Prognostic tools for identification of high risk in people with Crohn's disease: systematic review and cost-effectiveness study

Steven J Edwards,* Samantha Barton, Mariana Bacelar, Charlotta Karner, Peter Cain, Victoria Wakefield and Gemma Marceniuk

BMJ Technology Assessment Group, BMJ, London, UK

*Corresponding author sedwards@bmj.com

Declared competing interests of authors: none

Published March 2021
DOI: 10.3310/hta25230

Scientific summary

Prognostic tools for risk in people with Crohn's disease
Health Technology Assessment 2021; Vol. 25: No. 23
DOI: 10.3310/hta25230

NIHR Journals Library www.journalslibrary.nihr.ac.uk

Scientific summary

Background

Crohn's disease is characterised by inflammation of the gastrointestinal tract. Crohn's disease is a lifelong condition for which there is no cure. The course of Crohn's disease is characterised by recurring cycles of exacerbation (also referred to as flare) and remission, with the frequency of flare and duration of remission being highly variable. Some people are at a higher risk of following a more aggressive course of disease, which is typified by more frequent relapses and manifestation of penetrating or stricturing complications. The National Institute for Health and Care Excellence recommends that those with active disease receive treatment with a step-up approach, which involves initial treatment with a glucocorticosteroid and stepwise progression through a pathway of immunomodulator and, finally, biological therapy with or without immunomodulator, as determined by the response at each treatment step. However, research suggests that earlier aggressive treatment with the potent combination of biological therapy and immunomodulator could improve the clinical outcomes of those at high risk of developing severe Crohn's disease. No test is available in the NHS to stratify people with Crohn's disease by risk of following a severe course of the condition. Identifying those at a higher risk of developing complications of Crohn's disease could lead to the personalised management of an individual's condition.

Objectives

The aim of the diagnostic assessment review reported here was to assess the prognostic test accuracy, clinical impact and cost-effectiveness of two prognostic tools for inflammatory bowel disease in identifying those at high risk of following a severe course of Crohn's disease. To achieve the goal of the project:

- Systematic reviews of the literature were carried out to identify the evidence on the prognostic accuracy and clinical impact of IBDX® (Crohn's disease Prognosis Test; Glycominds Ltd, Lod, Israel) and PredictSURE-IBD™ (PredictImmune Ltd, Cambridge, UK) in stratifying those with Crohn's disease by risk of following a severe course of disease.
- An economic model was developed to assess the cost-effectiveness of using the IBDX and PredictSURE-IBD tools.

Methods

Assessment of prognostic accuracy and clinical impact

Electronic databases (MEDLINE, EMBASE, the Cochrane Database of Systematic Reviews and the Cochrane Central Register of Controlled Trials) were searched from inception to June 2019. These searches were carried out on 14 June 2019. Eligible studies assessed the prognostic accuracy or clinical impact of the IBDX (panel of six biomarkers) and PredictSURE-IBD tools in stratifying people at a higher risk of following a severe course of Crohn's disease. Two reviewers independently screened potentially relevant studies for inclusion against prespecified criteria, and assessed the quality of studies reporting prognostic accuracy using the Quality In Prognosis Studies in Systematic Reviews tool. One reviewer extracted data from the included studies, with a second reviewer validating the data.

Assessment of cost-effectiveness

The External Assessment Group developed a de novo economic model consisting of a decision tree to allocate patients to a response category after initial induction therapy in either the top-down or the step-up treatment arm. The decision tree was followed by a cohort model, in which patients' level of response to maintenance therapy was assessed.

Patients enter the decision tree model after being allocated to the test (with either PredictSURE-IBD, in the base case, or IBDX, in a scenario analysis) or no test (standard care) arm. In the test arm, patients are categorised as being at high risk or low risk of following a severe course of disease according to test results, whereas those in the no test arm are designated as being at high or low risk based on clinical judgement alone. Given that patients in the standard care arm of the model can receive the step-up treatment approach only and that the top-down treatment approach is assumed to be received by high-risk patients only, the economic model ultimately assesses the cost-effectiveness of top-down therapy compared with step-up therapy in high-risk patients.

After induction therapy, patients are classified as responders (an improvement in Crohn's Disease Activity Index score of > 70) or non-responders (a deterioration, no change, or an improvement in Crohn's Disease Activity Index score of < 70). The duration of induction therapy differs by class of treatment (i.e. immunomodulator, anti-tumour necrosis factor and second-line biologic). If patients respond to induction therapy, they move to the maintenance cohort model, whereas non-responders escalate to the next step of their allocated treatment strategy.

Responders to their first induction therapy enter the maintenance cohort model in remission (Crohn's Disease Activity Index < 150), mild (Crohn's Disease Activity Index 150–220), or moderate to severe (Crohn's Disease Activity Index 221–600) health states. Patients can move between these states during maintenance therapy, reflecting the different levels of response to treatment. The probability of patients transitioning between these states is also dependent on the treatment class received. Patients in the mild and moderate to severe states are at risk of escalating to the next treatment step.

