CURRENT OPINION

The End of Phase 3 Clinical Trials in Biosimilars Development?

Francois-Xavier Frapaise[1]

Abstract

Most patients still have limited or no access to life-changing therapeutic proteins in the treatment of their cancer or autoimmune disorders. The current clinical development model of biosimilars is expensive, and in most cases, large, phase 3 trials do not provide meaningful information on the clinical equivalence of biosimilars and reference compounds. At the same time, the development of state-of-the-art orthogonal analytical methods has enabled a better understanding of the structure and structure–function relationship of biotherapeutics. Hence, we suggest here that a solid chemistry, manufacturing, and controls (CMC) package and meaningful phase 1 studies will leave limited uncertainty on biosimilarity, which can be addressed—if needed—by post-approval, long-term follow-up studies (post-approval studies, pharmacovigilance, real world evidence data and registries, and possibly new post-approval models to be developed). We believe that this new approach may be more appropriate than 600- to 1000-patient, phase 3 trials in assessing biosimilarity and therapeutic equivalence, under the condition that the administered biosimilar given to individual patients can be clearly identified. Obviously, there will probably never be a "one size fits all" development model, and an individualized, risk-based approach to biosimilar development will always have to be considered and discussed early with regulators.

Key Points

Significant progresses in understanding the structure–function relationship of therapeutic proteins is likely to enable a different, more informative development approach for biosimilars.

As large and expensive phase 3 trials have not shown their ability to detect clinical differences between biosimilars and originators, it is fathomable that in the coming years, a stronger chemistry, manufacturing, and controls (CMC)/phase 1 package together with meaningful post-approval studies could replace the current development paradigm that is based on large, phase 3 studies.

1 Introduction

In his book, *The End of History and the Last Man*, published by the National Interest in 1992, Fukuyama [1] argues that the advent of Western liberal democracy may signal the endpoint of humanity's sociocultural evolution and the final form of human government; in other words, "the end of history". In the same vein, based on the recent development of sophisticated analytical methods and evolution of regulatory guidance, one should discuss the relevance of the current clinical development model for biosimilars and raise the following question: are we close to the end of history for large, phase 3 clinical trials, or, more realistically, could we not drastically reduce the financial and operational burden of biosimilar clinical trials in the foreseeable future?

Although we have recently come a long way in understanding biosimilar development and how to analyze biosimilars, as Fleischmann [2] states (concerning biosimilars), it is "not all white; there is still some grey and black." Indeed, proteins have unique, structural, organizational patterns and critical quality attributes (CQAs) that define their purity, potency, and safety.

Indeed, even biotherapeutics with identical amino acid sequences may have different biological effects due to structural folding differences, without even mentioning the effect

✉ Francois-Xavier Frapaise
  xavierfrapaise@yahoo.com

1   Paris, France

of post-translational modifications (PTMs). Thanks to the exponential development of multiple sophisticated analytical methods (so-called state-of-the-art orthogonal methods), enabling comparability assessment between originators and candidate biosimilars, both structurally and functionally, and considering the lack of sensitivity of many clinical models to detect meaningful differences between follow-on biologics and reference compounds and the recent evolution of the regulators' perspective on this matter, the relevance of the current clinical development model could be legitimately questioned. This question has actually already been raised in 2014 by McCamish and Woollett in their article "The Continuum of Comparability Extends to Biosimilarity: How Much is Enough and What Clinical Data are Necessary?" [3].

Obviously, one cannot ignore the difference between small proteins such as insulin, low molecular weight heparin, and peg-filgrastim (usually not requiring phase 3 clinical data for approval, and for which pharmacodynamic [PD] markers are available) and large, complex biotherapeutics such as monoclonal antibodies (mAbs), for which large, phase 3 trials are still deemed necessary.

2 Reference Compound or Reference Compounds?

One of the most frequent misconceptions among clinicians around biosimilars is the belief that reference compounds never change, when multiple reports have shown the contrary. Schneider [4] demonstrated in his 2013 paper that the number of changes in the manufacturing process for European Medicines Agency (EMA)-approved biologicals used in rheumatology exceeds 35 for Remicade®, 20 for Enbrel®, and 15 for Humira®.

