PredictSURE-IBD and IBDX to guide personalised treatment of Crohn's disease in adults

Addendum to the Diagnostic Assessment Report

This report was commissioned by the NIHR Systematic Reviews Programme as project number 128968/T

BMJ Technology Assessment Group

Title: PredictSURE-IBD and IBDX to guide personalised treatment of Crohn's disease in adults – addendum to the Diagnostic Assessment Report

Produced by: BMJ Technology Assessment Group (BMJ-TAG)

Authors: Steven J Edwards, Director of Health Technology Assessment, BMJ-TAG, London

Samantha Barton, Principal Health Technology Assessment Analyst, BMJ-TAG, London

Mariana Bacelar, Principal Health Economist, BMJ-TAG, London

Peter Cain, Senior Health Economist, BMJ-TAG, London

Correspondence to: Steve Edwards, BMJ-TAG [redacted]

Date completed: 13 March 2020

Source of funding: This report was commissioned by the NIHR Systematic Reviews Programme as project number 128968/T.

Declared competing interests of the authors:

No competing interests were declared which affect the impartiality of this report. BMJ Technology Assessment Group (BMJ-TAG) and the editorial team of The BMJ work

independently to one another. The views and opinions expressed in this report are those of the BMJ-TAG.

Rider on responsibility for addendum

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

This addendum should be referenced as follows: Edwards SJ, Barton S, Bacelar M, Cain P. PredictSURE-IBD and IBDX to guide personalised treatment of Crohn's disease in adults: Addendum to the Diagnostic Assessment Report. BMJ Technology Assessment Group, 2019.

SUMMARY

This addendum provides the results for the additional analyses undertaken by the External Assessment Group (EAG), as a result of the stakeholders' comments on the EAG's Diagnostic Assessment Report (DAR). The EAG also corrected an error in the economic model and provides the updated results. The additional analyses undertaken are explained in Section 1, while results are reported in Section 2.

Additionally, in order to mitigate some of the concerns raised by the stakeholders, the EAG reiterates the rationale underlying the clinical and economic analysis in detail, in Section 1.

The EAG used the updated economic model to reproduce the range of analyses (originally reported in the DAR) to test extreme scenarios around increasing the treatment effectiveness of the top-down (TD) approach, while decreasing the costs associated with TD.

The EAG's deterministic and probabilistic updated results show that there is a small difference in QALYs in favour of the step-up (SU) approach, suggesting that this strategy might be more beneficial than TD. However, the EAG notes that the difference in incremental QALYs is small (0.08 in the deterministic results and 0.03 in the probabilistic results), meaning that the final ICER is mainly driven by the difference in costs for TD (via PredictSURE IBD™) compared with SU (via the SC arm). This is portrayed in the cost-effectiveness plane shown in Figure A, where the dispersion of results is much greater around incremental costs than around incremental QALYs.

Figure A. Cost-effectiveness plane

Abbreviations in figure: WTP, willingness to pay.

The results of the scenario analyses conducted by the EAG have shown that the key drivers of the economic results remain the assumptions made around treatment discontinuation with biologics (and the impact this has on costs); and the assumptions made around the benefit of TD vs SU through the impact on time to treatment escalation and also through the proportion of patients in the SU arm who respond (and thus derive a benefit) to treatment with IMs.

The ICERs for using PredictSURE IBD™ to identify patients at high risk of complications (and so follow a TD rather than SU treatment strategy) are close to or below £30,000 when:

- the difference in discontinuation rates for biologic treatment increase (with higher discontinuation rates in the TD arm compared with the SU arm),
- and the benefit of TD compared with SU is increased (through delaying the time to treatment escalation in the TD arm compared to the base case analysis in combination with assuming that 0% of patients respond to IM treatment).

However, the EAG 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 where TD was assumed to be more effective than SU.

1 ANALYSIS UNDERTAKEN

1.1 Time dependent probability of treatment escalation

PredictImmune raised a concern that in the EAG's model time to treatment escalation (TTE) is dependent on time since starting the model and not time since starting a particular course of treatment. The implication of this assumption is that once patients escalate to their second (or further) treatment, the probability of treatment escalation does not reset to be the same as it was for when patients start their first treatment. In fact, the probability of treatment escalation decreases as time goes by, according to the KM data in Biasci *et al.* and the lognormal curve used to fit the latter. The company argues that TTE should reset every time a patient starts a new line of treatment.

The EAG's approach is based on the assumption that as patients escalate to more aggressive treatments, their probability of escalating to the next treatment (or the escalation hazard) diminishes, compared with the less-aggressive initial treatment received, to which patients end up losing response, and thus need to escalate.

However, the EAG agrees with the company that this is based on a clinical assumption, which the EAG considers to be as valid as the one the company proposes. However, given the company's concern was shared by other stakeholders, the EAG has implemented this assumption in their base case model and provides results in Section 2.

1.2 Error in the economic model

PredictImmune pointed to a potential mistake in the EAG's model, which was generating an "artificial" QALY gain for the step-up (SU) arm compared with the top-down (TD) arm. The EAG thanks the company for identifying this error. The EAG found the formulae mistake and corrected it in the model. This correction did not change the dominance of SU over TD in the EAG's base case. However, it impacted the results of the EAG's scenario analyses. These are reported in Section 2.

1.3 The benefit of step-up compared to top-down

As discussed in the DAR, the EAG considers that the evidence base in support of TD therapy improving clinical outcomes in CD is uncertain. The EAG did not identify any direct evidence on the effectiveness or the cost-effectiveness of a TD treatment sequence vs SU therapy for high-risk patients. The EAG found two main sources of evidence that could be used to model the relative effectiveness of the first step of TD compared with the first step of SU for TTE and time to surgery (TTS) outcomes. However,

no comparative or non-comparative data were found on either the TD or SU complete treatment sequences.

The long-term follow-up study Hoekman *et al.* (cited by the company as a source of evidence to demonstrate the relative benefit of TD vs SU) found no difference between TD and SU in 10-year clinical remission rate: endoscopic remission, hospitalisation, surgery or new fistulas. Furthermore, the study concluded that in the long-term a TD strategy had not proven to alter the natural history of CD. However, time to relapse (a proxy for time to treatment escalation) was found to be statistically significantly different across TD and SU arms in the 2-year analysis of the same data (D'Haens *et al.*). The EAG has therefore, incorporated this difference in treatment effect in its economic analysis.

The Biasci *et al*. data informed TTE and TTS according to high- and low-risk of CD complications (for the SU strategy) in the model; while the D'Haens *et al.* (and its 10-year follow-up study Hoekman *et al.*) informed TTE and TTS according to TD and SU treatments (for a population with mixed risk of disease complications). Combining these data was not ideal and created a patchwork network of evidence, introducing uncertainty in the economic results. However, the EAG did not find any alternative sources which could have mitigated these issues.

While the EAG notes that there is no evidence to support the benefit of TD vs SU, it is also aware of the concerns raised by different stakeholders around the fact that the EAG's economic analysis shows a benefit for SU compared to TD. The EAG reiterates that such benefit comes from the fact that the SU model arm has an additional initial treatment step with IMs (followed by treatment with anti-TNFs), whereas the TD strategy does not include the IM step as it begins with anti-TNF treatment. Even though the D'Haens *et al.* evidence is in support of the relative advantage of anti-TNFs vs IMs, there is still a proportion of patients with moderate to severe CD who will derive a temporary benefit from receiving treatment with IMs. These patients eventually escalate to treatment with anti-TNFs and further biologics, however they can respond to IMs (a less expensive treatment than biologics) for a period of time. This is in accordance with the evidence the EAG found and with clinical expert opinion provided to the EAG.

Therefore, the period of response to IMs in the model, yields a benefit (that of a response to treatment) at a much lower cost than patients who have a response to anti-TNFs. The EAG conducted scenario analyses to test the impact of decreasing patients' response to IMs in the model and provides results in Section 2.

Furthermore, the treatment effect of TD vs SU was also thoroughly varied in exploratory analysis reported in the DAR, which tested extreme scenarios around increasing the treatment effectiveness of

the TD approach while decreasing the costs associated with TD. These results are reported in Section 2 using the updated model.

The EAG considers that the PROFILE RCT, which is in progress, was designed to compare the efficacy of TD and SU therapy for high- and low-risk CD, and thus will provide robust evidence on whether early treatment with biologics yields a benefit compared with SU treatment.

1.4 The use of Biasci et al. *individual patient-level data*

The final 40 patients included in the EAG's analysis consisted of all patients in the Biasci *et al.* dataset who received corticosteroid treatment followed by treatment with IMs. The EAG removed 35 patients from the initial dataset of 88 newly diagnosed patients who never received a subsequent IM after corticosteroids. The EAG censored patients who did not have an escalation event after treatment with IM.

The EAG's clinical experts stated that patients with moderate to severe CD are highly unlikely to respond to treatment with corticosteroids. As such, as a model simplification, the EAG did not include this step in the model as the results would be the same in both strategies given that 100% of patients in the high-risk group (in both the TD and the SU arms) would receive initial induction treatment with corticosteroids and move to the next treatments step (with the treatment effect from D'Haens *et al.* only applied for the IM vs anti-TNF and subsequent treatment steps in the model). The EAG appreciates that this may result in a minor discrepancy in the costs associated with the two pathways, i.e. SU patients may receive a full course of corticosteroids and TD patients are likely to only receive a partial course of corticosteroids. However, given the uncertainty around length of treatment and the low cost of corticosteroids, the EAG considers this assumption would have a minimal impact on the results

Concerns were raised by the stakeholders that patients on the SU strategy receive a watchful waiting strategy with corticosteroids. PredictImmune noted that the EAG did not model those patients in the Biasci *et al.* dataset who never escalated from treatment with corticosteroids as they achieved a response with this treatment under the SU strategy. The EAG notes that the decision to exclude this step from the economic model was based on clinical experts advising the implausibility of moderate to severe CD patients responding to corticosteroids alone and also notes that if such approach had been taken, and if the EAG assumed that a proportion of patients in the SU strategy respond to initial treatment with corticosteroids (thus not needing to escalate to IM), then the benefit associated with the SU arm in the economic analysis would be even greater, as a proportion of patients could be successfully managed with a very inexpensive treatment in the SU arm compared with the TD arm.

