Published in final edited form as:

Inflamm Bowel Dis. 2011 January ; 17(1): 30–38. doi:10.1002/ibd.21386.

A Real-Time Tool to Display the Predicted Disease Course and Treatment Response for Children with Crohn's Disease

Corey A. Siegel¹, Lori S. Siegel², Jeffrey S. Hyams³, Subra Kugathasan⁴, James Markowitz⁵, Joel R. Rosh⁶, Neal Leleiko⁷, David R. Mack⁸, Wallace Crandall⁹, Jonathan Evans¹⁰, David J. Keljo¹¹, Anthony R. Otley¹², Maria Oliva-Hemker¹³, Sharmayne Farrior¹⁸, Christine R. Langton³, Iwona T. Wrobel¹⁴, Ghassan Wahbeh¹⁵, J. Antonio Quiros¹⁶, Gary Silber¹⁷, Ron J. Bahar¹⁸, Bruce E. Sands¹⁹, and Marla C. Dubinsky¹⁸

¹Dartmouth-Hitchcock Inflammatory Bowel Disease Center, Dartmouth-Hitchcock Medical Center, Lebanon, NH, USA.

²Siegel Environmental Dynamics, LLC, Hanover, NH, USA.

³Department of Pediatrics, Connecticut Children's Medical Center, Hartford, CT, USA.

⁴Children's Healthcare of Atlanta, Emory University, Atlanta, GA, USA.

⁵Schneider Children's Hospital, New Hyde Park, NY, USA.

⁶Department of Pediatrics, Morristown Memorial Hospital, Morristown, NJ, USA.

⁷Department of Pediatrics, Hasbro Children's Hospital, Providence, RI, USA.

⁸Department of Pediatrics, Children's Hospital of Eastern Ontario, Ottawa, ON, Canada.

⁹Department of Pediatrics, Nationwide Children's Hospital, Columbus, OH, USA.

¹⁰Department of Pediatrics, Nemours Children's Clinic, Jacksonville, FL, USA.

¹¹Children's Hospital of Pittsburgh, University of Pittsburgh Medical Center, Pittsburgh, PA, USA.

¹²Department of Pediatrics, Dalhousie University, Halifax, NS, Canada.

¹³Department of Pediatrics, Johns Hopkins Hospital, Baltimore, MD, USA.

¹⁴Department of Pediatrics, Alberta Children's Hospital, Calgary, AB, Canada.

¹⁵Department of Pediatrics, Seattle Children's Hospital, Seattle, WA, USA.

¹⁶California Pacific Medical Center, San Francisco, CA, USA.

¹⁷Department of Pediatrics, Phoenix Children's Hospital, Phoenix, AZ, USA.

¹⁸Department of Pediatrics, IBD Center, Cedars-Sinai Medical Center, Los Angeles, CA, USA.

Correspondence to: Dr. Corey A. Siegel, Inflammatory Bowel Disease Center, Section of Gastroenterology and Hepatology, Dartmouth-Hitchcock Medical Center, Lebanon, NH 03756, p (603) 650 8101, f (603) 650 5225, corey.a.siegel@hitchcock.org.

Disclosures: Dr. Siegel serves as a consultant to Abbott Laboratories and UCB. Dr. Dubinsky serves as a consultant to Abbott Laboratories, Centocor, Prometheus Labs, and UCB. Dr. Sands serves as a consultant to Abbott Laboratories, Centocor, Prometheus Labs, and UCB. Dr. Rosh serves as a consultant to Abbott Laboratories, Centocor and UCB, receives grant support from Abbott Laboratories and Centocor, and serves as a speaker for Abbott and Abbott Nutrition. Christine Langton receives salary support from Centocor. Dr. Hyams serves as a consultant to and receives research support from Abbott, Centocor, UCB. Dr. Markowitz serves as a consultant to and receives research support from Centocor and Prometheus Labs, and has received honoraria from Prometheus labs. Dr. Quiros serves as a consultant for UCB.

Dartmouth-Hitchcock Medical Center and Cedars-Sinai Medical Center have a United States patent pending on intellectual property related to this work, filed on March 24, 2010 as Application serial number 61/163,024. Inventors are CA Siegel, LS Siegel and MC Dubinsky.

¹⁹Gastrointestinal Unit, MGH Crohn's and Colitis Center, Massachusetts General Hospital, Boston, MA, USA.

Abstract

BACKGROUND—Immunomodulators and biologics are effective treatments for children with Crohn's disease (CD). The challenge of communicating the anticipated disease course with and without therapy to patients and parents is a barrier to the timely use of these agents. The aim of this project was to develop a tool to graphically display the predicted risks of CD and expected benefits of therapy.

