 2010;**16**:1898–904	C insufficient data
Yanai H, Hanauer SB. Assessing response and loss of response to biological therapies in IBD. *Am J Gastroenterol* 2011;**106**:685–98	Review without meta-analysis

C, correlation-type study; M, management-type study; RA, rheumatoid arthritis; SR, systematic review.

Excluded abstracts with reason

TABLE 47 Abstracts excluded from the review with reason for exclusion

Reference	Reason for exclusion
Abraham B, Chiorean M. False positive infliximab levels detected in patients treated with adalimumab for inflammatory bowel disease. *Am J Gastroenterol* 2012;**107**:S627	C insufficient data
Afif W, Loftus EV, Faubion WA, Hanson KA, Sandborn WJ. Clinical utility of measuring infliximab and human anti-chimeric antibody levels in patients with inflammatory bowel disease. *Gastroenterology* 2009;**136**(Suppl. 1):A147	Superseded by full text
Anonymous. New assay can detect infliximab levels and anti-infliximab antibodies from a single serum sample. *Clin Adv Hematol Oncol* 2012;**10**:27	Editorial no original data
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	M no algorithm specified/acted on

TABLE 47 Abstracts excluded from the review with reason for exclusion (*continued*)

Reference	Reason for exclusion
Arranz MDM, Arranz EM, Salcedo DP, De Diego C, Senent SG, Cordon JP, *et al.* Infliximab trough levels and antibodies: relationship with infusion reaction, immunomodulators and biological parameters. *Gastroenterology* 2014;**146**(Suppl. 1):243	C insufficient data
Baert FJ, Drobne D, Ballet V, Cleynen I, Compernolle G, Rutgeerts PJ, *et al.* Early trough levels and antibodies predict safety and success of restarting infliximab after long drug holiday. *Gastroenterology* 2011;**140**(Suppl. 1):62	C insufficient data
Baert FJ, Lockton S, Hauenstein S, Singh S, Gils A, Vermeire S. Antibodies to adalimumab predict inflammation in Crohn's patients on maintenance adalimumab therapy. *Gastroenterology* 2014;**146**(Suppl. 1):242	C insufficient data
Ben-Bassat O, Hauenstein S, Iacono A, Irwin SP, Singh S, Greenberg GR. Serum adalimumab and immunogenicity in IBD patients after 80 mg biweekly maintenance therapy. *Gastroenterology* 2013;**144**(Suppl. 1):771	C insufficient data
Ben-Horin S, Ungar B, Chowers Y, Yavzori M, Picard O, Fudim E, *et al.* The temporal evolution of anti-drug antibodies in IBD patients treated with infliximab. *J Gastroenterol Hepatol* 2013;**28**:145	C insufficient data
Bodini G, Savarino V, Dulbecco P, Baldissarro I, Savarino E. 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* 2014;**5**(Suppl. 1):238	C insufficient data
Bodini G, Savarino V, Dulbecco P, Baldissarro I, Savarino E. The influence of anti-adalimumab antibodies on adalimumab trough levels, TNF-alpha levels and clinical outcome. *J Crohns Colitis* 2014;**8**:S42	C insufficient data
Bodini G, Savarino V, Dulbecco P, Baldissarro I, Savarino EV. ELISA vs. HMSA: a comparison between two different methods for measuring adalimumab serum concentration and anti-adalimumab antibodies – preliminary data. *Dig Liver Dis* 2014;**46**:S67	Duplicate
Bodini G, Savarino V, Dulbecco P, Assandri L, Bruzzone L, Mazza F, *et al.* Correlation between adalimumab trough serum concentration, anti-adalimumab antibodies and TNF-alpha levels with clinical outcome in patients affected by Crohn's disease. *Gastroenterology* 2013;**144**(Suppl. 1):780	C insufficient data
Bodini G, Savarino V, Fazio V, Assandri L, Gemignani L, Dulbecco P, *et al.* Relationship between drug serum concentration and clinical activity in patients with Crohn's disease who achieved remission with adalimumab. *Dig Liver Dis* 2012;**44**:S69–70	Duplicate
Bodini G, Savarino V, Fazio V, Assandri L, Dulbecco P, Gemignani L, *et al.* Relationship between drug serum concentration and clinical activity in patients with Crohn's disease who achieved remission with adalimumab – a prospective study. *Gastroenterology* 2012;**142**(Suppl. 1):388	C insufficient data
Bortlik M, Duricova D, Malickova K, Komarek A, Machkova N, Bouzkova E, *et al.* Infliximab trough levels may predict sustained response to infliximab in patients with Crohn's disease: a single cohort study. *J Crohns Colitis* 2012;**6**:S153	Superseded by full text
Cardile S, Costa A, Loddo I, Morabito G, Pidone C, Romano C. Impact of measurement of infliximab and anti-infliximab antibodies levels in pediatric inflammatory bowel disease. *Dig Liver Dis* 2013;**45**:e294–5	C insufficient data
Chauhan U, Dutta U, Armstrong D, Greenwald E, Marshall J, Tse F, *et al.* Does measuring infliximab and human anti-chimeric antibody concentrations in patients with inflammatory bowel disease impact clinical management? A Canadian experience. *Inflamm Bowel Dis* 2012;**18**:S82–3	M no algorithm specified/acted on
Chauhan U, Dutta U, Armstrong D, Marshall J, Tse F, Greenwald E, *et al.* Does measuring IFX and human anti-chimeric antibody concentrations in patients with inflammatory bowel disease impact clinical management? A Canadian experience. *J Crohns Colitis* 2013;**7**:S228	Duplicate
Chollet-Martin S, Nicaise-Roland P, De Chaisemartin L, Grootenboer-Mignot S, Hayem G, Pelletier AL, *et al.* Simultaneous determination of anti-infliximab antibodies and residual infliximab levels to monitor anti-TNF therapy. *Ann Rheum Dis* 2013;**71**:666	Not M, C or ATC

continued

TABLE 47 Abstracts excluded from the review with reason for exclusion (*continued*)

Reference	Reason for exclusion
Church P, Guan J, Frost K, Muise A, Walters T, Griffiths A. Infliximab treatment for paediatric Crohn's disease: long-term outcomes at a single centre. *J Crohns Colitis* 2013;**7**:S198	Not M, C or ATC
Church P, Guan J, Salz L, Frost K, Muise A, Walters T, *et al.* Long-term outcomes with infliximab treatment in children with Crohn's disease at a single centre. *Inflamm Bowel Dis* 2012;**18**:S72–7	C insufficient data
Church P, Guan J, Salz L, Frost K, Muise A, Walters T, *et al.* Long-term outcomes with infliximab treatment in children with Crohn's disease at a single centre. *Inflamm Bowel Dis* 2012;**18**:S5–6	Duplicate
Cornillie F, Hanauer S, Diamond R, Wang J, Zelinger D, Xu Z, *et al.* 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. *Am J Gastroenterol* 2011;**106**:S462–3	C insufficient data
Corstjens PL, Wiesmeijer K, Wolbink GJ, Tanke J, Hommes DW, Fidder H. A rapid test for quantitative determination of infliximab trough levels in blood. *Gastroenterology* 2011;**14**:S276–7	C insufficient data
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. *Gastroenterology* 2013;**144**(Suppl. 1):781	C insufficient data
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. *J Crohns Colitis* 2013;**7**:S239	C insufficient data
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. *Dig Liver Dis* 2013;**45**:S109	C insufficient data
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 ImmunDiagnostik tests. *Gastroenterology* 2013;**144**(Suppl. 1):780	Duplicate
Daperno M, Fracchia M, Guiotto C, Germano L, Ercole E, Rigazio C, *et al.* Clinical implications and stability of determination of infliximab trough levels (IFX-TL) and antibodies to infliximab (ATI) in inflammatory bowel disease. *Dig Liver Dis* 2013;**45**:S145	C insufficient data
De Bruyn M, Bessissow T, Billiet T, Cleynen I, Kirkland R, Liu X, *et al.* Biomarker panel for prediction of mucosal healing in patients with Crohn's disease under infliximab therapy. *J Crohns Colitis* 2014;**8**:S45–6	C insufficient data
De Bruyn M, Bessissow T, Billiet T, Cleynen I, Kirkland R, Liu X, *et al.* Biomarker panel for prediction of mucosal healing in patients with Crohn's disease under infliximab therapy. *Gastroenterology* 2014;**146**(Suppl. 1):428	Duplicate
Dotan I, Yanai H, Ron Y, Kariv R, Fishman S, Yahav L, *et al.* Population pharmacokinetic evaluation of adalimumab reveals patient factors that increase adalimumab clearance and shorten half-life in inflammatory bowel disease patients. *Gastroenterology* 2014;**146**(Suppl. 1):243	C insufficient data
Dotan I, Ron Y, Yanai H, Becker SA, Fishman S, Yahav L, *et al.* Population pharmacokinetic evaluation of infliximab reveals patient factors that increase infliximab clearance and shorten half-life in inflammatory bowel disease patients. *Gastroenterology* 2013;**144**(Suppl. 1):774	C insufficient data
Drastich P, Kozeluhova J, Jaresova M, Spicak J. Infliximab serum trough levels and deep remission in patients with IBD. *Gastroenterology* 2011;**140**(Suppl. 1):292	C insufficient data
Drobne D, Bossuyt P, Breynaert C, Vande Casteele N, Compernolle G, Juergens M, *et al.* Long term evolution and impact of immunomodulator co-treatment and withdrawal on infliximab trough levels in 223 patients with Crohn's disease. *J Crohns Colitis* 2011;**5**:S10–11	M no algorithm specified/acted on

TABLE 47 Abstracts excluded from the review with reason for exclusion (*continued*)

Reference	Reason for exclusion
Drobne D, Bossuyt PJ, Breynaert C, Casteele NV, Compernolle G, Jurgens M, *et al.* Crohn's disease: infliximab trough levels and CRP during infliximab–immunomodulator combination treatment are associated with clinical outcome after immunomodulator withdrawal. *Gastroenterology* 2011;**140**(Suppl. 1):62	C insufficient data
Duricova D, Malickova K, Bortlik M, Machkova N, Komarek V, Bouzkova E, *et al.* Predictors of sustained response to infliximab in patients with Crohn's disease: a single cohort study. *Gastroenterology* 2011;**140**(Suppl. 1):593	C insufficient data
Echarri A, Ferreiro R, Fraga-Iriso R, Barreiro-De Acosta M, Cid J, De-Castro L, *et al.* Drug trough levels and primary nonresponse to antiTNF therapy in moderate-severe Crohn disease. Results of the optimiza study. *Gastroenterology* 2014;**146**(Suppl. 1):247	C insufficient data
Eser A, Primas C, Shringarpure R, Hauenstein S, Wang SL, Reinisch W. 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. *Am J Gastroenterol* 2012;**107**:S657	Duplicate
Fasanmade AA, Wagner C, Davis H, Graham M, Everitt D, Gottlieb A. Comparison of the pharmacokinetics of infliximab in patients with psoriasis or Crohn's disease not receiving concomitant immunosuppressants or corticosteroids. *J Invest Dermatol* 2002;**119**:243	Not C, M or ATC
Fasanmade AA, Marsters P, Munsanje E, Graham MA, Davis HM, Van Deventer S. Infliximab pharmacokinetics and improvement in fistulizing Crohn's Disease. *Gastroenterology* 2003;**124**:A61	C insufficient data
Fasanmade AA, Zhu YW, Wagner C, Pendley C, Davis HM. Population pharmacokinetics of single dose infliximab in patients with Crohn's disease. *Clin Pharmacol Ther* 2002;**71**:P66	Not C, M or ATC
Fasanmade A, Olson A, Bao W, Pendley C, Davis H, Mayer L. Relationship between infliximab pharmacokinetics and improvement in Crohns disease. *Gastroenterology* 2002;**122**:A617–18	C insufficient data
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. *Ann Rheum Dis* 2012;**71**(Suppl. 3):634–5	Superseded by full paper
Garces S, Demengeot J, Wolbink GJ, Aarden L, Benito-Garcia E. 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 Rheum* 2011;**63**(Suppl. 10):464	Superseded by full paper
Garces S, Demengeot J, Da Silva JC, Aarden L. Bridging ELISA as a screening assay to monitor immunogenicity in routine clinical practice. *Arthritis Rheum* 2011;**63**(Suppl. 10):1841	< 50% CD
Garces S, Demengeot J, Canas-da-Silva J, Aarden L. Bridging ELISA as a screening assay to monitor immunogenicity in routine clinical practice. *Ann Rheum Dis* 2013;**71**:711	Duplicate
Garces S, Freitas J, Canas-Silva J, Aarden L, Demengeot J. The impact of immunogenicity on drug safety profile. *Ann Rheum Dis* 2013;**72**(Suppl. 3):A436	C insufficient data
Garimella TS, Peng JZ, Beck K, Noertersheuser PA, Lomax KG, Paulson SK, *et al.* 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* 2006;**130**:A481	C insufficient data
Guilday C, Eastwood D, Zadvornova Y, Stein D, Naik AS, Best K, *et al.* Concomitant use of immunomodulator therapy results in higher serum infliximab levels compared to monotherapy without lowering serum haca levels. *Gastroenterology* 2013;**144**(Suppl. 1):429	C insufficient data
Guiotto C, Germano L, Vizzini M, Cerruti R, Frigerio F, Daperno M, *et al.* Determination of infliximab trough levels (IFX-TL) and antibodies to infliximab (ATI) in inflammatory bowel disease. *Biochim Clin* 2013;**37**:S475	Superseded

continued

TABLE 47 Abstracts excluded from the review with reason for exclusion (*continued*)

Reference	Reason for exclusion
Hadigan CBR, Braegger CP, Vasilauskis E, Escher JC, Sinaasappel M, Ferry GD, *et al.* Pharmacokinetics of infliximab (Anti-TNFx) in children with Crohn's disease: a multicenter trial. *J Pediatr Gastroenterol Nutr* 1999;**29**;525	C insufficient data
Hauenstein S, Salbato J, Lockton S, Singh S. Characterization of neutralizing anti-drug antibody response in patients with loss of response to anti-TNF therapy. *Gastroenterology* 2013;**144**(Suppl. 1):418	C insufficient data
Hayes Inc. *Use of Anti-Infliximab Antibody Levels to Monitor Infliximab Treatment in Patients with Inflammatory Bowel Disease (IBD)*. Health Technology Brief Publication. Lansdale, PA: Hayes Inc.; 2013	Not available
Hester KD, Liu X, Salbato J, Lockton S, Hauenstein S, Singh S. Improved homogeneous mobility shift assay (HMSA) for the detection of neutralizing antibodies (NAB) in IBD patients treated with infliximab or adalimumab. *Gastroenterology* 2014;**146**(Suppl. 1):248	Not M, C or ATC
Hibi T, Sakuraba A, Watanabe M, Motoya S, Ito H, Sato N, *et al.* 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* 2012;**142**(Suppl. 1):388	C insufficient data
Hoekman DR, Brandse JF, De Meij T, Hummel T, Lowenberg M, Benninga MA, *et al.* Large variation in infliximab trough levels is associated with disease activity in paediatric inflammatory bowel disease. *Gastroenterology* 2014;**146**(Suppl. 1):782	C insufficient data
Hoekman D, Brandse H, De Meij T, Hummel T, Lowenberg M, Benninga M, *et al.* Large variation in infliximab trough levels is associated with disease activity in paediatric inflammatory bowel disease. *J Crohns Colitis* 2014;**8**:S35	C insufficient data
Huang VW, Prosser C, Shalapay C, Fedorak DK, Dhami N, Wang H, *et al.* In IBD outpatients knowledge of fecal calprotectin and infliximab trough levels significantly enhances infliximab dose escalation decision making. *J Crohns Colitis* 2014;**8**:S255	M no algorithm specified/ acted on
Huang VW, Dhami N, Fedorak DK, Prosser C, Shalapay C, Kroeker KI, *et al.* Disparity between infliximab trough level and infliximab associated adverse events. *J Crohns Colitis* 2014;**8**:S282	C insufficient data
Huang V, Kroeker KI, Wang H, Prosser C, Carol S, Dhami N, *et al.* In IBD outpatients knowledge of fecal calprotectin and infliximab trough levels significantly alters clinical decision making. *Gastroenterology* 2014;**146**(Suppl. 1):241–2	Duplicate
Huang V, Dhami N, Fedorak DK, Prosser C, Carol S, Kroeker KI, *et al.* Infliximab trough levels are correlated with infliximab-associated adverse events. *Gastroenterology* 2014;**146**(Suppl. 1):1	Not M, C or ATC
Imaeda H, Andoh A, Ban H, Bamba S, Sasaki M, Tsujikawa T, *et al.* 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* 2012;**142**(Suppl. 1):349	Superseded by full text
Imaeda H, Andoh A, Takahashi K, Fujimoto T, Ban H, Bamba S, *et al.* Serum infliximab trough levels above 1.0 mg/ml are required to obtain clinical efficacy in patients with Crohn's disease. *Inflamm Bowel Dis* 2012;**18**:S59–60	C insufficient data
Imaeda H, Andoh A, Bamba S, Tsujikawa T, Fujiyama Y. Development of a new immunoassay for the accurate determination of anti-infliximab antibodies in Crohn's disease. *Inflamm Bowel Dis* 2011;**17**:S42	Duplicate
Imaeda H, Takahashi K, Fujimoto T, Bamba S, Sasaki M, Tsujikawa T, *et al.* Accurate determination of serum adalimumab and anti-adalimumab antibodies levels during maintenance therapy for Crohn's disease. *Gastroenterology* 2013;**144**(Suppl. 1):431	Not M, C or ATC
Irving PM, Arkir Z, Duncan J, Sastrillo M, Anderson S, Sanderson J. Initial experience with infliximab levels in a tertiary IBD centre. *Gut* 2012;**61**:A238	C insufficient data
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. *J Crohns Colitis* 2013;**7**:S202–3	C insufficient data

TABLE 47 Abstracts excluded from the review with reason for exclusion (*continued*)

Reference	Reason for exclusion
Juan G, Alvarino 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. *J Crohns Colitis* 2014;**8**:S190	C insufficient data; M no algorithm specified/acted on
Karmiris K, Paintaud G, Degenne D, Ferrante M, Duveau AC, Noman M, *et al.* 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? *Gastroenterol* 2008;**134**:A68	C insufficient data
Karmiris K, Noman M, Paintaud G, Ferrante M, Duveau AC, Degenne D, *et al.* 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* 2008;**134**:A640	C insufficient data
Karsan SS, Cohen ER, Targan SR, Ippoliti A, Shih DQ, Vasiliauskas EA, *et al.* 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* 2012;**142**(Suppl. 1):264	C insufficient data M no algorithm specified/acted on
Kerr J, Nair A, Hinds R. Variable practice in children with inflammatory bowel disease requiring infliximab infusions across Australia. *J Gastroenterol Hepatol* 2013;**28**:141	M no algorithm specified/acted on
Kong JY, Bundell C, Pawlik J, Hollingsworth P, Forbes G. Low trough serum infliximab and antibodies to infliximab in smokers. *Inflamm Bowel Dis* 2013;**19**:E35–6	C insufficient data
Kong JY, Bundell CS, Pawlik J, Hollingsworth PN, Forbes GM. Smoking is associated with low trough serum infliximab levels and presence of anti-infliximab antibody in maintenance treatment of inflammatory bowel disease (IBD). *J Gastroenterol Hepatol* 2011;**26**:59	C insufficient data
Lamblin C, Aubourg A, Ternant D, Picon L, Lecomte T, Paintaud G. Concentration effect relationship of infliximab in Crohn's disease: results of a cohort study. *J Crohns Colitis* 2012;**6**:S142–3	C insufficient data
Leclerc M, Marotte H, Paul S, Del Tedesco E, Gonzalo P, Phelip JM, *et al.* Persistence of antibodies to infliximab for more than two months strongly predicts loss of response to infliximab in inflammatory bowel diseases. *J Crohns Colitis* 2014;**8**:S226–7	C insufficient data
Li JL, Paulson SK, Chiu YL, Robinson A, Lomax KG, Pollack PF. Evaluation of potential correlations between serum adalimumab concentration and remission in patients with Crohn's disease in classic I and II. *Gastroenterology* 2010;**138**(Suppl. 1):101	C insufficient data
Lowenberg M, Brandse J, Vos L, Ponsioen C, Van Den Brink G, D'Haens G. High infliximab trough levels are associated with impaired quality of life in IBD patients in clinical and biochemical remission on maintenance IFX therapy. *J Crohns Colitis* 2014;**8**:S262–3	C insufficient data
Lowenberg M, Brandse JF, Vos LM, Ponsioen C, Van Den Brink GR, D'Haens GR. High infliximab trough levels are associated with impaired quality of life in IBD patients in clinical and biochemical remission on maintenance IFX therapy. *Gastroenterology* 2014;**146**(Suppl. 1):450	C insufficient data
Lukas M, Malickova K, Bortlik M, Duricova D. Anti-infliximab antibodies in routine clinical practice – is it worth to assess them? *Gastroenterology* 2009;**136**(Suppl. 1):A679	C insufficient data
Malickova K, Janatkova I, Duricova D, Bortlik M, Lukas M. Serum infliximab levels, antibodies to infliximab and albumin concentrations during infliximab treatment in patients with inflammatory bowel disease. *Clin Exp Rheum* 2011;**29**:213	C insufficient data
Martin Arranz MD, Martin Arranz E, Pascual-Salcedo D, De Diego C, Jaquotot M, Gomez Senent S, *et al.* Infliximab trough levels and antibodies: relationship with infusion reaction, immunomodulators and biological parameters. *J Crohns Colitis* 2014;**8**:S251	C insufficient data
Mazor Y, Koplov U, Ben Hur D, 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. *J Crohns Colitis* 2013;**7**:S217	C insufficient data
McTigue M, Sandborn W, Levesque B, Patel D. Infliximab therapeutic drug monitoring in clinical practice: indications and utility. *Am J Gastroenterol* 2013;**108**:S512	C insufficient data

continued

TABLE 47 Abstracts excluded from the review with reason for exclusion (*continued*)

