CORRESPONDENCE

Dear Editors,

We read with great interest the article by Harris et al \(^1\) reporting the Competition for Clinical Trials in Inflammatory Bowel Diseases. The authors provide an alarming overview of the situation, with a decrease in the last 20 years (1998–2018) of the average recruitment rate from 0.32 to 0.13 patients per site per month in moderate-to-severe ulcerative colitis, and from 0.65 to 0.1 in moderate to severe Crohn’s disease. They present an in-depth analysis of the causes of this scourge and propose innovative solutions to try to overcome it, as well as opening new centers in new countries. However, this last point will further spread the problem and increase the already astronomical cost of clinical trials for inflammatory bowel disease (IBD). In addition, their proposals mainly focus on the level of complexity of the studies, whereas it is well-known that there are 3 others bottlenecks in patient recruitment in clinical trials, at the level of the patient, the doctor, and the participating center.\(^2\)

As a national study group in IBD, the GETAID (Groupe Etude Therapeutique des Affections Inflammatoires Digestives) was previously threatened of bankruptcy when it was confronted to a lower than expected recruitment rate for a major trial we accepted to promote.\(^3\) Considering that, as a physician, it is very challenging to be aware of all ongoing trials on-site, to precisely know inclusion and exclusion criteria for these trials, and to find time available for recruitment tasks, we developed the CT-Scout solution, a multidevice web application that aims to optimize patient recruitment by enabling all physicians from one center, including noninvestigators, to identify potentially eligible patients. Identifying these patient, in real time, by answering a simple and short questions in just a few clicks, allows physicians to check whether the patient matches with an ongoing computed tomography (CT) scans on site and, in that case, to send a notification to the research team, which takes over the recruitment process.\(^4\) Importantly, this solution, customized per site and including potentially all active trials, provides real-time data on recruitment in each site, increasing the visibility for all stakeholders from detection to randomization.

We recently conducted a study aiming to compare the number of patients enrolled in 2 phase III clinical trials HICKORY\(^5\) and BERGAMOT\(^6\) evaluating the efficacy and safety of etrolizumab in ulcerative colitis and Crohn’s disease in sites equipped or not with CT-Scout.\(^7\) These results were presented as an oral communication during the United European Gastroenterology Week in 2019. Briefly, it was a multicenter, prospective, open-label, observational study including all sites opened for >6 months. Recruitment figures were provided by the sponsor, which considered all French sites equipped with CT-Scout and sites in other countries not equipped with CT-Scout at the time of study launch. The primary endpoint was the mean number of patients randomized per site in both studies. Secondary endpoints were the mean number of patients randomized in each study. Patients who signed informed consent form (screened) and those finally randomized were compared in sites equipped or not with CT-Scout using a 1-way analysis of variance followed by post hoc Tukey and Mann-Whitney tests. During the observational period of 40 months (September 2015 to December 2018), 644 and 289 patients were screened and randomized in 134 sites in both trials, respectively. Twenty-one sites in France were equipped with CT-Scout and were compared with the 113 sites non-equipped with the app, located in Belgium (n = 14), Germany (n = 41), Spain (n = 19), the UK (n = 26), and Israel (n = 13). There were 307 and 149 patients in 78 sites for Hickory, and 337 and 140 patients for Bergamot in 102 sites. The mean number of patients screened and randomized per site in sites equipped and nonequipped in both studies was 7.55 and 3.05 (P < .001) and 3.79 and 1.27 (P < .001), respectively. For Hickory, they were 9.17 and 3.14 (P < .001) and 5.17 and 1.28 (P < .001), respectively. For Bergamot, they were 5.94 and 2.97 (P = .003) and 2.41 and 1.26 (P = .009), respectively. The mean number of patients detected and selected with the app was 13.9 and 15.9 and 9.2 and 5.9, for Hickory and Bergamot, respectively. So, the switch from prescreened to screened patients was lower in Bergamot, raising interesting questions that need to be addressed, and in phase with the results reported by Harris et al.\(^1\) We therefore demonstrated a significant increase in patient recruitment in IBD clinical trials, with randomization rates 2–4 times higher in equipped sites compared with nonequipped ones.

We work therefore on a similar platform accessible from patients’ association web-site (Totem4me) giving to each patient the opportunity to test himself and in case of matching, to be referenced to a site having the ongoing recruitment trial using geolocation. Based on these findings, we believe that such innovative solutions should be extended worldwide to contribute to solve the challenging issue of insufficient patient recruitment in IBD and non-IBD clinical trials.