The SCMs raised a concern for the potential risk of additional complications associated with the SU strategy given the delay for initiating treatment with biologics. The EAG notes that Hoekman *et al.* concluded that in the long-term (10 year follow up) there was no difference found in complications, such as new fistulas or surgery, across the TD and SU arms. Furthermore, even though not based on comparative evidence, the Biasci *et al.* data reported only very few events that required surgery, and no patients had more than one surgery within their follow up period while receiving a SU strategy.

Therefore, the EAG considers that the general view that early biologics are better than later biologics may apply only to those who do not respond to treatment with IMs. Nonetheless, the EAG varied the rate of response to IM treatment in the model, including a scenario where 0% of moderate to severe CD patients do not respond to treatment with IMs. Although the EAG did not find any evidence to support this scenario, it portrays the extreme case where treatment with IMs does not have any benefit compared with treatment with biologics. Results are described in Section 2.

1.5 Differences in modelling approaches

In response to concerns raised by NICE around the differences in the EAG's and the company's modelling results, the EAG lists here the main differences likely to be driving the benefit estimated for TD in the company's model (relative to SU), and the lack thereof in the EAG's model. The EAG also notes that methodological and structural differences have been listed in the DAR.

1. Differences in treatment sequences modelled: The TD strategy in the company's model has the IM step as the last treatment option after treatment with biologics, hence TD patients have the opportunity to respond to IMs, which is not the case in the TD arm in EAG's model. Therefore, the IM step at the beginning of SU in the company's model "cancels out" (not entirely because of deaths but these are few) with the IM step at the end of TD. The EAG has enquired this thoroughly with its three clinical experts, who have all said that IMs would not be given after biologics. Therefore, the EAG's model does not include this as the last step at the end of TD. Nonetheless, as a response to a request made by NICE, the EAG has undertaken an additional scenario analysis where the IM step was included as the last treatment option in the TD treatment sequence (results are provided in section 2.2.1).

   Furthermore, the company's model has a prednisolone step before IMs in the SU group. As prednisolone is less effective than IMs, the first treatment step in the SU arm of the company's model (prednisolone) is less effective than the first treatment step in SU arm of the EAG's model (IMs). Hence the EAG's additional step in the model generates a bigger benefit than the additional step in the company's model;

2. Model inputs: The company has assumed a constant relative risk of treatment escalation of 0.4 (at least for the first 10 years) for TD vs SU, whereas the EAG's modelling implies that the relative effect diminishes over time (i.e. the relative risk gets closer to 1 and TD becomes as effective as SU as time goes by in the model). The relative risk in the EAG's model starts below 0.4 but rises above that after less than 3 months. After a year the relative risk in the EAG's model is at 0.7 and continues increasing after that. Therefore, the company's effectiveness estimates, based on a very simplified approach are potentially overestimating the effect of TD vs SU;

   Furthermore, the company applies the relative risk of treatment escalation of 0.4 for every step in the sequence which means that the probability of escalating at each cycle for each treatment step in TD vs SU. As discussed in the DAR and in the addendum, there is no available evidence to suggest that TD is more effective than SU as an entire treatment sequence, with the only treatment effect available in literature being the D'Haens *et al.* estimate for a proxy of the anti-TNF vs IM step, which is the only step in the EAG's base case model for which a treatment effect is applied.

2 RESULTS

The incremental cost-effectiveness results of the EAG's original base case analysis, with the correction described in Section 1.2 applied, together with the assumption described in Section 1.1, are presented in Table 1. The results of the scenario analyses using various price discounts for the anti-TNF and second-line biologic treatments (25%, 50% and 75% discounts) based on the corrected model are given in

Table 2.

The results of one-way sensitivity analyses (OWSAs) alongside the upper and lower inputs tested in each analysis are given in Table 3. A tornado plot showing the analyses that had the greatest impact in terms of incremental net monetary benefit at a threshold of £30,000 per QALY, is given in Figure 1. ICERs are also provided alongside each OWSA displayed in the plot.

Table 1. Corrected original base case fully incremental cost effectiveness results (discounted)

Intervention	Total Costs	Total QALYs	Incremental costs	Incremental QALYs	ICER
Standard of Care	£201,925	15.86	–	–	–
IBDX®	£210,106	15.79	£8,181	-0.08	Dominated
PredictSURE IBD™	£211,009	15.79	£903	0	Dominated
Abbreviations: ICER, incremental cost effectiveness ratio; QALY, quality adjusted life year.					

Table 2. Drug price discount scenarios

Intervention	Total Costs	Total QALYs	Incremental costs	Incremental QALYs	ICER
Biologic discount: 25%					
Standard of Care	£185,539	15.86	–	–	–
PredictSURE IBD™	£193,276	15.79	£7,737	-0.08	Dominated
Biologic discount: 50%					
Standard of Care	£169,153	15.86	–	–	–
PredictSURE IBD™	£175,543	15.79	£6,390	-0.08	Dominated
Biologic discount: 75%					
Standard of Care	£152,767	15.86	–	–	–
PredictSURE IBD™	£157,809	15.79	£5,043	-0.08	Dominated
Anti-TNF discount: 25%					

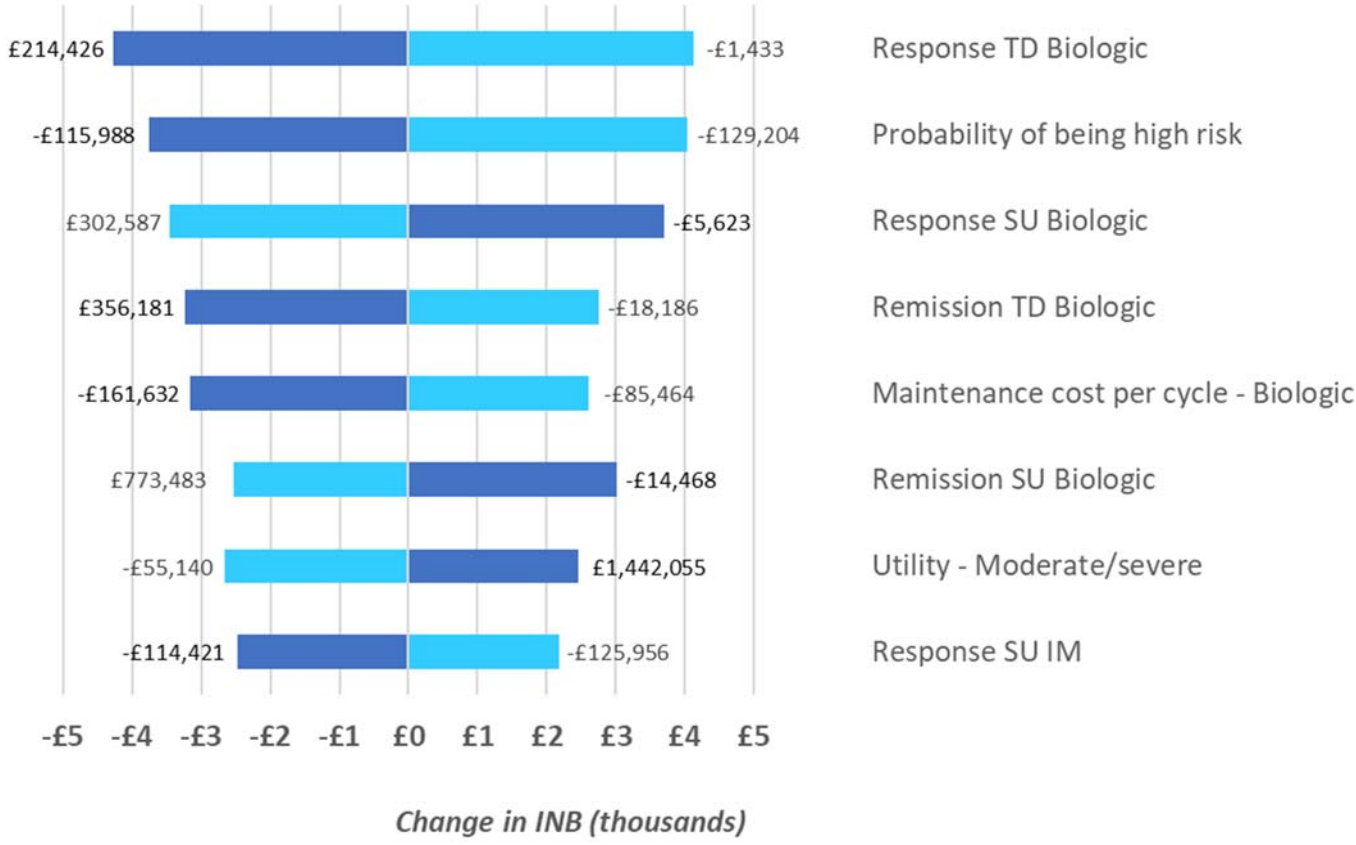
Standard of Care	£193,410	15.86	–	–	–
PredictSURE IBD™	£202,514	15.79	£9,104	-0.08	Dominated
Anti-TNF discount: 50%					
Standard of Care	£184,894	15.86	–	–	–
PredictSURE IBD™	£194,018	15.79	£9,124	-0.08	Dominated
Anti-TNF discount: 75%					
Standard of Care	£176,378	15.86	–	–	–
PredictSURE IBD™	£185,523	15.79	£9,144	-0.08	Dominated
Biologic and Anti-TNF discount: 25%					
Standard of Care	£177,023	15.86	–	–	–
PredictSURE IBD™	£184,780	15.79	£7,757	-0.08	Dominated
Biologic and Anti-TNF discount: 50%					
Standard of Care	£152,122	15.86	–	–	–
PredictSURE IBD™	£158,552	15.79	£6,430	-0.08	Dominated
Biologic and Anti-TNF discount: 75%					
Standard of Care	£127,220	15.86	–	–	–
PredictSURE IBD™	£132,323	15.79	£5,103	-0.08	Dominated
Abbreviations: ICER, incremental cost effectiveness ratio; QALY, quality adjusted life year.					