These changes may impact the biological activities of these reference compounds. As mentioned by Lamanna et al. [5], different lots of Aranesp® conveyed varying charges resulting from varying numbers of sialic acids per molecule and different lots of Rituxan/Mabthera® and Enbrel® showed changes in N- and C-terminal heterogeneity (the former also showed variation in antibody-dependent, cell-mediated cytotoxicity [ADCC] activity among batches). For biosimilars, the crucial step is the acquisition of a sufficient number of unique originator drug product lots to measure the variability in the originator drug manufacturing process and provide sufficient statistical power for the analytical data comparisons. The situation is made more complex for biosimilar manufacturers when the reference compound formulation and/or dosage strength changes during the course of the clinical development of a biosimilar. This obviously represents a significant challenge for biosimilar developers, as they are shooting at a moving target. As George Bernard Shaw once stated, "the only man who behaved sensibly was my tailor; he took my measure anew every time he saw me, while all the rest went on with their old measurements and expected them to fit me" [6].

3 The Recent Development of Sophisticated Analytical Methods

Several techniques are now available to characterize the structure and biological activities of biosimilars, which enables the determination of the protein sequence, the identification and quantification of enzymatic and non-enzymatic PTMs, the analysis of biological functionality, and the establishment of the variability in product quality attributes (QAs).

As mentioned by Dipaola [7], the primary structure of proteins and possible PTMs can reliably be characterized by different methods such as peptide mapping by liquid chromatography–mass spectrometry (LC–MS), electrospray mass ionization–mass spectrometry (ESMI-MS), and N-terminal sequence by automated Edman degradation. The bioactivity of these proteins can be assessed by measuring cell uptake/proliferation, as well as cytotoxicity: ADCC, antibody-dependent cellular phagocytosis (ADCP), and complement-dependent cytotoxicity (CDC). Many different techniques enable the assessment of target binding: radioligand binding, binding analysis by surface plasmon resonance, binding analysis by fluorescence anisotropy immunoassays, and radioimmunoassay (RIA).

Protein content/concentration can readily be measured by absorbance at 280 nm, colorimetric assays, chromatography (reverse-phase, high-performance liquid chromatography [RP-HPLC], and size-exclusion chromatography [SEC]).

The presence of high molecular weight (HMW) species and/or aggregates can be detected by SEC-HPLC, size-exclusion chromatography–multi-angle light scattering (SEC-MALS), native MS sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) (reduced and non-reduced), capillary electrophoresis (CE)-SDS (reduced and non-reduced), analytical ultracentrifugation, field-flow fractionation, and dynamic light scattering, just to name a few.

Higher order structures can be assessed by different methods, such as far and near ultraviolet (UV) circular dichroism, Fourier-transform infrared spectroscopy, antibody array, intrinsic/extrinsic fluorescence, differential scanning calorimetry, hydrogen/deuterium exchange MS, ion-mobility MS, 1D and 2D nuclear magnetic resonance (NMR), and X-ray crystallography. State-of-the-art methods are also available to measure charge distribution, non-enzymatic

PTMs, N- or C-terminal truncation, glycosylation, and process-related impurities.

In parallel, statistical approaches have been recommended by Tsong et al. [8] for analytical similarity assessment; namely, an equivalence test for the CQAs in tier 1, a quality range approach for QAs in tier 2, and a side-by-side graphic comparison approach for QAs in tier 3. The September 2017 draft guidance issued by the Food and Drug Administration (FDA), "Statistical Approaches to Evaluate Analytical Similarity—Guidance for Industry," should be taken into account in biosimilar development strategies [9]. As this is a fast evolving topic and still a subject of debate, discussions with regulators should take place as early as possible to ensure this does not become a major review issue.

Interestingly, it has been recently suggested by Zeng et al. [10] that one could improve the power to establish clinical similarity in a phase 3 efficacy trial by incorporating prior evidence of analytical and pharmacokinetic (PK) similarity. This approach, according to the authors, would enable a power increase of clinical investigations meant to eliminate possible remaining uncertainties. In other words, one could decrease the sample size of confirmatory phase 3 trials by generating more analytical and pharmacological data on the products to be compared for similarity assessment.