Reference	Reason for exclusion
Morgenstern J, Baestlein E, Leifeld L, Nguyen P, Stein J, Kruis W. Infliximab drug levels in Crohn's disease responding to the treatment. *J Crohns Colitis* 2012;**6**:S125	C insufficient data
Noman M, Baert F, Vermeire S, Van Assche G, D'Haens G, Carbonez A, *et al.* Post infusion infliximab levels determine duration of response in Crohn's disease and are directly related to infusion reactions. *Gastroenterology* 2002;**122**:A100	C insufficient data
O'Donnell S, Stempak JM, Silverberg MS. Is there a higher rate of infliximab dose optimization in initial responders between UC and CD cases? *Gastroenterology* 2014;**146**(Suppl. 1):462–3	C insufficient data; M no algorithm specified/acted on
Papamichail K, Casteele NV, Hauenstein S, Princen F, Singh S, Ferrante M, *et al.* Prediction of sustained remission after discontinuation of infliximab in patients with Crohn's disease. *Gastroenterology* 2014;**146**(Suppl. 1):457	C insufficient data
Pariente B, De Chambrun GP, Desroches M, De Cassan C, Gornet JM, Desreumaux P, *et al.* Clinical value of measuring trough levels and human anti-chimeric antibodies in patients with inflammatory bowel disease who lost response to infliximab therapy. *Gastroenterology* 2011;**140**(Suppl. 1):277	Superseded by full text
Pariente B, Pineton De Chambrun G, Desroches M, De Cassan C, Gornet J, Desreumaux P, *et al.* Clinical value of measuring trough levels and human anti-chimeric antibodies in patients with inflammatory bowel disease who lost response to infliximab therapy. *J Crohns Colitis* 2011;**5**:S111–12	Duplicate
Paul S, Del Tedesco E, Marotte H, Clavel L, Phelip JM, Peyrin-Biroulet L, *et al.* Therapeutic drug monitoring of infliximab and mucosal healing in inflammatory bowel disease: a prospective study. *Gastroenterol* 2013;**144**(Suppl. 1):92	Superseded by full text
Paul S, Del Tedesco E, Marotte H, Rinaudo-Gaujous M, Phelip JM, Peyrin-Biroulet L, *et al.* Infliximab concentration is associated with mucosal healing in intestinal bowel disease (IBD). *J Crohns Colitis* 2013;**7**:S199–200	C insufficient data
Paulson SK, Nocrtersheuser P, Pollack PF, Hoffman RS. Pharmacokinetics of adalimumab from classic, a randomized phase 3 trial for the induction of clinical remission in patients with Crohn's. *Gastroenterology* 2005;**128**:A585	C insufficient data
Pradhan RS, Sharma S, Thakkar R, Robinson A, Hyams JS, Rosh JR, *et al.* Relationship between adalimumab concentration and efficacy for the induction of clinical remission in pediatric patients with moderate to severe Crohn's disease. *Gastroenterology* 2013;**144**(Suppl. 1):230–1	Duplicate
Pradhan R, Sharma S, Thakkar R, Robinson A, Hyams J, Rosh J, *et al.* Relationship between adalimumab concentration and efficacy for the induction of clinical remission in pediatric patients with moderate to severe Crohn's disease. *J Crohns Colitis* 2013;**7**:S164	C insufficient data
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	M no algorithm specified/acted on
Rekvig M, Gedde Dahl M, Bratlie J, Bolstad N, Moum B, Jahnsen J, *et al.* Anti-TNFalpha drug level measurements in IBD patients. *J Crohns Colitis* 2014;**8**:S301–2	C insufficient data
Roblin X, Marotte H, Del Tedesco E, Rinaudo-gaujous M, Phelip JM, Paul S. Residual adalimumab trough levels are associated with clinical remission and mucosal healing in IBD. *Gastroenterology* 2013;**144**(Suppl. 1):778	C insufficient data; M no algorithm specified/acted on
Roblin X, Rinaudo M, Del Tedesco E, Phelip JM, Peyrin Biroulet L, Paul S. Development of an algorithm incorporating pharmacokinetics of adalimumab in inflammatory bowel diseases. *J Crohns Colitis* 2014;**8**:S41	C insufficient data; M no algorithm specified/acted on
Roblin X, Rinaudo-gaujous 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. *Gastroenterology* 2014;**146**(Suppl. 1):150	Duplicate
Rosenthal C, Melmed G, Tripuraneni B, Gebbia J, Callejas S, Farrior S, *et al.* Early infliximab trough levels predict remission at one year in pediatric IBD patients. *Inflamm Bowel Dis* 2012;**18**:S81	C insufficient data

TABLE 47 Abstracts excluded from the review with reason for exclusion (*continued*)

Reference	Reason for exclusion
Rubin D, Hauenstein S, Singh S. Post-marketing review of serum adalimumab and antibodies to adalimumab using the mobility shift assay platform. *Am J Gastroenterol* 2013;**108**:S532	C insufficient data
Scaldaferri F, Pecere S, Petito V, Cammarota G, Campanale MC, Rapaccini GL, *et al.* Infliximab and TNF alfa measurement in intestinal mucosa of IBD patients: a new tool for the clinic? *Dig Liver Dis* 2012;**44**:S189–90	C insufficient data
Schatz SB, Prell C, Freudenberg F, Schwerd T, Bufler P, Koletzko S. Correlation of infliximab levels and antibodies with clinical outcome in children with IBD. *J Pediatr Gastroenterol Nutr* 2011;**52**:E45	C insufficient data
Semmler JM, Pilch A, Armbruster FP, Dignass A, Kruis W, Stein J. Development of a new immunoassay for the accurate determination of anti-infliximab antibodies in inflammatory bowel disease. *J Crohns Colitis* 2014;**8**:S41	Duplicate
Semmler J, Pilch A, Armbruster FP, 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* 2014;**52**:S1008	Duplicate
Settesoldi A, Giannotta M, Milla M, Genise S, Santini A, Bagnoli S, *et al.* Loss of efficacy and adverse drug reactions during infliximab therapy in IBD patients are related to the appearance of anti-infliximab antibodies. *J Crohns Colitis* 2012;**6**:S151	Duplicate
Settesoldi A, Giannotta M, Genise S, Santini A, Bagnoli S, Milla M, *et al.* Loss of efficacy and adverse drug reactions during infliximab therapy in IBD patients are related to the appearance of anti-infliximab antibodies. *Dig Liver Dis* 2012;**44**:S194	C insufficient data
Sharma S, Pradhan R, Thakkar R, Robinson A, Hyams J, Rosh J, *et al.* Relationship between adalimumab concentration and efficacy for the maintenance of clinical remission in pediatric patients with moderate to severe Crohn's disease. *J Crohns Colitis* 2013;**7**:S163	C insufficient data
Sharma S, Pradhan RS, Thakkar R, Robinson A, Hyams JS, Rosh JR, *et al.* Relationship between adalimumab concentration and efficacy for the maintenance of clinical remission in pediatric patients with moderate to severe Crohn's disease. *Gastroenterology* 2013;**144**(Suppl. 1):231	C insufficient data
Sorrentino D, Hauenstein S, Marino M, Lockton S, Zarifi D, Del Bianco T, *et al.* 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* 2013;**144**(Suppl. 1):777	C insufficient data
Steenholdt C, Brynskov J, Thomsen O, Munck LK, Fallingborg J, Christensen LA, *et al.* 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* 2013;**144**(Suppl. 1):22	Superseded by full paper
Steenholdt C, Brynskov J, Thomsen O, Munck LK, Fallingborg J, Christensen LA, *et al.* Secondary infliximab treatment failure in Crohn's disease: therapeutic implications of measuring drug and anti-drug antibodies by three different binding assays. *Gastroenterology* 2013;**144**(Suppl. 1):773	Duplicate
Steenholdt C, Brynskov J, Thomsen O, Munck L, Fallingborg J, Christensen L, *et al.* Secondary infliximab treatment failure in Crohn's disease: therapeutic implications of measuring drug and anti-drug antibodies by three different binding assays. *J Crohns Colitis* 2013;**7**:S159	Superseded by full text
Steenholdt C, Brynskov J, Thomsen O, Munck L, Fallingborg J, Christensen L, *et al.* Comparison of techniques for monitoring infliximab and antibodies to infliximab in Crohn's disease patients with infliximab treatment failure. *Am J Gastroenterol* 2012;**107**:S622	Superseded by full text
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. *J Crohns Colitis* 2014;**8**:S291	Duplicate

continued

TABLE 47 Abstracts excluded from the review with reason for exclusion (*continued*)

Reference	Reason for exclusion
Steenholdt C, Bendtzen K, Brynskov J, Thomsen O, Ainsworth MA. Clinical implications of measuring drug and anti-drug antibodies by different assays when optimizing infliximab treatment failure in Crohn's disease. *Gastroenterology* 2014;**146**(Suppl. 1):240	Superseded by full text
Steenholdt C, Bendtzen K, Thomsen OO, Brynskov J, Ainsworth M. 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. *Scand J Gastroenterol* 2010;**45**:59	Superseded by full text
Steenholdt C, Svenson M, Bendtzen K, Thomsen O, Brynskov J, Ainsworth MA. Can measurements of anti-infliximab antibodies predict acute severe infusion reactions to infliximab? *Gastroenterology* 2011;**140**(Suppl. 1):774	C insufficient data
Steenholdt C, Svenson M, Ainsworth MA, Thomsen O, Brynskov J, Bendtzen K. Comparison of techniques for monitoring infliximab bioavailability and immunogenicity in Crohn's disease. *Gastroenterology* 2012;**142**(Suppl. 1):781	ATC insufficient data
Steenholdt C, Thomsen OO, Brynskov J, Bendtzen K, Ainsworth MA. 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* 2010;**138**(Suppl. 1):687–8	Insufficient data
Steenholdt C, Palarasah Y, Bendtzen K. Teisner A, Teisner B, Brynskov J, *et al.* Pre-existing IgG antibodies to the Fab region of infliximab predict efficacy and safety in IBD patients naive to anti-TNF agents. *J Crohns Colitis* 2013;**7**:S6–S6	C insufficient data
Steenholdt C, Palarasah Y, Bendtzen K, Teisner A, Brynskov J, Teisner B, *et al.* Pre-existing IGG antibodies to the fab region of infliximab predict efficacy and safety in IBD patients naive to anti-TNF agents. *Scand J Immunol* 2013;**77**:333	C insufficient data
Szepes Z, Kunstar E, Farkas K, Nagy F, Gyulai R, Kui R, *et al.* 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. *J Crohns Colitis* 2013;**7**:S118–19	C insufficient data
Tang J, Gao X, Zhi M, Zhou H, Chen H, Zhang M, *et al.* Serum infliximab levels and early mucosal healing in Crohn's disease. *J Crohns Colitis* 2014;**8**:S209–10	C insufficient data
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;**144**(Suppl. 1):531	M no algorithm specified/ acted on
Ungar B, Anafy A, Yavzori M, Picard O, Fudim E, Kopylov U, *et al.* The clinical and immunological significance of low level of infliximab in the absence of anti-infliximab antibodies in patients with IBD. *J Crohns Colitis* 2014;**8**:S113	Duplicate
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	C insufficient data
Ussia V, Ceccarelli L, Maltinti S, Di Fluri G, Mumolo MG, Bolognesi V, *et al.* A prospective assessment of anti-drug antibody response over time by a new ELISA in patients with IBD treated with infliximab. *J Crohns Colitis* 2014;**8**:S298–9	C insufficient data
Van Der Woude CJ, Bultman E, Deuring J, West R, Zelinkova Z, Peppelenbosch M. Adalimumab trough levels in a prospective cohort of Crohn's disease patients. *J Crohns Colitis* 2013;**7**:S250	C insufficient data
Van Der Woude CJ, Deuring JJ, West R, Zelinkova Z, Peppelenbosch MP. Adalimumab trough levels in a prospective cohort of Crohn's disease patients. *Gastroenterology* 2013;**144**(Suppl. 1):567	Duplicate
Van Moerkercke W, Ackaert C, Compernolle G, Jurgens M, Cleynen I, Van Assche GA, *et al.* High infliximab trough levels are associated with mucosal healing in Crohn's disease. *Gastroenterology* 2010;**138**(Suppl. 1):60	C insufficient data
Van Moerkercke W, Compernolle G, Ackaert C, Gils A, Vermeire S, Jurgens M, *et al.* Mucosal healing in Crohn's disease is associated with high infliximab trough levels. *J Crohns Colitis Suppl* 2010;**4**:30–1	C insufficient data

TABLE 47 Abstracts excluded from the review with reason for exclusion (*continued*)

Reference	Reason for exclusion
Vande Casteele N, Gils A, Compernolle G, Ballet V, Peeters M, Van Steen K, *et al.* Drug level versus clinically based dosing of infliximab maintenance therapy in IBD: final results of the randomized controlled taxit trial. *Inflamm Bowel Dis* 2013;**19**:S2–3	Superseded by full text
Vande Casteele N, Compernolle G, Ballet V, Van Assche G, Gils A, Vermeire S, *et al.* Results on the optimisation phase of the prospective controlled trough level adapted infliximab treatment (TAXIT) trial. *Gastroenterology* 2012;**142**(Suppl. 1):211–12	Superseded by full text
Vande Casteele N, Compernolle G, Ballet V, Van Assche G, Gils A, Vermeire S, *et al.* Individualised infliximab treatment using therapeutic drug monitoring: a prospective controlled Trough level Adapted infliXImab Treatment (TAXIT) trial. *J Crohns Colitis* 2012;**6**:S6	Duplicate
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;**146**(Suppl. 1):242	C insufficient data M no algorithm specified/ acted on
Vande Casteele N, Cuypers L, Singh S, Ohrmund L, Hauenstein S, Van Assche G, *et al.* Antibodies to infliximab can either be persistent or transient: a retrospective case–control study in IBD patients treated with infliximab maintenance therapy. *Gastroenterology* 2012;**142**(Suppl. 1):114	M no algorithm specified/ acted on
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	M no algorithm specified/ acted on
Vande Casteele N, Peeters M, Ferrante M, Compernolle G, Van Assche G, Vermeire S, *et al.* Functional cellular based assay reveals neutralising anti-drug antibodies in IBD patients treated with maintenance adalimumab. *J Crohns Colitis* 2014;**8**:S268–9	Duplicate
Vaughn BP, Martinez-Vazquez M, Patwardhan V, Moss AC, Sandborn WJ, Cheifetz AS. A pilot study of optimized monotherapy with infliximab for patients with inflammatory bowel disease. *Gastroenterology* 2014;**146**(Suppl. 1):55	Superseded by full paper
Vaughn BP, Martinez-Vazquez M, Patwardhan V, Moss AC, Sandborn WJ, Cheifetz AS. Prospective therapeutic drug monitoring to optimizing infliximab (IFX) maintenance therapy in patients with inflammatory bowel disease (IBD). *Gastroenterology* 2014;**146**(Suppl. 1):54	Superseded by full paper
Vaughn B, Matinez-Vazquez M, Cheifetz A. Infliximab dosing changes based on trough levels in a cohort of IBD patients in clinical remission. *Inflamm Bowel Dis* 2013;**19**:S59	M no algorithm specified/ acted on
Velayos FS, Sheibani S, Lockton S, Hauenstein S, Singh S, Terdiman JP, *et al.* 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. *Gastroenterol* 2013;**144**(Suppl. 1):91	C insufficient data
Veres G, Kaplan JL, De Greef E, Chuang E, Szabo D, Molnar K, *et al.* 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* 2012;**142**(Suppl. 1):386	C insufficient data
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. *Arthritis Rheum* 2011;**63**(Suppl. 10):1266	Duplicate
Wang SL, Hauenstein S, Ohrmund L, Shringarpure R, Wolf DC, Diab IA, *et al.* Influence of trough serum drug level and immunogenicity on the lack of response to adalimumab therapy in inflammatory bowel disease patients. *Arthritis Rheum* 2012;**64**:S819–20	C insufficient data
Wang SL, Hauenstein S, Ohrmund L, Shringarpure R, Wolf D, Diab I, *et al.* Influence of trough serum drug level and immunogenicity on the lack of response to adalimumab therapy in IBD patients. *Am J Gastroenterol* 2012;**107**:S680	Duplicate
Wolf DC, Hauenstein S, Lockton S, Singh S. Mechanisms of loss of response to adalimumab in Crohn's disease. *Gastroenterology* 2013;**144**(Suppl. 1):775	C insufficient data

continued

TABLE 47 Abstracts excluded from the review with reason for exclusion (*continued*)

Reference	Reason for exclusion
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;**144**(Suppl. 1):423	M no algorithm specified/ acted on
Wolf D, Shringarpure R, Lockton S, Corey R, Woods S, Aguilar H, *et al.* 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. *Am J Gastroenterol* 2012;**107**:S658	C insufficient data
Yamada A, Sono K, Takeuchi K, Suzuki Y. Clinical and basic studies to understand factors associated with the loss of response to infliximab in patients with Crohn's disease. *J Crohns Colitis* 2013;**7**:S239	C insufficient data
Yanai H, Lichtenstein L, Assa A, Mazor Y, Weiss B, Levine A, *et al.* Anti-TNF and anti-drug antibodies levels predict the outcomes of interventions after loss of response to adalimumab and infliximab. *Gastroenterology* 2014;**146**(Suppl. 1):381	C insufficient data; M no algorithm specified/acted on
Yarur A, Trivella JP, Sussman DA, Drake K, Barkin JS, Hauenstein S, *et al.* 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* 2014;**146**(Suppl. 1):243–4	C insufficient data
Yarur A, Drake K, Kubiliun M, Dauer RM, Sussman DA, Hauenstein S, *et al.* Anti-tumor necrosis factor levels are not associated with intestinal extent of mucosal inflammation in patients with inflammatory bowel diseases. *Gastroenterology* 2014;**146**(Suppl. 1):244	Not M, C or ATC
Zelinkova Z, Peppelenbosch MP, Van Liere-Baron A, De Haar C, Van Der Woude CJ. Naturally-occurring autoantibodies against TNF-alpha are present in sera of inflammatory bowel disease patients and influence the response to adalimumab. *Gastroenterology* 2011;**140**(Suppl. 1):62	C insufficient data

ATC, assay-type comparison study; C, correlation-type study; M, management-type study.

Appendix 7 Ongoing trials

Ongoing trials using an algorithm to determine treatment change according to test results

TABLE 48 Ongoing trials using an algorithm to direct treatment

Title (acronym)	Status; start date; estimated completion date	URL
Pediatric Crohn's disease Adallmumab Level-based Optimization Treatment (PAILOT)	Ongoing – not yet open for participant recruitment; start: November 2014; primary completion due: July 2018	http://clinicaltrials.gov/ct2/show/NCT02256462 (accessed 11 November 2014)
A randomized controlled trial investigating tailored treatment with infliximab for active luminal Crohn's disease (TAILORIX)	Ongoing; start: March 2012; primary completion due: June 2015	www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2011-003038-14 (accessed 11 November 2014)
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: October 2013; primary completion due: December 2014	www.trialregister.nl/trialreg/admin/rctview.asp?TC=4067 (accessed 11 November 2014)

Ongoing correlation studies

TABLE 49 Ongoing studies aiming to correlate test results and clinical status

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 (accessed 11 November 2014)
Improving treatment of inflammatory bowel diseases through better understanding infliximab drug and antibody levels (OPTIMISE)	Ongoing; study start: March 2014; study primary completion due: March 2015	https://clinicaltrials.gov/ct2/show/NCT01787786 (accessed 11 November 2014)
Personalised anti-TNF therapy in Crohn's disease (PANTS)	Ongoing; currently recruiting	http://public.ukcrn.org.uk/Search/StudyDetail.aspx?StudyID=14175 (accessed 11 November 2014)
Utilising drug levels and anti-drug antibodies to predict response to treatment in patients with Inflammatory Bowel Disease	Ongoing; start: October 2012	www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2011-006084-22 (accessed 11 November 2014)

Appendix 8 Excluded assay-type comparison studies

TABLE 50 Studies of assay comparisons excluded from the review

Reference	Reason for exclusion
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	Irrelevant comparison
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**:7367–75	Irrelevant comparison
Greathead L, Kelleher P, Steel A. Development and validation of ELISA to measure serum anti TNFa levels. *J Crohns Colitis* 2014;**8**:S97–8	Irrelevant comparison
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**:136–43	Irrelevant comparison
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**:100–9	Irrelevant comparison
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**:1628–33	Irrelevant comparison
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	Irrelevant comparison
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**:eA27–8	Irrelevant comparison
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**:530–8	Irrelevant comparison
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;**146**(Suppl. 1):245	Irrelevant comparison
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;**146**(Suppl. 1):242	Irrelevant comparison

Appendix 9 Summary of studies evaluating the clinical utility of measuring levels of anti-tumour necrosis factor alpha and its antibodies

This appendix summarises the studies by Afif *et al.*,[56] Pariente *et al.*,[59] Roblin *et al.*[57] and Paul *et al.*[58] and details the studies' proposed algorithms.

TABLE 51 Overview of study characteristics of studies evaluating the clinical utility of measuring levels of anti-TNF-α and its antibodies

Item	Study (first author, year)			
	Afif *et al.*, 2010[56]	Pariente *et al.*, 2012[59]	Roblin *et al.*, 2014[57]	Paul *et al.*, 2013[58]
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	IFX	IFX	ADA	IFX
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 diarrhoea and abdominal pain, complete closure of all fistulas	Clinical response: (1) decrease of at least 2 points of HBI or (2) overall assessment by treating clinician	Clinical remission: CDAI score of < 150, FC < 250 µg/g	Clinical remission: CDAI score of < 150, mucosal healing: FC < 250 µg/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 ADA: *n* = 5 (7%) Surgery: *n* = 1 (1%)	(1) ADA 40 mg e.o.w. to ADA 40 mg e.w. (2) Switch to IFX 5 mg/kg at 0, 2 and 6 weeks	IFX 5 mg/kg every 8 weeks to IFX 10 mg/kg every 8 weeks
Algorithm proposed	Yes	No	Yes	Yes

e.o.w., every other week; e.w., every week.

The two retrospective studies[56,59] 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[57,58] 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.*[56] concluded that measurement of drug and anti-drug antibodies impacts management and is clinically useful; Paul *et al.*[58] concluded that measurement of drug and anti-drug antibodies predicts clinical remission and may guide clinical decisions in practice; Roblin *et al.*[57] concluded that knowledge of drug and anti-drug antibody levels may have a strong impact on the management of IBD patients with LOR; but Pariente *et al.*[59] concluded that clinical response to drug intensification cannot be accurately predicted by measurement of drug and anti-drug antibody levels. The authors reported that a considerable number of patients (70%) 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, cotreatment with immunosuppressants, disease duration and time of disease assessment following anti-TNF-α optimisation. The proposed algorithms varied considerably in terms of drug and anti-drug antibody levels used to predict response; however, the proposed treatment changes were comparable but differed in detail.

Algorithm proposed by Afif *et al.*[56]

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-TNF-α agent and switch class of drug if symptoms persist
- therapeutic IFX concentrations (> 12 µg/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
- subtherapeutic IFX concentrations (< 12 µg/ml at 4 weeks or undetectable trough level < 1.4 µg/ml) – IFX 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 have subtherapeutic or therapeutic IFX 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-TNF-α drug, whereas those with therapeutic IFX levels responded if they stayed on the same dose of IFX. Patients with subtherapeutic IFX levels responded to dose intensification of IFX. The study concluded that use of IFX and anti-drug antibody tests can potentially avoid inappropriate management.

Algorithm proposed by Paul *et al.*[58]

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:

- subtherapeutic IFX levels (< 2 µg/ml) and anti-drug antibodies > 200 ng/ml – switch to another anti-TNF-α agent, or optimise IFX and introduce immunosuppressant
- subtherapeutic IFX levels (< 2 µg/ml) and anti-drug antibodies < 200 ng/ml – IFX intensification
- therapeutic IFX levels (> 2 µg/ml) and anti-drug antibodies < 10 ng/ml – IFX intensification
- therapeutic IFX levels (> 2 µg/ml) and anti-drug antibodies > 10 ng/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 IFX dose intensification.

Algorithm proposed by Roblin *et al.*[57]

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 ADA concentrations (< 4.9 µg/ml) and detectable anti-drug antibodies (> 10 ng/ml) – switch to another anti-TNF-α agent
- high trough ADA concentrations (> 4.9 µg/ml) – switch class of drug
- low trough ADA concentrations (< 4.9 µg/ml) and no detectable anti-drug antibodies (< 10 ng/ml) – ADA intensification (40 mg every week).

The findings of the study suggested that those with low trough levels of anti-TNF-α drug and undetectable levels of antibodies or high trough levels of anti-TNF-α drug had the greatest chance of achieving clinical remission following anti-TNF-α drug optimisation, whereas those with low levels of anti-TNF-α drug levels and high levels of antibodies had the lowest chance of achieving clinical remission.

Appendix 10 Quality appraisal of included management studies

Cochrane's tool for assessing risk of bias for a randomised controlled trial (adapted from Higgins *et al.*[70])

First author surname and year of publication: Steenholdt 2014[123] and 2015[124]

Name of first reviewer: Paul Sutcliffe.