METHODS—Using prospectively collected data from 796 pediatric CD patients we developed a model using system dynamics analysis (SDA). The primary model outcome is the probability of developing a CD related complication. Input variables include patient and disease characteristics, magnitude of serologic immune responses expressed as the quartile sum score (QSS), and exposure to medical treatments.

RESULTS—Multivariate Cox proportional analyses show variables contributing a significant increase in the hazard ratio (HR) for a disease complication include female gender, older age at diagnosis, small bowel or perianal disease, and a higher QSS. As QSS increases, the HR for early use of corticosteroids increases, in contrast to a decreasing HR with early use of immunomodulators, early or late biologics, and early combination therapy. The concordance index for the model is 0.81. Using SDA, results of the Cox analyses are transformed into a simple graph displaying a real-time individualized probability of disease complication and treatment response.

CONCLUSIONS—We have developed a tool to predict and communicate individualized risks of CD complications and how this is modified by treatment. Once validated, it can be used at the bedside to facilitate patient decision making.

Keywords

Crohn's; complication; prediction; model; pediatric

Crohn's disease is a chronic inflammatory bowel disease (IBD) with a high rate of complications that frequently lead to a decreased quality of life. Despite our better understanding of disease pathophysiology and treatment over the past three decades, the risk of intestinal complications and need for resultant surgery have not changed significantly.¹ Fortunately, newer treatments and more aggressive treatment algorithms are showing promise with higher rates of clinical remission. Specifically, the earlier use of anti-tumor necrosis factor (TNF) and immunomodulator (IM) therapy may be able to alter the natural history of Crohn's disease and prevent complications before they develop.^{2–4}

Enthusiasm over the use of these agents earlier in the disease course is tempered by the concern of side-effects associated with immune suppression. Patients and parents appear willing to accept risks of adverse events including life-threatening infections and lymphoma, but this is dependent upon the expectation that the disease is severe (i.e., they “deserve the treatment”), and that treatment will be effective.^{5, 6} If we were able to identify the patients who are at the most risk from their disease, then it would allow us to more easily justify using our most effective therapies sooner. Conversely, patients at low risk can be spared exposure to potentially toxic medications. Clinical and serologic factors have been identified that appear to serve as predictors of disease severity.^{7, 8} Although useful in helping to risk stratify patients, the results of these studies are presented as relative statistical comparisons (e.g., odds ratios, hazard ratios, p-values), which are difficult to translate clearly for patients so that they understand the implications of their disease. A tool to present more clearly

individual absolute disease risk and predicted response to therapy in real-time would enhance a provider's ability to communicate with patients and their families.

System dynamics analysis (SDA) is a methodology that addresses the inherent dynamic complexity of interactions between variables. Used more commonly in the fields of economics and engineering, SDA has had limited use in medicine.^{9, 10} Advantages of SDA over traditional statistical methods are that it can provide real-time individualized predictions of outcomes, it uses a straightforward control panel to input variables, and it can create simple graphs to convey predicted outcomes over time. Applied to medicine, SDA may have the ability to translate complex clinical data into patient friendly results.

Using a cohort of pediatric Crohn's disease patients with detailed clinical and serologic profiles we developed a model to predict the individualized risk of complicated Crohn's disease and how this risk is modified by treatment. SDA provides a graphic display of these results for a real-time visual explanation predicting disease outcome with or without medical therapy.

METHODS

Patient population and data collection

The patient population for analysis included pediatric Crohn's disease patients recruited from 21 participating sites. The details of the sites, eligibility for enrollment, diagnostic criteria for inclusion and processing of blood for serological testing has been previously described in the publication by Dubinsky and colleagues.⁷ Data collected included: age at diagnosis, gender, follow-up duration, time to complication, disease location, disease phenotype, antibody immune responses and medication exposure (including date of treatment in relationship to a disease complication). Each site had Internal Review Board approval.

Model variables—Disease location was defined by the extent of bowel involvement at the time of their initial presentation, and based upon radiologic, endoscopic and histologic evidence of inflammation. **Serologic inputs** were based upon the antibody immune responses from blood samples taken at the time of patient enrollment into the prospective cohort. Sera were analyzed quantitatively for anti-*Saccharomyces cerevisiae* antibody (ASCA) IgA and IgG, anti-flagellin (anti-CBir1), anti-outer-membrane porin C (OmpC) of *E. coli*, and perinuclear anti-neutrophil antibody (pANCA).⁷ For ASCA, the numerical titer of IgA and IgG were added together and assigned a quartile (1–4) in relationship to the range of all patient results. Each patient was also assigned a quartile for anti-CBir1 and anti-OmpC. The quartile scores for each of these markers were added together to yield a quartile sum score (QSS).⁷ pANCA was handled as dichotomous based on a cutoff value provided by the reference lab. Therefore, each patient was assigned a QSS (range 3–12) and pANCA positive or negative. QSS groups were developed, with 1 representing those with the lowest QSS and 4 representing those with the highest, to allow for categorical analyses.