4 A Better Understanding of the Structure–Activity Relationship

Not only are sophisticated analytical methods now available, but the relevance of their possible analytical differences is also much better understood; many recent investigations have indeed enabled a better understanding of the structure–activity relationships of biotherapeutics. Thus, it has been shown by Liu [11] that the glycan profile can have a significant impact on the PK and PD of mAbs. Mannosylated glycans can impact the PK of the molecule, leading to reduced exposure and potentially lower efficacy. The level of the sialic acid, N-acetylneuraminic acid (NANA), can also have a significant impact on the PK of Fc-fusion molecules. Core fucose in the glycan structure reduces immunoglobulin G (IgG) antibody binding to IgG Fc receptor IIIa relative to IgG lacking fucose, potentially resulting in decreased ADCC activities.

A recent article by Lee et al. [12] also shows that the glycosylation profiles and biological activities of several infliximab-based compounds (Remicade®, Remsima®, and Flixabi®) are not, indeed, identical. However, key parameters, namely the Fab-related activity of the three compounds, were highly similar, even though, compared with Remicade®, Flixabi® had a lower percentage of charged glycan, and Remsima® had a higher percentage of galactosylated glycan and a lower percentage of afucose plus a higher percentage of mannosylated glycans. However, it is acknowledged that where a difference in biological activity (in vitro) is detected (as a result of a difference in, for example, glycan structures), it is not always clear which levels of difference in biological activity are clinically meaningful and which are not.

The relationship between the relative content of HMW species detected by the SEC method and the incidence of anti-drug antibodies (ADAs) reported from clinical trials of Remsima® has been evaluated. From analyses using the SEC method, the HMW level of the infliximab biosimilar was higher than that of the reference product. However, in the clinical trials of the infliximab biosimilar, the ADA incidence was comparable to that of the reference compound. However, for the infliximab biosimilar Flixabi®, the ADA incidence was higher than that for Remicade® in both a phase 1 volunteer study and a phase 3 randomized controlled trial in rheumatoid arthritis (RA) patients.

Terminal Gal in a mAs is important in terms of the CDC in that lower levels of Gal reduce CDC activity. The glycan profile can also impact the safety of mAbs.

5 Much Can Be Learnt from Well-Designed Phase 1 Trials

Comparative phase 1 PK/PD studies, designed to demonstrate that the biosimilar has a similar PK profile to the reference product with regard to key PK parameters, are an essential part of the biosimilar development program. A range of 80–125% is generally used to demonstrate equivalence at the 90% confidence level for PK/PD evaluations. The relevance of phase 1 trials is obviously stronger when reliable biomarkers are available, which is the case for peg-filgrastim, though, unfortunately, it is not the case with anti–tumor necrosis factors (anti-TNFs) and anti–vascular endothelial growth factors (anti-VEGFs). These phase 1 trials can also address the issue of immunogenicity, a key concern for immunogenic biotherapeutics. Biosimilar bioequivalence studies are generally large in size, expose volunteers to the tested compounds for up to a 1 year (e.g., peg-filgrastim studies), and can involve up to hundreds of subjects, depending on the tested molecule.

Phase 1 studies usually enroll healthy subjects, comprising a homogeneous population of immune-competent subjects not receiving concomitant medications, thus allowing a sensitive comparison of PK and immunogenicity. Some variability, be it in the biosimilar or reference product, is expected between different batches of product. Such variability can be detrimental to the outcome of a phase 1 PK/PD study; indeed, differences can even be observed in the PK profiles of reference compounds sourced from the USA

and Europe (based on the author's own experience in developing adalimumab biosimilars).

6 Evolution of the Regulatory Landscape

Those who regularly interact with key regulatory agencies and follow the evolution of international guidelines [13–16] have observed new trends that are compatible with the long-term objective suggested by this article's title. These are as follows, at least for smaller therapeutic proteins: no need to perform animal studies, and healthy volunteer PK/PD studies alone may be relied on to support the registration of some biosimilar compounds (without testing the candidate biosimilar in patients); wider acceptance by the FDA/EMA of extrapolation from one indication to all other approved indications of the reference compound (e.g., infliximab); issuance of guidance on interchangeability assessment; acceptance, in some cases, of asymmetric margins for bioequivalence testing; and agencies' willingness, if not encouragement, to test clinical biosimilarity in indications that were not approved for the originator compound. One should refer to the recently issued FDA draft guidance "Considerations in Demonstrating Interchangeability with a Reference Product: Guidance for Industry" and the recent "Reflection Paper on Statistical Methodology for the Comparative Assessment of Quality Attributes in Drug Development" of the EMA [16–18].