Table 3. Inputs and results of OWSAs

Model Parameter	Lower bound	Upper bound	Lower ICER	Upper ICER
Age	22.7	50.0	-£113,635	-£136,727
Crohn's disease expected body weight	46.2	102.0	-£120,774	-£118,762
Proportion of males	0.23	0.53	-£119,367	-£120,394
Probability of being high risk	0.34	0.79	-£129,204	-£115,988
Proportion on infliximab in anti-TNF biologics class	0.25	0.56	-£119,684	-£120,059
Proportion on vedolizumab in non-anti-TNF biologics class	0.31	0.69	-£116,088	-£123,642
Proportion on azathioprine for immunomodulators	0.41	0.99	-£120,489	-£119,559
Proportion of 6-mercaptopurine for immunomodulators	0.06	0.14	-£119,906	-£119,841
Proportion of anti-TNF with IM bundle	0.19	0.42	-£119,874	-£119,856
Proportion of Biologics with IM bundle	0.13	0.28	-£119,735	-£120,016
Response TD Biologic	0.20	0.44	-£1,433	£214,426
Remission TD Biologic	0.08	0.19	-£18,186	£356,181
Response TD anti-TNF	0.16	0.36	-£97,736	-£185,310
Remission TD anti-TNF	0.23	0.52	-£89,429	-£266,559
Response SU Biologic	0.20	0.44	£302,587	-£5,623
Remission SU Biologic	0.08	0.19	£773,483	-£14,468
Response SU anti-TNF	0.16	0.36	-£203,848	-£71,062
Remission SU anti-TNF	0.23	0.52	-£280,003	-£58,931

Response SU IM	0.14	0.32	-£125,956	-£114,421
Remission SU IM	0.10	0.22	-£129,351	-£111,917
Probability of death following surgery	0.0010	0.0021	-£118,576	-£121,465
Health state cost - Remission	£11	£24	-£120,223	-£119,432
Health state cost - Mild	£17	£38	-£120,554	-£119,029
Health state cost - Moderate/severe	£79	£174	-£113,183	-£127,978
Health state cost - No response	£79	£174	-£120,757	-£118,783
Induction cost per cycle - Anti TNF	£982	£2,169	-£119,157	-£120,725
Induction cost per cycle - Biologic	£1,000	£2,207	-£117,755	-£122,428
Induction cost per cycle - Immunomodulator	£3	£6	-£119,924	-£119,794
Maintenance cost per cycle - Anti TNF	£346	£765	-£121,205	-£118,239
Maintenance cost per cycle - Biologic	£425	£938	-£85,464	-£161,632
Maintenance cost per cycle - Immunomodulator	£8	£17	-£121,021	-£118,462
IV administration first attendance	£129	£284	-£119,607	-£120,179
IV administration follow-up	£137	£303	-£109,074	-£132,967
Cost of surgery	£5,704	£12,589	-£122,801	-£116,301
Utility - Remission	0.40	1.00	-£2,438,927	-£109,748
Utility - Mild	0.40	0.95	-£1,283,717	-£83,963
Utility - Moderate/severe	0.34	0.78	-£55,140	£1,442,055
Disutility for surgery	0.03	0.06	-£120,664	-£118,917

Figure 1. Tornado plot showing OWSAs that have the greatest impact on incremental net monetary benefit (ICERs given at the top and lower end of bars)



Abbreviations in figure: INB, incremental net benefit; SU, step up; TD, top down.

Table 4 presents the deterministic base-case incremental cost-effectiveness ratio (ICER) for PredictSURE IBD™ compared with SC in the updated model, with the correction described in Section 1.2 and the clinical assumption in Section 1.1. both applied. The results show that the TD strategy (via the use of PredictSURE IBD™ in the model) is dominated by SU (via the SC arm of the model), with an additional cost of £9,084 and a QALY loss of 0.08.

Table 4. Base case deterministic cost effectiveness results (discounted)

Intervention	Total Costs	Total QALYs	Incremental costs	Incremental QALYs	ICER
Standard of Care	£201,925	15.86	–	–	–
PredictSURE IBD™	£211,009	15.79	£9,084	-0.08	Dominated
Abbreviations: ICER, incremental cost effectiveness ratio; QALY, quality adjusted life year.					

The EAG conducted a probabilistic sensitivity analysis (PSA) to assess the impact of the combined uncertainty from all parameters in the model. The methodology undertaken to run the PSA is described in Section 5.1 of the DAR. The probabilistic ICER is reported in Table 5. Figure 2 displays the scatterplot showing the spread of results from the individual samples. The incremental costs and QALYs relative to SC are shown in the cost-effectiveness plane in Figure 3, while the cost-effectiveness acceptability curves (CEACs) showing the probability of PredictSURE IBD™ being cost-effective against SC over a range of willingness to pay thresholds, are given in Figure 4.

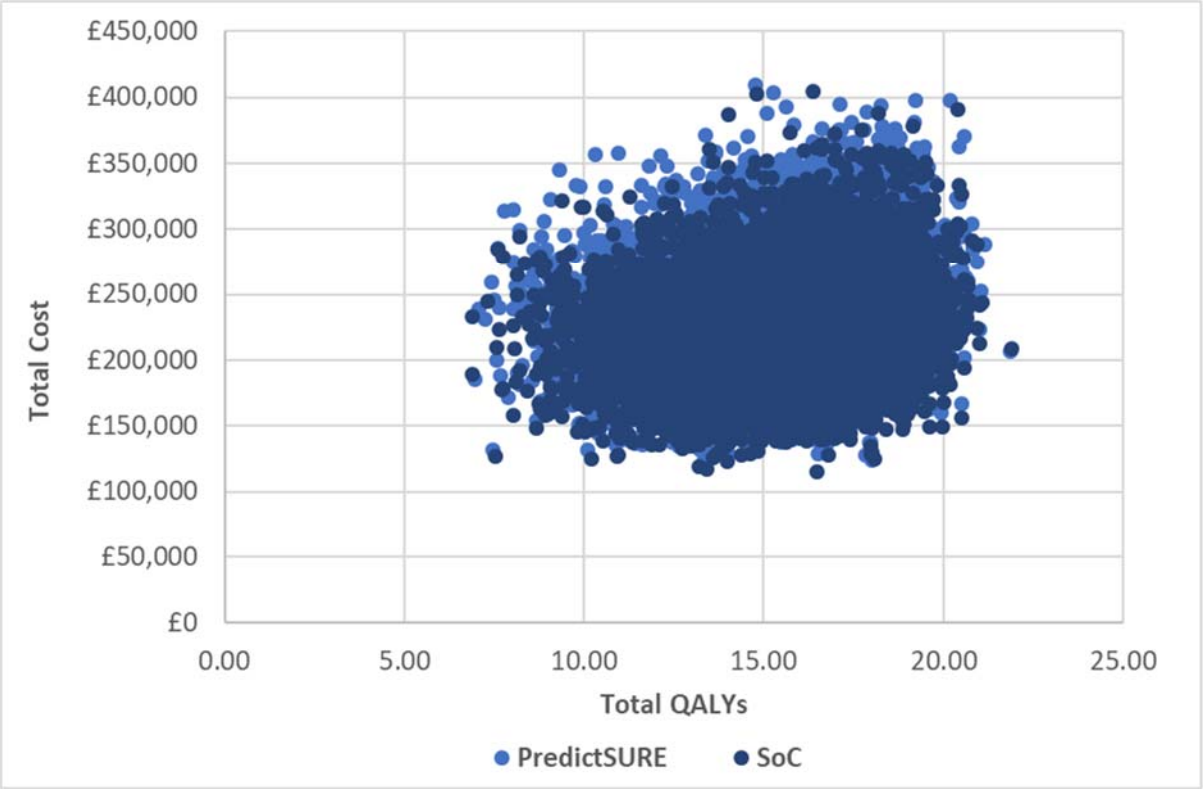
The probabilistic ICER is consistent with the deterministic ICER showing PredictSURE IBD™ is dominated by SC. The CEACs show that the prognostic test has a 0% probability of being cost-effective against SC at the £20,000 – £30,000 ICER threshold used by NICE. However, the EAG notes that the difference in incremental QALYs is small (0.03), meaning that the final ICER is mainly driven by the difference in costs for TD (via PredictSURE IBD™) compared with SU (via the SC arm).

The EAG varied the willingness to pay threshold to assess when the CEACs would begin to converge and at a threshold of £500,000 per QALY gained, the probability of PredictSURE IBD™ being cost-effective was just below 35% against approximately 65% for the SC arm.

Table 5. Base case probabilistic cost effectiveness results (discounted)

Intervention	Total Costs	Total QALYs	Incremental costs	Incremental QALYs	ICER
Standard of Care	£224,904	15.70	–	–	–
PredictSURE IBD™	£237,036	15.67	£12,132	-0.03	Dominated
Abbreviations: ICER, incremental cost effectiveness ratio; QALY, quality adjusted life year.					