Although the model is set up to accept genetic inputs including the NOD2 variants SNP8, SNP12 and SNP13, these data were not available for a number of subjects included in the final analyses. To avoid having to exclude these patients because of missing data and lose overall statistical power, we chose not to include genetics in the final model. However, the model has the ability to incorporate these results when they become consistently available from this patient cohort or others.

Specific treatment data were collected for corticosteroids, IM (6-mercaptopurine, azathioprine), and anti-TNF agents. For IM and anti-TNF agents, the actual date of initiation

was recorded. Where dates of initiation are not available, data were available to determine if they were administered within 30 days of diagnosis, between 30–90 days, or after 90 days. The exact date of corticosteroid administration was not recorded, but data were available to determine if they were administered within 30 days of diagnosis, between 30–90 days, or after 90 days. The model tests early IM and early anti-TNF treatment versus late IM and late anti-TNF treatment, with the cutoff being 90 days. The model only tests early corticosteroids, since data for the timing of late corticosteroids were of variable precision and it would be difficult to interpret their influence on clinical outcomes.

Main outcome of the model

The dependent variable of the model is the probability of a complication of Crohn's disease. A complication was defined as the development of either internal penetrating (IP) or stricturing (S) Crohn's disease. Internal penetrating is further clarified as evidence of an entero-enteric or entero-vesicular fistula, intra-abdominal abscess or intestinal perforation. Stricturing disease specifically referred to the occurrence of a persistent luminal narrowing demonstrated by radiologic, endoscopic or surgical examination combined with pre-stenotic dilatation and/or obstructive signs or symptoms.

Exclusion criteria

Patients were excluded from the final model analysis if data were incomplete. An exception was if only one of the serologic markers were missing, in which case the missing antimicrobial antibody titer was imputed as the average of the two other markers. Patients were also excluded if the sequence of medical therapy and the complication were unclear. For example, if a patient had a complication within the first month of enrollment in the study and received anti-TNF therapy that same month, it was impossible to know if they received the treatment because of a complication, or developed the complication despite the treatment (bias by indication). In addition, if a patient receiving IM therapy had a complication within the first three months of this treatment, they were excluded since it is unclear if IMs have any effect before three full months of exposure.

Statistical analysis

Both univariate and multivariate Cox proportional analyses were performed to examine the association of patient characteristics and individual treatments on the risk of a complication. Independent variables included: age at diagnosis, gender, disease phenotype, serologic immune responses, and exposure to medications (corticosteroids, immunomodulators, and anti-TNF agents).

Some patients had a complication within the first month of entering the cohort. Since it was not clear if these complications already existed at the time of diagnosis, we did not want to include them in the longitudinal analysis. To avoid excluding these patients entirely, two separate analyses were completed. The first analysis examined how patient variables influence the risk of a complication within this first month ("initial analysis"), developing hazard ratios (HRs) for each significant variable. The second analysis was performed to develop HRs for the longitudinal follow-up period out to five years ("5 year analysis"). Each of these analyses results in an independent hazard function. To study the effect of treatment as a factor of the serologic quartile, interaction terms were created to report the HRs for each treatment within QSS group, and also to calculate the statistical trend from QSS group one through four.

The final model was calibrated to obtain the lowest Akaike information criteria (AIC) and significant likelihood ratio (LR) chi-square. The model was internally validated using bootstrapping (internal cross-validation between multiple subsets of patients) and by

evaluating the Harrell's C concordance index.¹¹ Analyses were performed using Stata/IC 11.0 (College Station, TX).

System dynamics analysis

The SDA model is built based on the results from the Cox proportional analyses. The model integrates the two hazard functions to bridge the initial risk of complication within the first month (initial analysis) and the risk of complication from the first month to 5 years since diagnosis (5 year analysis). The risk of an individual patient is calculated according to the HR of each variable. The overall unit hazard function is the risk of a complication at any given time. Equation 1 shows that it is the product of the baseline function and the expression derived from the calculated HR of each variable (HR_i) for the value of the variable, i.e.,

$$h(t|x)=h(t|0)*\prod HR_i^{\text{Value}_i} \quad (\text{Equation 1})$$

This expression is simplified as follows:

$$h(t|x)=h(t|0)*\exp\left(\sum \text{LN}(HR_i)*\text{Value}_i\right) \quad (\text{Equation 2})$$

where the expression $\text{LN}(HR)$ yields the coefficient of each variable (Equation 2). Model simulations run from month 0, i.e., when a patient enters the analysis, through 5 years. However, graphical output presents only three years from the time the patient enters the analysis, allowing for two years since diagnosis to have elapsed. Model values are updated every 0.25 months. The model stores and can plot and print the output for every time step or for other time intervals, as desired. Figure 1 shows an overview of the system dynamics model. Relationships between variables are demonstrated by arrows, and the boxed text represents the variables that accumulate risk over time. Multiple sub-models are imbedded within this overall model view but are not displayed. Definitions of the variables are shown in Table 1 along with the formulas that define the functional relationships. The model underwent multiple iterations before arriving at the current model structure. SDA were performed using Vensim (Harvard, MA).