Name of second reviewer: Martin Connock.

Domain	Description	Review authors' judgement
Selection bias: sequence generation	The authors 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 authors state: *Patients were blinded* *Physicians were blinded to test results in the IFX escalation group* *Physicians were not blinded to test results in algorithm group* *Results of analyses of serum IFX and IFX antibodies were used in the treatment of patients in this 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 reporting of the outcome, subgroups or analysis	The study appears to be free of any selective outcome reporting of outcome, subgroups or analysis	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	We note that 42% of patients were not treated in accordance with the algorithm, resulting in patients 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[73]

Name of first reviewer: Paul Sutcliffe.

Name of second reviewer: Martin Connock.

Domain	Description	Review authors' judgement
Selection bias: sequence generation	The authors 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 authors state: *Both patients and treating physicians were blinded to individual IFX trough and ATI concentrations* This is unclear. No further information is provided; this limits our rating of whether or not the knowledge of the allocated intervention was adequately prevented during the study	Unclear risk of bias
Detection bias: blinding of outcome assessors	The authors 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 optimisation phase because of 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 appear 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

continued

Domain	Description	Review authors' judgement
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 randomised maintenance phase was only 1 year, which prevents the analysis of long-term clinical and pharmacoeconomic outcomes	Low risk of bias

ATI, antibodies to IFX.

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[71] for non-randomised primary clinical studies

First author (year) study identification number: Vaughn 2014[128]

Name of first reviewer: Paul Sutcliffe.

Name of second reviewer: Martin Connock.

Criteria	Rating
Reporting	
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

continued

Criteria	Rating
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 interquartile range of results should be reported. In normally distributed data the standard error, SD or CIs 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	
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	
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

Criteria	Rating
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 non-parametric 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)	
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 ITT; 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 non-randomised 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	
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 < 5%? (Yes/no/unable to determine)	No

Quality Assessment of Diagnostic Accuracy Studies-2[69] tool with index questions adapted to the review for studies comparing performance of different tests: Steenholdt[122,123]

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)	Patients with CD (adults and children) receiving IFX or ADA, either whose disease responds to treatment with a TNF inhibitor, or who experience secondary LOR 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

Quality Assessment of Diagnostic Accuracy Studies-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	Study[122] included 66 patients with CD with secondary LOR to IFX, which were all part of a RCT. The RCT paper[123] 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)	Patients with CD with secondary LOR to IFX
Range of drug/antibody concentrations	From patients
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	
Is there concern that the intervention test, its conduct, or interpretation differs from the review question?	Concern: high

RGA, reporter gene assay.

Domain 3: reference standard

A. Risk of bias	
Describe the reference standard and how it was conducted and interpreted	RIA on samples stored at room temperature
Were the reference standard results interpreted without knowledge of the results of the index test?	Yes
Is the reference standard likely to correctly classify the target condition?	No
Could the comparison test, its conduct or its interpretation have introduced bias?	Risk: high
B. Concerns regarding applicability	
Same test and threshold as used 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

Quality Assessment of Diagnostic Accuracy Studies-2[69] tool with index questions adapted to the review for studies comparing performance of different tests: Vande Casteele 2013[126]

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)	Patients with CD (adults and children) receiving IFX or ADA, either whose disease responds to treatment with TNF inhibitor or who experience secondary LOR during maintenance treatment with a 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

Quality Assessment of Diagnostic Accuracy Studies-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 anti-drug 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)	Crohn's and UC patients selected on basis of index test results
Range of drug/antibody concentrations	From patients
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 Vande Casteele *et al.* 2012[67]
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	
Is there concern that the intervention test, its conduct or interpretation differs 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 Laboratories from biobanked samples
Is the comparison test likely to correctly classify the target condition (only matters if doing more than correlation studies)?	No
Were the reference standard results interpreted without knowledge of the results of the index test?	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 can't 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

Quality Assessment of Diagnostic Accuracy Studies-2[69] tool with index questions adapted to the review for studies comparing performance of different tests: Vande Casteele 2012[67]

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)	Patients with CD (adults and children) receiving IFX or ADA, either whose disease responds to treatment with a TNF inhibitor or who experience secondary LOR during maintenance treatment with a 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

Quality Assessment of Diagnostic Accuracy Studies-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	Drug levels up to 30 mg/l from visual inspection of plots. Antibody levels NR
Is there concern that the included patients or range of drug/antibody concentrations do not match the review question?	Concern: high

NR, not reported.

Domain 2: index test(s)

A. Risk of bias	
Describe the intervention test and how it was conducted and interpreted	LISA-TRACKER assays 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 how sample was stored before intervention test
Is there concern that the intervention test, its conduct or interpretation differs 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
Were the reference standard results interpreted without knowledge of the results of the index test?	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	Unclear
Describe the time interval and any interventions between intervention test and comparison test(s)	Unclear
Was there an appropriate interval between intervention test and comparison test(s)? (Ideally conducted at same time so samples can't 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

Quality Assessment of Diagnostic Accuracy Studies-2[69] tool with index questions adapted to the review for studies comparing performance of different tests: Wang 2012[131]

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)	Patients with CD (adults and children) receiving IFX or ADA, either whose disease responds to treatment with a TNF inhibitor or who experience secondary LOR during maintenance treatment with a 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

Quality Assessment of Diagnostic Accuracy Studies-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	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 UC
Range of drug/antibody concentrations	From patients
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	
Is there concern that the intervention test, its conduct, or interpretation differs 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 Laboratories
Were the reference standard results interpreted without knowledge of the results of the index test?	Unclear
Is the reference standard likely to correctly classify the target condition?	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)	Unclear
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

Appendix 11 Parametric modelling for Vaughn[128] and the Trough level Adapted infliXImab Treatment trial

Parametric models were fitted to reconstructed IPD of time to treatment failure for proactive drug monitoring and control patients in remission who commenced maintenance IFX at the start of 2009.[128] This was done so that treatment failure could be modelled to 10 years (the time horizon of the economic model) with potential for use in the model.

According to the Akaike information criterion and Bayesian information criterion, the information criteria that best fitted the data were provided by log-normal, log-logistic and Weibull models (*Table 52*); a gamma model could not be fitted for the standard care arm.

Log-normal and Weibull models are shown in *Figure 36*.

Trough level Adapted infliXImab Treatment 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.

TABLE 52 Akaike information criterion and Bayesian information criterion values for parametric models for time-to-treatment failure

Model	Observations	ll (model)	df	Akaike information criterion	Bayesian information criterion
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
Log-normal	68	–51.55	2	107.0999	111.5389
Log-logistic	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
Log-normal	39	–18.8709	2	41.74174	45.06886
Log-logistic	39	–19.1771	2	42.35412	45.68124

df, degrees of freedom.

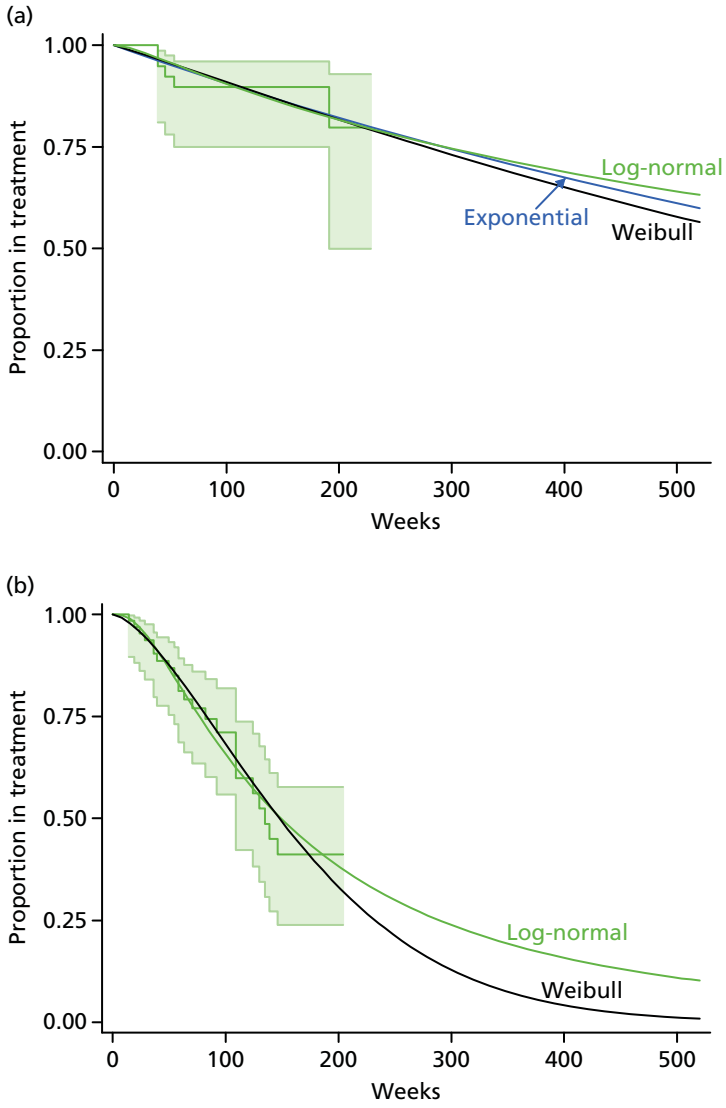


FIGURE 36 Log-normal and Weibull models extended to 10 years' follow-up [Vaughn *et al.*:[128] (a) with therapeutic drug concentration monitoring; and (b) without therapeutic drug concentration monitoring study].

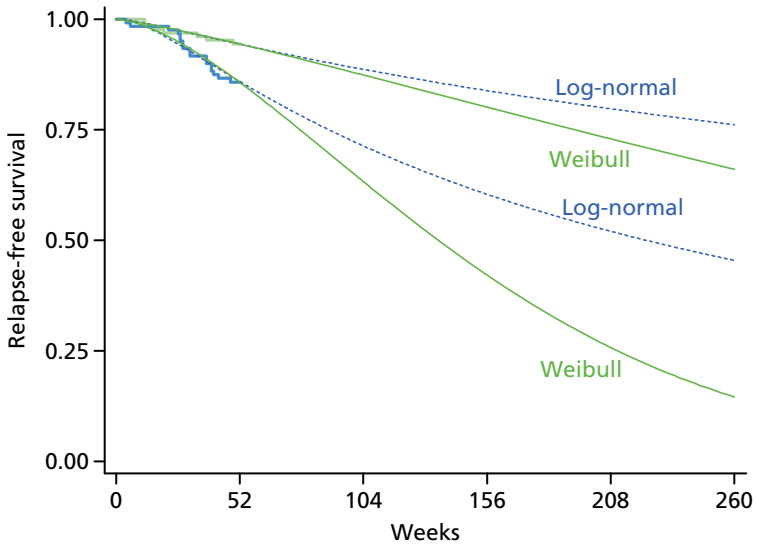


FIGURE 37 Parametric modelling of time to relapse based on the TAXIT study.

Appendix 12 Meta-analysis results

Correlation studies that permitted extraction of a 2 × 2 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 53*. Forest plots of sensitivity and specificity for prediction of loss or lack response and summary ROC plots are presented according to the test applied.

Infliximab trough level tests for loss of response or lack of regaining response

Eleven studies were included,[38,47,78,82,85,98,99,103,120,123,134] of which two were reported as abstracts.[78,134] Sensitivity and specificity pairs are summarised in *Figure 38*.

Hierarchical meta-analysis yielded the test accuracy results summarised in *Table 54* and *Figure 39*.

Subgroup analyses examining responder populations only, and ELISA studies only, had little effect on pooled estimates.

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 55*.

TABLE 53 Major features of studies included for hierarchical meta-analysis

Study (first author and year)	Drug	Diagnosis	Resp/LOR	Test	Res-def
IFX trough level as predictor of loss or lack of response					
Ainsworth *et al.*, 2008[47]	IFX	CD	LOR	RIA	PJ
Ben-Basset *et al.*, 2013[78] (abstract)	IFX	IBD ≈0.93 CD	Resp	HMSA	HBI
Bortlik *et al.*, 2013[82]	IFX	CD	Resp	ELISA	PJ
Cornillie *et al.*, 2014[85]	IFX	CD	Resp	ELISA	CDAI
Hibi *et al.*, 2014[98]	IFX	CD	Resp	ELISA	CDAI
Imaeda *et al.*, 2012[99]	IFX	CD	Resp	ELISA	CDAI
Kopylov *et al.*, 2012[103]	IFX	CD	Resp	ELISA	PJ
Maser *et al.*, 2006[38]	IFX	CD	Resp	Unclear	HBI
Steenholdt *et al.*, 2011[120]	IFX	CD	Resp	RIA	PJ
Steenholdt *et al.*, 2014[123]	IFX	CD	LOR	RIA	CDAI
Yanai *et al.*, 2012[134] (abstract)	IFX	CD	Resp	ELISA	PJ
Trough antibodies to IFX as predictor of loss or lack of response					
Ainsworth *et al.*, 2008[47]	IFX	CD	LOR	RIA	PJ
Baert *et al.*, 2014[77]	IFX	IBD ≈0.8 CD	LOR	HMSA	PJ
Ben-Horin *et al.*, 2011[79]	IFX	IBD ≈0.82 CD	Resp	NR	ST

continued

TABLE 53 Major features of studies included for hierarchical meta-analysis (*continued*)

Study (first author and year)	Drug	Diagnosis	Resp/LOR	Test	Res-def
Ben-Horin *et al.*, 2012[80]	IFX ADA	IBD ≈0.9 CD	LOR	ELISA	PJ
Bodini *et al.*, 2014[81] (abstract)	IFX	CD	Resp	HMSA	HBI
Candon *et al.*, 2006[83]	IFX	CD	LOR	ELISA	UC
Dauer *et al.*, 2013[88] (abstract)	IFX	CD ≈0.83 CD	Resp	NR	PJ
Farrell *et al.*, 2003[92]	IFX	CD	Resp	ELISA	PJ
Hanauer *et al.*, 2004[40]	IFX	CD	Resp	ELISA	CDAI
Imaeda *et al.*, 2012[99]	IFX	CD	Resp	ELISA	CDAI
Kong *et al.*, 2011[102] (abstract)	IFX	IBD ≈0.83 CD	Resp	ELISA	PJ
Kopylov *et al.*, 2012[103]	IFX	CD	Resp	ELISA	PJ
Marzo *et al.*, 2014[106] (abstract)	IFX	NR	Resp	ELISA	CDAI
Nagore *et al.*, 2015[110] (abstract)	IFX	IBD ≈0.86 CD	Resp	ELISA	PJ
Pariente *et al.*, 2012[59]	IFX	CD and UC	LOR	ELISA	PJ or HBI
Steenholdt *et al.*, 2011[120]	IFX	CD	Resp	RIA	PJ ST
Steenholdt *et al.*, 2013[52]	IFX	CD	Resp	ELISA	PJ
Steenholdt *et al.*, 2014[123]	IFX	CD	LOR	RIA	CDAI
Vande Casteele *et al.*, 2013[126]	IFX	IBD ≈0.70 CD	LOR	HMSA	CRP TC
Vande Casteele *et al.*, 2013[126]	IFX	IBD ≈0.70 CD	Resp	HMSA	CRP TC
ADA trough level as predictor of loss or lack of response					
Chiu *et al.*, 2013[84]	ADA	CD	LOR	ELISA	CDAI
Frederiksen *et al.*, 2014[94]	ADA	IBD	Resp	RIA	PJ BM
Imaeda *et al.*, 2014[100]	ADA	CD	Resp	ELISA	CRP
Mazor *et al.*, 2014[108]	ADA	CD	Resp	ELISA	PJ
Roblin *et al.*, 2014[115]	ADA	IBD ≈0.55 CD	Resp	ELISA	CDAI
Trough antibodies to ADA as predictor of loss or lack of response					
Frederiksen *et al.*, 2014[94]	ADA	IBD	Resp	RIA	PJ BM
Imaeda *et al.*, 2014[100]	ADA	CD	Resp	ELISA	CRP
Mazor *et al.*, 2014[108]	ADA	CD	Resp	ELISA	PJ
West *et al.*, 2008[133]	ADA	CD	Resp	RIA	PJ
Ben-Horin *et al.*, 2012[80]	IFX ADA	IBD ≈0.9 CD	LOR	ELISA	SA
Roblin *et al.*, 2014[115]	ADA	CD	Resp	ELISA	CDAI

NR, not reported; PJ, physicians' judgement; PJ BM, physicians' judgement and biological measure; resp, responding patients; SA, switch anti-TNF-α; ST, stop anti-TNF-α; TC, Trough Concentration.
Diagnosis refers to study patient population. Res-def refers to the method used for defining clinical response.

FIGURE 38 Trough IFX for predicting LOR or failure to regain response. Response (Res) refers to method for estimating clinical response. FN, false negative; FP; false positive; PJ, physicians' judgement; Pop, study patient population; R, responders; Res, response; TN, true negative; TP; true positive; UC, unclear.

TABLE 54 Test accuracy results from hierarchical meta-analysis

Studies included	Parameter	Point estimate	95% CI
All 11 studies	Sens	0.657232	0.546288 to 0.753299
	Spec	0.80625	0.744166 to 0.85618
	DOR	7.978975	4.119972 to 15.45254
	LR+	3.392169	2.35152 to 4.893351
	LR–	0.425139	0.305104 to 0.592398
	1/LR–	2.352175	1.688056 to 3.277573
Responder populations only	Sens	0.681452	0.592117 to 0.759178
	Spec	0.790873	0.723301 to 0.845468
	DOR	8.090128	4.353039 to 15.03551
	LR+	3.258549	2.287802 to 4.641198
	LR–	0.402781	0.298559 to 0.543385
	1/LR–	2.482739	1.840315 to 3.349423
ELISA studies only	Sens	0.652104	0.564027 to 0.730877
	Spec	0.789041	0.691592 to 0.861849
	DOR	7.010794	3.450232 to 14.24578
	LR+	3.091133	1.959085 to 4.877331
	LR–	0.440911	0.329778 to 0.589495
	1/LR–	2.268033	1.696367 to 3.032348

LR, likelihood ratio; Sens, sensitivity; Spec, specificity.

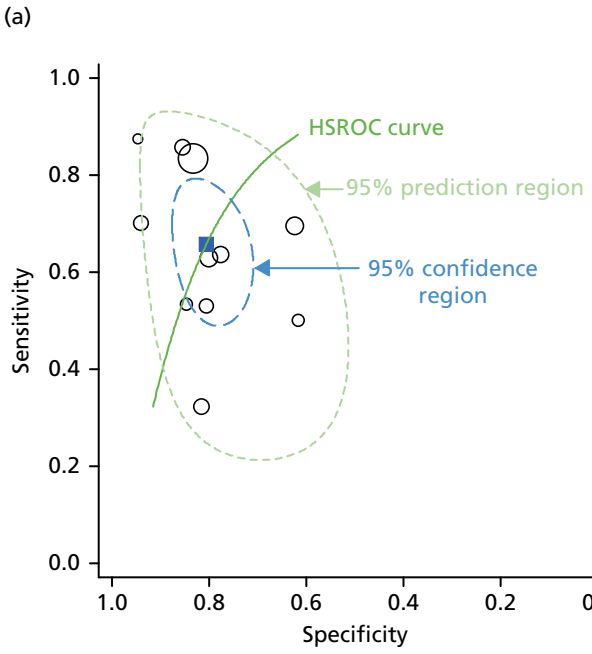


FIGURE 39 Trough IFX levels for predicting LOR; hierarchical meta-analysis of test accuracy. The blue square represents the summary point estimate on the HSROC curve. (a) All 11 studies; and (b) responder studies only. HSROC, hierarchical summary receiver operating characteristic. (*continued*)

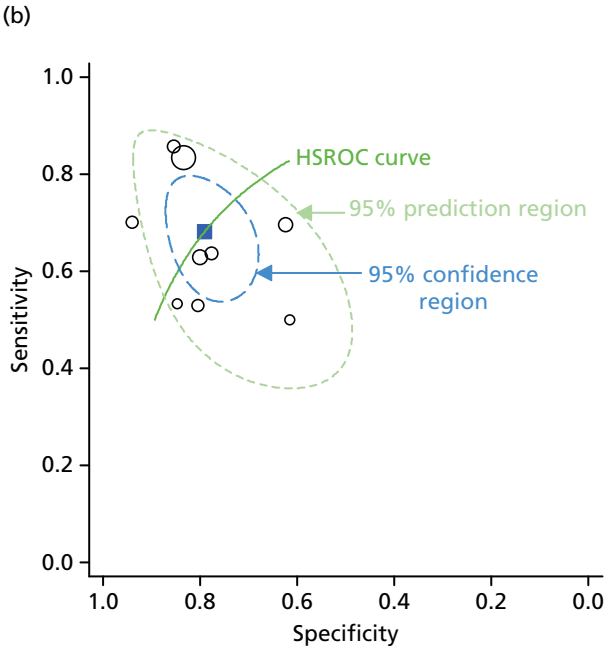


FIGURE 39 Trough IFX levels for predicting LOR; hierarchical meta-analysis of test accuracy. The blue square represents the summary point estimate on the HSROC curve. (a) All 11 studies; and (b) responder studies only. HSROC, hierarchical summary receiver operating characteristic.

TABLE 55 Probability of a positive and negative test result (range based on 95% CI prevalence)

Item	Calculation	Value
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)

Sens, sensitivity; Spec, specificity.

Antibodies to infliximab tests for loss of response or lack of regaining response

Nineteen studies were included,[40,47,52,59,77,79–81,83,88,92,99,102,103,106,120,123,126,182] of which five were reported as abstracts.[81,88,102,106,110] Sensitivity and specificity pairs are summarised in *Figure 40*.

Hierarchical meta-analysis yielded test accuracy results summarised in *Table 56* and *Figure 41*.

Subgroup analyses removing two outlier studies, examining responder populations only and ELISA studies only, had little effect on pooled summary point estimates.

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 to 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 57*.