Figure 2. Scatterplot of the 10,000 PSA samples of costs and QALYs

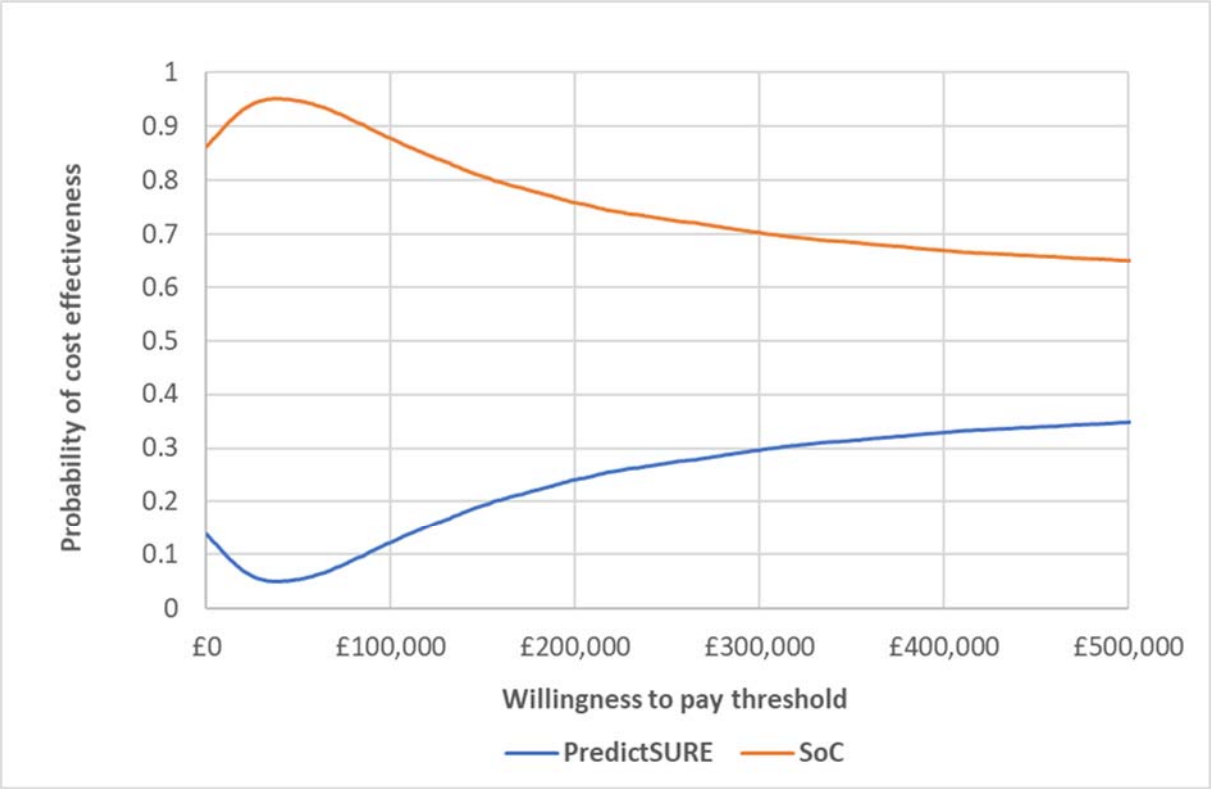


Abbreviations in figure: SoC, standard of care.

Figure 3. Cost-effectiveness plane

Abbreviations in figure: WTP, willingness to pay.

Figure 4. Cost-effectiveness acceptability curve



Abbreviations in figure: SoC, standard of care.

2.1 Scenario analyses

The EAG conducted scenario analyses to assess the potential impact of the uncertainty around some of the assumptions made in the model. Results are reported in Table 6.

1. The EAG ran the economic model using the IBDX® cost (reported in Section 4.2.6 of the DAR). The EAG notes that the clinical input parameters in the base case economic model for PredictSURE IBD™ and in the scenario analysis for IBDX® are the same;

2. The EAG used the utility values in TA456 in a scenario analysis;

3. The EAG applied the treatment effectiveness (i.e. induction vectors and transition probabilities) from TA352 studies;

4. As an exploratory analysis, the EAG assumed that TTS is the same in the TD and the SU arms for high-risk patients;

5. The EAG removed the age and sex utility adjustments from the economic analysis;

6. As a scenario analysis, the EAG used the minimum induction period from the treatment class in the model to estimate induction costs;

7. The EAG assumed that 100% of high-risk patients who receive SU do not respond to treatment and therefore escalate to anti-TNF after induction with IMs.

All of the scenario analyses undertaken produced dominated ICERs against PredictSURE-IBD™ compared to SC. The only exception was scenario 7, where the EAG assumed that 100% of high-risk patients who receive SU therapy do not respond to IMs (therefore not deriving any benefit from response to this treatment). The ICER for PredictSURE-IBD™ compared to SU changed from dominated (against the prognostic tool) to £170,180. To note is that the EAG tested the impact of varying the proportion of patients who do not respond to IM treatment in the analysis. When the EAG assumed that 97% of high-risk patients who receive SU therapy do not respond to IMs (therefore not deriving any benefit from response to this treatment), the two strategies (TD and SU) became clinically equivalent.

Table 6. Results of scenario analyses

Intervention	Total Costs	Total QALYs	Incremental costs	Incremental QALYs	ICER
Scenario 1: Applying IBDX cost					
Standard of Care	£201,925	15.86	–	–	–
IBDX	£210,106	15.79	£8,181	-0.08	Dominated
Scenario 2: Applying utilities from TA456					
Standard of Care	£201,925	15.57	–	–	–
PredictSURE IBD™	£211,009	15.50	£9,084	-0.08	Dominated
Scenario 3: Applying induction vectors and transition probabilities based on TA352 studies					
Standard of Care	£201,695	15.86	–	–	–
PredictSURE IBD™	£210,841	15.78	£9,146	-0.08	Dominated
Scenario 4: Applying equivalent TTS curves for top down and step up					
Standard of Care	£201,925	15.86	–	–	–
PredictSURE IBD™	£211,575	15.78	£9,650	-0.08	Dominated
Scenario 5: Removing Ara & Brazier utility adjustment					
Standard of Care	£201,925	15.92	–	–	–

PredictSURE IBD™	£211,009	15.84	£9,084	-0.08	Dominated
Scenario 6: Use the minimum induction period from the treatment class to estimate induction costs					
Standard of Care	£196,077	15.84	–	–	–
PredictSURE IBD™	£204,704	15.76	£8,627	-0.08	Dominated
Scenario 7: 100% of high-risk patients who receive SU do not respond to IM treatment					
Standard of Care	£209,797	15.78	–	–	–
PredictSURE IBD™	£211,009	15.79	£1,212	0.01	£170,180
Abbreviations: ICER, incremental cost effectiveness ratio; QALY, quality adjusted life year; TTS, time-to-surgery.					

As discussed throughout the DAR and the addendum, the EAG conducted a range of additional analyses to test extreme scenarios around increasing the treatment effectiveness of the TD approach while decreasing the costs associated with TD. These scenarios are described below, together with the respective results.

2.1.1 Accounting for the cost-effectiveness of misdiagnosed cases

The prognostic test accuracy in the base case economic model for PredictSURE IBD™ and in the scenario analysis included in the DAR for IBDX® was the same and assumed to be 100%. This is unlikely to reflect the tests' actual accuracy in clinical practice; however, no robust data were found to inform this in the analysis.

In the absence of real data to inform the costs and consequences of misdiagnosing patients according to their risk of disease severity, the EAG has undertaken a theoretical scenario analysis. The EAG assumed that both prognostic tools are 75% accurate and therefore, 25% of CD cases are assumed to be misdiagnosed in the analysis. The results of this scenario are presented in Table 8 and more details on this scenario analysis can be found in Section 5.2.1 of the DAR.

2.1.2 Varying the assumptions around the measure of relative treatment effectiveness for time to treatment escalation

To aid the interpretation of this scenario analysis, the EAG reproduced the modelled treatment sequences and respective application relative treatment effects in the EAG's model in Figure 5 and in Figure 6 for the TD and the SU strategies, respectively.

The relative treatment effect of TD vs SU was applied only in the IM vs anti-TNF step in the EAG's model and taken from D'Haens *et al.* (in the form of a hazard function applied to TTD and TTS SU data). As some high-risk patients who receive SU treatment respond to IM treatment, having the additional IM step in the SU strategy is advantageous to patients in the EAG's base case analysis as patients in the SU still subsequently receive treatment with biologics, which are assumed to have the same benefit as biologics is the TD arm (see Figure 5 and in Figure 6).

The ERG varied these assumptions in two scenario analyses:

a) High-risk patients on anti-TNF after IM (second step on SU arm) do not do as well as high-risk patients on first-line anti-TNF (first step on TD arm) and thus, the former escalate treatment quicker than the latter. This assumes that anti-TNF treatment is less effective in the SU strategy than in the TD strategy. Given that the EAG did not find any data to support this reduction in relative treatment effect across strategies, a theoretical assumption was made and varied:

    i. Half of the risk of relapse from D'Haens *et al.* for TD (anti-TNF) vs SU (IMs) was assumed for anti-TNFs in the TD approach vs the risk of relapse with anti-TNFs in the SU approach (thus making anti-TNFs more effective in TD than in SU);

    ii. The difference in risk of relapse identified in D'Haens for TD (anti-TNF) vs SU (IMs) was applied to anti-TNFs in TD vs anti-TNFs in SU (thus making anti-TNFs more effective in TD than in SU).

Scenario a also assumes that the benefit in the anti-TNF step of the TD strategy compared to the anti-TNF step in the SU strategy carries through to the next treatment steps. Therefore, patients on second line biologic treatment in the TD strategy receive an increase in benefit comparatively to second line biologic treatment in the SU arm (as do patients on third line biologics). It is also assumed that second and third line biologic treatment is as effective as anti-TNF treatment within the respective TD and SU arms, and thus there is a benefit associated with biologic treatment in the TD arm compared to biologic treatment in the SU arm (see Figure 5 and in Figure 6 and Table 7).

b) Same assumptions as in scenario a with regards to the benefit of anti-TNF in TD and SU, with the exception that once patients have moved on to second and third-line biologics, there is no further benefit for TD vs SU. In the base case treatment with anti-TNF and second and third-line biologics are assumed to be equally effective. However, as an alternative to scenario a, where the increased benefit of TD vs SU carries through all of these treatment steps, scenario b assumes that the increased benefit only applies to treatment with ant-TNF (i.e. second and third-line

biologics are considered equally effective to the same treatments in the SU strategy) (see Figure 5 and in Figure 6 and Table 7).

Results for these scenarios are presented in Table 8.