RESULTS

Patient population

The entire cohort included 796 patients and reasons for exclusion are displayed in Figure 2. Sixty-seven were excluded before analysis based on missing data and an unclear sequence of complication and treatment. Initially, 729 patients were included for analysis of 1 month outcomes. After 1 month, 150 patients were excluded, primarily for an unclear sequence of complication and treatment, leaving a total of 579 patients for inclusion for the 5 year analysis. Table 2 shows the characteristics of the patients included in the initial and 5 year analyses, the distribution of disease location and phenotype, frequency of antibody responses and medication exposure.

Cox proportional analysis

Univariate analyses were performed both for the initial and 5 year analyses. Older age (HR 1.15, 95% CI 1.01–1.21), colonic disease (HR 0.18, 95% CI 0.09–0.39) and increasing QSS (HR 1.4, 95% CI 1.15–1.64) were statistically significant for an association with a complication within the first month. Table 3 shows the HR and 95% confidence intervals for variables included in the 5 year univariate analysis.

The multivariate analysis showed a significant association between older age at diagnosis (HR 1.1, 95% CI 0.99–1.29), colonic or upper tract disease (HR 0.28, 95% CI 0.15–0.51), and increasing QSS (HR 1.3, 95% CI 1.12–1.58) for the risk of a complication within the first month. Table 3 displays the multivariate analysis for the risk of a complication over the 5 year follow-up period. Results show that female gender, small bowel or perianal disease and increasing QSS were significant predictors of a disease complication; older age at diagnosis showed a trend towards an association with disease complications. A positive pANCA was protective against a disease complication, and colonic or upper tract disease showed a trend for a protective benefit.

To explore how medication effect might be influenced by other variables, interaction terms were developed between QSS groups and different treatments. As the QSS group increases, the HR for the early use of corticosteroids increases, in contrast to a decreasing HR with early use of IM, early or late biologics, and early combination therapy (Table 4). Late IM is shown without the interaction term as the association was stronger as a multivariate factor on its own.

The Cox model passes the tests for proportionality for the overall model and the individual parameters included in the multivariate analysis. The model with the lowest AIC has an LR Chi-square $p < 0.0001$. Bootstrapping showed consistent results within the 100 different run permutations. The Harrell's C for the full model is 0.81.

System dynamics model

The results of the Cox model drive the system dynamics model. The system dynamics model allows all input variables to be entered using a simple computer control panel (Figure 3a), and then predicted outcomes to be displayed real-time in a graphical format. An individual patient's characteristics are input using the toggle switches above, and then the treatment options below can be switched on or off. The model output shows the probability for that specific patient to have a complication of Crohn's disease over the next three years. A sample patient is displayed in Figure 3b. This figure also shows how early anti-TNF therapy decreases this patient's risk over time. Figure 3c shows how this same patient's risk increases if treated with corticosteroids. A different patient example is shown in Figure 3d, representing a low-risk patient with the maximum risk of a complication around 13% at three years. Although there is a 50% relative risk reduction with early IM therapy, the absolute change (approximately 7%) is very small.

DISCUSSION

Using system dynamics analysis, we have developed a tool to predict and communicate individualized risks of Crohn's disease complications and how this is modified by treatment. This model translates complex statistical results into a patient friendly graphical output. It offers patients and their families a real-time presentation of the risks of their disease and what they can expect as a result of treatment.

Children with moderate-severely active Crohn's disease benefit greatly from the early use of immunomodulators¹² and from anti-TNF therapy.³ Although a decision to administer these treatments may be straight-forward for physicians based on their understanding of clinical trial data and knowledge of the natural history of Crohn's disease, acceptance of these therapies by patients and their parents is more difficult. The difficulty is likely based on their fear of side-effects of immune suppression (e.g., lymphomas), and their lack of appreciation of the high likelihood of experiencing complicated Crohn's disease and the substantial benefit offered by these treatments. A tool to help show patients and their parents why a

specific therapy is being recommended will most likely allow for better communication and more informed decisions.