Study	TP	FP	FN	TN	Assay	Pop	Res	Sensitivity (95% CI)	Specificity (95% CI)
Ben-Horin 2012[80]	3	2	11	11	ELISA	LOR	PJ	0.21 (0.05 to 0.51)	0.85 (0.55 to 0.98)
Pariente 2012[59]	4	6	8	21	ELISA	LOR	PJ or HBI	0.33 (0.10 to 0.65)	0.78 (0.48 to 0.91)
Candon 2006[83]	6	3	2	11	ELISA	LOR	UC	0.75 (0.35 to 0.97)	0.79 (0.49 to 0.95)
Vande Casteele 2013 LOR[126]	12	31	15	2	HMSA	LOR	CRP TC	0.44 (0.25 to 0.65)	0.06 (0.01 to 0.20)
Baert 2014[71]	9	40	2	81	HMSA	LOR	PJ	0.82 (0.48 to 0.98)	0.67 (0.58 to 0.75)
Steenholdt 2014[123]	9	9	22	29	RIA	LOR	CDAI	0.29 (0.14 to 0.48)	0.76 (0.60 to 0.89)
Ainsworth 2008[47]	8	10	0	9	RIA	LOR	PJ	1.00 (0.63 to 1.00)	0.47 (0.24 to 0.71)
Hanauer 2004[40]	16	13	58	137	ELISA	R	CDAI	0.22 (0.13 to 0.33)	0.91 (0.86 to 0.95)
Imaeda 2012[99]	12	4	5	37	ELISA	R	CDAI	0.71 (0.44 to 0.90)	0.90 (0.77 to 0.97)
Marzo 2014[106]	9	9	13	51	ELISA	R	CDAI	0.41 (0.21 to 0.64)	0.85 (0.73 to 0.93)
Steenholdt 2013[52]	16	2	5	6	ELISA	R	PJ	0.76 (0.53 to 0.92)	0.75 (0.35 to 0.97)
Farrell 2003[92]	19	0	13	21	ELISA	R	PJ	0.59 (0.41 to 0.76)	1.00 (0.84 to 1.00)
Kong 2011[102]	4	2	10	14	ELISA	R	PJ	0.29 (0.08 to 0.58)	0.88 (0.62 to 0.98)
Kopylov 2012[103]	17	5	13	28	ELISA	R	PJ	0.57 (0.37 to 0.75)	0.85 (0.68 to 0.95)
Nagore 2015[182]	4	3	3	40	ELISA	R	PJ	0.57 (0.18 to 0.90)	0.93 (0.81 to 0.99)
Vande Casteele 2013 R[126]	43	10	17	20	HMSA	R	CRP TC	0.72 (0.59 to 0.83)	0.67 (0.47 to 0.83)
Bodini 2014[81]	5	3	8	5	HMSA	R	HBI	0.38 (0.14 to 0.68)	0.63 (0.24 to 0.91)
Steenholdt 2011[120]	21	9	5	50	RIA	R	PJ ST	0.81 (0.61 to 0.93)	0.85 (0.73 to 0.93)
Dauer 2013[88]	3	1	3	10	UC	R	PJ	0.50 (0.12 to 0.88)	0.91 (0.59 to 1.00)
Ben-Horin 2011[79]	10	19	2	31	UC	R	ST	0.83 (0.52 to 0.98)	0.62 (0.47 to 0.75)

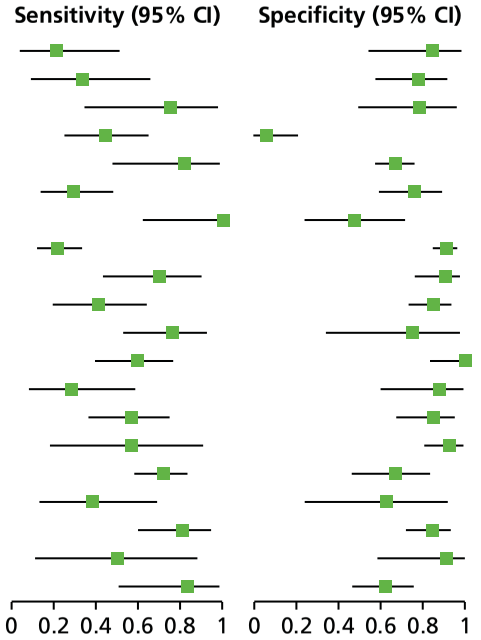


FIGURE 40 Sensitivity and specificity of tests of antibodies to IFX for prediction of LOR or failure to regain response. Response (Res) refers to method for estimating clinical response. Note that Bodini *et al.*,[81] Dauer *et al.*,[88] Kong *et al.*,[102] Marzo *et al.*[106] and Nagore *et al.*[110] were published in brief as abstracts or conference proceedings. FN, false negative; FP, false positive; PJ, physicians' judgement; Pop, study patient population; R, responders; Res, response; ST, stop IFX therapy; TC, treatment change; TN, true negative; TP, true positive; UC, unclear.

TABLE 56 Test accuracy results from hierarchical meta-analysis

Studies included	Parameter	Point estimate	95% CI
All 20 studies	Sens	0.559745	0.444812 to 0.668611
	Spec	0.792243	0.688105 to 0.868267
	DOR	4.848283	2.519589 to 9.329239
	LR+	2.694226	1.72293 to 4.213088
	LR–	0.555707	0.426575 to 0.72393
	1/LR–	1.799509	1.38135 to 2.344251
All studies minus outliers	Sens	0.597	0.477 to 0.707
	Spec	0.807	0.742 to 0.859
	DOR	6.183	3.805 to 10.050
	LR+	3.088	2.311 to 4.127
	LR–	0.500	0.381 to 0.655
	1/LR–	2.002	1.528 to 2.623
Responder populations only	Sens	0.570	0.445 to 0.687
	Spec	0.849	0.787 to 0.896
	DOR	7.460	4.544 to 12.250
	LR+	3.778	2.722 to 5.244
	LR–	0.506	0.388 to 0.660
	1/LR–	1.974	1.514 to 2.574
ELISA studies only	Sens	0.482	0.355 to 0.611
	Spec	0.880	0.841 to 0.911
	DOR	6.830	3.872 to 12.050
	LR+	4.022	2.805 to 5.768
	LR–	0.589	0.459 to 0.755
	1/LR–	1.698	1.324 to 2.178

LR, likelihood ratio; Sens, sensitivity; Spec, specificity.

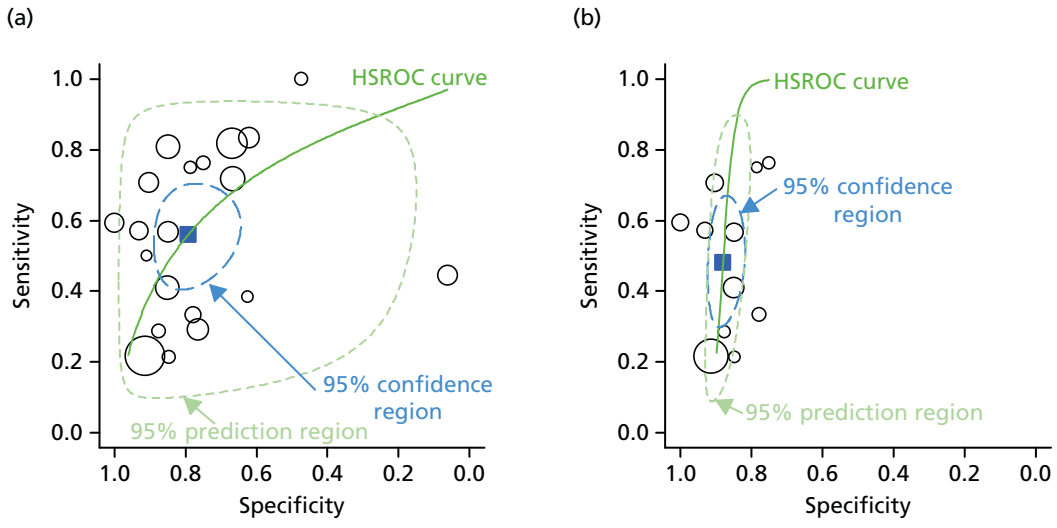


FIGURE 41 Antibodies to IFX for predicting LOR; hierarchical meta-analysis of test accuracy. The blue square represents the summary point estimate on the HSROC curve. (a) All 20 studies; (b) ELISA studies only; (c) all studies excluding two influential outliers; and (d) responder populations only. HSROC, hierarchical summary receiver operating characteristic. (*continued*)

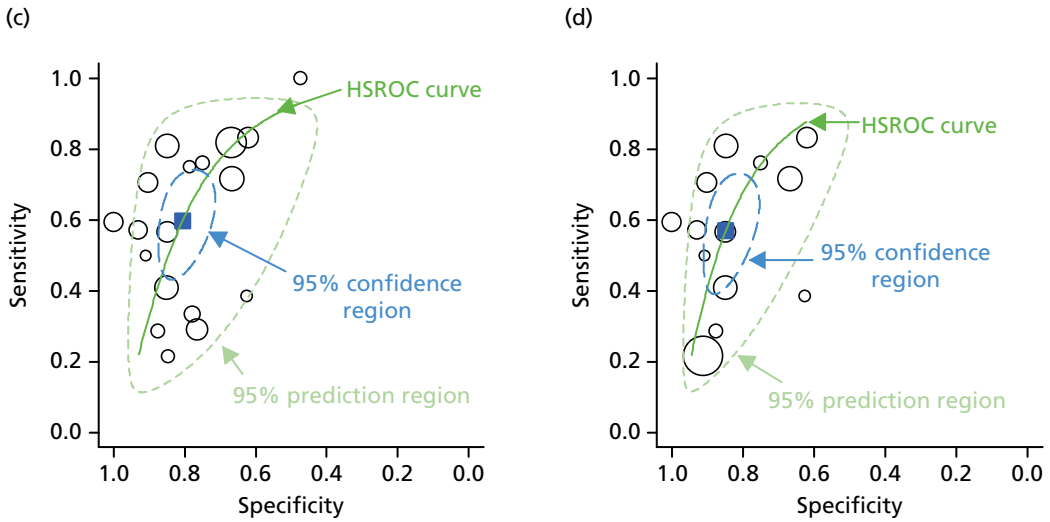


FIGURE 41 Antibodies to IFX for predicting LOR; hierarchical meta-analysis of test accuracy. The blue square represents the summary point estimate on the HSROC curve. (a) All 20 studies; (b) ELISA studies only; (c) all studies excluding two influential outliers; and (d) responder populations only. HSROC, hierarchical summary receiver operating characteristic.

TABLE 57 Probability of a positive and negative test result, all studies (range based on 95% CI prevalence)

Item	Calculation	Value
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)

Sens, sensitivity; Spec, specificity.

Adalimumab trough level test for loss of response or lack of regaining response

Four studies of responders were included.[94,100,108,115] The study of Roblin *et al.*[115] included 18 UC and 22 CD patients. Mazor *et al.*[108] reported results by test rather than by patient (there were 118 tests in 71 patients; authors stated that using the first test result for each patient did not alter the results). Sensitivity and specificity pairs are summarised in *Figure 42*.

A single study of patients with LOR was identified[84] (as shown in *Figure 43*). This study appeared to be an outlier and meta-analysis was restricted to responder populations).

Hierarchical meta-analysis yielded test accuracy results summarised in *Table 58* and *Figure 44*.

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 overestimate owing to double-counting of patients from the Mazor *et al.*[108] study. Given the meta-analysis values and the sensitivity, specificity and prevalence values, the point estimate for the probability of positive and negative test results is 0.444 (range 0.389–0.499) and 0.556 (range 0.501–0.611), respectively.

Study	TP	FP	FN	TN	Assay	Pop	Res	Sensitivity (95% CI)	Specificity (95% CI)
Frederiksen 2014[94]	12	9	6	20	RIA	R	PJ BM	0.67 (0.41 to 0.87)	0.69 (0.49 to 0.85)
Imaeda 2014[100]	10	2	5	23	ELISA	R	CRP	0.67 (0.38 to 0.88)	0.92 (0.74 to 0.99)
Mazor 2014[108]	48	16	20	34	ELISA	R	PJ CRP	0.71 (0.58 to 0.81)	0.68 (0.53 to 0.80)
Roblin 2014[115]	16	2	8	14	ELISA	R	MH	0.67 (0.45 to 0.84)	0.88 (0.62 to 0.98)

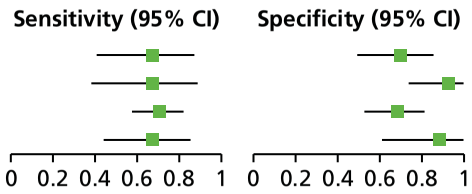


FIGURE 42 Trough ADA for predicting LOR in responders. Response (Res) refers to method for estimating clinical response. CRP level is > 3 mg/ml. FN, false negative; FP, false positive; Pop, study patient population; R, responders; Res, response; PJ BM, physicians' judgement and biological measure; PJ CRP, physicians' judgement and CRP level; TN, true negative; TP, true positive.

Study	TP	FP	FN	TN	Assay	Pop	Res	Sensitivity (95% CI)	Specificity (95% CI)
Chiu 2013[84]	7	7	56	98	ELISA	LOR	CDAI100	0.11 (0.05 to 0.22)	0.93 (0.87 to 0.97)
Roblin 2014[115]	16	2	8	14	ELISA	R	MH	0.67 (0.45 to 0.84)	0.88 (0.62 to 0.98)
Mazor 2014[108]	48	16	20	34	ELISA	R	PJ CRP	0.71 (0.58 to 0.81)	0.68 (0.53 to 0.80)
Imaeda 2014[100]	10	2	5	23	ELISA	R	CRP	0.67 (0.38 to 0.88)	0.92 (0.74 to 0.99)
Frederiksen 2014[94]	12	9	6	20	RIA	R	PJ BM	0.67 (0.41 to 0.87)	0.69 (0.49 to 0.85)

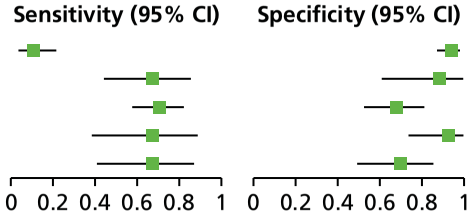


FIGURE 43 Trough ADA for predicting LOR or failure to regain response. Response (Res) refers to method for estimating clinical response. CRP level is > 3 mg/ml. CDAI100, CDAI score reduction of 100; FN, false negative; FP, false positive; Pop, study patient population; R, responders; Res, response; PJ BM, physicians' judgement and biological measure; PJ CRP, physicians' judgement and CRP level; TN, true negative; TP, true positive.

TABLE 58 Test accuracy results from hierarchical meta-analysis (four studies)

Parameter	Point estimate	95% CI
Sens	0.684	0.591 to 0.764
Spec	0.786	0.643 to 0.883
DOR	7.971	3.646 to 17.428
LR+	3.201	1.822 to 5.623
LR–	0.402	0.297 to 0.542
1/LR–	2.490	1.844 to 3.363

LR, likelihood ratio; Sens, sensitivity; Spec, specificity.

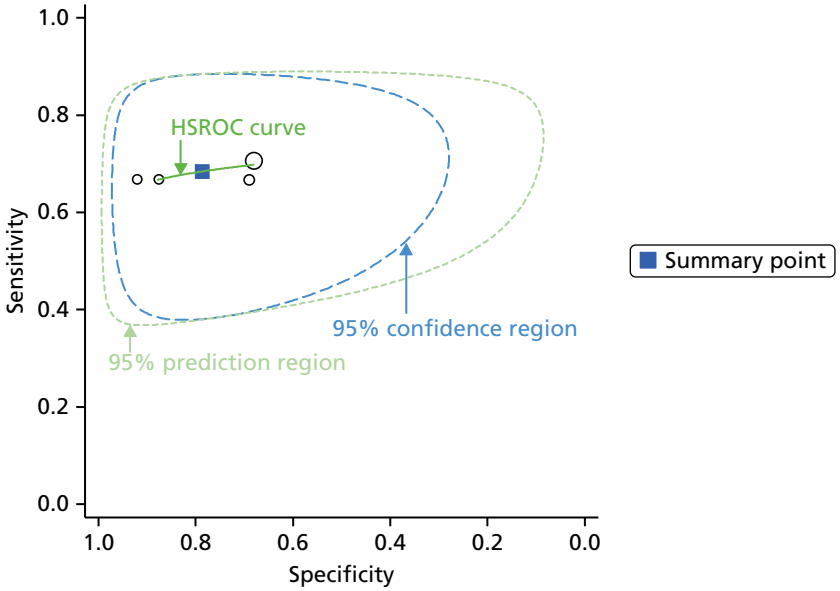


FIGURE 44 Trough ADA levels for predicting LOR; hierarchical meta-analysis of test accuracy. Only responder studies are included. HSROC, hierarchical summary receiver operating characteristic.

Antibodies to adalimumab as test for loss of response or lack of regaining response

Six studies of responders or secondary starters were included.[80,94,100,108,115,133] Mazor *et al.*[108] reported results by test rather than by patients (there were 118 tests in 71 patients; authors stated that using the first test result for each patient did not alter the results). Sensitivity and specificity pairs are summarised in *Figure 45*.

Hierarchical meta-analysis yielded test accuracy results summarised in *Table 59* and *Figure 46*.

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 overestimate owing to double-counting of patients from the Mazor *et al.*[108] study. Given the meta-analysis values and the sensitivity, specificity and prevalence values, the point estimate for the probability of positive and negative test results is 0.253 (range 0.212–0.293) and 0.747 (range 0.707–0.788), respectively.

Study	TP	FP	FN	TN	Assay	Res	Pop	Sensitivity (95% CI)	Specificity (95% CI)
Ben-Horin 2012[80]	3	2	11	11	ELISA	SA	RS	0.21 (0.05 to 0.51)	0.85 (0.55 to 0.98)
Frederiksen 2014[94]	9	1	9	28	RIA	PJ BM	R	0.50 (0.26 to 0.74)	0.97 (0.82 to 1.00)
Imaeda 2014[100]	10	4	5	21	ELISA	CRP	R	0.67 (0.38 to 0.88)	0.84 (0.64 to 0.95)
Mazor 2014[108]	14	1	49	54	ELISA	PJ	R	0.22 (0.13 to 0.34)	0.98 (0.90 to 1.00)
Roblin 2014[115]	10	4	3	5	ELISA	CDAI	R	0.77 (0.46 to 0.95)	0.56 (0.21 to 0.86)
West 2008[133]	4	1	3	22	RIA	PJ	R	0.57 (0.18 to 0.90)	0.96 (0.78 to 1.00)

FIGURE 45 Antibodies to ADA for predicting LOR. Response (Res) refers to method for estimating clinical response. FN, false negative; FP, false positive; PJ BM, physicians' judgement and biological measure; Pop, study patient population; R, responders; Res, response; RS, restarters on ADA; SA, stop anti-TNF-α; TN, true negative; TP, true positive.

TABLE 59 Test accuracy results from hierarchical meta-analysis (five studies)

Parameter	Point estimate	95% CI
Sens	0.471206	0.2903357 to 0.66
Spec	0.915467	0.7939073 to 0.968
DOR	9.65022	4.387759 to 21.22
LR+	5.574189	2.646268 to 11.74
LR–	0.577623	0.4208713 to 0.793
1/LR–	1.731233	1.261422 to 2.376

LR, likelihood ratio; Sens, sensitivity; Spec, specificity.

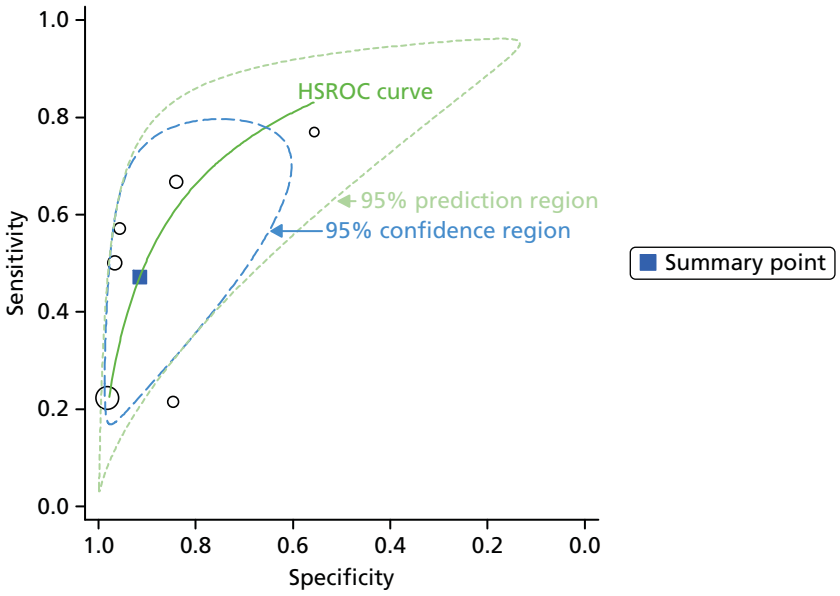


FIGURE 46 Antibodies to ADA for predicting LOR; hierarchical meta-analysis of test accuracy. HSROC, hierarchical summary receiver operating characteristic.

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.

The predictive values are indicative of moderate test accuracy so that between about 20% and 30% of positive and negative test results are likely to be incorrect.

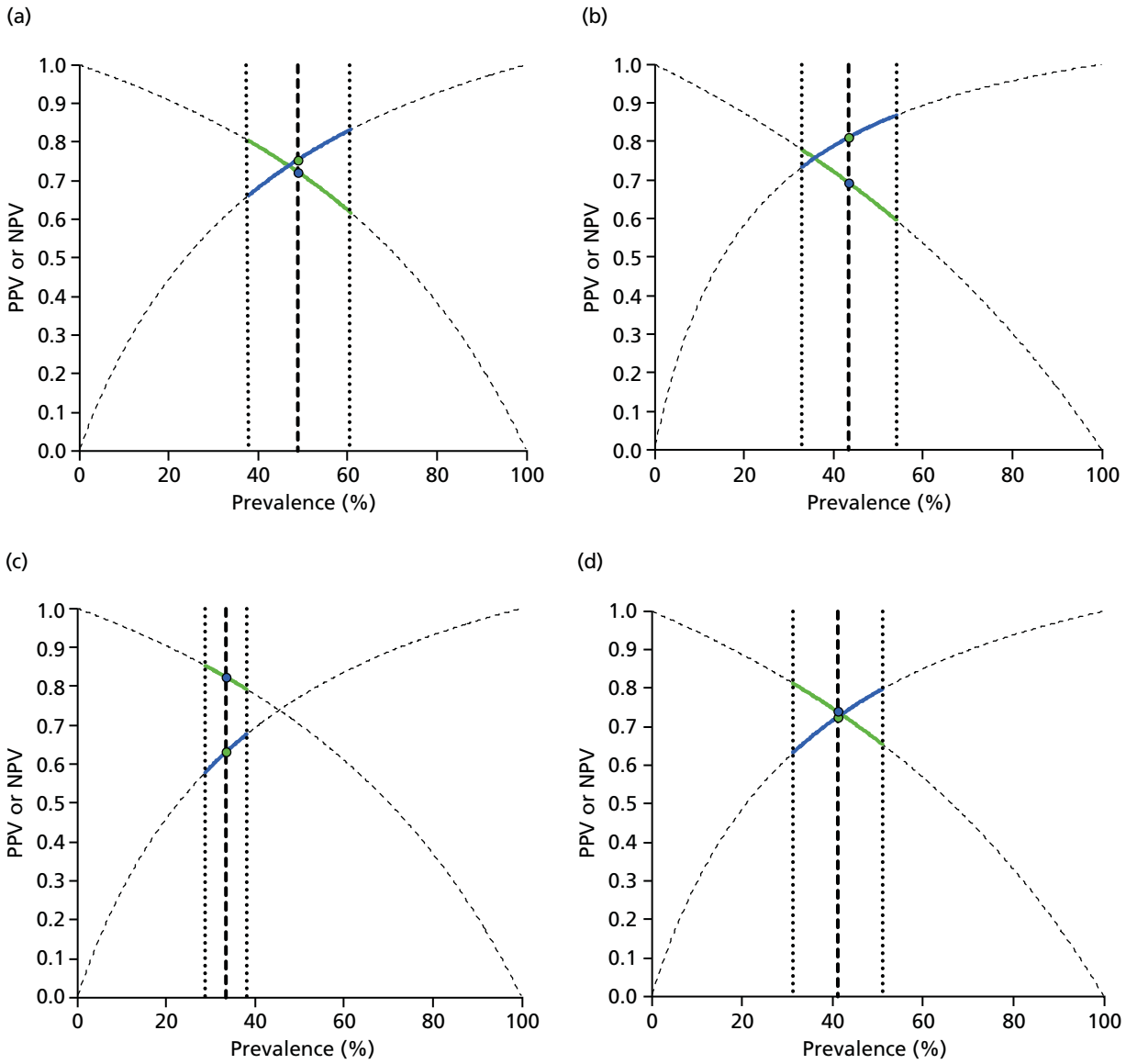


FIGURE 47 Positive predictive values and NPVs according to prevalence of LOR (or inability to regain response) at the sROC model estimate of sensitivity and specificity; (a) ADA; (b) antibodies to ADA; (c) IFX; and (d) antibodies to IFX. As prevalence increases PPV increases and NPV decreases. The data points are PPV and NPV at sROC sensitivity and specificity and pooled prevalence. Dashed vertical lines are pooled prevalence and 95% CI. Thick curves are PPV and NPV at hierarchical model sensitivity and specificity across at pooled prevalence and 95% CI.

