A Data Driven Approach to Support Tailored Clinical Programs for Biosimilar Monoclonal Antibodies

Elena Guillen[1,2,*], Niklas Ekman[3], Sean Barry[4], Martina Weise[5] and Elena Wolff-Holz[6]

Biosimilar monoclonal antibodies (mAbs) have been approved in the European Union since 2013 and have been demonstrated to reduce healthcare costs and to expand patient access. Biosimilarity is mainly established on the basis of demonstrated similarity of relevant quality attributes (QAs), determined by comprehensive physiochemical and functional analyses, and demonstration of bioequivalence. In addition, comparative efficacy/safety studies have been requested for all approved biosimilar mAbs so far, although the European Medicines Agency (EMA) Guidelines state that such confirmatory clinical trials may not be necessary in specific circumstances. In order to evaluate the degree of analytical similarity, how residual uncertainty regarding biosimilarity was resolved, and the value of clinical data, we analyzed the quality and clinical data packages for authorized adalimumab (7 products) and bevacizumab (5 products) biosimilars. The percentage of biosimilar batches meeting the similarity range for QAs, as defined by the biosimilar manufacturer based on a comprehensive characterization of the EU reference product (RP), was determined and clinical data were reviewed. Our analyses show that QAs of approved adalimumab and bevacizumab biosimilars have varying concordance with the EU-RP similarity range. In this study, we found that clinical efficacy data played a limited role in addressing quality concerns. Therefore, we encourage a regulatory review of the standards for clinical data requirements for mAb and fusion protein biosimilars. This study outlines a quality data driven approach for facilitating tailored clinical programs for biosimilars.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

☑ Biosimilarity of monoclonal biosimilar antibodies (mAbs) is established on the basis of analytical and functional (quality) similarity and demonstration of bioequivalence. In the European Union, marketing authorization approval also requires a confirmatory comparative efficacy/safety study. Whereas the European Medicines Agency (EMA) Guideline states that this confirmatory clinical trial may not be necessary in specific circumstances, to date, all approved mAbs have included one large equivalence trial.

WHAT QUESTION DID THIS STUDY ADDRESS?

☑ How similar are biosimilar mAbs compared to their respective reference products? If there is < 100% of biosimilar batches within the reference similarity range, how do regulators decide whether the product may be viewed as biosimilar? What role does data from clinical efficacy trials play in reaching conclusion on biosimilarity?

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?

☑ Over 90% (mostly 100%) of the batches of adalimumab and bevacizumab biosimilars met the EU reference product similarity range for critical quality attributes (QAs). For critical QAs where < 100% of batches were inside the similarity ranges, further evidence of similarity was gathered and queries resolved at the quality level.

HOW MIGHT THIS CHANGE CLINICAL PHARMACOLOGY OR TRANSLATIONAL SCIENCE?

☑ This study shows that comparable clinical performance, in most cases, can be predicted on the basis of quality and clinical pharmacokinetic data as analyzed for two substances (adalimumab and bevacizumab). This outlines a quality data driven approach for facilitating tailored clinical programs for biosimilars.

Biosimilars are biological medicinal products that contain a highly similar version of the active substance of an already authorized original biological product (reference product (RP)). They differ from generic drugs due to their biological source, in the size of the active substance, their complexity, and the nature of the manufacturing process. The mainstay of any biosimilar development is the comprehensive

[1]Department of Pharmacology, Therapeutics and Toxicology, Universitat Autònoma de Barcelona, Bellaterra, Spain; [2]Clinical Pharmacology Service, University Hospital Vall d'Hebron, Barcelona, Spain; [3]Finnish Medicines Agency Fimea, Helsinki, Finland; [4]Health Products Regulatory Authority, Dublin, Ireland; [5]Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM), Bonn, Germany; [6]Paul-Ehrlich-Institut PEI Str 51-58, Langen, Germany. *Correspondence: Elena Wolff-Holz (elena.wolff-holz@pei.de)

Received August 31, 2022; accepted October 23, 2022. doi:10.1002/cpt.2785

demonstration of close physicochemical and functional similarity as well as bioequivalence with their RP. In addition, comparative efficacy/safety studies have so far been requested for all approved biosimilar monoclonal antibodies (mAbs) to confirm the absence of clinically meaningful differences compared with the RP.[1,2] Typically, confirmation of comparable clinical efficacy in one "model" indication is required and other indications of the RP can be extrapolated,[3] which leads to reduced development costs[4] and allows for competitive price reductions for biosimilars thus facilitating patient access.[5]

In this paper, the terms comparability and (bio-)similarity exercise are used synonymously.