The External Assessment Group estimated surgical events as a standalone outcome in the model. This means that patients do not explicitly leave their health state in a specific cycle to move to the surgery state. Instead, in every model cycle, a proportion of surgeries is estimated and the associated costs and impact on the patients' quality of life is calculated.

The economic assessment was undertaken from the perspective of the NHS and Personal Social Services, and both costs and benefits were discounted at 3.5% per annum. The cycle length in the model was 2 weeks, and the time horizon of the model was 65 years.

Results

Searches of electronic database searches retrieved 6258 unique records. The initial screening of titles and abstracts led to the identification of 36 publications for review of full texts. Of the 36 articles evaluated, 16 publications, including systematic reviews, were deemed relevant to the review of prognostic accuracy. Additionally, documents supplied by the companies marketing the prognostic tools were reviewed. Included studies were assessed for risk of bias and applicability using the QUIPS (QUality In Prognosis Studies) tool. Most studies reporting results for the IBDX tool were determined to be at a moderate risk of bias for the population domain, as the studies included those with a recent diagnosis and those with an established diagnosis of Crohn's disease, and, in some studies, those with severe disease at baseline. Data were not analysed separately for the individual subgroups. Most studies were considered to be at a low risk of bias for attrition and for measurement of prognostic factors because all samples taken were analysed with the relevant tool and the results were generated as per each company's individual protocols. Additionally, many studies were deemed to be at low risk of bias for outcome assessment as the clinicians in the studies were masked to the results of the biomarker assessment.

Prognostic test accuracy

Twelve publications, describing eight studies, were included in the assessment of the prognostic accuracy of the tests. Seven of the studies reported results on the utility of the IBDX kit, and one study provided data on PredictSURE-IBD for stratifying those at high risk of following a severe course of Crohn's disease. Limited evidence is available from the included full-text publications on the prognostic accuracy of PredictSURE-IBD, and no evidence is available on the prognostic accuracy of IBDX, as determined by measures such as sensitivity and specificity. Most of the evidence on the utility of the two tools is derived from observational studies that report estimates of the risk of experiencing a clinical outcome associated with an aggressive course of Crohn's disease, for example the need for treatment escalation, the development of a complication or surgery. No retrieved study reported the clinical impact of using IBDX or PredictSURE-IBD in terms of influencing the treatments given in the management of active Crohn's disease.

IBDX

Two studies reported an effect estimate for the risk of experiencing a complication and need for surgery by number of biomarkers testing positive. Both studies prospectively followed a cohort of people with an established diagnosis of Crohn's disease. The two studies reported an increased risk of experiencing a complication or requiring surgery in those with positive status for at least two or three biomarkers out of the six constituting the IBDX panel. A third study identified a trend towards a larger proportion of people requiring surgery with increasing number of biomarkers testing positive, with a statistically significant difference across the categories assessed ($p < 0.0001$).

PredictSURE-IBD

One observational study reported a sensitivity and specificity for predicting the need for multiple escalations within the first 18 months of 72.7% and 73.2%, respectively, where a cut-off point of two or more treatment escalations was applied to categorise people as having followed a more aggressive course of Crohn's disease. A negative predictive value of 90.9% was reported for PredictSURE-IBD for predicting multiple escalations within the first 18 months. The study additionally reported that those categorised as at high risk of following a severe course of Crohn's disease had a statistically significantly higher risk of first treatment escalation than those designated as at low risk, with a hazard ratio of 2.65 (95% confidence interval 1.32 to 5.34; $p = 0.006$).

Cost-effectiveness

As no robust evidence was identified on the prognostic accuracy of the biomarker stratification tools, the development of an economic model to accurately assess the cost-effectiveness of the diagnostic tools was not possible. Instead, the economic model that was developed sets a structural framework for analysing future available data on prognostic accuracy, and assesses the costs and consequences of treating high- and low-risk patients with both top-down and step-up strategies.

The clinical input parameters in the base-case economic model for PredictSURE-IBD and the scenario analysis for IBDX are the same. The only difference in the cost-effectiveness analyses of the two diagnostic tests is the cost of the tests.

The External Assessment Group found two main sources of evidence that could be used to model the time to treatment escalation and time to surgery. Nevertheless, each source could only partially inform the time to treatment escalation and time to surgery analyses in the model. Therefore, clinical data informing the analysis had to be derived from multiple sources. This approach is not ideal and creates a patchwork network of evidence, introducing uncertainty to the economic results. The External Assessment Group anticipates that this problem will be (at least partially) overcome when the results of the PROFILE (PRedicting Outcomes For Crohn's dIsease using a moLecular biomarkEr) trial are available to populate the economic model [Parkes M, Noor NM, Pombal DR, Hou M, Lewis N, *et al.* PRedicting Outcomes For Crohn's dIsease using a moLecular biomarkEr (PROFILE): protocol for a multicentre, randomised, biomarker-stratified trial. *BMJ Open* 2018;**8**:e026767].