Appendix 13 List of excluded cost-effectiveness studies with reason

TABLE 60 List of excluded studies from the literature review

Reference	Reason(s) for exclusion
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**:77–85	No testing kits used to monitor anti-TNF-α or antibodies to anti-TNF-α levels
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**:265–74	No testing kits used to monitor anti-TNF-α or antibodies to anti-TNF-α levels
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. *J Crohns Colitis* 2011;**5**:301–16	No testing kits used to monitor anti-TNF-α or antibodies to anti-TNF-α levels
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)	No testing kits used to monitor anti-TNF-α or antibodies to anti-TNF-α levels
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**:1509–20	No testing kits used to monitor anti-TNF-α or antibodies to anti-TNF-α levels

Appendix 14 Data extraction sheets of included health economic studies

Name of first reviewer: Hema Mistry.

Name of second reviewer: Peter Auguste.

Study details	Notes
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[157]
First author	Fernando S Velayos
Coauthors	Ames G Kahn, William J Sandborn and Brian G Feagan
Source of publication: journal year;volume:pp.	*Clinical Gastroenterology and Hepatology* 2013;**11**:654–66
Language	English
Publication type	Journal article
Inclusion criteria/study eligibility/PICOS	
Population	Patients with CD who become unresponsive to therapy with TNF antagonists – IFX
Intervention(s)	Testing-based strategy
Comparator(s)	Empiric dose escalation strategy
Outcome(s)	Cost per QALY gained
Study design	Cost-effectiveness analysis
Methods	
Target population and subgroups	Patients with moderate to severe active CD
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	QALYs
Measurement and valuation of preference-based outcomes	Not reported
Resource use and costs	Direct medical costs included: cost of the interventions – IFX, ADA, certolizumab, natalizumab and surgery; and the cost of diagnostics: anti-IFX antibody/serum IFX measurement, computerised tomography enterography and colonoscopy
Currency, price date and conversion	US dollars
Model type	Decision-analytical model

Study details	Notes
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 IFX dose escalation was assumed to be equal to that of ADA switching The presence of drug antibody, drug concentration and inflammation accurately categorises the mechanism for LOR and the proposed interventions represent the best approach to remedy a given mechanism
Analytical methods	ICERs were presented. Extensive one-way sensitivity analyses were conducted and PSAs using 10,000 simulations determined uncertainty in model results
Results	
Study parameters	Proportion with mild/minimal inflammation with symptoms • Initial response: switching to ADA, anti-IFX antibody present, subtherapeutic and therapeutic IFX concentrations, IFX increase to 10 mg/kg, anti-IFX antibody present, subtherapeutic and therapeutic IFX concentration • Sustained response at 1 year: ADA switch, IFX increase to 10 mg/kg, ADA increase to 40 mg every week, IFX 5 mg/kg maintenance, surgery switch, sustained responders in remission, restart biological for post-operative 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 to the empiric strategy (0.801 vs. 0.800, respectively), but was less expensive (US$31,870 vs. US$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 fivefold more than the base case. PSAs 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 TNF-α switching and IFX 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 LOR
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	

PICOS, participants, interventions, comparisons, outcomes, study design.

Name of first reviewer: Hema Mistry.

Name of second reviewer: Peter Auguste.

Study details	Notes
Study title	Individualised therapy is more cost-effective than dose intensification in patients with Crohn's disease who lose response to anti-TNF-α treatment: a randomised, controlled trial[123]
First author	Casper Steenholdt
Coauthors	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 and Mark Andrew Ainsworth
Source of publication: journal year;volume:pp.	*Gut* 2014;**63**:919–27
Language	English
Publication type	Journal article
Inclusion criteria/study eligibility/PICOS	
Population	Eligible adult patients with CD
Intervention(s)	Receive treatment based on serum concentrations of IFX and IFX antibodies at the time of IFX treatment failure in accordance with the algorithm
Comparator(s)	Receive IFX at an increased dose frequency of 5 mg/kg every 4 weeks
Outcome(s)	Cost per ITT and cost PP population
Study design	Randomised, controlled, single-blind, clinical trial
Methods	
Target population and subgroups	All patients had secondary IFX treatment failure on IFX maintenance therapy defined as recurrence of active disease with a CDAI score of ≥ 220 and/or a minimum of one draining perianal fistula Subgroup analyses included: proposed mechanisms for therapeutic failure, assessment of coprimary end points in patients stratified for 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 – LOR to IFX 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 relating to treatment of CD such as diagnoses and diagnostic and treatment procedures were recorded, as well as standardised IFX doses. Expenses related to CD in the 12 months before inclusion were comparable between randomisation groups
Currency, price date and conversion	Danish kroner and converted into Euros. Price date 1 January 2012
Model type	Not applicable
Assumptions	Not applicable

Study details	Notes
Analytical methods	Costs were compared using arithmetic means and were assessed by non-parametric bootstrap analysis to determine statistical significance. Data were analysed by ITT and PP population. One-way sensitivity analyses of key primary and secondary end points conducted
Results	
Study parameters	Primary end points: costs of CD and clinical response Secondary end points included CDAI 100 response, clinical remission, CDAI decrease, PDAI decrease and IBDQ increase
Incremental costs and outcomes	Costs were significantly lower in the algorithm group than in the IFX intensification group in both the ITT population (mean difference per patient –€3141) and the PP population (mean difference per patient –€5116) Response rates in the ITT population were 58% in the algorithm group and 53% in the IFX intensification group (RR 1.091, 95% CI 0.713 to 1.673). The difference between response rates was 5% in favour of the algorithm group. In the PP population, 47% in the algorithm group and 53% in the IFX intensification group showed a clinical response (RR 0.898, 95% CI 0.510 to 1.580) ICERs were not reported
Characterising uncertainty	One-way sensitivity analyses included (1) estimated administrative costs for biological drugs, (2) use of actual IFX 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 or not 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 IFX treatment failure by an algorithm based on serum IFX and IFX antibodies to define the mechanistic basis and corresponding interventions is more cost-effective than an intensified IFX regimen	
Reviewer's conclusion	
Although patient numbers were small, the authors used appropriate trial evidence to demonstrate the cost-effectiveness of an algorithm-based strategy compared with an intensified dose strategy	

PICOS, participants, interventions, comparisons, outcomes, study design.

Name of first reviewer: Hema Mistry.

Name of second reviewer: Peter Auguste.

Study details	Notes
Study title	Trough concentrations of infliximab guide dosing for patients with inflammatory bowel disease[73]
First author	Niels Vande Casteele
Coauthors	Marc Ferrante, Gert Van Assche, Vera Ballet, Griet Compernolle, Kristel Van Steen, Steven Simoens, Paul Rutgeerts, Ann Gils and Séverine Vermeire
Source of publication: journal year;volume:pp.	*Gastroenterology* 2015;**148**:1320–9
Language	English
Publication type	Journal article
Inclusion criteria/study eligibility/PICOS	
Population	Patients with a diagnosis of moderate to severe CD or UC
Intervention(s)	Concentration-based IFX dosing
Comparator(s)	Clinically based IFX dosing
Outcome(s)	Cost per QALY
Study design	RCT
Methods	
Target population and subgroups	Cohort of CD (and UC) responder patients. Patients needed to be treated with maintenance IFX 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	QALYs
Measurement and valuation of preference-based outcomes	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
	ICERs were presented. Uncertainty in incremental QALYs and costs was determined by non-parametric bootstrapping consisting of 1000 iterations and plotted onto a cost-effectiveness plane
Results	
Study parameters	Primary end points: clinical and biochemical remission at 1 year after the optimisation phase (increasing and maintaining remission)
	Secondary end points: durable remission, relapse, IFX trough concentration within the optimal interval, antibodies to IFX positivity, total cost of IFX treatment and QALYs

Study details	Notes
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
	ICER = €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	
PICOS, participants, interventions, comparisons, outcomes, study design.	

Name of first reviewer: Hema Mistry.

Name of second reviewer: Peter Auguste.

Study details	Notes
Study title	Individualized therapy is a long-term cost-effective method compared to dose intensification in Crohn's disease patients failing infliximab[124]
First author	Casper Steenholdt
Coauthors	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 and Mark A Ainsworth
Source of publication: journal year;volume:pp.	*Dig Dis Sci* 2015;**60**:2762–70
Language	English
Publication type	Journal article
Inclusion criteria/study eligibility/PICOS	
Population	Eligible adult patients with CD
Intervention(s)	Receive treatment based on serum concentrations of IFX and IFX antibodies at the time of IFX treatment failure in accordance with the algorithm

Study details	Notes
Comparator(s)	Receive IFX at an increased dose frequency of 5 mg/kg every 4 weeks
Outcome(s)	Cost per ITT and cost PP population
Study design	Randomised, controlled, single-blind, clinical trial
Methods	
Target population and subgroups	All patients had secondary IFX treatment failure on IFX maintenance therapy defined as recurrence of active disease with a CDAI score of ≥ 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 a ≥ 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 a CDAI score of ≤ 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 relating to treatment of CD such as diagnoses and diagnostic and treatment procedures were recorded, as well as standardised IFX and anti-IFX doses
Currency, price date and conversion	Danish kroner and converted into US dollars. 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 ITT, PP population, PP completion at end of trial week 12, and PP completion at end of follow-up week 20. One-way sensitivity analyses of key primary and secondary end points conducted
Results	
Study parameters	End points: costs of CD, 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 • ITT: –US$5296 • PP: –US$8494 • PP end of trial week 12: –US$8546 • PP end of follow-up week 20: –US$10,720 1 year • ITT: –US$7006 • PP: –US$13,383 • PP end of trial week 12: –US$13,265 PP end of follow-up week 20: –US$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-week and 1-year time frames

Study details	Notes
Discussion	
Study findings	The algorithm group had significantly lower costs than 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 IFX treatment failure based on monitoring of IFX and anti-IFX antibodies are long-term cost-effective method compared with IFX 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	

PICOS, participants, interventions, comparisons, outcomes, study design.

Name of first reviewer: Hema Mistry.

Name of second reviewer: Peter Auguste.

Study details	Notes
Study title	A systematic review and economic evaluation of the use of tumour necrosis factor-alpha (TNF-α) inhibitors, adalimumab and infliximab, for Crohn's disease[5]
First author	J Dretzke
Coauthors	R Edlin, J Round, M Connock, C Hulme, J Czeczot, A Fry-Smith, C McCabe and C Meads
Source of publication: journal year;volume:pp.	*Health Technology Assessment* 2011;**15**(6)
Language	English
Publication type	Monograph
Inclusion criteria/study eligibility/PICOS	
Population	Adult patients with moderate to severe CD
Intervention(s)	Anti-TNF-α therapy for CD – IFX and ADA
Comparator(s)	Standard care for CD
Outcome(s)	Cost per QALY gained
Study design	Cost-effectiveness analysis
Methods	
Target population and subgroups	Adult patients with moderate to severe CD 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

Study details	Notes
Discount rate	Not reported
Measurement of effectiveness	QALYs
Measurement and valuation of preference-based outcomes	Choice-based time-trade off measure providing utility value
Resource use and costs	Cost of anti-TNF-α 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, whereas 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 2005–6
Model type	Markov model
Assumptions	Model did not take into account mortality. Used Silverstein *et al.*[150] for all transition probabilities in the intervention arm
Analytical methods	ICERs and cost-effectiveness acceptability curves were presented. One-way sensitivity analyses and PSAs 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 CD, both ADA and IFX dominated standard care (i.e. cheaper and more effective). For maintenance therapy for severe CD, neither drug was cost-effective (well above NICE thresholds) For moderate CD, for maintenance therapy for both drugs and induction therapy for IFX, these were not cost-effective (well above NICE thresholds); however, for induction therapy for ADA dominated standard care
Characterising uncertainty	Patients who had severe disease, IFX induction treatment was found to be cost-effective relative to maintenance treatment and standard care in > 99% of cases at all points up to £100,000 per QALY. Likewise, ADA 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 ADA and IFX were cost-effective (dominant relative to standard care) for severe CD and that ADA was cost-effective (dominant relative to standard care) for moderate CD. Induction therapy with IFX was not cost-effective for moderate CD. 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.*[150] data had its own problems, that is 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

Study details	Notes
Authors' conclusion	
IFX is not likely to be cost-effective in the management of moderate CD. Although ADA may be cost-effective, there is uncertainty regarding the ICERs value. Neither of these therapies is likely to be as 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-TNF-α drug therapies compared with standard care; although there are some limitations in terms of how the transition probabilities and utility values were estimated	

PICOS, participants, interventions, comparisons, outcomes, study design.

Appendix 15 Quality assessment of included health economic studies

Consolidated Health Economic Evaluation Reporting Standards quality assessment checklist for economic evaluation studies

Assessment	Study (first author and year of publication)				
	Velayos *et al.*, 2007[157]	Steenholdt *et al.*, 2014[123]	Vande Casteele *et al.*, 2015[73]	Steenholdt *et al.*, 2015[124]	Dretzke *et al.*, 2011[5]
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
Generalisability	P	P	N	P	N
Other					
Source of funding	Y	Y	Y	Y	Y
Conflicts of interest	Y	Y	Y	Y	N

N/A, not applicable; N, no; P, partial; UNC, unclear; Y, yes.

Philips' quality assessment checklist for studies that included an economic model

Philips' criteria	Studies (first author and publication year)	
	Velayos *et al.*, 2007[157]	Dretzke *et al.*, 2011[5]
Structure		
Is there a clear statement of the decision problem?	Y	Y
Is the objective of the model specified and consistent with the stated decision problem?	Y	Y
Is the primary decision-maker specified?	N	Y
Is the perspective of the model stated clearly?	Y	Y
Are the model inputs consistent with the stated perspective?	Y	Y
Has the scope of the model been stated and justified?	Y	Y
Are the outcomes of the model consistent with the perspective, scope and overall objective of the model?	Y	Y
Is the structure of the model consistent with a coherent theory of the health condition under evaluation?	Y	Y
Are the sources of the data used to develop the structure of the model specified?	Y	Y
Are the causal relationships described by the model structure justified appropriately?	UNC	Y
Are the structural assumptions transparent and justified?	Y	Y
Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Y	Y
Is there a clear definition of the options under evaluation?	Y	Y
Have all feasible and practical options been evaluated?	Y	Y
Is there justification for the exclusion of feasible options?	Y	Y
Is the chosen model type appropriate given the decision problem and specified casual relationships within the model?	Y	Y
Is the time horizon of the model sufficient to reflect all important differences between the options?	Y	Y
Are the time horizon of the model, the duration of treatment and the duration of treatment described and justified?	Y	Y
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
Is the cycle length defined and justified in terms of the natural history of disease?	Y	Y
Data		
Are the data identification methods transparent and appropriate given the objectives of the model?	Y	Y
Where choices have been made between data sources are these justified appropriately?	UNC	Y
Has particular attention been paid to identifying data for the important parameters of the model?	Y	Y
Has the quality of the data been assessed appropriately?	Y	Y
Where expert opinion has been used are the methods described and justified?	UNC	N/A

Philips' criteria	Studies (first author and publication year)	
	Velayos *et al.*, 2007[157]	Dretzke *et al.*, 2011[5]
Is the data modelling methodology based on justifiable statistical and epidemiological techniques?	UNC	Y
Is the choice of baseline data described and justified?	UNC	Y
Are transition probabilities calculated appropriately?	UNC	N
Has a half-cycle correction been applied to both costs and outcomes?	N	N
If not, has the omission been justified?	N	N
If relative treatment effects have been derived from trial data, have they been synthesised using appropriate techniques?	N/A	Y
Have the methods and assumptions used to extrapolate short-term results to final outcomes been documented and justified?	UNC	Y
Have alternative extrapolation assumptions been explored through sensitivity analysis?	Y	Y
Have assumptions regarding the continuing effect of treatment once treatment is complete been documented and justified?	UNC	Y
Have alternative assumptions regarding the continuing effect of treatment been explored through sensitivity analysis	UNC	Y
Are the costs incorporated into the model justified?	Y	Y
Has the source for all costs been described?	Y	Y
Have discount rates been described and justified given the target decision-maker?	N	N
Are the utilities incorporated into the model appropriate?	Y	Y
Is the source of utility weights referenced?	N	Y
Are the methods of derivation for the utility weights justified?	N	Y
Have all data incorporated into the model been described and referenced in sufficient detail?	Y	Y
Has the use of mutually inconsistent data been justified (i.e. are assumptions and choices appropriate?)	Y	Y
Is the process of data incorporation transparent?	Y	Y
If data have been incorporated as distributions, has the choice of distributions for each parameter been described and justified?	Y	Y
If data have been incorporated as distributions, is it clear that second order uncertainty is reflected?	UNC	Y
Have the four principal types of uncertainty been addressed?	N	N
If not, has the omission of particular forms of uncertainty been justified?	N	N
Have methodological uncertainties been addressed by running alternative versions of the model with different methodological assumptions?	N	Y
Is there evidence that structural uncertainties have been addressed via sensitivity analysis?	N	Y
Has heterogeneity been dealt with by running the model separately for different subgroups?	Y	Y
Are the methods of assessment of parameter uncertainty appropriate?	Y	Y
If data are incorporated as point estimates, are the ranges used for sensitivity analysis stated clearly and justified?	Y	Y

Philips' criteria	Studies (first author and publication year)	
	Velayos *et al.*, 2007[157]	Dretzke *et al.*, 2011[5]
Is there evidence that the mathematical logic of the model has been tested thoroughly before use?	N	Y
Are any counterintuitive results from the model explained and justified?	Y	Y
If the model has been calibrated against independent data, have any differences been explained and justified?	N	UNC
Have the results been compared with those of previous models and any differences in results explained?	N	N

N/A, not applicable; N, no; UNC, unclear; Y, yes.

Appendix 16 Decision tree structure for the responders' model

This appendix summarises the underlying decision tree structure of the model for responders to anti-TNF-α therapy in several figures for:

- concurrent testing (*Figures 48–53*)
- no testing (*Figures 54* and *55*)
- reflexing testing (*Figures 56–59*).

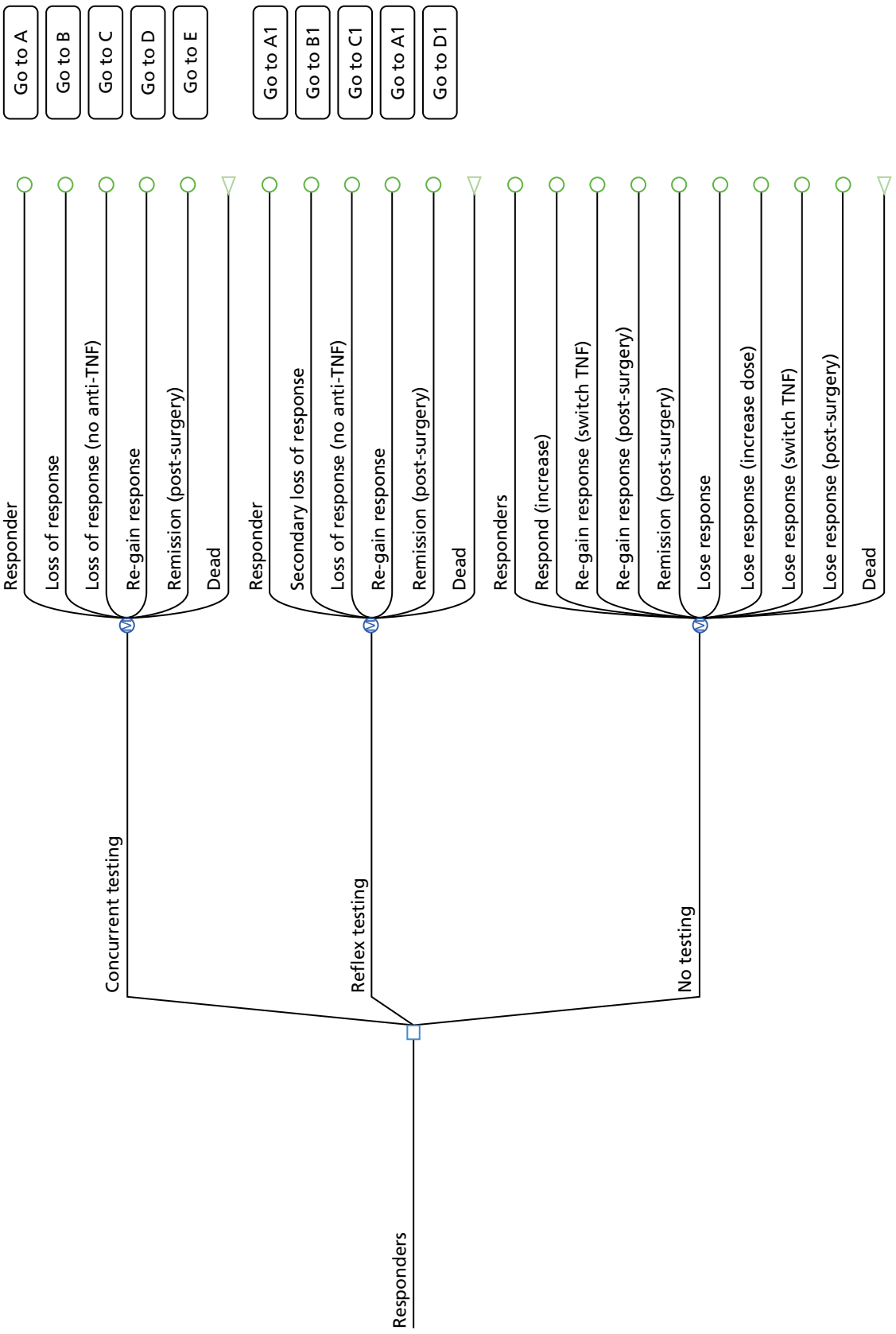


FIGURE 48 Decision tree structure for the responders' model for concurrent testing.