7 The Evolution of Clinicians' Perspectives on Biosimilars

When the first anti-TNF biosimilars were developed in RA and ankylosing spondylitis, many gastroenterologists were very reluctant to extrapolate clinical data obtained in rheumatology and prescribe these biosimilars to patients with inflammatory bowel disease (IBD), with the European Crohn's and Colitis Organisation (ECCO) stating that, "Switching from an established biologic to a biosimilar to save costs is likely to be as inappropriate and ineffective as switching between current biologics that act on the same target, except when there is loss of response" [19]. This ECCO position was immediately challenged by the European regulators. A survey of 307 ECCO members in 2014 [20] showed that IBD specialists were reasonably informed on biologic agents, regarded cost sparing (89%) as the main advantage, and listed immunogenicity (67%) as their main concern. Sixty-four percent disagreed with automatic replacement of originator biologic agents with a biosimilar by a pharmacist, although 18% supported substitution for new prescriptions, and only 6% felt that

biosimilars were interchangeable. However, with the accumulation of evidence of efficacy and safety of anti-TNF biosimilars in IBD patients, as reported by Avila-Ribeiro et al. [21] (23 observational studies, 12 of them assessing switching from infliximab originator to a biosimilar and 17 assessing induction therapy with infliximab biosimilar), the position of most gastroenterologists has significantly evolved, and the ECCO has recently revised their position and provided the following statements [22]:

> Biosimilarity is more sensitively characterized by performing suitable in vitro assays than clinical studies. Clinical studies of equivalence in the most sensitive indication can provide the basis for extrapolation. Therefore data for the usage of biosimilars in IBD can be extrapolated from another sensitive indication. Demonstration of safety of biosimilars requires large observational studies with long-term follow up in IBD patients.
> Adverse events and loss of response due to immunogenicity to a biologic drug cannot be expected to be overcome with a biosimilar of the same molecule.
> As for all biologics, traceability should be based on a robust pharmacovigilance system and the manufacturing risk management plan.
> Switching from the originator to a biosimilar in patients with IBD is acceptable. Studies of switching can provide valuable evidence for safety and efficacy. Scientific and clinical evidence is lacking regarding reverse switching, multiple switching, and cross-switching among biosimilars in IBD patients.

This does not mean that an important education effort should not be undertaken by biosimilar manufacturers and independent institutions, especially when one considers that many patients in 2018 are still reluctant to accept generic prescriptions.

8 The Limitations and Challenges of the Current Biosimilar Phase 3 Trials Paradigm

Biosimilar phase 3 clinical trials have multiple limitations and challenges. First, they are not powered to detect meaningful differences in the safety profiles of biosimilars, and when numerical imbalances in adverse events are observed during clinical development of a biosimilar, the interpretation of limited differences is very difficult; only large cohort studies may possibly detect differences, if there are any, in safety parameters. The probability of detecting differences in efficacy is also very limited; many phase 3 biosimilar trials have enrolled 500–800 patients, and none of them have detected significant differences in efficacy parameters, even

when differences in biological parameters such as ADCC, considered by some regulators as a predictor of efficacy of anti-TNF biologics, were present. It is also worth noting that the chances of observing differences in immunogenicity in RA patients, many of them receiving methotrexate—an immunosuppressant—are, at best, remote. Hence, we believe that a different approach to confirm similarity observed in chemistry, manufacturing, and controls (CMC)/PK data would consist of undertaking well-designed, randomized, post-marketing studies and generating "real world evidence" data.

It is worth mentioning that the 90 switching studies recently analyzed by Cohen et al. in *Drugs* [23], enrolling 14,225 patients, failed to show any meaningful clinical effect of single switches. In this respect, the recently presented post-approval NOR-SWITCH study [24], co-funded by the Norwegian government, a 52-week, randomized, double-blind, non-inferiority, phase 4 trial in which patients with six different diagnoses were randomly assigned to receive an originator or biosimilar medication, showed similar treatment-emergent events and similar immunogenicity with and without switching, as well as comparable long-term safety and efficacy of the biosimilar after switching from the originator. The large discounts offered by biosimilar companies in Norway has enabled rapid adoption of infliximab biosimilars and is generating massive real world evidence data that further convinced regulators, clinicians, and patients that it safe and very cost-effective to switch patients from originator compound to biosimilars.