Figure 5. Top-down treatment strategy

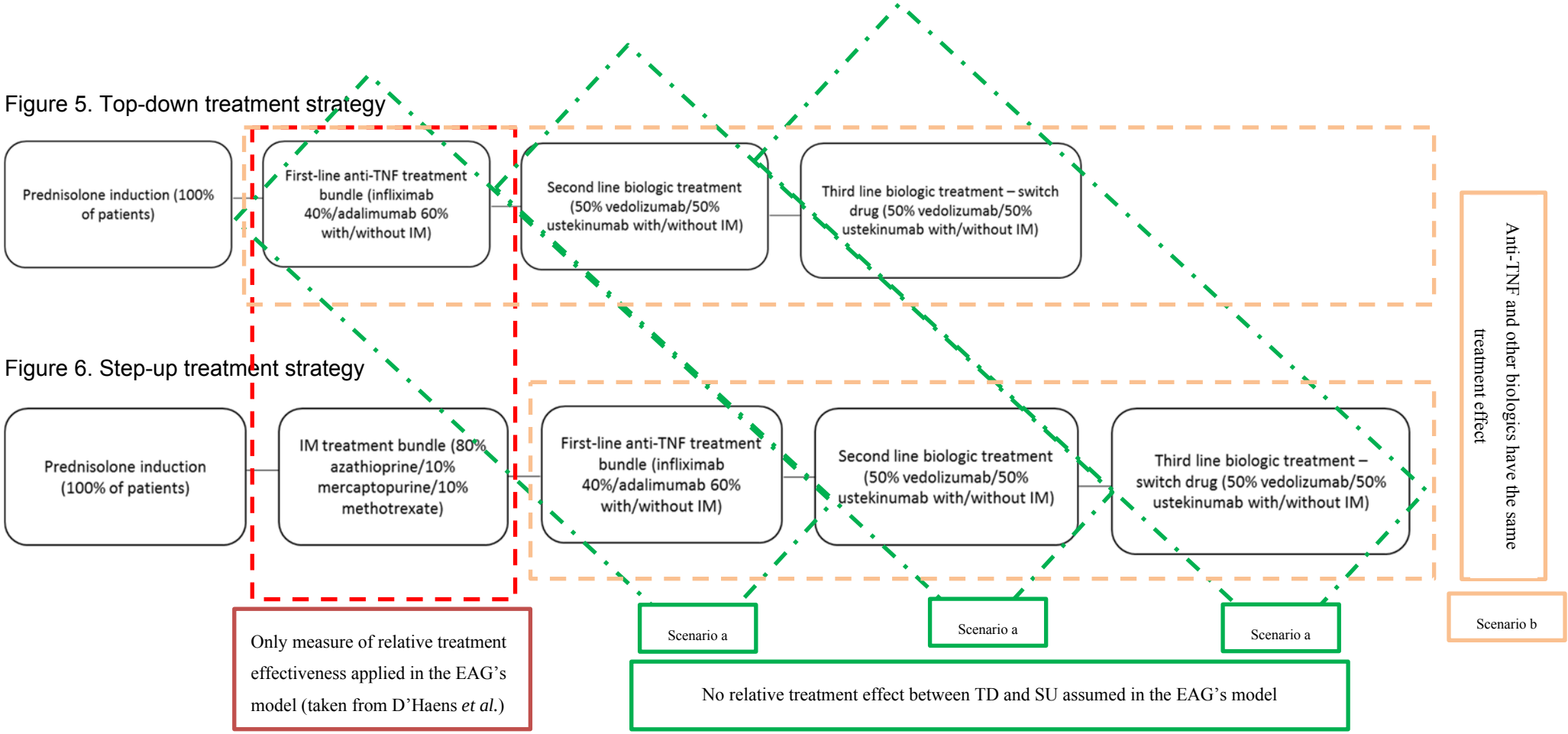


Figure 6. Step-up treatment strategy

Table 7. Summary of exploratory analyses

Steps in the model	Base case	Scenario a	Scenario b
Anti-TNF (TD) vs IM (SU)	Risk of relapse identified in D'Haens *et al*. for TD (anti-TNF) vs SU (IMs)	Same as base case	Same as base case
Anti-TNF (TD) vs anti-TNF (SU)	No relative benefit	i)Half of the risk of relapse from D'Haens *et al.* for TD (anti-TNF) vs SU (IMs) was assumed for anti-TNFs in the TD approach vs the risk of relapse with anti-TNFs in the SU approach; ii) The difference in risk of relapse identified in D'Haens for TD (anti-TNF) vs SU (IMs) was applied to anti-TNFs in TD vs anti-TNFs in SU	Same as scenario a
Second and third line biologic (TD) vs second and third line biologic (SU)	No relative benefit	i)Half of the risk of relapse from D'Haens *et al.* for TD (anti-TNF) vs SU (IMs) was assumed for biologics in the TD approach vs the risk of relapse with biologics in the SU approach; ii) The difference in risk of relapse identified in D'Haens for TD (anti-TNF) vs SU (IMs) was applied to biologics in TD vs biologics in SU	No relative benefit
Second and third line biologic (TD) vs anti-TNF (TD)	No relative benefit	No relative benefit	i)Half of the risk of relapse from D'Haens *et al.* for TD (anti-TNF) vs SU (IMs) was assumed for biologics in the TD approach vs the risk of relapse with anti-TNFs in the TD approach; ii) The difference in risk of relapse identified in D'Haens for TD (anti-TNF) vs SU (IMs) was applied to biologics in TD vs anti-TNFs in TD
Second and third line biologic (SU) vs anti-TNF (SU)	No relative benefit	No relative benefit	No relative benefit

2.1.3 Assumptions around treatment discontinuation in the model

a) The EAG assumed that after 2 years in remission with any biologic treatment, a proportion of patients experience mucosal healing and therefore, stop treatment permanently. The EAG used the Marchetti *et al.* paper to inform this scenario. The study reports that after 2 years in remission, 76% of patients in the TD strategy experience mucosal healing, while 40% of patients in the SU arm experience the same outcome (which the EAG has ran in the model in scenario 2.1.3 a i).

   The EAG also varied the Marchetti *et al.* assumptions and explored the possibility of TD and SU therapies having the same impact on the 2-year probability of mucosal healing. Therefore, the EAG assumed that both TD and SU arms would experience the same probability (either 76% in scenario 2.1.3 a ii or 40% in scenario 2.1.3 a iii) of mucosal healing.

   The EAG notes that Hoekman *et al.* concluded that in their 10-year follow-up study, "*mucosal healing 2 years after the start of treatment was associated with a reduced use of anti-TNF treatment during long-term follow-up. Other outcomes, however, did not differ significantly between patients with and without mucosal healing 2 years after the start of treatment, which is in contrast to a recent meta-analysis of 12 studies with 673 patients that showed that mucosal healing is associated with an increased likelihood of long-term clinical remission*." Furthermore, Hoekman *et al.* also reported that another study has shown that 2–4 years after randomisation, mucosal healing at week 104 after randomisation, but not treatment allocation, was associated with stable, corticosteroid-free remission (Baert *et al*.).

   Therefore, while there is some evidence supporting that 2-year endoscopic mucosal healing is associated with long-term, corticosteroid-free clinical remission, there does not seem to be any evidence supporting that mucosal healing at 2 years differs according to TD or SU treatment. To note is that estimates used in Marchetti *et al.* were taken from another study, which the EAG did not have access to (Baert *et al*.).

b) The company in TA352 assumed that patients discontinued treatment with biologic agents approximately 1 year after maintenance treatment. The ERG in TA352 was concerned that a discontinuation rule may not have been appropriate for patients who are not in remission as the NICE recommendation for infliximab and adalimumab suggests that, "*specialists should discuss the risks and benefits of continued treatment with patients and consider a trial withdrawal from treatment for all patients who are in stable clinical remission. People who continue treatment with infliximab or adalimumab should have their disease reassessed at least*

every 12 months to determine whether ongoing treatment is still clinically appropriate. People whose disease relapses after treatment is stopped should have the option to start treatment again". The EAG notes that duration of treatment with biologics in clinical practice remains uncertain. The clinical experts advising the EAG reported that treatment with anti-TNF and second-line biologics would be given as long as patients continue to show a response.

For completeness, the EAG ran an additional scenario analysis assuming that 100% of patients in continuous remission for 12 months with maintenance treatment of any biologic (i.e. anti-TNF, second- or third-line biologics), discontinue treatment.

Results for these scenarios are presented in Table 8.

2.1.4 Surgery as a final treatment step in the economic model

a) The clinical experts advising the EAG explained that once patients exhaust all the biologic treatments available, they receive surgery. Therefore, the EAG ran a scenario analysis where patients escalating from third-line biologic treatment in the model receive surgery. The EAG assumed that surgery had a temporary "curative" effect of 2 years, where patients experience the costs and utility associated with being in the remission state. After 2 years it was assumed that patients revert to the moderate to severe state, where they remain for the rest of the model;

b) To test the sensitivity of the results of the model to assumptions relating to surgery, the EAG ran a separate scenario analysis excluding surgeries from the model.

Results for these scenarios are presented in Table 8.

2.2 Results of individual scenario analysis

Results of the individual scenario analysis are reported in Table 8. The EAG notes that the majority scenarios ran individually did not change the dominance of SU (via the SC arm of the model) over the TD strategy (via the use of PredictSURE IBD™ in the model). However, the EAG considers these scenarios to have more impact when considered in combination with each other. The next subsection discusses such scenarios, together with results.

Scenario 2.1.1 produced an ICER of £64,876 per QALY gained, with PredictSURE IBD™ being more costly than SC but generating a QALY gain of 0.15. Even though this scenario assumes lower test accuracy, the assumed consequences of misdiagnosis produced a QALY gain for the prognostic tool. This is related to the assumption of allocating low-risk patients (misdiagnosed as high-risk) to the anti-

TNF state in the model, without any further need for further escalation. Given that treatment with anti-TNF holds the highest remission rate in the EAG's analysis, and that 62% of high-risk patients (misdiagnosed as low-risk) in the SU arm were assumed to not derive any benefit from treatment with IMs, the results produced positive incremental QALYs for the prognostic tool (thus, for the TD strategy).