In addition to the development of the risk communication tool, our analysis also showed some interesting and clinically relevant results. Based on Equations 1 and 2, the risk of a complication is dependent on the time at which a patient presents for treatment. For example, if a patient has had Crohn's disease for many months or years and has not had a complication, their chance of a future complication is lower than someone presenting with new onset disease (i.e., a patient with a longer time without a complication has proven to have less aggressive disease). It has been previously reported that the risk of a complication increases as the serologic QSS increases,⁷ but we show a novel finding that treatment response is also dependent on the QSS. As QSS increases, our data indicate that IM and anti-TNF therapy have an increasing benefit, and steroids are more detrimental. Conversely, patients with a low QSS have a lower benefit of IM and anti-TNF therapy, and less risk associated with steroid use. Using QSS and other variables, the ability to focus early intensive therapy on certain patients is becoming clearer.

As with any complex disease model, there are limitations to these results. First, the model has not yet been externally validated. Although all models should ultimately be tested in a separate group of patients, statistical techniques such as bootstrapping can be used to see how the model behaves with multiple different subgroups of patients. In this model, despite using 100 different patient subsets, the model results did not change significantly. Second, statistical power was likely lost since some complication and treatment scenarios included only a few patients for analysis. For example, we are able to report associations with individual medications, but only for a few of the many combination treatment regimens. Although a clinically relevant question, in its current state the model is not powered to address all treatment permutations. Next, it is possible that the influence of therapy on the risk of a complication is confounded by disease severity or the indication for therapy. We accounted for this with our exclusion criteria regarding the timing of a complication in relationship to initiation of therapy, but it is still possible that we are over- or underestimating the true impact of medical treatment. Finally, this model does not incorporate data on the risks of therapy. Fortunately, rates of serious adverse events related to treatment are too low to show in a meaningful way using this type of SDA model. Considerations of the risks of therapy should not be ignored, but reviewed with patients in another manner at the time of displaying their projected disease course.

Previous studies have also shown that disease characteristics, and serologic and genetic markers can predict a disabling or complicated disease course.^{7, 8, 13, 14} The novel contribution of our study is the translation of complicated data to a format that can be easily shared with patients. Work in risk communication focuses on a number of techniques to enhance patient communication.¹⁵ Many of these ideas focus on the use of visual tools to aid in comprehension, avoiding relative rates (e.g., hazard ratios, relative risk), and allowing for individualized as opposed to generic predictions. In this work, we believe we address many of the hurdles in accurate patient communication and take a great stride forward in being able to engage our patients in the medical decision making process. Then, our patients can make informed, and preference based decisions taking into account how much risk of their disease they are willing to tolerate, and how much benefit they demand if they are going to agree to take a medication with concerning side-effects.

The goal will be to bring this risk prediction and patient communication tool to the bedside. To achieve this goal, an important next step is to externally validate this model in a separate patient cohort to ensure generalizability. In addition, to make sure that this is clinically useful, we need to test patients' and parents' comprehension of the model output, and to

determine how this tool influences their decisions. Other future plans include adding results of genome wide association studies for these patients, which we expect will improve the precision of the predictions. Other variables such as new serologic markers, endoscopic results and clinical indicators can also be added if they enhance the value of the model. The ideal practical use of the model would be to have the physician input the disease characteristics and serologic results, and then allow patients to change the treatment toggle switches themselves to see how individual treatments might impact their future disease course. Ultimately, this model will be the centerpiece of a “decision aid” that also includes information for patients on Crohn’s disease itself, the risks of therapy, and tools to help patients and their families make decisions about treatment based on their personal preferences. An adult model and decision aid are also being developed.

With the dramatic shift taking place in the care of IBD patients focusing on early intervention with IM, anti-TNF and eventually other biologic therapies, it will be even more important to properly show patients what to expect with or without treatment. Concern over the risks of therapy is often a barrier to the appropriate use of medications. Although treatment related risks still need to be carefully reviewed, this model allows it to be balanced against the substantial risks associated with the disease. This will not only help patients feel more comfortable with their treatments, but likely will also help physicians become more efficient in discussing prognosis, increase compliance, and lead to an overall improvement in their quality of life. This model will become increasingly more valuable as data and time evolve.

Acknowledgments

Grant Support: Dr. Siegel is supported by a CCFA career development award and by Grant Number K23DK078678 from the National Institute of Diabetes and Digestive and Kidney Diseases. Dr. Dubinsky is supported by the Abe and Claire Levine Chair in Pediatric IBD, and Grant Number K23DK066248 from the National Institute of Diabetes and Digestive and Kidney Diseases. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Diabetes and Digestive and Kidney Diseases or the National Institutes of Health. Dr. Ghassan is supported in part by the Truman Katz Foundation.