This is in line with Paramsothy et al.'s conclusion on the use of biosimilars in the treatment of IBD [25, 26]: "the data currently available are positive in regard to the bioequivalence of these agents in the de novo setting, although interchangeability has not been adequately established…. Ongoing post marketing studies are essential to clearly define the safety, efficacy, and immunogenicity profiles of biosimilar agents in IBD…"

The need for post-approval trials/registries/pharmacovigilance activities will have to be discussed with regulators, particularly in case there is "remaining uncertainty on biosimilarity" based on CMC/phase 1 data. These post-approval commitments are actually part of the usual regulatory process for all new chemical and biological entities.

9 The Financial Burden of Phase 3 Biosimilar Clinical Trials

Figure 1 is often used to illustrate the difference between originator and biosimilar development pathways; it is, however, relatively misleading, at least in terms of financing, as the clinical phase of biosimilar development still represents the lion's share of necessary investments.
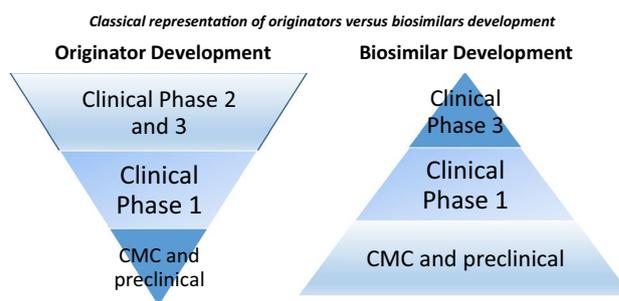

Classical representation of originators versus biosimilars development

Fig. 1 The biosimilar development paradigm shift

The time requirement and the financial burden of running phase 3 clinical trials (at least US$50 million for a 600-patient trial) is indeed a barrier, if not a deterrent, for many mid-size biosimilar companies (some of whom may have developed quasi-fingerprint biosimilars) to develop and commercialize their compounds. The burden of running such trials is made worse by the cost of sourcing reference compounds and, at times, the extreme difficulties in sourcing these from the originator. Obtaining large quantities of a single batch of reference product can actually prove to be extremely challenging. Hence, it represents an obstacle to wider biosimilar adoption and the reduction of the healthcare burden. Other challenges should be mentioned, e.g., the understandable request by many international review boards, investigators, and patients to provide patients with the biosimilar—free of charge—at the end of the trial, sometimes for as long as the compound is not approved. Hence, considering the financial and operational burden of running phase 3 pivotal trials based on clinical endpoints and the likely acceptability of surrogate markers, one can predict that clinical trial designs and endpoints for biosimilars are likely to significantly evolve in the coming years.

10 Conclusion

One cannot ignore the difference between small proteins, such as insulin, low molecular weight heparin, and pegfilgrastim (usually not requiring phase 3 clinical data for approval) and the complex, large biotherapeutics such as mAbs for which the jury is still out. While the concept of no longer performing large, phase 3 clinical trials may not be quite ready for prime time, the scientific community, most clinicians, the regulators, and the biosimilar industry now understand how to best develop, assess, and predict the performance of biosimilars. Newer, sophisticated, state-of-the-art orthogonal methods have largely enabled this evolution. Hence, we believe it is now time to re-assess whether the current clinical development paradigm really makes sense from a scientific and economic perspective while millions

of patients still have limited or, more often, no access to life-changing/life-saving mAbs. As stated by Low [27], "if we don't figure out a way to make this work for more folks, we are going to end up in a situation where we have miracles to offer, but no way to pay for them." Therefore, we suggest here that in many instances, a solid CMC package and meaningful phase 1 studies and—in case there are remaining uncertainties on biosimilarity—well-designed, post-approval, long-term, follow-up studies (pharmacovigilance studies, real world evidence data and registries) could be more appropriate than 600-patient, phase 3 trials in "eliminating residual uncertainties on biosimilarity and therapeutic equivalence," under the condition that administered biosimilars, after commercialization, can be clearly identified. Obviously, regular consultations with regulators are essential all along the biosimilar development process.

Compliance with Ethical Standards

Funding No funding was received for the preparation of this review.

Conflict of interest Dr Frapaise has no conflicts of interest.