The biosimilar regulatory framework was initially developed with the conservative stance that one comparative efficacy study will always be required as a safeguard and precautionary measure to ensure that biosimilarity demonstrated at the analytical and functional (quality) level indeed translates into biosimilarity at the clinical level. However, in recent years, due to the advancement in the analytical sciences and the vast experience gained, the extent and usefulness of this clinical confirmation has been questioned, and regulators have started to adopt a more flexible approach where the extent of clinical data required can vary depending on the product class.[6–9] Recent guidelines[10–13] state that a pre-licensing efficacy study may be waived in case biosimilarity can be convincingly concluded based on physicochemical and functional characterization studies using sensitive, orthogonal, and state-of-the-art analytical methods, together with comparison of the pharmacokinetic (PK) and/or pharmacodynamic (PD) profiles of the biosimilar and the RP.

In the particular case of mAbs, this has been considered challenging, given their relative complexity.[14] However, since assessment and EU marketing authorization (MA) of the first biosimilar mAb in 2013,[15] the physiochemical and functional assays have continued to evolve, with greater understanding of the relevant quality attributes (QAs) and increased sensitivity of analytical methods to detect relevant differences.[16,17] Therefore, the extent of analytical data routinely provided in biosimilar dossiers currently may give sufficient assurance that a biosimilar is indeed highly similar to the RP such that no difference in clinical performance is expected. As such, the default requirement for confirmatory clinical studies could be questioned also for mAbs, and alternative regulatory pathways and/or guidance may be warranted. Streamlining developments to become more cost and time efficient and sparing patients from entering unnecessary and redundant clinical trials is of foremost importance from an ethical point of view and at a time when public and patient resources are becoming increasingly strained.[18,19]

In an effort to provide a deeper understanding of the magnitude and strength of the analytical and functional similarity data available for mAbs, data of two biosimilar mAb classes were analyzed: seven approved adalimumab and five approved bevacizumab biosimilars. In addition, clinical efficacy and safety comparability data were reviewed on a product basis by studying the European Product Assessment Reports (EPARs).

The aim of this study was to analyze whether, and to which degree, QAs were within the similarity ranges established by the biosimilar developer based on a comprehensive characterization of the RP and what role clinical data played in the final conclusion of biosimilarity.

METHODS

We reviewed, categorized, and anonymized the analytical and functional similarity data and analyzed the clinical data packages for approved adalimumab and bevacizumab biosimilars. The analysis included seven adalimumab biosimilars (Amgevita/Solymbic, Imraldi, Hyrimoz/Halimatoz/Hefiya, Hulio, Idacio/Kromeya, Amsparity, and Yuflyma)[20–26] and five bevacizumab biosimilars (Mvasi, Zirabev, Aybintio/Onbevzi, Alymsys/Oyavas, and Abevmy/Lextemy).[27–31] The data lock point for the analysis was September 2021. We only included biosimilars that were authorized at the time of study analysis. Adalimumab and bevacizumab were selected as representative examples of widely used biologicals which cover different therapeutic areas (autoimmune and oncologic indications).

Comparison of analytical biosimilarity across products

QA characterization data were extracted from raw data of the biosimilar product dossiers submitted to the European Medicines Agency (EMA) for MA approval. The data were anonymized due to confidentiality.

The comparative QA data extracted were categorized into four pattern and color-coded categories: depending on the degree of similarity with the RP (see **Table 1**). This categorization was performed considering the percentage of analyzed biosimilar batches with values within the similarity range: solid dark green for QAs with 100% biosimilar batches within the similarity range; light green horizontal stripes for QAs with 90–99%, light blue diagonal stripes for QAs with 50–89%, and dark blue dots for QAs with < 50% of the batches within the similarity range or when the data was lacking. This crude categorization was chosen by the authors to allow for meaningful differentiation of similarity ranges, without losing the anonymity of products.

Some assays are product specific (e.g., human umbilical vein endothelial cells (HUVECs) antiproliferation and human vascular endothelial growth factor (VEGF) binding for bevacizumab) and are therefore represented in the gray grid in **Table 1**.