Yoram Bouhnik
Hôpital Beaujon
Gastroentérologie
MCI et Assistance Nutritive
Clichy, France

Edouard Louis
CHU Sart Tilman
Service d’hépato-gastroentérologie
Liège, Belgique

Laurent Peyrin-Biroulet
CHU de Nancy
Hôpital de Nancy

1. Harris et al. \(\text{Gastroenterology 2020;158:2306–2314}\)
2. Peyrin-Biroulet et al. \(\text{Gastroenterology 2018;156:939–941}\)
3. Peyrin-Biroulet et al. \(\text{Gastroenterology 2017;153:910–912}\)
4. Peyrin-Biroulet et al. \(\text{Gastroenterology 2016;151:1736–1738}\)
5. Peyrin-Biroulet et al. \(\text{Gastroenterology 2015;159:2306–2314}\)
6. Peyrin-Biroulet et al. \(\text{Gastroenterology 2014;167:2306–2314}\)
7. Peyrin-Biroulet et al. \(\text{Gastroenterology 2013;155:1736–1738}\)
Service d’hépato-gastro-entérologie
Nancy Cedex, France

References

Acknowledgements
The CT-SCOUT platform has been developed and commercialized by a private company, CTMA (Clinical trials Mobile Application, Asnières, France).

Conflicts of interest
The authors have made the following disclosures: Yoram Bouhnik reports honoraria from AbbVie, Biogaran, Biogen, Boehringer Ingelheim, Celgene, Ferring, Gilead, Hospira, Janssen, Mayol Spindler, MSD, Norgine, Pfizer, Roche, Samsung Bioepis, Sandoz, Sanofi, Shire, Takeda, and UCB; grants from Pfizer and Takeda; and stock options with CTMA (family members). Edouard Louis reports research grants from Takeda, Pfizer, Janssen; educational grants from Abbvie, MSD, Takeda, and Janssen; speaker fees from Abbvie, Ferring, MSD, Falk, Takeda, Hospira, Janssen, and Pfizer; advisory board membership for Abbvie, Ferring, MSD, Takeda, Celgene, Hospira, Janssen, and Pfizer; consultant for Abbvie; and stock options with CTMA. Laurent Peyrin-Biroulet reports honoraria from AbbVie, Janssen, Hospira, Janssen, and P; advisory board membership for Abbvie, Ferring, MSD, Takeda, Celgene, Hospira, Janssen, and Pfizer; consultant for Abbvie; and stock options with CTMA.

RE: Validation of a Machine Learning Model That Outperforms Clinical Risk Scoring Systems for Upper Gastrointestinal Bleeding

Dear Editors,

We read with interest the article titled “Validation of a Machine Learning Model That Outperforms Clinical Risk Scoring Systems for Upper Gastrointestinal Bleeding” by Shung et al 1 published in the January issue of the Journal. The authors concluded that the machine learning model had a greater area under the curve and had higher levels of specificity at 100% sensitivity in comparison to other risk scoring systems such as the Glasgow-Blatchford score, admission Rockall score, and AIMS65 score. However, we would like to ask the authors for clarification of the external validation set (n = 800) that is mentioned in the visual abstract and discussion.

The article explained that the dataset was separated into the training set (n = 1958) and the external validation set (n = 399). The external validation set is also known as the testing dataset, which is designed for testing the final machine learning model on an independent dataset to assess the performance, accuracy, and generalizability of the machine learning model. 2 Although the authors clearly stated in the Methods that the external validation was performed on the external validation sample of 399 cases, the visual abstract showed the performance of the machine learning model and the Glasgow-Blatchford score model tested in 800 patients. Hence, it is important to clarify whether this dataset of 800 patients represents a different dataset used for external validation than the one described in the Methods and Results. Our concern is that the external validation size of 800 patients may represent the testing dataset and a portion of the training dataset. By using the dataset that machine learning model has already learned for testing its predictive performance, the machine learning model will have almost the perfect chance of making the correct prediction, which will create a bias toward better performance.

We would respectfully ask the authors to clarify the number shown in the visual abstract and discussion. In case that the dataset used for external validation is a separate dataset, we would like to recommend using the training dataset size shown in the abstract and discussion.