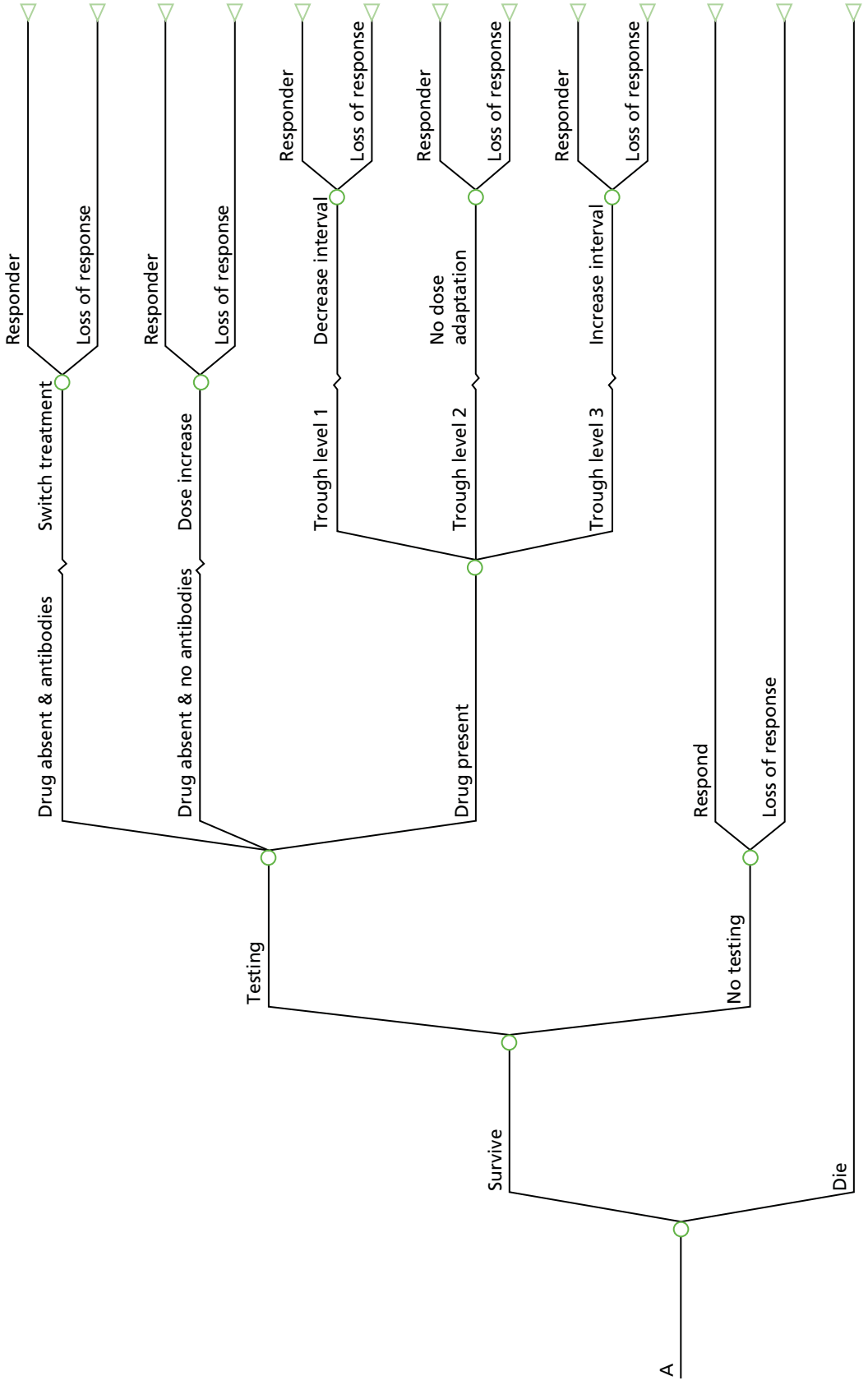


FIGURE 49 Decision tree structure for the responders' model for concurrent testing (A).

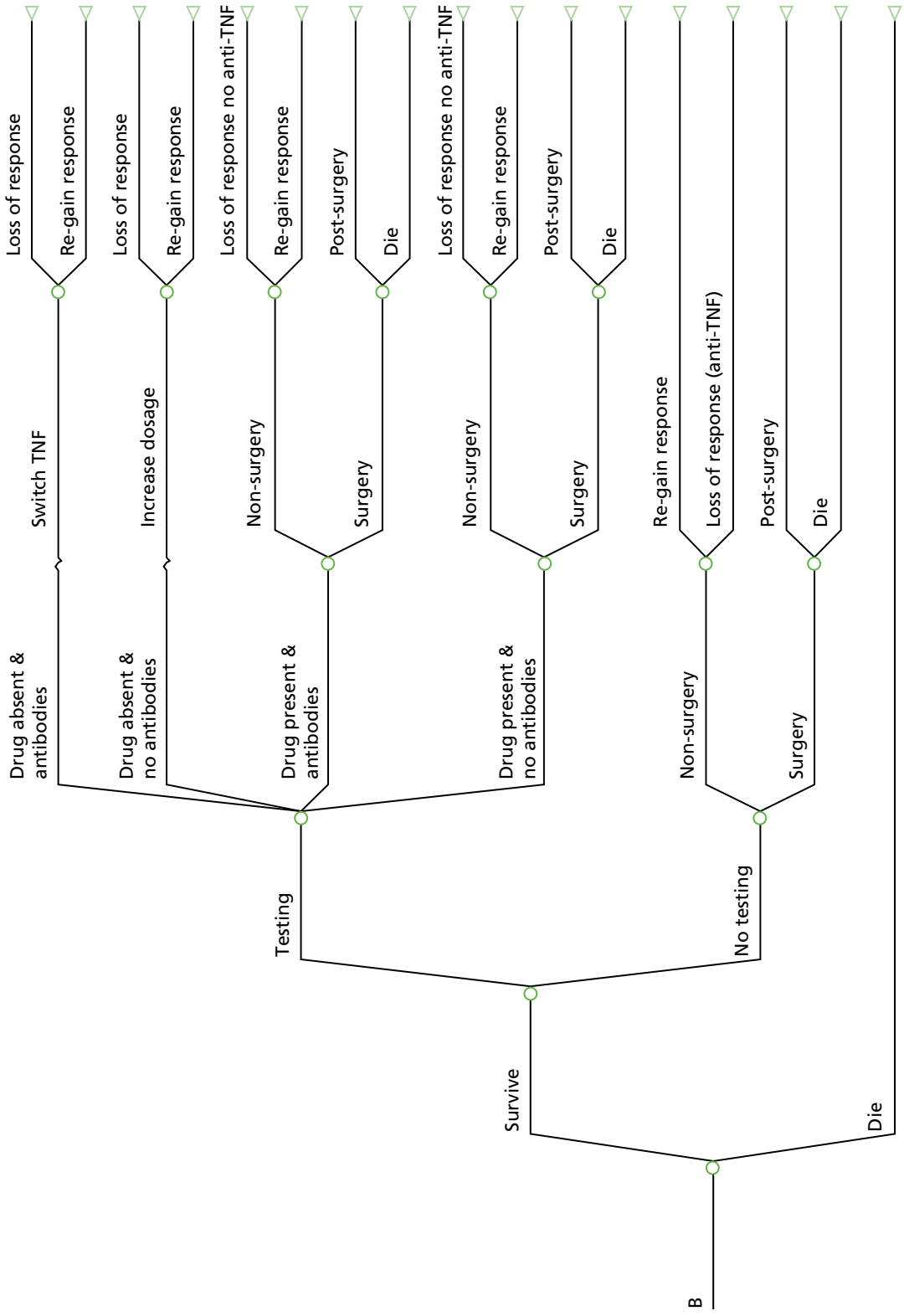


FIGURE 50 Decision tree structure for the responders' model concurrent testing (B).

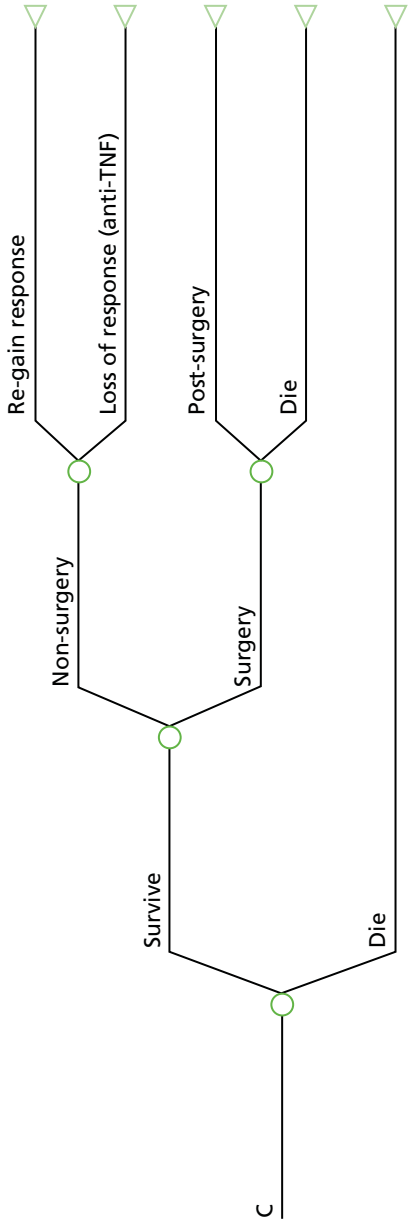


FIGURE 51 Decision tree structure for the responders' model concurrent testing (C).

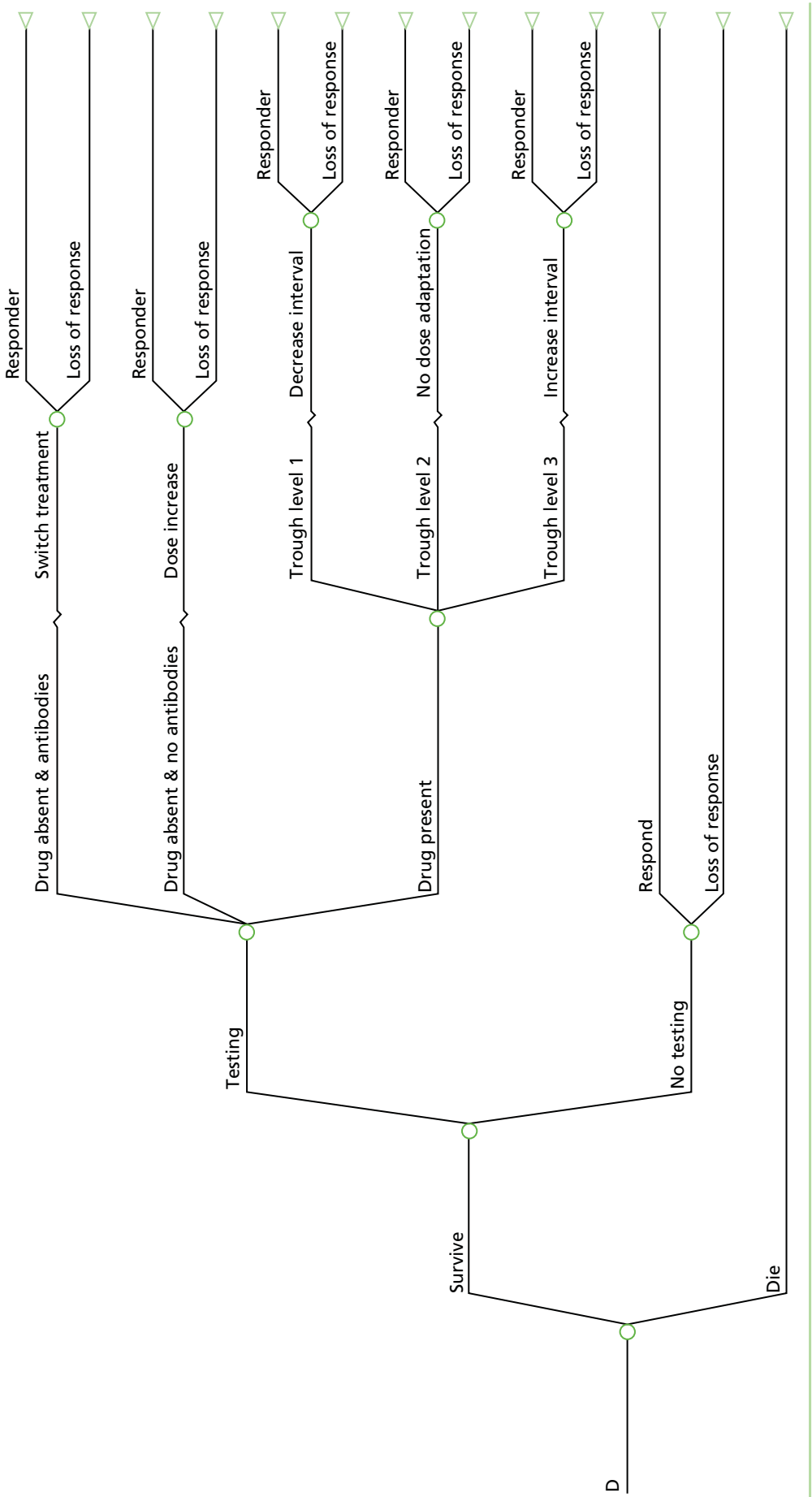


FIGURE 52 Decision tree structure for the responders' model concurrent testing (D).

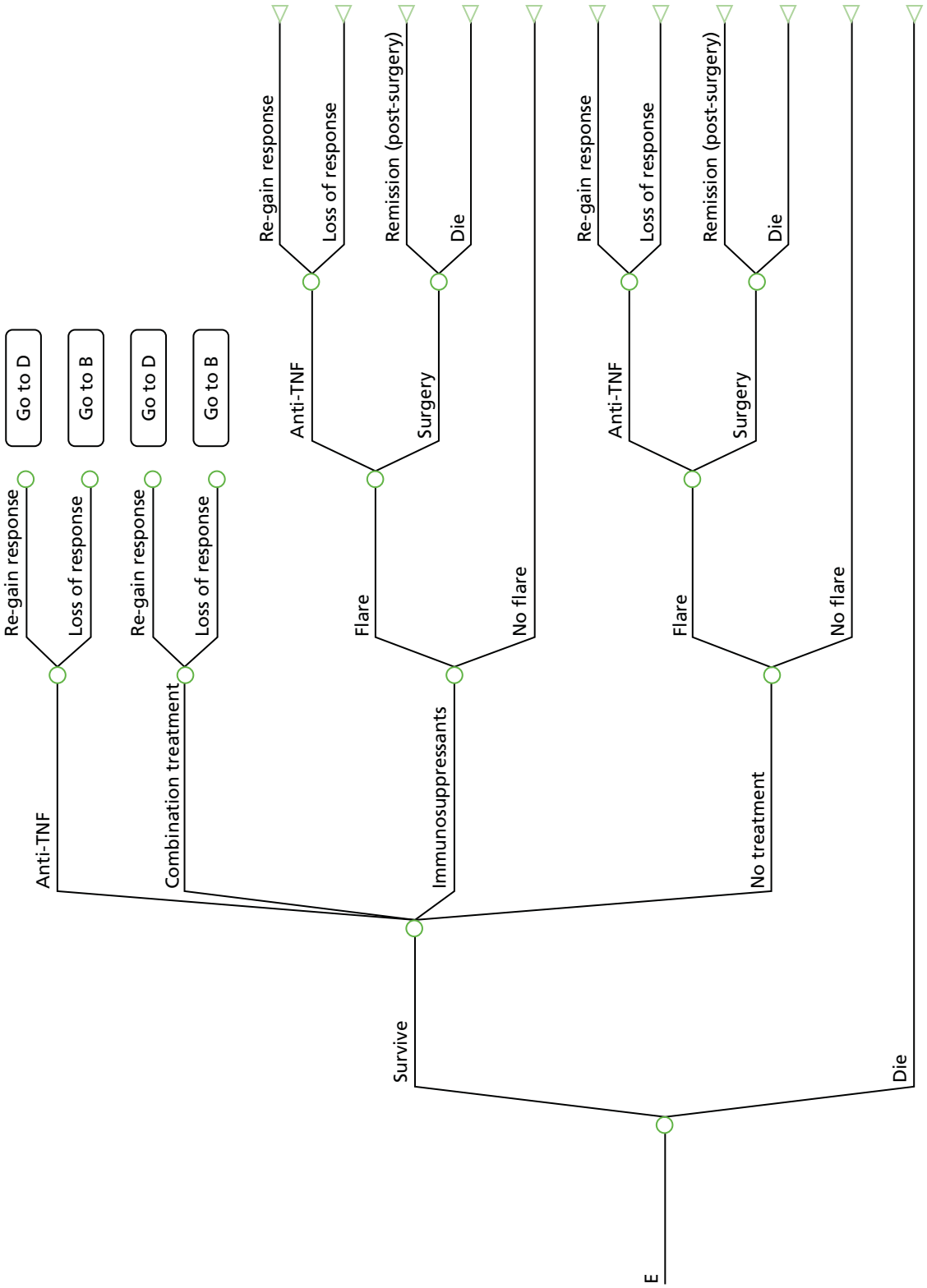


FIGURE 53 Decision tree structure for the responders' model concurrent testing (E).

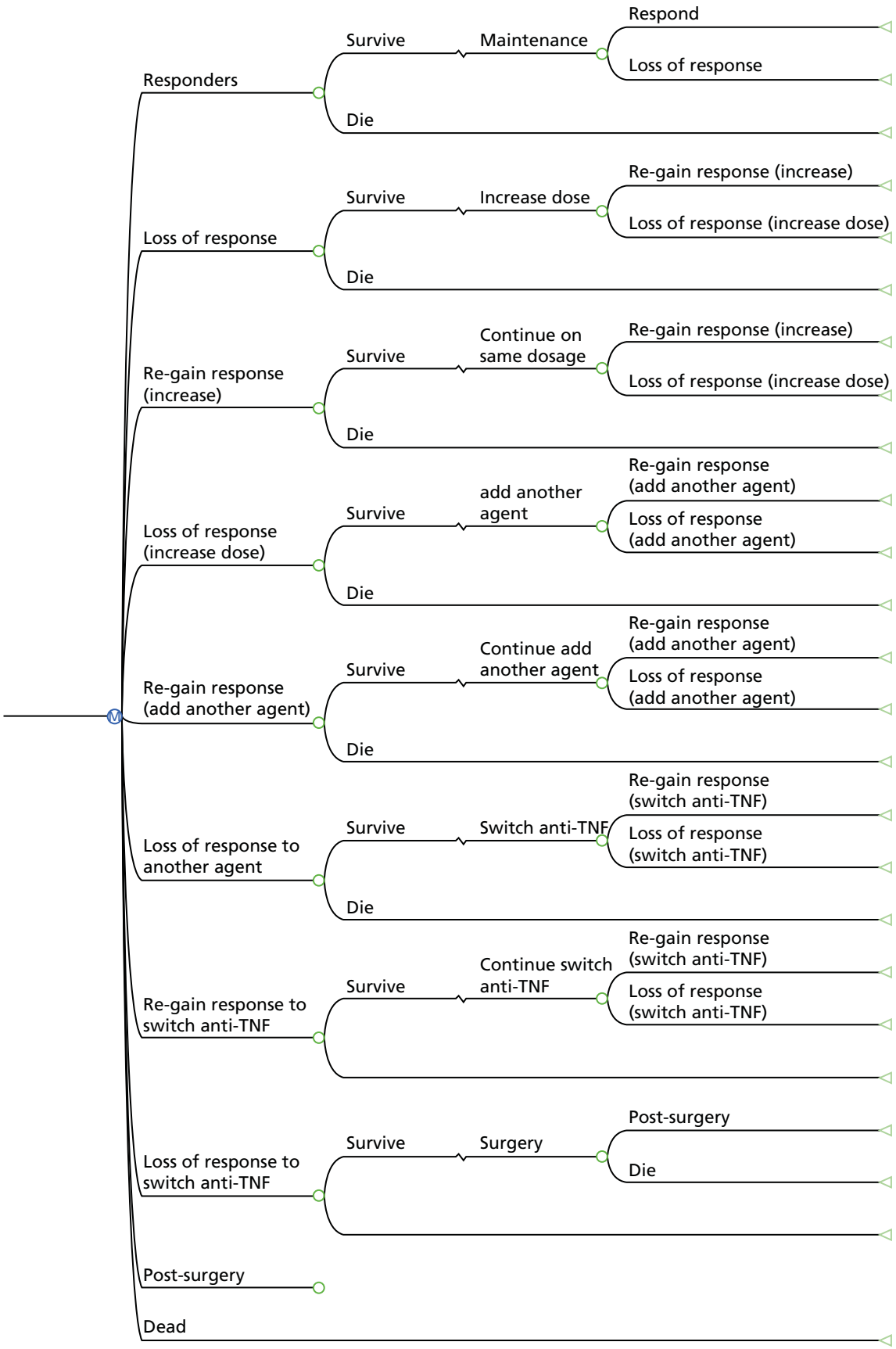


FIGURE 54 Decision tree structure for the no-testing strategy.

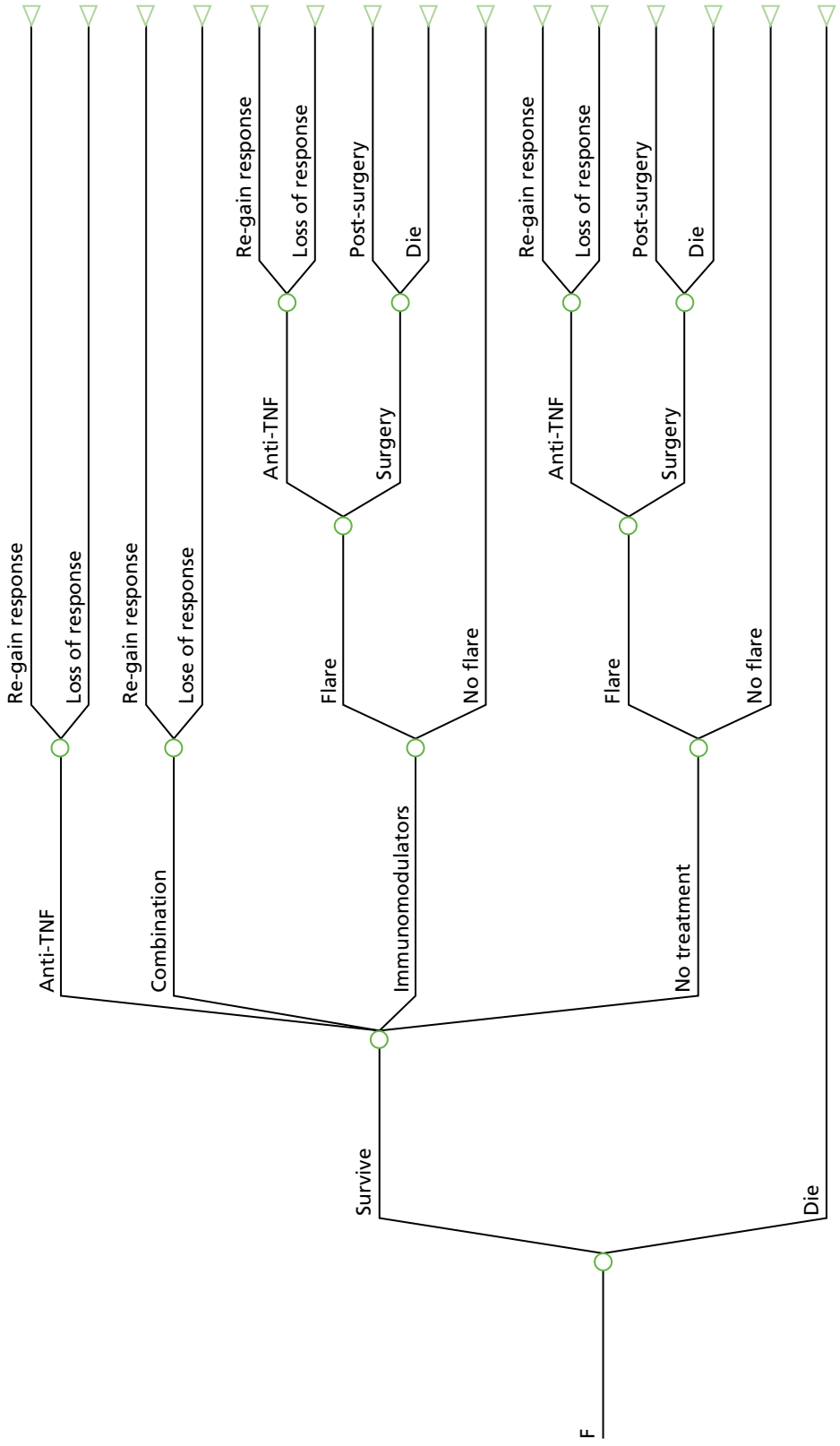


FIGURE 55 Patient pathway for patients in the post-surgery health state.