Out of all variations of scenario 2.1.2, scenario a ii assumed the highest benefit for TD vs SU in terms of TTE, as it used the difference in risk of relapse identified in D'Haens for TD (anti-TNF) vs SU (IMs) to be applied to anti-TNFs in TD vs anti-TNFs in SU (thus making anti-TNFs more effective in TD than in SU). This scenario still generated a dominated ICER against TD, with a very small QALY loss of 0.002. Scenario 2.1.2 a i assumed half of the risk of relapse from D'Haens et al. for anti-TNFs in the TD approach vs the risk of relapse with anti-TNFs in the SU approach also resulting in dominated ICER against PredictSURE IBD™ vs SC and a QALY loss of 0.04. Finally, scenario b i and ii also assumed a smaller relative benefit across TD and SU than scenario a ii, as these scenarios varied the assumption of the effectiveness of anti-TNF compared with second- and third-line biologics, rather than the relative effectiveness across treatment arms. These scenarios produced dominated ICERs against PredictSURE IBD™.

Scenario 2.1.3 a i produced an ICER of £46,263 for SC vs PredictSURE IBD™, meaning that the prognostic tool is less expensive than SC (by £3,506) but also less effective (0.08 QALY loss). This scenario reduced the costs of biologic treatment in the TD arm, by assuming that a higher proportion of patients in the TD arm achieve mucosal healing and thus stop treatment. Even though these patients were "kept" in the remission state, the QALYs generated with this assumption were not enough to produce a QALY gain compared with the benefit patients derive from initial treatment with IMs in SU. The EAG notes that scenario 2.1.3 a i can also be interpreted as a proxy for a scenario assuming de-escalation from biologic treatment in the TD arm to IMs. This is because the scenario reduced treatment costs (by stopping treatment with biologics), which would be similar to replacing treatment with biologics with IMs in the model due to the low cost of IM treatment.

The other variations of scenario 2.1.3, where the same proportion of patients were assumed to achieve mucosal healing in the TD and SU arms, produced dominated ICERs against the prognostic tool (and thus TD). The EAG notes that Hoekman *et al.* did not show a difference in mucosal healing for TD vs SU (although it is not clear if the authors investigated the impact that the strategies had on this outcome). Notwithstanding, the authors reported that the rate of mucosal healing reported in another study (Baert *et al.*) had shown that 2–4 years after randomisation treatment allocation was associated with stable, treatment-free remission.

Scenario 2.1.4 a shows that assuming surgery is the last treatment option for patients in the model, through which they achieve remission for 2 years, increased the total QALYs in both treatment arms (as the alternative option for these patients is to be in the moderate to severe health state in the base case) and increased total costs in both arms. The ICER remained dominated against PredictSURE IBD™.

Finally, scenario 2.1.4 b shows that removing surgeries altogether from the model increased the total QALYs in both treatment arms (as surgeries have a negative impact in the model) and decreased total costs in both arms, with the highest decrease relative to the base case costs observed for PredictSURE IBD™, as there were more surgeries in this arm of the model. The ICER remained dominated against PredictSURE IBD™.

Table 8. Results of scenario analyses

Intervention	Total Costs	Total QALYs	Incremental costs	Incremental QALYs	ICER
Scenario 2.1.1 Misdiagnosis					
Standard of Care	£201,925	15.86	–	–	–
PredictSURE IBD™	£211,782	16.01	£9,856	0.15	£64,876
Scenario 2.1.2 a i - Assuming half of the base case risk of relapse (in the first treatment steps) for TD vs SU for second and subsequent treatment steps					
Standard of Care	£197,986	15.78	–	–	–
PredictSURE IBD™	£208,878	15.75	£10,892	-0.04	Dominated
Scenario 2.1.2 a ii - Assuming the same as base case risk of relapse (in the first treatment steps) for TD vs SU for second and subsequent treatment steps					
Standard of Care	£193,282	15.70	–	–	–
PredictSURE IBD™	£205,961	15.70	£12,679	-0.002	Dominated
Scenario 2.1.2 b i - Assuming half of the base case risk of relapse (in the first treatment steps) for TD vs SU for anti-TNF vs biologics in TD					
Standard of Care	£197,986	15.78	–	–	–

PredictSURE IBD™	£207,699	15.73	£9,713	-0.06	Dominated
Scenario 2.1.2 b ii - Assuming the same as base case risk of relapse (in the first treatment steps) for TD vs SU for anti-TNF vs biologics in TD					
Standard of Care	£193,282	15.70	–	–	–
PredictSURE IBD™	£203,599	15.66	£10,317	-0.04	Dominated
Scenario 2.1.3 a i – Assuming discontinuation of biologic treatment for 76% TD; 40% SU.					
Standard of Care	£181,522	15.86	–	–	–
PredictSURE IBD™	£178,016	15.79	-£3,506	-0.08	£46,263
Scenario 2.1.3 a ii - Assuming discontinuation of biologic treatment for 76% TD; 76% SU.					
Standard of Care	£163,159	15.86	–	–	–
PredictSURE IBD™	£169,238	15.79	£6,079	-0.08	Dominated
Scenario 2.1.3 a iii - Assuming discontinuation of biologic treatment for 40% TD; 40% SU.					
Standard of Care	£181,522	15.86	–	–	–
PredictSURE IBD™	£189,024	15.79	£7,502	-0.08	Dominated
Scenario 2.1.3 b - Assuming discontinuation of biologic treatment for 100% TD; 100% SU.					
Standard of Care	£150,917	15.86	–	–	–
PredictSURE IBD™	£156,047	15.79	£5,130	-0.08	Dominated
Scenario 2.1.4 a – Assuming surgery as last treatment step					
Standard of Care	£203,916	16.13	–	–	–
PredictSURE IBD™	£213,060	16.06	£9,144	-0.07	Dominated
Scenario 2.1.4 b – Removing surgery from the model					
Standard of Care	£197,827	15.88	–	–	–

PredictSURE IBD™	£207,497	15.80	£9,670	-0.08	Dominated
Abbreviations: ICER, incremental cost effectiveness ratio; QALY, quality adjusted life year; TTS, time-to-surgery. *This ICER is for SC vs PredictSURE IBD™, meaning that the prognostic tool is cheaper than SC but also less effective.					

2.2.1 Adding an additional step of treatment with immunomodulators at the end of the top-down arm

As per the request from NICE, the EAG has conducted a scenario analysis where patients in the TD arm of the model had the option to receive IMs at the end of the treatment pathway (after relapsing on second line biologics). However, the EAG reiterates that according to its clinical experts' opinion, this is not a clinically realistic treatment pathway.

Given the lack of alternative data, the EAG assumed that patients on IMs as the last treatment step of the TD arm have the same probability of remission and relapse as patients receiving IMs on the first treatment step in the SU approach. When patients relapse on IMs there are no more treatment options and so these are assumed to remain in the moderate to severe health state of the model.

When this option is implemented in the model, the EAG's deterministic base case ICER (dominated against TD) changes to £105,148 per QALY gained, with TD (via the use of PredictSURE IBD™) generating 0.07 additional QALYs compared to SU (15.93 TD vs 15.86 SU), at an additional cost of £7,502 (£209,427 TD vs £201,925 SU). This scenario results in an increase in the total QALYs associated with TD and decrease in total costs compared with the EAG's base case ICER, while the costs and QALYs associated with SC remain unchanged. The overall increase in QALYs and decrease in costs associated with TD is due to the IM treatment costs being lower than the costs associated with patients staying in the alternative moderate to severe health state after they relapsed on second line biologics. Similarly, the IM state is associated with a probability of remission and mild disease and both of these health states yield a higher utility value than the moderate to severe states.

2.3 Combined scenario analysis

The EAG combined a range of the scenarios described above in order to assess the impact of increasing the effectiveness of the TD strategy while decreasing costs with biologic treatments. These combinations are described, in turn, below and results are reported in the text and summarised in Table 9.

2.3.1 Accounting for the cost-effectiveness of misdiagnosed cases and assumptions around treatment discontinuation in the model

a) The EAG combined the misdiagnosis scenario 2.1.1 with scenario 2.1.3 a i, where it was assumed that after 2 years in remission, 76% of patients in the TD strategy experience mucosal healing, while 40% of patients in the SU arm experience the same outcome.

b) The EAG also combined scenario 2.1.1 with scenario 2.1.3 a ii, where it was assumed that after 2 years in remission, 76% of patients in both the TD and the SU strategies experience mucosal healing.

c) The EAG also combined scenario 2.1.1 with scenario 2.1.3 a iii, where it was assumed that after 2 years in remission, 40% of patients in both the TD and the SU strategies experience mucosal healing.

2.3.2 Varying the assumptions around the measure of relative treatment effectiveness on time to treatment escalation and assumptions around treatment discontinuation in the model

The EAG explored the impact of combining scenario 2.1.3 (where costs associated with biologics were decreased) with changing the effectiveness of TD through the assumptions made for TTE in the model. The EAG used scenario 2.1.2. a ii for all the analyses as this is the scenario that assumes the highest benefit for TD vs SU in terms of TTE.

a) The EAG combined scenario 2.1.2 a ii with scenario 2.1.3 a i, where it was assumed that after 2 years in remission, 76% of patients in the TD strategy experience mucosal healing, while 40% of patients in the SU arm experience the same outcome.

b) The EAG also combined scenario 2.1.2 a ii with scenario 2.1.3 a ii, where it was assumed that after 2 years in remission, 76% of patients in the TD and the SU strategies experience mucosal healing.

c) The EAG also combined scenario 2.1.2 a ii with scenario 2.1.3 a iii, where it was assumed that after 2 years in remission, 40% of patients in the TD and the SU strategies experience mucosal healing.