Abbreviations used

AIC	Akaike information criteria
ASCA	anti-saccharomyces cerevisiae antibody
CBir1	anti-flagellin
CD	Crohn’s disease
DX	diagnosis
HR	hazard ratio
IBD	inflammatory bowel disease
IM	immunomodulators
IP	internal penetrating
LB	large bowel
LR	likelihood ratio
OmpC	out-membrane porin C
PA	perianal

pANCA	perinuclear antineutrophil antibody
QSS	quartile sum score
S	stricturing
SB	small bowel
SDA	system dynamics analysis
SNP	single nucleotide polymorphism
TNF	tumor necrosis factor
UT	upper tract

REFERENCES

1. Cosnes J, Nion-Larmurier I, Beaugerie L, et al. Impact of the increasing use of immunosuppressants in Crohn's disease on the need for intestinal surgery. *Gut*. 2005; 54:237–241. [PubMed: 15647188]
2. D'Haens G, Baert F, van Assche G, et al. Early combined immunosuppression or conventional management in patients with newly diagnosed Crohn's disease: an open randomised trial. *Lancet*. 2008; 371:660–667. [PubMed: 18295023]
3. Hyams J, Crandall W, Kugathasan S, et al. Induction and maintenance infliximab therapy for the treatment of moderate-to-severe Crohn's disease in children. *Gastroenterology*. 2007; 132:863–873. quiz 1165–6. [PubMed: 17324398]
4. Sandborn W, Rutgeerts P, Reinisch W, et al. SONIC: A randomized, double-blind, controlled trial comparing infliximab and infliximab plus azathioprine to azathioprine in patients with Crohn's disease naive to immunomodulators and biologic therapy. *Am J Gastroenterol*. 2008; 103:S436.
5. Johnson FR, Ozdemir S, Mansfield C, et al. Crohn's disease patients' risk-benefit preferences: serious adverse event risks versus treatment efficacy. *Gastroenterology*. 2007; 133:769–779. [PubMed: 17628557]
6. Johnson FR, Ozdemir S, Mansfield C, et al. Are adult patients more tolerant of treatment risks than parents of juvenile patients? *Risk Anal*. 2009; 29:121–136. [PubMed: 18826414]
7. Dubinsky MC, Kugathasan S, Mei L, et al. Increased immune reactivity predicts aggressive complicating Crohn's disease in children. *Clin Gastroenterol Hepatol*. 2008; 6:1105–1111. [PubMed: 18619921]
8. Beaugerie L, Seksik P, Nion-Larmurier I, et al. Predictors of Crohn's disease. *Gastroenterology*. 2006; 130:650–656. [PubMed: 16530505]
9. Ford, A. Modeling the environment: An introduction to system dynamics modeling of environmental systems. Island Press; 1999.
10. Jones AP, Homer JB, Murphy DL, et al. Understanding diabetes population dynamics through simulation modeling and experimentation. *Am J Public Health*. 2006; 96:488–494. [PubMed: 16449587]
11. Harrell, FE. Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis. Springer; 2001. p. 87-287.
12. Markowitz J, Grancher K, Kohn N, et al. A multicenter trial of 6-mercaptopurine and prednisone in children with newly diagnosed Crohn's disease. *Gastroenterology*. 2000; 119:895–902. [PubMed: 11040176]
13. Henckaerts L, et al. NOD2/CARD15 and risk of complicating disease behavior in Crohn's disease. *Gastroenterology*. 2007; 132:A-17.
14. Ferrante M, Henckaerts L, Joossens M, et al. New serological markers in inflammatory bowel disease are associated with complicated disease behaviour. *Gut*. 2007
15. Lipkus IM. Numeric, verbal, and visual formats of conveying health risks: suggested best practices and future recommendations. *Med Decis Making*. 2007; 27:696–713. [PubMed: 17873259]

Figure 1. Model overview

The model overview shows the structure of the system dynamics model. Red text represent the main outcome of interest. Text in black are variables that are defined within this sector of the model. Text in grey are defined in other sub-models within this model overview. Blue arrows represent a functional relationship between two variables. Boxed text represents the variables that accumulate risk over time. Arrows with black arrowhead represent the flow of accumulating risk. The shape at the tail of this large arrow indicates that the model is an open system, with risk accumulating based on other contributing variables.

Figure 2. Flow chart of patients included in the final analysis

Flow chart of patients from entire cohort to only those included in the model based on inclusion and exclusion criteria.

Figure 3a.