The incremental analysis of cost-effectiveness demonstrates that the top-down strategy (via the use of PredictSURE-IBD in the model) is dominated by the step-up strategy (via the standard care arm of the model).

Conclusions

Despite extensive systematic searches of the literature, no robust evidence was identified on the prognostic accuracy of the biomarker stratification tools IBDX and PredictSURE-IBD. In terms of sensitivity and specificity for the estimate of prognostic accuracy, the External Assessment Group is unaware of a validated definition for determining whether or not a person has followed a severe course of Crohn's disease, for example a set number of treatment escalations or the development of a complication or a need for surgery. Thus, the External Assessment Group considers the criterion for a true positive or false positive result using IBDX and PredictSURE-IBD to be unclear. The External Assessment Group considers that it would be challenging to ascertain an accurate estimate of prognostic accuracy of the tools in stratifying course of Crohn's disease and that to do so would require carrying out a prospective study that included a group or groups that received only step-up treatment after the determination of their risk of a severe course of Crohn's disease. The ongoing PROFILE randomised controlled trial randomises people to accelerated step-up or top-down treatment after they are determined to be at high or low risk of following a severe course of Crohn's disease, and so this trial will provide additional data to inform estimates of prognostic accuracy (Parkes *et al.* 2018).

One of the key underlying assumptions in the External Assessment Group's base-case economic analysis is that high-risk patients who initiate treatment with immunomodulators escalate treatment quicker than high-risk patients who initiate treatment with anti-tumour necrosis factor (supported by the data presented in D'Haens G, Baert F, van Assche G, Caenepeel P, Vergauwe P, Tuynman H, *et al.* Early combined immunosuppression or conventional management in patients with newly diagnosed Crohn's disease: an open randomised trial. *Lancet* 2008;**371**:660–7). However, once these patients initiate subsequent treatment with an anti-tumour necrosis factor (their second treatment step), they 'catch up' with patients on the top-down treatment strategy. As some high-risk patients who receive step-up treatment respond to immunomodulator treatment, the additional immunomodulator step in the step-up strategy is advantageous to patients in the External Assessment Group's base-case analysis as the patients still subsequently receive treatment with biologics, which are assumed to have the same effect as biologics in the top-down arm. Given the paucity of data to substantiate any further benefits of subsequent treatment steps on top-down approaches compared with step-up approaches, the External Assessment Group considered this to be the most conservative modelling approach.

The External Assessment Group's analysis has shown that too high a level of uncertainty remains around the potential benefits of top-down treatment for high-risk patients. The cost-effectiveness of a top-down strategy compared with a step-up strategy in high-risk patients is highly dependent on two unanswered questions: (1) do some high-risk patients derive a benefit from receiving immunomodulator treatment before moving on to biologic treatment?; and (2) do step-up high-risk patients have the same benefits as top-down high-risk patients once they start the top-down treatment pathway (i.e. treatment with anti-tumour necrosis factor)? In the External Assessment Group's model, the potential disadvantage of waiting to start treatment with anti-tumour necrosis factor was based on the increased risk of surgery in the step-up arm only; however, the negative impact of surgery in the analysis was not enough to offset the advantages of initial treatment with immunomodulator for step-up patients.

The External Assessment Group conducted a range of analyses to test extreme scenarios around increasing the relative treatment effectiveness of the top-down approach while decreasing the relative costs associated with top down. The incremental cost-effectiveness ratios for PredictSURE-IBD (and top down) compared with standard care (and step up) fell below £30,000 in the analysis. However, the External Assessment Group notes that these results need to be interpreted with extreme caution,

as the assumptions made in these scenarios were designed to test extreme clinical scenarios and were not evidence based. The External Assessment Group concludes that its base-case analysis showing that top down is dominated by step up remains the most conservative assessment of the relative cost-effectiveness of these treatment strategies.

Study registration

This study is registered as PROSPERO CRD42019138737.

Funding

This project was funded by the National Institute for Health Research (NIHR) Evidence Synthesis programme and will be published in full in *Health Technology Assessment*; Vol. 25, No. 23. See the NIHR Journals Library website for further project information.

Health Technology Assessment

NICE TAR and DAR

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 3.370

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, the Cochrane Library and Clarivate Analytics 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: journals.library@nihr.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

Health Technology Assessment (HTA) research is undertaken where some evidence already exists to show that a technology can be effective and this needs to be compared to the current standard intervention to see which works best. Research can evaluate any intervention used in the treatment, prevention or diagnosis of disease, provided the study outcomes lead to findings that have the potential to be of direct benefit to NHS patients. Technologies in this context mean any method used to promote health; prevent and treat disease; and improve rehabilitation or long-term care. They are not confined to new drugs and include any intervention used in the treatment, prevention or diagnosis of disease.

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.

This report

The research reported in this issue of the journal was commissioned and funded by the Evidence Synthesis Programme on behalf of NICE as project number NIHR128968. The protocol was agreed in June 2019. The assessment report began editorial review in January 2019 and was accepted for publication in July 2020. 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.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health and Social Care.

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