2.3.3 Varying the proportion of patients who respond to IM and varying the assumptions around the measure of relative treatment effectiveness for time to treatment escalation

One of the scenario analyses carried out by the EAG assumed that 100% of high-risk patients fail to respond to IMs (therefore not deriving any benefit from response to this treatment). This scenario intended to portray an extreme clinical reality where high-risk patients need treatment with a biologic for a response and its impact on the final ICER. The ICER for PredictSURE-IBD™ compared to SU changed from the EAG's base case of dominated (against the prognostic tool) to £34,578.

Therefore, the EAG combined scenario 2.1.2 a ii with varying the proportion of high-risk patients who receive SU therapy and do not respond to IMs thus, increasing the benefit of TD and decreasing the effectiveness of SU, both in terms of TTE and the probability of response and remission in the model.

The EAG tested the assumption that 100% of patients do not respond to IM and varied this percentage to assess the impact on the final ICERs.

2.3.4 Varying the proportion of patients who respond to IM; varying the assumptions around the measure of relative treatment effectiveness for time to treatment escalation; and varying treatment discontinuation assumptions

a) The EAG combined scenario 2.3.2 a with varying the proportion of high-risk patients who receive SU therapy and do not respond to IMs (therefore not deriving any benefit from response to this treatment).

b) The EAG combined scenario 2.3.2 b with varying the proportion of high-risk patients who receive SU therapy and do not respond to IMs (therefore not deriving any benefit from response to this treatment).

c) The EAG combined scenario 2.3.2 c with varying the proportion of high-risk patients who receive SU therapy and do not respond to IMs (therefore not deriving any benefit from response to this treatment).

All the scenarios increased the benefit of TD in terms of TTE and decreased the costs associated biologic treatment (to different amounts). For all scenarios, the EAG tested the assumption that 100% of patients do not respond to IM and varied this percentage to assess the impact on the final ICERs.

2.4 Results of combined scenario analysis

Results of the EAG's scenario analyses are reported in Table 9. The combined analyses produced a wide range of results, with ICERs going from dominant in favour of TD (via the use of PredictSURE IBD™) to being dominated against TD.

Scenario 2.3.1 a resulted in a dominant ICER for PredictSURE IBD™ (and TD), with the prognostic tool being associated with less costs and higher QALYs than SC (and SU). This scenario combines modelling misdiagnosed cases with reducing the costs associated with TD, therefore generating additional QALYs for the prognostic tool at a lower cost, given the assumption that a proportion of patients on TD enter a permanent stage of remission. Given that scenario 2.3.1 a assumes a difference in the rate of treatment discontinuation for biologics (whereby TD patients have a higher probability of discontinuing treatment – due to mucosal healing – than SU patients), this scenario produced the highest cost savings for TD. Scenarios 2.3.1 b and c produced higher ICERs as the relative costs associated with treatment with biologics (and the prognostic tool) increased; however scenario 2.3.1 b resulted in an ICER of £32,875per QALY gained, therefore close to the upper threshold (£30,000) typically used in the NICE decision-making process.

Scenario 2.3.2 a, b and c, explored increasing the effectiveness of TD vs SU with respect to time to treatment escalation (TTE), combined with decreasing the treatment costs with biologics. Scenario a resulted in an ICER of £330,616, with the prognostic tool being associated with less costs but also lower QALYs than SC (and SU). Scenarios 2.3.2 b and c resulted in dominated ICERs against the prognostic tool. As scenario 2.1.2 a ii was used throughout these three scenarios, (thus assuming the highest benefit for TD vs SU in terms of TTE), all three scenarios generated a very small QALY loss of 0.002 for TD (via the use of PredictSURE IBD™). Scenario a generated the cost savings as in this scenario TD patients have a higher probability of discontinuing treatment – due to mucosal healing – than SU patients.

Scenario 2.3.3 and scenario 2.3.4 explored increasing the effectiveness of TD vs SU with respect to time to treatment escalation (TTE), combined with decreasing the treatment costs with biologics and with varying the assumption around the rate of response to IM treatment in the SU strategy.

Scenario 2.3.3 shows that when the relative benefit of the TD strategy compared with SU is increased and when 100% of SU patients are assumed to not respond to treatment with IM, the ICER amounts to £57,757 per QALY gained. Therefore, even when 100% of high-risk patients do not respond to IMs, the ICER for the prognostic tool (and TD) compared to SC (and SU) is still above an ICER threshold of £30,000.

Scenario 2.3.4 a shows that when the benefit of the TD strategy compared with SD is increased (scenario 2.1.2 a ii); when a higher proportion of patients in the TD arm achieves mucosal healing (scenario 2.1.3 a i); and when 100% of SU patients are assumed to not respond to treatment with IM, the final ICER becomes dominant for PredictSURE IBD™ (and TD), with the prognostic tool being associated with less costs and higher QALYs than SC (and SU). The prognostic tool remains dominant up to when the assumption around the proportion of high-risk SU patients not responding to IM treatment is decreased from 100% to 62%. To note is that the EAG's base case analysis estimates that 62% of high-risk patients do not respond to initial treatment with IMs.

Scenario 2.3.4 b and c show that when the benefit of the TD strategy compared with SD is increased (scenario 2.1.2 a ii); when the same proportion of patients in the TD and SU arms achieves mucosal healing (scenario 2.1.3 a ii for 76% and 40%, respectively); and when 100% of SU patients are assumed to not respond to treatment with IM, the final ICERs are £29,225 and £42,740, respectively. Both scenarios generate a QALY gain for the prognostic tool (and TD) compared to SC (and SU); however the additional costs associated with TD are higher in scenario c (40% of patients in remission stop treatment with biologics in both the TD and SU arms) than in scenario b (76% of patients in remission stop treatment with biologics in both the TD and SU arms).

The EAG has produced plots to demonstrate the impact of reducing the percentage of high-risk patients who do not respond to IM from 100% to zero for scenario 2.3.4 a (where PredictSURE IBD™ is dominant). The plot in Figure 6 shows the changes in the incremental costs and QALYs on the cost-effectiveness plane and demonstrates the ICER changing from dominant at 100% non-response to IMs, moving into the south-west quadrant (less costly and less effective for TD) at 62%, then becoming dominated from below 57%. Figure 7 shows the resulting final ICERs, and the drastic variation in these at 62% non-response, when the incremental QALYs become close to zero.

Table 9. Results of scenario analyses

Intervention	Total Costs	Total QALYs	Incremental costs	Incremental QALYs	ICER
Scenario 2.3.1 a (Scenario 2.1.1 Misdiagnosis + assuming discontinuation of biologic treatment for 76% TD; 40% SU)					
Standard of Care	£181,522	15.86	–	–	–
PredictSURE IBD™	£176,541	16.01	-£4,981	0.15	Dominant
Scenario 2.3.1 b (Scenario 2.1.1 Misdiagnosis + assuming discontinuation of biologic treatment for 76% TD; 76% SU)					

Standard of Care	£163,159	15.86	–	–	–
PredictSURE IBD™	£168,153	16.01	£4,995	0.15	£32,875
Scenario 2.3.1 c (Scenario 2.1.1 Misdiagnosis + assuming discontinuation of biologic treatment for 40% TD; 40% SU)					
Standard of Care	£181,522	15.86	–	–	–
PredictSURE IBD™	£188,819	16.01	£7,298	0.15	£48,034
Scenario 2.3.2 a (Assuming the same as base case risk of relapse for second and subsequent treatment steps + assuming discontinuation of biologic treatment for 76% TD; 40% SU)					
Standard of Care	£174,162	15.70	–	–	–
PredictSURE IBD™	£173,517	15.70	-£645	-0.002	£330,616
Scenario 2.3.2 b Assuming the same as base case risk of relapse for second and subsequent treatment steps + assuming discontinuation of biologic treatment for 76% TD; 76% SU)					
Standard of Care	£156,954	15.70	–	–	–
PredictSURE IBD™	£165,233	15.70	£8,279	-0.002	Dominated
Scenario 2.3.2 c (Assuming the same as base case risk of relapse for second and subsequent treatment steps + assuming discontinuation of biologic treatment for 40% TD; 40% SU)					
Standard of Care	£174,162	15.70	–	–	–
PredictSURE IBD™	£184,525	15.70	£10,363	-0.002	Dominated
Scenario 2.3.3 (Assuming the same as base case risk of relapse for second and subsequent treatment steps + assuming that 100% of SU patients do not respond to IM)					
Standard of Care	£201,178	15.61	–	–	–
PredictSURE IBD™	£205,961	15.70	£4,782	0.08	£57,757
Scenario 2.3.4 a (Assuming the same as base case risk of relapse for second and subsequent treatment steps + assuming discontinuation of biologic treatment for 76% TD; 40% SU + assuming that 100% of SU patients do not respond to IM)					

Standard of Care	£180,986	15.61	–	–	–
PredictSURE IBD™	£173,517	15.70	-£7,469	0.08	Dominant
Scenario 2.3.4 b (Assuming the same as base case risk of relapse for second and subsequent treatment steps + assuming discontinuation of biologic treatment for 76% TD; 76% SU + assuming that 100% of SU patients do not respond to IM)					
Standard of Care	£162,813	15.61	–	–	–
PredictSURE IBD™	£165,233	15.70	£2,420	0.08	£29,225
Scenario 2.3.4 c (Assuming the same as base case risk of relapse for second and subsequent treatment steps + assuming discontinuation of biologic treatment for 40% TD; 40% SU + assuming that 100% of SU patients do not respond to IM)					
Standard of Care	£180,986	15.61	–	–	–
PredictSURE IBD™	£184,525	15.70	£3,539	0.08	£42,740
Abbreviations: ICER, incremental cost effectiveness ratio; QALY, quality adjusted life year; TTS, time-to-surgery. *This ICER is for SC vs PredictSURE IBD™, meaning that the prognostic tool is cheaper than SC but also less effective.					