Dx=diagnosis; SB=small bowel; LB=large bowel; UT=upper tract; PA=perianal; ASCA=anti-*Saccharomyces cerevisiae* antibody; anti-CBir1=anti-flagellin; OmpC=anti-outer-membrane porin C of *E. coli*; pANCA=perinuclear anti-neutrophil antibody; SNP=single nucleotide polymorphism; TNF=tumor necrosis factor; IM=immunomodulator

Figure 3a. Model control panel

Figure 3a. The model control panel is an interactive computer-based tool such that you can move the individual toggle switches on the right to match individual patient characteristics to show the probability of a disease complication over three years, and how this is modified by different treatments. The y-axis is the probability of a complication (0%-100%) and the x-axis is time from baseline to three years.

Figure 3b.

Figure 3b. Model output for a high-risk 16-year-old girl treated with early anti-TNF therapy

Figure 3b. The model is set for a high-risk female patient and shows her probability of a disease complication of 100% by 1.5 years. The early anti-TNF treatment toggle switch is turned on to demonstrate how the risk is attenuated with this treatment.

Figure 3c. **Figure 3c. Model output for a high-risk 16-year-old girl treated with prednisone**

Figure 3c. The model is set for the same high-risk female patient as in 3b, but now the corticosteroids toggle switch is turned on and demonstrates how the use of corticosteroids predicts a disease complication sooner than without treatment.

Figure 3d.

Figure 3d. Model output for a low-risk 6-year-old boy treated with an early immunomodulator
Figure 3d. This model is set for a low-risk male patient who has a much lower chance of a complication than the patient in Figure 3b. It also demonstrates the value of showing absolute as opposed to relative risk reduction, since a 50% relative risk reduction with the use of early immunomodulators has limited clinical value in this low-risk patient.

Table 1

Variable definitions and formulations

Variable	Definition	Formula
IP or S risk	The risk that a patient will have a complication of IP or S within the period of time, accumulating the unit hazards over time but only once the patient enters the analysis, i.e., time elapsed prior to that point do not accumulate risk.	$\min(("100 \text{ percent"} - \text{IP or S risk}) / \text{TIME STEP}, \text{if then else}(\text{Time} > \text{Months since Dx} : \text{AND: Time} = \text{time of IP or S event}, \text{Hazard of IP or S} / \text{TIME STEP} * "100 \text{ percent}", 0))$
Change in risk	The addition of unit hazards over time, avoiding exceeding 100% of risk	$\min(("100 \text{ percent"} - \text{IP or S risk}) / \text{TIME STEP}, \text{if then else}(\text{Time} \geq \text{Months since Dx}, \text{Actual hazard of IP or S} * "100 \text{ percent}", 0))$
Risk of IP or S	The IP or S risk trajectory smoothed over a smoothing period.	$\text{smooth}(\text{if then else}(\text{Genetic information} = 0, \text{IP or S risk}, \text{Alternate IP or S risk}), \text{smoothing time})$
Smoothing time	The time period over which to smooth the risk trajectory.	2 months
IP or S risk without treatment	Same as IP or S risk, but calculated with assuming all treatment data = 0	$\min(("100 \text{ percent"} - \text{IP or S risk}) / \text{TIME STEP}, \text{if then else}(\text{Time} > \text{Months since Dx} : \text{AND: Time} = \text{time of IP or S event}, \text{Hazard of IP or S without treatment} / \text{TIME STEP} * "100 \text{ percent}", 0))$
Change in risk without treatment	The addition of unit hazards over time calculated with assuming all treatment data = 0, avoiding exceeding 100% of risk	$\min(("100 \text{ percent"} - \text{IP or S risk without treatment}) / \text{TIME STEP}, \text{if then else}(\text{Time} \geq \text{Months since Dx}, \text{Actual hazard of IP or S without treatment} * "100 \text{ percent}", 0))$
Months since Dx	Specifies time since diagnosis when patient enters analysis	Excludes unit hazards that apply to given times that have already elapsed without the occurrence of complication
Overall HR	Provides comprehensive measure of HR for individual patient, incorporating HRs of all independent variables	$\text{if then else}(\text{Time} \leq 1, \text{zidz}(\text{Hazard of IP or S}, \text{Baseline if less than 1 month}), \text{zidz}(\text{Hazard of IP or S}, \text{Baseline IP or S hazard}))$
Benefits of Therapy	Calculates the change in risk between with and without treatment at a specified time threshold	$\text{Sample if true}(\text{Current risk time} = \text{Time threshold}, \text{zidz}(\text{Three year IP or S risk without treatment} - \text{Three year IP or S risk}, \text{Three year IP or S risk without treatment}), 0)$