References

1. Fukuyama F. The End of History and the Last Man. Perennial, 1992.
2. Fleischmann R. Editorial: The American College of Rheumatology White Paper on Biosimilars: it isn't all white-there is some gray and black. Arthritis Rheumatol. 2018;70(3):323–5.
3. McCamish M, Woollett G. The continuum of comparability extends to biosimilarity: how much is enough and what clinical data are necessary? Clin Pharmacol Ther. 2013;93(4):315–7.
4. Schneider CK. Biosimilars in rheumatology: the wind of change. Ann Rheum Dis. 2013;(72)3:315–318.
5. Lamanna WC, et al. The structure-function relationship of disulfide bonds in etanercept. Sci Rep. 2017;7:3951. https://doi.org/10.1038/s41598-017-04320-5.
6. George Bernard Shaw, Laurence DH. Man and Superman. UK: Penguin Adult; 2000. ISBN 0140437886
7. Dipaola M. Analytical strategy in the development of biosimilars. BioPharm Int. 2017;30(8):38–43, 46.
8. Tsong Y, et al. Development of statistical methods for analytical similarity assessment. J Biopharm Stat. 2016;27:197–205.
9. FDA. US Department of Health and Human Services. CDER/CBER, Biosimilars; 2017. https://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm576786.pdf.
10. Zeng D, et al. Improving the power to establish clinical similarity in a phase 3 efficacy trial by incorporating prior evidence of analytical and pharmacokinetic similarity. J Biopharm Stat. 2017;28(2):320–32.
11. Liu L. Antibody glycosylation and its impact on the pharmacokinetics and pharmacodynamics of monoclonal antibodies and Fc-fusion proteins. J Pharm Sci. 2015;104(6):1866–84.
12. Lee C, et al. Glycosylation profile and biological activity of Remicade® compared with Flixabi® and Remsima®. MAbs., U.S. National Library of Medicine, 2017;(6):968–77.
13. Committee for Medicinal Products for Human Use. (CHMP) guideline on data monitoring committees. Stat Med. 2006;25(10):1639–45.
14. FDA releases program for approval of biosimilars. draft guidance for industry from FDA: scientific considerations in demonstrating biosimilarity to a reference product: draft guidance for industry from FDA: quality considerations in demonstrating biosimilarity to a reference protein product. Biotechnol Law Rep. 2012;31(2):142–3.
15. FDA releases program for approval of biosimilars. Document: draft guidance for industry from FDA: scientific considerations in demonstrating biosimilarity to a reference product. Document: Draft guidance for industry from FDA: quality considerations in demonstrating biosimilarity to a reference protein product. Biotechnol Law Rep. 2012;31(2):142–3.
16. European Medicines Agency. Reflection paper on statistical methodology for the comparative assessment of quality attributes in drug development. EMA, 2017.
17. FDA. US Department for Health and Human Services. CDER/CBER. Biosimilars; 2017. https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM537135.pdf
18. EMA/CHMP/138502/2017. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2017/03/WC500224995.pdf
19. Silvio D, Gomollon F. ECCO position statement: the use of biosimilar medicines in the treatment of inflammatory bowel disease (IBD). J Crohn's Colitis. 2013;(7)7:586–9.
20. ECCO survey highlights lack of confidence in biosimilar MAbs. Posted 22/08/2014.
21. Avila-Ribeiro P, et al. The experience with biosimilars of infliximab in inflammatory bowel disease. Curr Pharm Des. 2018;23(44):6759–69.
22. Danese et al. ECCO Position statement on the use of biosimilars for inflammatory bowel disease—an update. J Crohn's Colitis, Oxford Academic. OUP Academic, Oxford University Press, 2016.
23. Cohen, Hillel P, et al. Switching reference medicines to biosimilars: a systematic literature review of clinical outcomes. Drugs. 2018; 78(4):463–7.
24. Biosimilar infliximab (CT-P13) is not inferior to originator infliximab: results from a 52-week randomized switch trial in Norway. ACR Meet Abstr. 2016.
25. Paramsothy S, Cleveland NK, Zmeter N, Rubin DT. The role of biosimilars in inflammatory bowel disease. Gastroenterology Hepatology. 2016;12(12):741–51.
26. Fiocchi C. Inflammatory bowel disease pathogenesis: where are we? J Gastroenterol Hepatol. 2015;30:12–8.
27. Low M. The next drugs: the future for biosimilars an Atlantic policy briefing. The Atlantic: Atlantic Media Company; 2017.