Figure 7. Incremental costs and QALYs as percentage of high risk IM non-responders varies

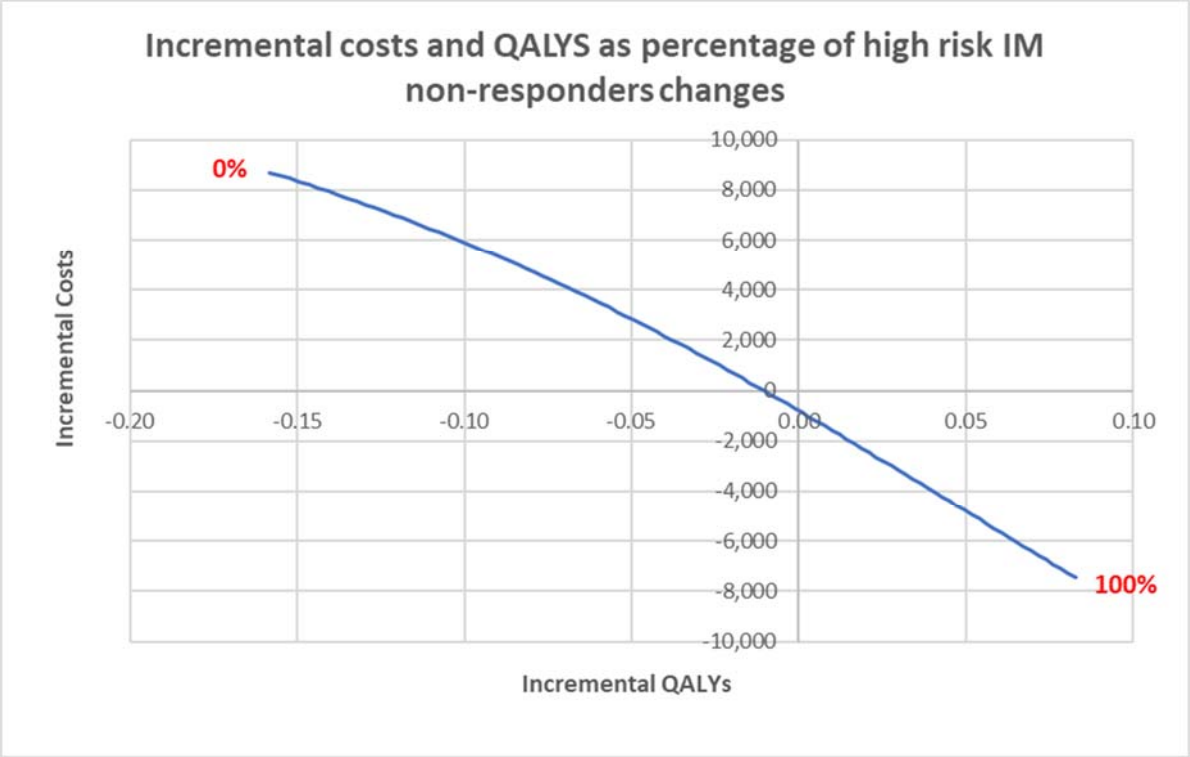
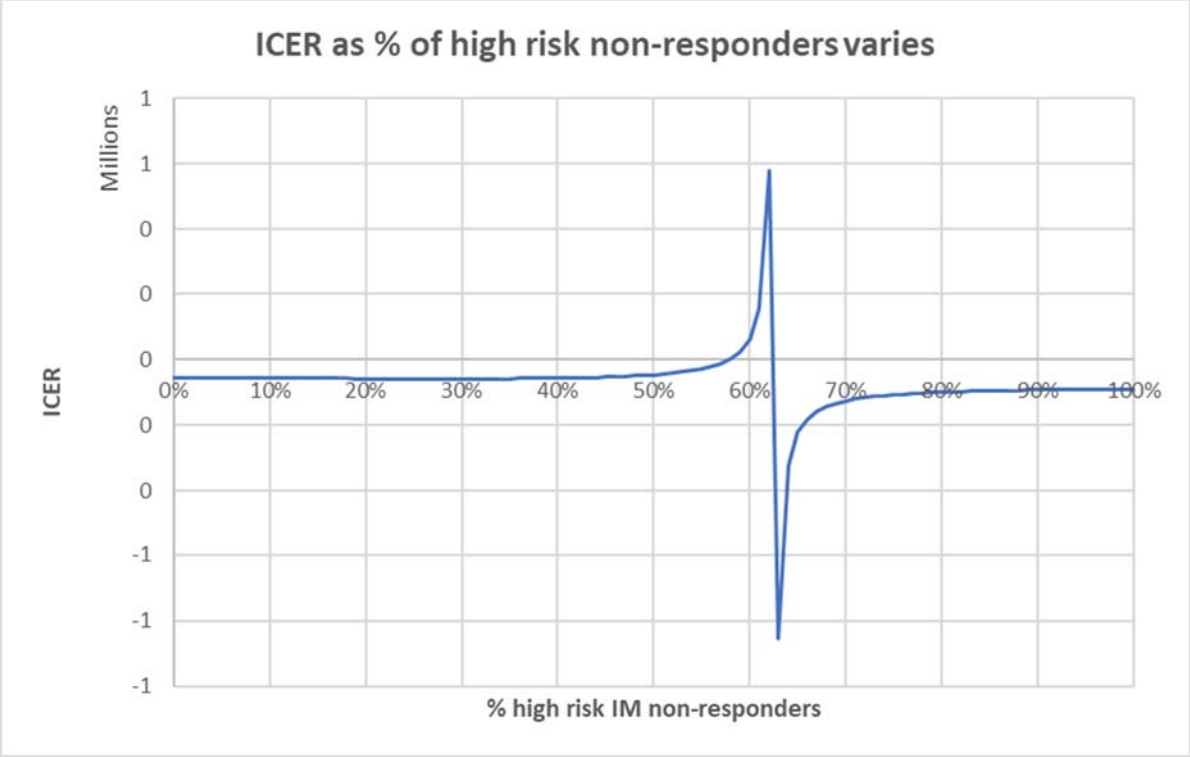


Figure 8. Resulting ICERs as the percentage of high risk IM non-responders varies



2.4.1 Conclusions

The EAG base case and PSA show that there is a small difference in QALYs in favour of SU, suggesting that this strategy might be more beneficial than TD. However, the EAG notes that the difference in incremental QALYs is small, meaning that the final ICER is mainly driven by the difference in costs for TD (via PredictSURE IBD™) compared with SU (via the SC arm).

The EAG conducted extensive scenario analyses and concluded the following:

1. None of the individual scenario analysis led to a change that approximated the final ICER to the upper threshold typically used in the decision-making process by NICE of £30,000 per QALY gained. Excluding surgeries from the model did not have an impact on the dominance of SC over TD, and neither did assuming that surgery has a curative effect for 2-years (scenario 2.1.4);

2. The key drivers of the economic results are the assumptions made around treatment discontinuation with biologics (and the impact this has on costs) combined with the assumptions made around the benefit of TD vs SU through the impact on time to treatment escalation and also through the proportion of patients in the SU arm who respond (and thus derive a benefit) to treatment with IMs;

3. There is one scenario that is above, however close, to a £30,000 ICER threshold. This consists on combining the misdiagnosis scenario with decreasing the costs associated with biologic treatment (through assuming a 76% rate of mucosal healing leading to remission in TD and SU);

4. One scenario is below a £30,000 ICER threshold, generating a QALY gain for PredictSURE IBD™ (via TD) at a higher cost when compared with SU. This consisted of increasing the relative effectiveness of TD on TTE and additionally reducing the effectiveness of SU (through assuming a 0% probability of response to IM treatment for high-risk patients) combined with assuming a 76% rate of mucosal healing leading to remission in TD and SU);

5. There were two scenarios that generated a QALY gain for PredictSURE IBD™ (via TD) at a smaller cost than SU (i.e. SU was dominated by TD). These consisted of: 1) combining the misdiagnosis scenario with decreasing the costs associated with biologic treatment (through assuming different rates of mucosal healing leading to remission in TD and SU - 76% TD and 40% SU, respectively); and 2) increasing the relative effectiveness of TD on TTE and additionally reducing the effectiveness of SU (through assuming a 0% probability of response

to IM treatment for high-risk patients combined with assuming different rates of mucosal healing leading to remission in TD and SU - 76% TD and 40% SU, respectively).

In conclusion, the cost associated with the prognostic tool (and TD) decreases as the difference in discontinuation rates for biologic treatment increase (with higher discontinuation rates in the TD arm compared with the SU arm). Furthermore, assuming an increase in the benefit of TD compared with SU (through delaying the time to treatment escalation in the TD arm compared to the base case analysis in combination with assuming that 0% of patients respond to IM treatment) generates a QALY gain for TD vs SU.

However, the EAG 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 where TD was assumed to be more effective than SU.

3 REFERENCES

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.

International Standard Randomised Controlled Trials Number Registry. PROFILE - personalised medicine in Crohn's disease. 2019. Available from: http://www.isrctn.com/ISRCTN11808228. Date accessed: 9 Oct 2019.

Hoekman D, Stibbe J, Baert F, Caenepeel P, Vergauwe P, De VM, et al. Long-term Outcome of Early Combined Immunosuppression Versus Conventional Management in Newly Diagnosed Crohn's Disease. 2018. Available from: https://www.cochranelibrary.com/central/doi/10.1002/central/CN-01656067/full. Date accessed.

National Institute for Health and Care Excellence (NICE). Vedolizumab for treating moderately to severely active Crohn's disease after prior therapy. Technology appraisal guidance [TA352]. 2015. Available from: https://www.nice.org.uk/guidance/ta352. Date accessed: July 2019

Marchetti M, Liberato NL, Di Sabatino A, Corazza GR. Cost-effectiveness analysis of top-down versus step-up strategies in patients with newly diagnosed active luminal Crohn's disease. *European Journal of Health Economics* 2013; 14: 853-61.

Baert, F.J., Moortgat, L., Van Assche, G.A., Caenepeel, P., Vergauwe, P.L., De Vos, M., Stokkers, P.C., Hommes, D.W., Vermeire, S., Ritgeerts, P.J., Feagan, B.G., D'Haens, G.: Mucosal healing predicts sustained clinical remission in early Crohn's disease. Gastroenterology 134(suppl. 1), A640 (2008)