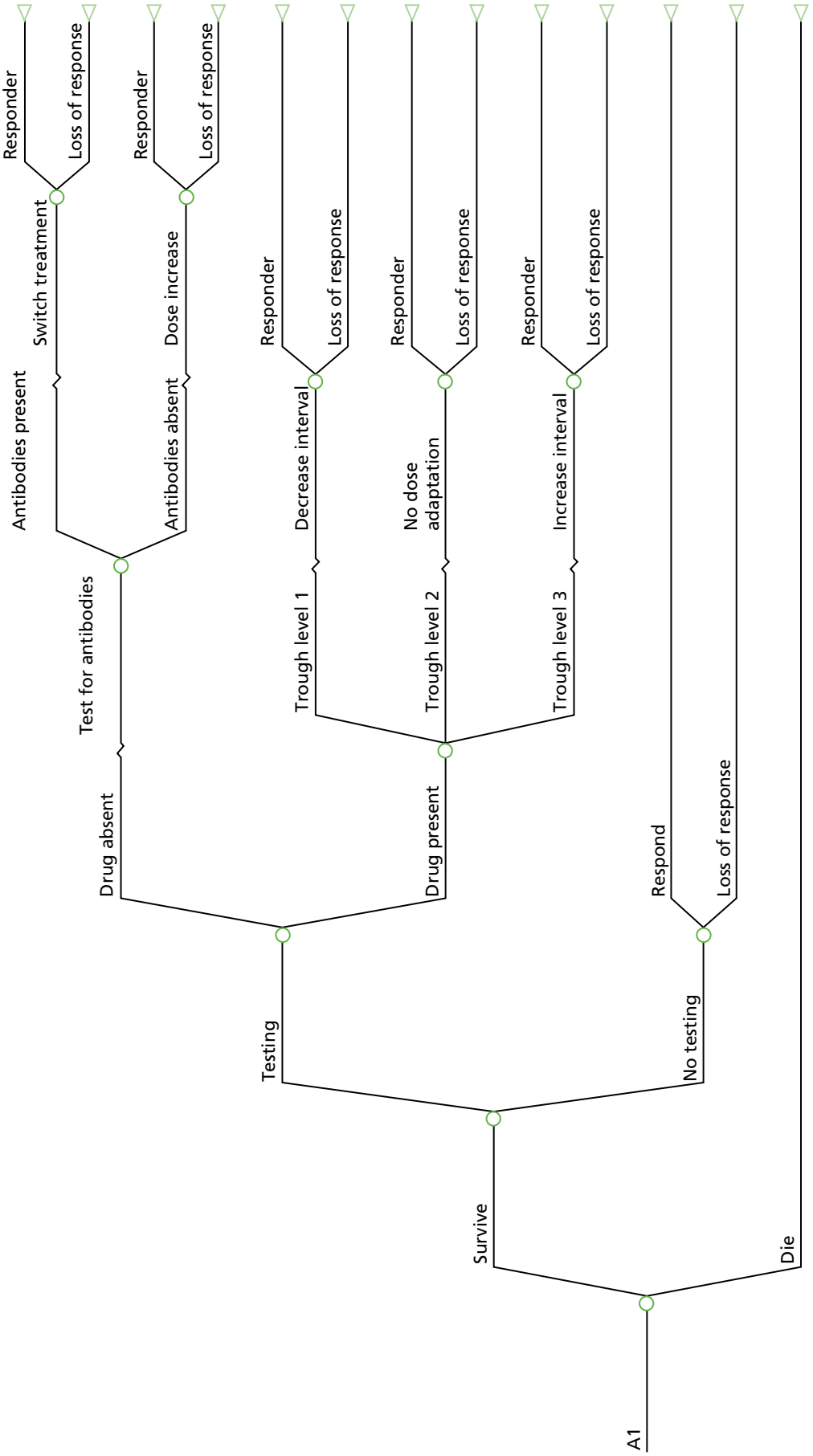


FIGURE 56 Decision tree structure for responders' model for reflex testing (A1).

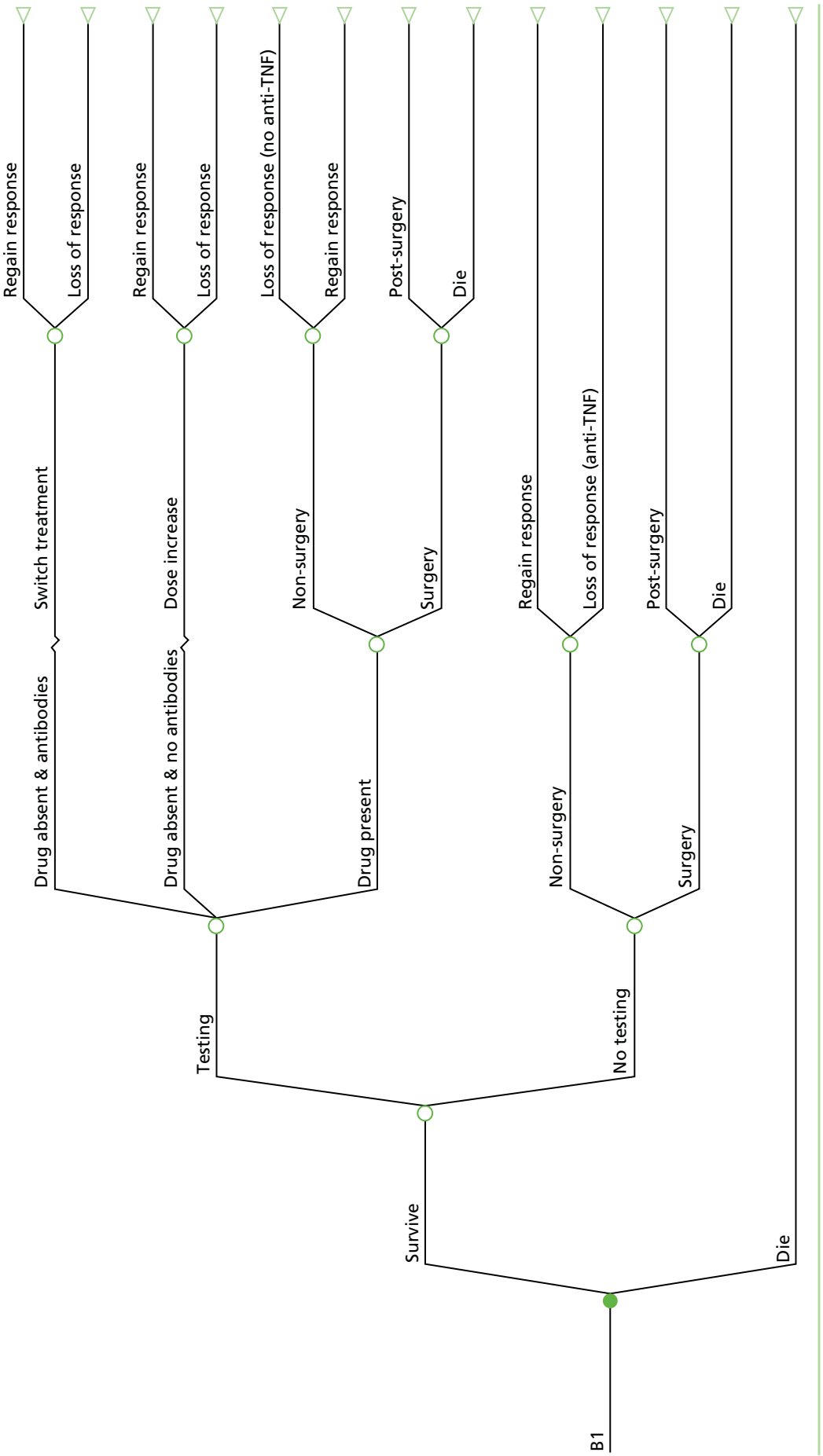


FIGURE 57 Decision tree structure for responders' model for reflex testing (B1).

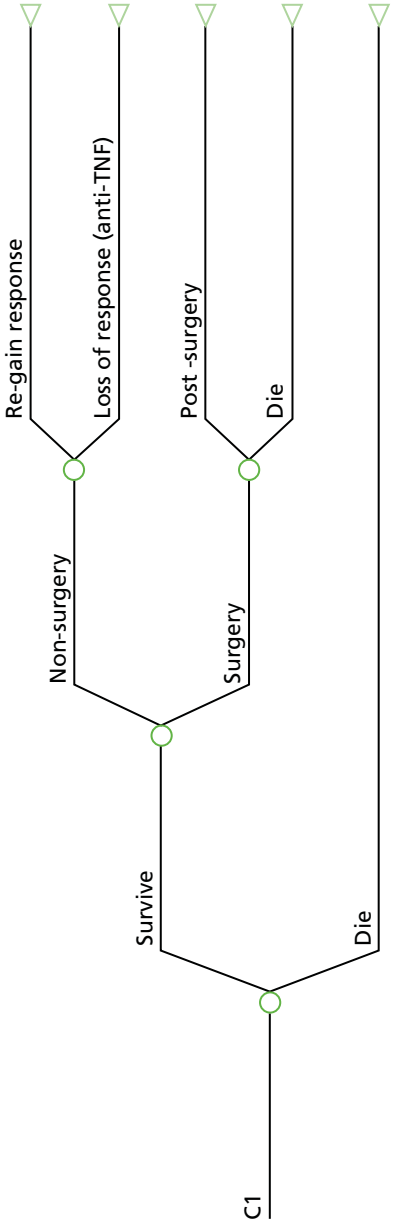


FIGURE 58 Decision tree structure for responders' model for reflex testing (C1).

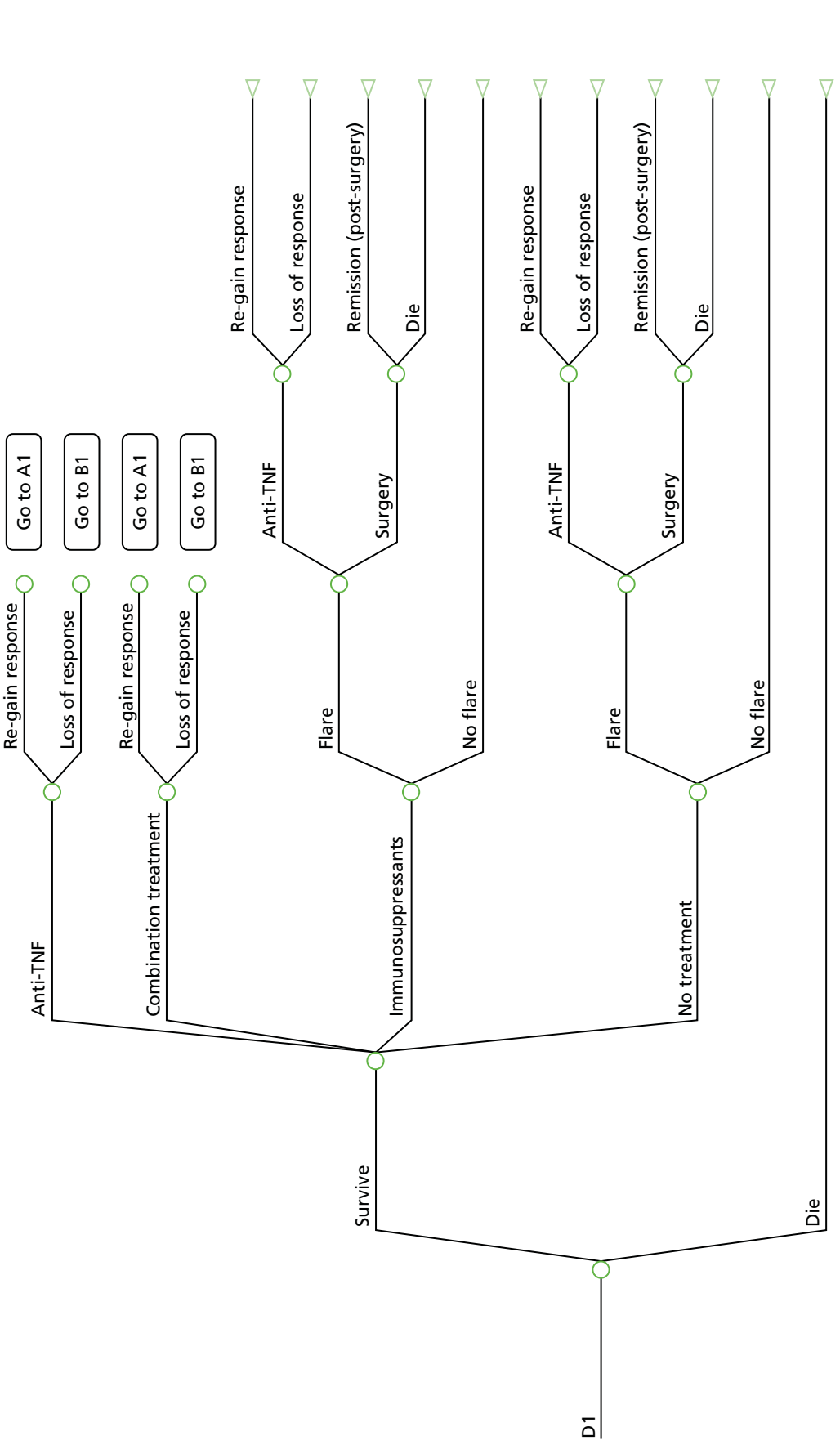


FIGURE 59 Decision tree structure for the responders' model for reflex testing (D1).

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 ADA test-based treatment strategy. The model therefore addresses patients responding to IFX maintenance therapy (the transition probabilities used are summarised in *Table 35*). It should be emphasised that there were no prospective or other test-directed management studies describing outcomes for IFX responders followed from maintenance treatment through to treatments subsequent to LOR to maintenance. Therefore, by necessity, model structure for the intervention arm is based on the algorithms used in the two identified RCTs describing test-based patient management, specifically the TAXIT trial[73] for responders and Steenholdt *et al.*[123] for patients with LOR to maintenance IFX (see *Chapter 3, Objective B: description of algorithms prescribing patient management following test outcomes for drug and/or anti-drug antibody levels*); we aimed to use as many data from these RCTs as possible to populate the model. Unfortunately the control arm in the TAXIT trial does not provide information for the model's standard care arm (a no-test management strategy) because all patients in the TAXIT trial 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 that reported reasonable quality data for time to LOR or to cessation of IFX treatment for patients on maintenance treatment with IFX were identified.[82,128,159] Reconstructed Kaplan–Meier plots with candidate parametric models are shown in *Figure 60*. For the Juillerat *et al.*[159] study, several models provided reasonable fit to 130 cycles.

These three studies generate fairly 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 *et al.* study[159] was selected for model inputs. In Juillerat *et al.*,[159] 21% of patients received dose escalation, but the time to escalation was not reported. However, Ma *et al.*[161] have reported the time to LOR requiring dose escalation for patients with CD on IFX maintenance therapy; Weibull and Gompertz models provided best fits to the Ma *et al.*[161] data.

Figure 61 shows both Juillerat *et al.*[159] and Ma *et al.*[161] 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: (1) untreated with IFX, (2) in standard dose treatment with IFX and (3) in escalated dose treatment with IFX. The resulting percentages were 35.6%, 24.0% and 40.4%, respectively.

Standard care arm treatment after loss of response to infliximab

On failure of response to IFX maintenance (with or without dose escalation) it is assumed patients are switched to ADA induction therapy followed by maintenance on ADA for those responding to induction. We classify those that fail induction as patients who have lost response during the first cycle of treatment. We have taken this from the Gauging Adalimumab Efficacy in Infliximab Nonresponders (GAIN) RCT,[162] which investigated ADA for patients who had failed IFX. Thereafter, the transition probability for LOR to

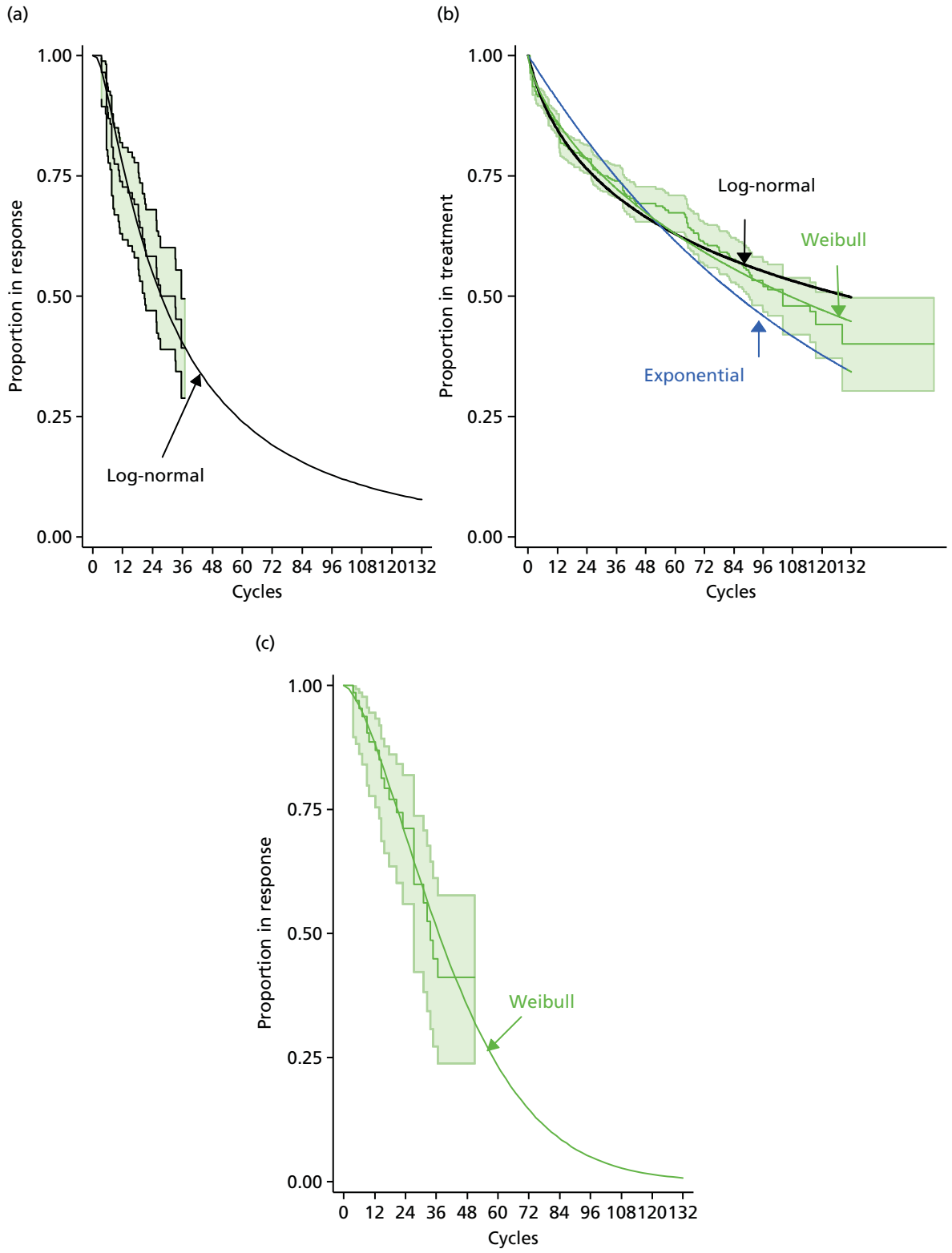


FIGURE 60 Reconstructed Kaplan–Meier plots for time to LOR or to cessation of treatment of responders on maintenance IFX therapy by 4-week cycle. (a) Bortlik *et al.*[82] ($n = 84$); (b) Juillerat *et al.*[159] ($n = 1014$); and (c) Vaughn *et al.*[128] ($n = 68$).

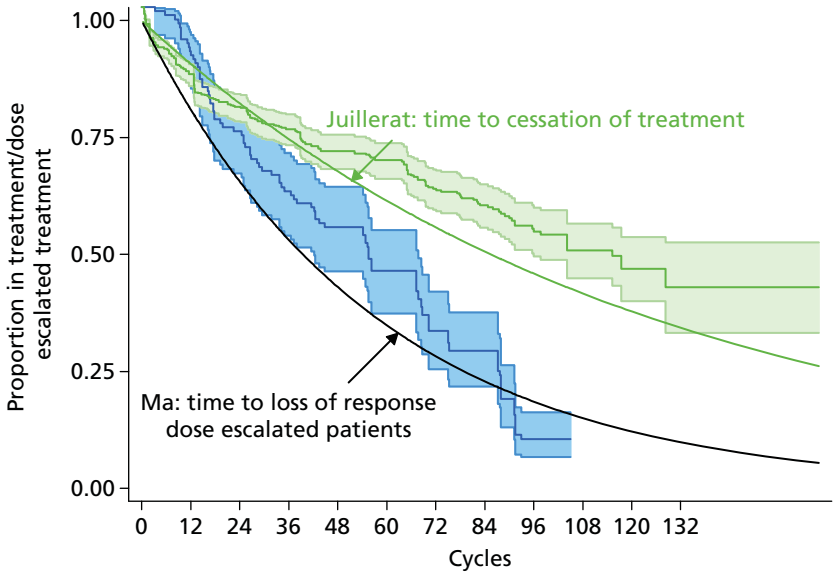


FIGURE 61 Reconstructed Kaplan–Meier plots and exponential fits for time to cessation of IFX treatment and time to LOR requiring dose escalation of IFX by 4-week cycle (studies of Juillerat *et al.*[159] and Ma *et al.*[161]).

ADA was derived from the study by Karmiris *et al.*[48] of 152 CD responders receiving ADA followed up prospectively (*Figure 62*). Exponential and Weibull distributions provided a reasonable fit to reconstructed IPD. Combined data from Sandborn *et al.*[162] for induction failure on ADA (after IFX failure) and Karmiris *et al.*[48] for failure after successful induction with ADA, provided a transition probability of 0.058553 per cycle that was used in the model.

After failure of ADA we have assumed patients remain in a LOR 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.[163]

Time to surgery

No data were found for time to surgery for patients who experience LOR or a failure to regain response after a treatment switch aimed to reinstate a response. We identified three studies[163,183,184] that provided time from diagnosis to surgery for recent cohorts of patients with CD (i.e. coincident with the era of

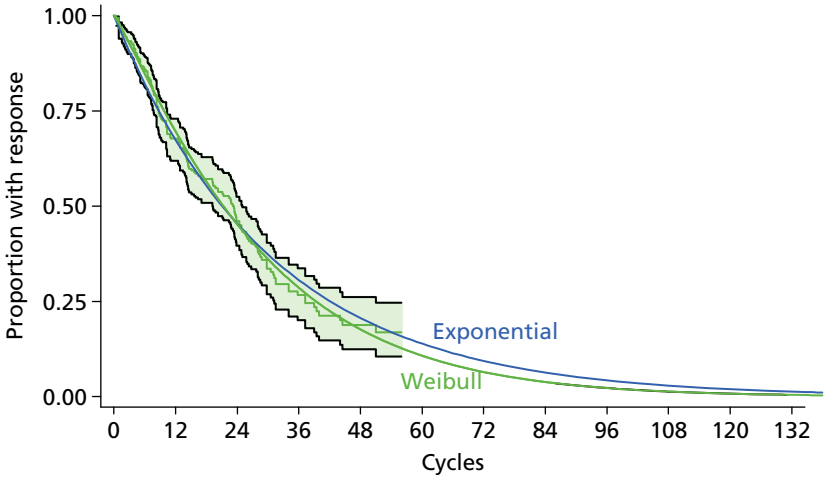


FIGURE 62 Reconstructed Kaplan–Meier and Weibull model for time to LOR for patients with CD on maintenance therapy with ADA by 4-week cycle.[48]

anti-TNF-α therapies for CD). Vester-Anderson *et al.*[183] reported surgical relapse rates of 6%, 18% and 23% at 1, 5 and 7 years (91 cycles) after diagnosis, respectively; similarly a UK study[175] that included 137 patients observed approximately 24% primary surgery 5 years after diagnosis (*Figure 63*) and a larger Canadian study[163] included > 1000 patients and also data for recurrent surgery (see *Figure 63*). *Figure 63* shows the time to primary surgery was similar in the UK and Canadian studies. As a result of its size and because it provided data for both primary and recurrent surgery, the Canadian study was used in the economic modelling for both primary and recurrent surgery.