Table 2

Patient Characteristics

Characteristic	Initial analysis (N=729)	5 year analysis (N=579)
Demographics		
Age at diagnosis (median, range)	12 yrs (2 mo-19 yrs)	12 yrs, (2 mo-19 yrs)
Female	42%	40%
Follow-up duration (median, range)	-	33 mo (1 mo-235 mo)
Time to complication (median, range)	-	36 mo (1 mo-174 mo)
Disease Location	N(%)	N (%)
Small bowel	488 (67)	390 (67)
Colonic	624 (86)	504 (87)
Upper tract	264 (36)	217 (38)
Perianal	171 (24)	132 (23)
Disease Phenotype		
Non-penetrating/non-stricturing	-	446 (77)
Internal penetrating	-	25 (4)
Stricturing	-	42 (7)
Perianal fistulizing	-	84 (15)
Antibody immune responses		
ASCA IgA	269 (37)	-
ASCG IgG	263 (36)	-
anti-OmpC	130 (18)	-
Anti-CBir1	361 (52)	-
pANCA	160 (22)	-
Medication Exposure		
Early steroids	-	361 (62)
Early IM	-	299 (52)
Late IM	-	198 (34)
Early anti-TNF	-	25 (4)
Late anti-TNF	-	206 (36)

mo = months; yrs = years; IM = immunomodulators; TNF=tumor necrosis factor; ASCA = anti-saccharomyces cerevisiae antibody; OmpC=out-membrane porin C; CBir1=anti-flagellin; pANCA = perinuclear antineutrophil antibody; Early = $\leq$ 90 days from diagnosis, Late = $>$ 90 days from diagnosis

Table 3

Cox proportional analyses for the risk of a complication over 5 years

Univariate analysis			
Variable	Hazard Ratio	95% Confidence Interval	p
Patient Characteristic			
Older age at diagnosis	1.1	1.01–1.24	0.02
Female	1.8	0.99–3.34	0.05
Colonic disease	0.4	0.20–0.82	0.02
Small bowel disease	3.9	1.55–10.03	<0.001
Upper tract disease	0.8	0.39–1.44	0.37
Increasing QSS	1.3	1.12–1.46	<0.001
Positive pANCA	0.2	0.04–0.70	0.001
Treatment			
Early steroids	1.1	0.6–2.14	0.69
Early IM	1.3	0.7–2.44	0.38
Late IM	0.4	0.19–0.89	0.02
Early anti-TNF	1.3	0.30–5.20	0.76
Late anti-TNF	0.2	0.08–0.49	<0.001
Multivariate analysis			
Variable	Hazard Ratio	95% Confidence Interval	p
Older age at diagnosis	1.1	0.98–1.23	0.1
Female	1.9	1.0–3.59	0.05
Colonic or Upper tract	0.7	0.41–1.08	0.1
Small bowel or Perianal	1.9	1.12–3.21	0.02
Increasing QSS	1.3	1.13–1.54	0.001
Positive ANCA	0.2	0.05–0.91	0.04

QSS = quartile sum score; pANCA = perinuclear antineutrophil antibody; IM = immunomodulator; TNF = tumor necrosis factor; Early = $\leq$ 90 days from diagnosis, Late = $>$ 90 days from diagnosis

Table 4

Multivariate analysis for the risk of a complication as a factor of treatment and quartile sum score group

Variable	Hazard Ratio	95% Confidence Interval	p (trend)
QSS *Early Steroids			0.083
QSS group 1	1.2	0.98–1.52	
QSS group 2	1.5	1.07–2.73	
QSS group 3	1.8	1.10–4.51	
QSS group 4	2.2	1.14–7.45	
QSS *Early IM (mono)			0.048
QSS group 1	0.72	0.57–0.99	
QSS group 2	0.52	0.33–0.99	
QSS group 3	0.36	0.19–0.99	
QSS group 4	0.27	0.11–0.99	
QSS *Early anti-TNF (mono)			0.064
QSS group 1	0.16	0.0217–1.11	
QSS group 2	0.02	0.0005–1.259	
QSS group 3	0.0038	0–1.38	
QSS group 4	0.0006	0–1.54	
Early IM *Early anti-TNF (combo)			0.027
QSS group 1	14.04	0.20–1001.56	
QSS group 2	1.58	0.02–112.41	
QSS group 3	0.18	0.0025–12.61	
QSS group 4	0.02	0.0003–1.41	
Late IM	0.28	0.10–0.74	0.01
QSS *Late anti-TNF			<0.001
QSS group 1	0.58	0.42–0.78	
QSS group 2	0.33	0.18–0.62	
QSS group 3	0.19	0.08–0.49	
QSS group 4	0.11	0.03–0.38	

QSS = quartile sum score; IM=immunomodulator; TNF = tumor necrosis factor; Early = ≤ 90 days from diagnosis, Late = > 90 days from diagnosis

* denotes that QSS group is an interaction term with the medical treatment regimen