Crohn's disease patients in the TAXIT trial[73] and Steenholdt *et al.*[123,124] management studies varied considerably in the time from diagnosis to study entry and also in if they had experienced previous surgery (e.g. TAXIT trial[73] patients, on average, were diagnosed 13.7 years prior to entry and 70% had received previous surgery; in Steenholdt *et al.*[123,124] patients were diagnosed, on average, 9 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 *et al.*[123,124] and 1/251 by week 52 in the TAXIT trial[73]). It appears that during the short follow-up periods observed the use of surgery was a relatively rare event. The Weibull and Gompertz parametric models (see *Figure 63*) of time to surgery generated probabilities of progressing to surgery that varied considerably according to time from diagnosis; the economic model was not based on newly diagnosed patients, and data for other health states used in modelling were not based on newly diagnosed patients; in the absence of more appropriate data transition

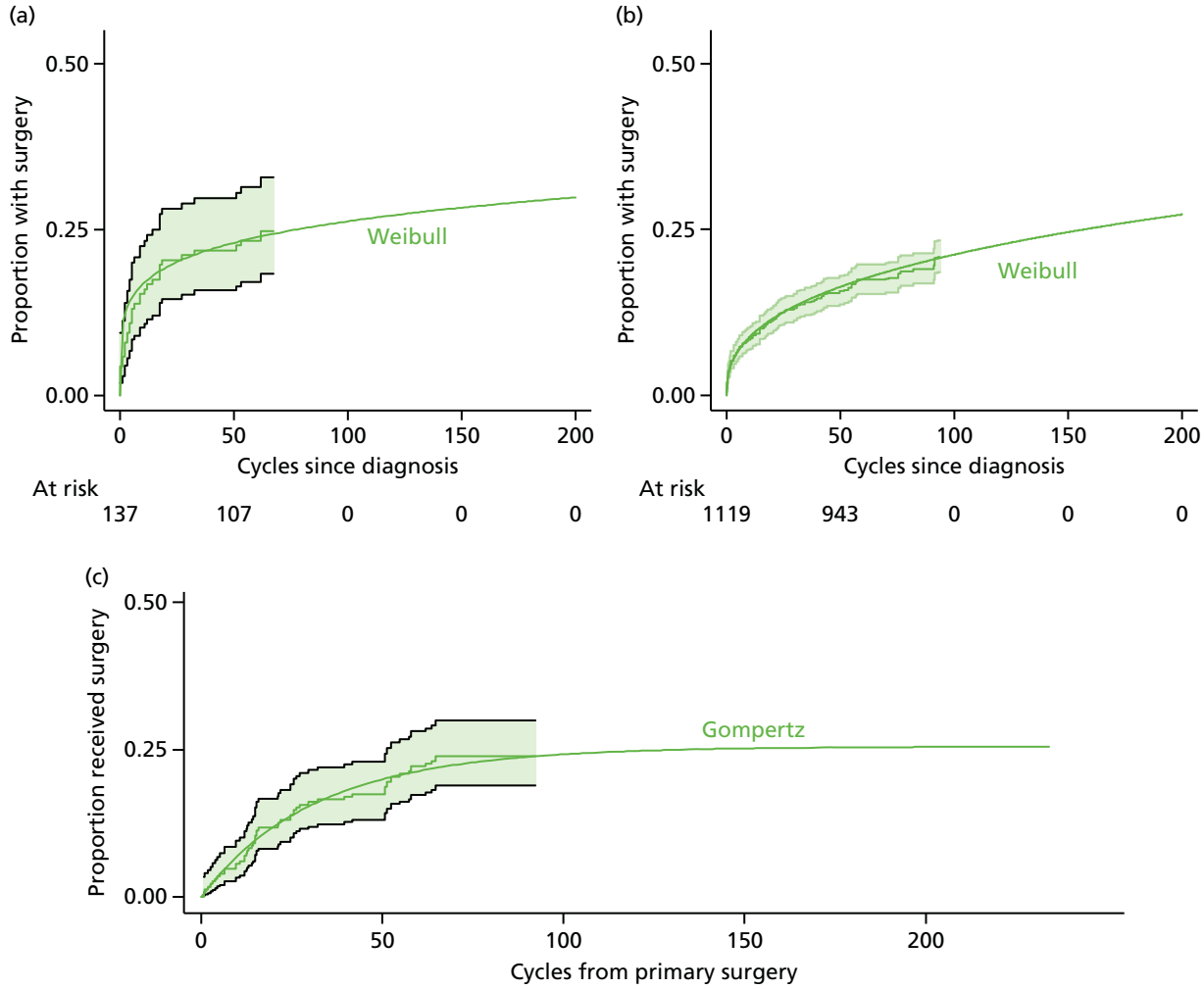


FIGURE 63 Time from diagnosis of CD (a) to primary surgery in the UK; (b) to primary surgery in Canada; and (c) time from first surgery to second surgery in Canada.

probability to surgery we used an exponential (constant hazard) fit to the Canadian data so as to capture an approximate average transition probability (see *Table 37*). It is recognised that these selections are somewhat arbitrary and that modelling extends beyond the observed data.

Maintenance of surgery-induced remission

The scant evidence about maintenance of surgically induced remission in CD was reviewed by Gordon *et al.*[164] 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 2 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 among 265 patients. Assuming a constant hazard the estimated transition probability to post-surgical clinical relapse is 0.023971 per cycle (95% CI 0.025398 to 0.035624 per cycle). In the economic model, this was taken to apply for both therapies (5-ASA and purine analogues). Relative to purine analogues, the review data suggest that 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 per cycle (95% CI 0.033248 to 0.108412 per cycle); this was used in the model for the group given no therapy. One study[185] included in the Gordon *et al.*[164] review found a RR for clinical relapse of 0.5 for IFX versus purine analogues; this study observed only three events among 22 patients giving, on assumption of constant hazard, a transition probability to clinical relapse for IFX-treated patients of 0.0119855 per cycle.

In view of the considerable uncertainty necessarily associated with this estimate of response loss with IFX, and the lack of information on timing of events, we looked for alternative data. Baert *et al.*[77] reported time-to-event data for reintroduction of IFX following at least 15 months after LOR despite dose optimisation. During the ≥ 15-month IFX holiday some patients received surgery. Time to LOR after IFX reintroduction is shown in *Figure 64* together with the exponential fit used to estimate transition probabilities for the economic model. Owing to a lack of data, we have assumed the same transition probabilities for patients receiving anti-TNF-α in combination with immunosuppressants to be the same as that for IFX alone.

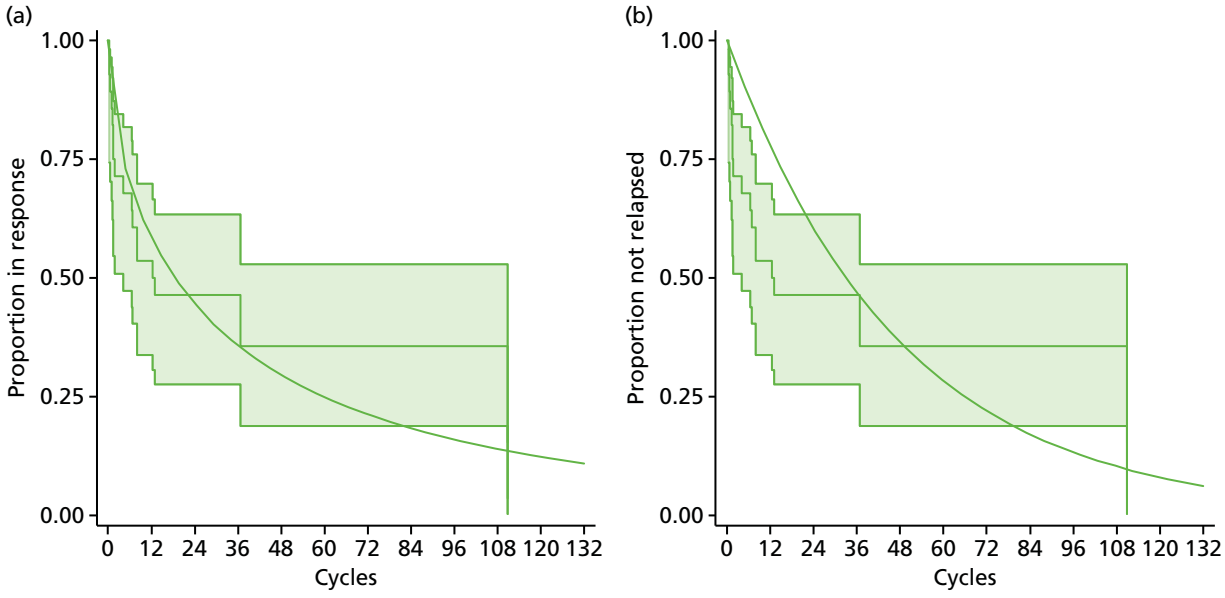


FIGURE 64 Reconstructed Kaplan–Meier plot and Weibull fit for time to LOR after reintroduction of IFX after surgery by 4-week cycle (based on data from Baert *et al.*[77]). (a) Weibull; and (b) exponential.

Intervention arm: loss of response to test-directed infliximab maintenance

Only two management studies of IFX responders were found and one of these, Vaughn *et al.*,[128] 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 (*Figure 65*). The TAXIT[73] randomised controlled management study of responders to IFX maintenance did not report time to LOR. 'Durable remission' among the TAXIT trial 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 LOR 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 patients with CD in the TAXIT trial 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 per cycle (*Figure 66*). This represents a very severe test for LOR because patients without clinical remission are likely to be retained in anti-TNF-α treatment because of a partial response. The retrospective management study of Vaughn *et al.*[128] reported vastly superior performance for 39 IBD patients receiving a test algorithm strategy relative to 68 patients given a 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 per cycle is generated (see *Figure 66*). These transition probabilities are very different and it is doubtful that either generates an appropriate input for the economic model.

In the TAXIT study before dose optimisation, 131 out of 178 (73.59%) patients with CD were in clinical remission; after dose optimisation with a test-directed dose adjustments, 138 out of 173 (79.77%) were in remission (five patients with CD could not be optimised to target trough level). According to ITT analysis, this represents a 3.9% improvement. With continued test-directed dosing post randomisation, 62.6% of patients with CD 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

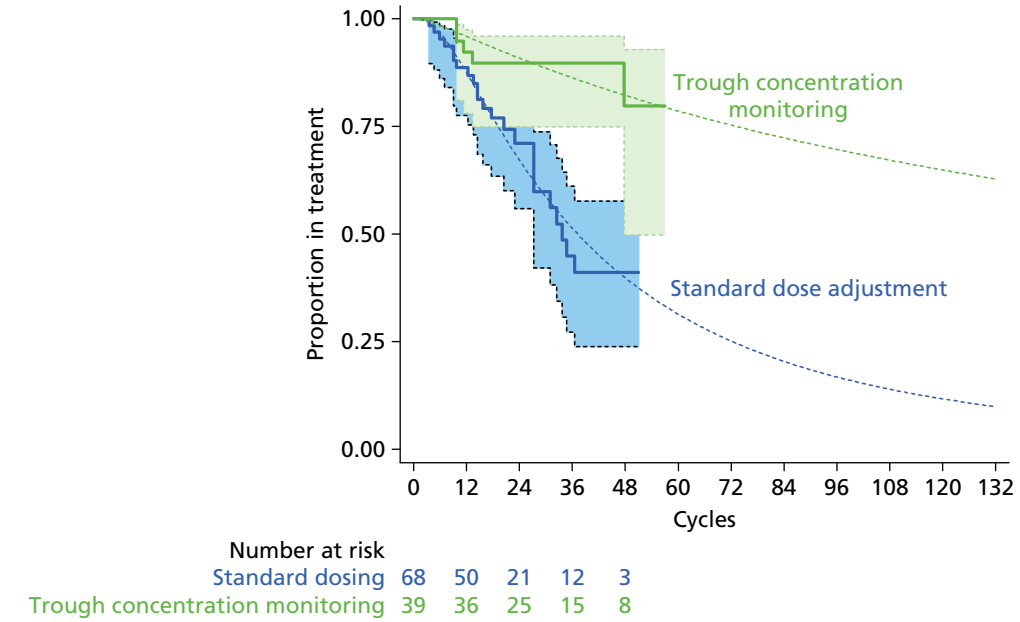


FIGURE 65 Log-normal models for retention in IFX maintenance therapy for IFX responders (based on Vaughn *et al.*[128] and used in sensitivity analysis).

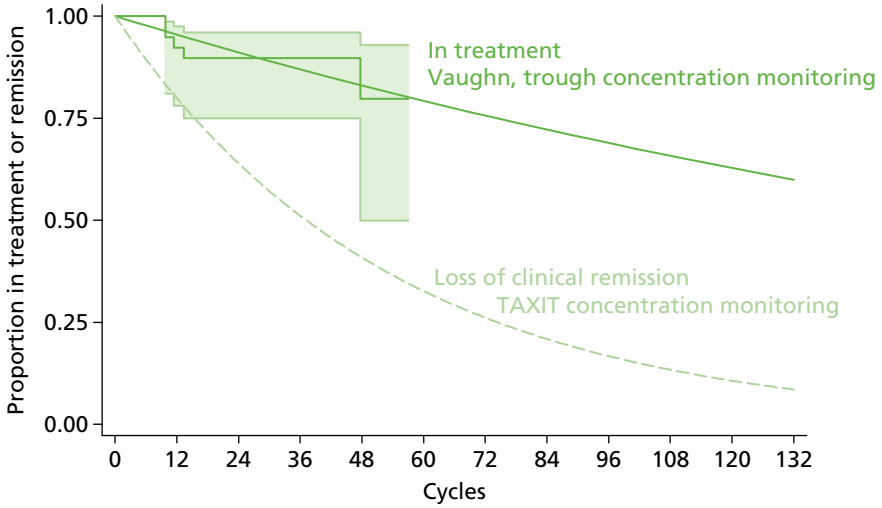


FIGURE 66 Time to event for responders receiving a test algorithm strategy. Time to clinical remission in the TAXIT trial and retention in treatment in Vaughn *et al.*[128] by 4-week cycle.

other evidence demonstrating an advantage for a test algorithm-based strategy the model uses the same probability for LOR to IFX as used for the standard care arm (fit to Juillerat *et al.*[159] data).

Intervention arm: regain of response with test-directed treatments following response loss to maintenance infliximab

The treatments for patients with LOR to maintenance IFX were informed by the management study of Steenholdt *et al.*[123] [*Chapter 3*, *Objective C2: studies relating test results to clinical state of patients (correlation studies)*]. Patients enrolled in this study had failed IFX maintenance in which patients received 'regular infusions of 5 mg/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,[73] in which dose was variously adjusted to bring trough IFX to a target range. In Steenholdt *et al.*[123] patients received concurrent testing at the time of IFX 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) IFX negative and antibodies positive ($n = 5$; 15.15%); (2) IFX negative and antibodies negative ($n = 1$; 3.03%); (3) IFX positive and antibodies negative ($n = 26$; 78.79%); and (4) IFX positive and antibodies positive ($n = 1$; 3.03%). The study reported the proportion who regained a response by 12 weeks, but time-to-event data were not reported. We have assumed that those who had not regained response by week 12 have lost response at a rapid rate over three cycles and remained in the non-response state (until surgery was implemented), and those who 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 3 was small, and so outcomes are associated with great uncertainty. We have assumed that the single group 4 individual (positive test results for both IFX and antibodies to IFX) had the test results confirmed and was subsumed according to the treatment algorithm into group 3, which then accounts for 27 out of 33 (81.8%) of intervention patients. Unfortunately, the various treatments used for the group 3 patients were insufficiently prescribed to be usable (e.g. surgery 'should be considered').

For intervention group 1 patients (15.15% of IFX failures), the algorithm-prescribed treatment was a switch to maintenance therapy with ADA: at 12 weeks 2 out of 5 had regained response. This is a poor response rate, but is based on only five patients and is uncertain. We have therefore used the same transition probabilities for these patients as for ADA-treated patients in the standard care arm (based on the GAIN RCT and on the study by Karmiris *et al.*[48] described in *Figure 62*).

The single group 2 patient (3.03% of intervention patients) received IFX intensification and failed to regain response by week 12. However, all control arm patients in the trial also received IFX intensification and at 12 weeks 19 out of 36 had regained response; which, when combined with the single group 2 patient, provides an estimate of 19 out of 37 (51.3%) in response at week 12. We assume the last patients move to non-response at constant hazard over the first three cycles (12 weeks) providing a transition probability of 0.19948 per cycle. However, using this transition probability would considerably disadvantage the model intervention arm relative to the control arm, and is based on a single time point estimate for a small group of patients. Therefore, the rate of loss of regained response for group 2 was assumed to be the same as that for dose-escalated IFX patients described by Ma *et al.*,[161] which is based on 6 years of time-to-event data (see *Figure 61* for the model based on data from Ma *et al.*[161]).

In groups 3 and 4 (81.81% of IFX failures), 16 out of 27 had regained a response at 12 weeks and 11 out of 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 per cycle. As the treatment for group 3 patients was not prescribed, other than that it lacked anti-TNF-α, we have assumed that after cycle 4 (12 weeks) LOR occurs at constant hazard based on the Rutgeerts *et al.* RCT[165] placebo arm (background therapies including purine analogues, steroids, methotrexate and 5-ASA), in which about half of patients had previously received previous anti-TNF-α therapy. This suggested a transition probability of 0.08617343 per cycle.

In the Steenholdt *et al.* study[123] about half of group 3 patients likely received IFX in contradiction to the specified treatment according to the algorithm.

Of these, 12 patients continued IFX (nine patients from group 3 and one patient from group 4). The applied IFX regimen was (all received 5 mg/kg):

- IFX q8 regimen (two infusions during the trial, i.e. weeks 0 and 8) – $n = 5$
- IFX q4 regimen (four infusions during the trial, i.e. weeks 0, 4, 8 and 12) – $n = 2$
- IFX q4 regimen but not throughout the entire trial (three infusions during the trial) – $n = 1$
- IFX q4 regimen but not throughout the entire trial (two infusions during the trial) – $n = 2$
- IFX q4 regimen but not throughout the entire trial (one infusions during the trial) – $n = 2$.

The remaining two patients had been switched to ADL because of misinterpretation of test results (see *Figure 2*). Both patients were in group 3.

The applied ADL regimen was:

- ADL induction (160 mg–80 mg–40 mg) and followed by 40 mg every other week.

This indicates the various treatments received by the 14 patients in the intervention arm who did not receive algorithm-directed treatments. In view of these difficulties it is difficult to discern how treatment received relates to response observed at cycle 3.

The patients with LOR from all groups remain on palliative care in a LOR state until surgery. It is possible that some of these patients (and also those patients with LOR after ADA in the standard care arm), at some time may be reintroduced to IFX (or possibly ADA) 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.

Appendix 18 Resource-use data

In this appendix, we report on the unit costs derived for monitoring IFX and antibodies to IFX, treatment costs for patients receiving IFX maintenance therapy and cost of a surgical procedure.

TABLE 61 Unit costs for monitoring IFX and antibodies to IFX

Resource use	Quantity	Description	Unit costs (£, 2014)	Source
LISA-TRACKER assays for monitoring IFX and antibodies to IFX (concurrent testing)				
Assay kit used to monitor IFX and antibodies to IFX (concurrent testing)	1	Total cost of kits for monitoring IFX and antibodies to IFX is £1568. Number of patient samples per kit is 42	37.33	Sarah Bond (NICE, 2014, personal communication)
Laboratory technician	1	Assay takes 3 hours to perform in the laboratory. Based on a clinical support worker as a proxy (£21 per hour)	1.50	PSSRU[170]
LISA-TRACKER assays for monitoring IFX and antibodies to IFX (reflex testing)				
Assay kit used to monitor IFX	1	Total cost of kit for monitoring IFX is £850. Number of patient samples per kit is 42	20.24	Sarah Bond (NICE, 2014, personal communication)
Assay kit used to monitor antibodies to IFX	1	Total cost of kit for monitoring antibodies to IFX is £850. Number of patient samples per kit is 42	20.24	Sarah Bond (NICE, 2014, personal communication)
Laboratory technician	1	Assay takes 3 hours to perform in the laboratory. Based on a clinical support worker as a proxy (£21 per hour)	1.50	PSSRU[170]
Estimated total cost for monitoring IFX and antibodies, per person (concurrent testing)			38.83	
Estimated total cost for monitoring IFX and antibodies to IFX, per person (reflex testing)			43.48	
Estimated total cost for monitoring IFX, per person			21.74	

PSSRU, Personal Social Services Research Unit.

TABLE 62 Treatment of CD with IFX and ADA

Resource use	Quantity	Description	Unit costs (£, 2014)	Source
IFX treatment				
IFX (Remicade)[a]	1	5 mg/kg intravenous infusion over a 2-hour period every 8 weeks 100 mg/vial = £419.62. Four vials required 4 × £419.62 = £1678.48	1678.48	BNF 2013/14[166]
Administration cost	1		287.93	PSSRU 2014[170]
Estimated cost per individual receiving IFX maintenance therapy every 8 weeks			1966.41	
ADA treatment				
ADA (Humira)	1	40 mg every 2 weeks	352.14	BNF 2013/14[166]

PSSRU, Personal Social Services Research Unit.
a Patients on maintenance therapy receiving IFX treatment are given a 5 mg/kg intravenous infusion over a 2-hour period every 8 weeks. We assumed that patients are, on average, weighing 70kg.

TABLE 63 Cost of a surgical procedure

Resource use	Quantity	Description	Unit costs (£, 2014)	Source
Investigations				
Laparoscopic ileocolic resection	1	FZ74F elective inpatients – complex large intestine procedures, aged ≥ 19 years, with a CC score of 0–2	6803	*NHS Reference Costs 2013 to 2014*[167]
Outpatient visits (follow-up consultation)	1	WF01A colorectal surgery – consultant-led outpatient attendance non-admitted	105	*NHS Reference Costs 2013 to 2014*[167]
Cost of laparoscopic ileocolic resection			6908	

CC, complications or comorbidities.

TABLE 64 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 to 2014*[167] 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	*NHS Reference Costs 2013 to 2014*[167] and expert opinion
MRI	1	Outpatient RA01A MRI	145	*NHS Reference Costs 2013 to 2014*[167] and expert opinion
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 to 2014*[167] 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	145	
Cost for the regain response health state			725.69	
LOR				
Outpatient visits	2	WF01A colorectal surgery – consultant-led outpatient attendance non-admitted	105	*NHS Reference Costs 2013 to 2014*[167] and expert opinion
Colonoscopy	2	Weighted average of NHS reference cost outpatient for FZ51Z diagnostic colonoscopy without biopsy or FZ52Z diagnostic colonoscopy with biopsy	370.69	
MRI	2	Outpatient RA01A MRI	145	
Cost for the LOR health state			1241.38	

TABLE 64 Additional costs associated with occupying health states (*continued*)

Health state	Quantity	Description	Unit costs (£, 2014)	Source
Post surgery (remission)				
Outpatient visits	4	WF01A colorectal surgery – consultant-led outpatient attendance non-admitted	105	*NHS Reference Costs 2013 to 2014*[167] 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	

MRI, magnetic resonance imaging.

EME
HS&DR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

Published by the NIHR Journals Library