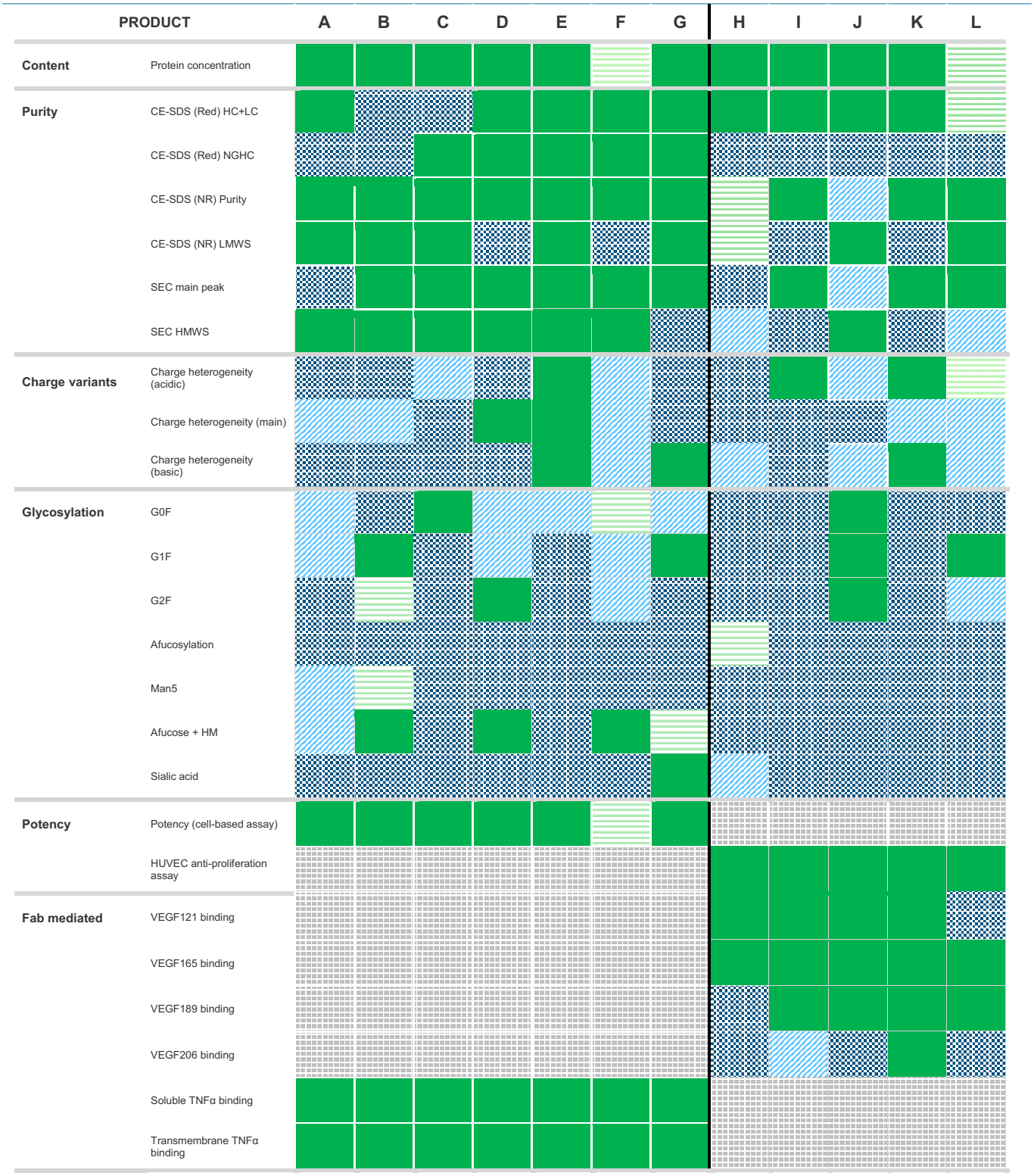
The reference (similarity) range for establishing analytical similarity is determined by the biosimilar manufacturer based on characterization data of the RP. Similarity ranges are usually calculated based on statistical analysis of the RP dataset and may be based on ranges such as mean ± 3 × SD (standard deviation), tolerance intervals, or a minimum-maximum range.[32] The approach for setting similarity ranges may vary between products, however, all the statistical approaches used were individually justified and assessed during the respective MA procedures.

Batch results outside the similarity range were conservatively counted as being "non-similar" regardless of how far outside of the similarity range the results were. The number of biosimilar batches analyzed per product varied between 8 and 20, for most QAs. **Table 1** includes mainly the analysis of quantitative QAs. In some cases, QAs were presented graphically in the MA, if the profiles were considered to be similar (i.e., the profiles of the biosimilar and RP overlap and are visually comparable), in addition, these were categorized in solid dark green.

For purity/impurity-related QAs, one-sided similarity ranges were considered, that is, if the biosimilar exhibited higher level of purity/lower level of impurities compared to the RP, this was considered *de facto* to be comparable. In such cases, 100% biosimilar batches were considered to be within the similarity range (=solid dark green).

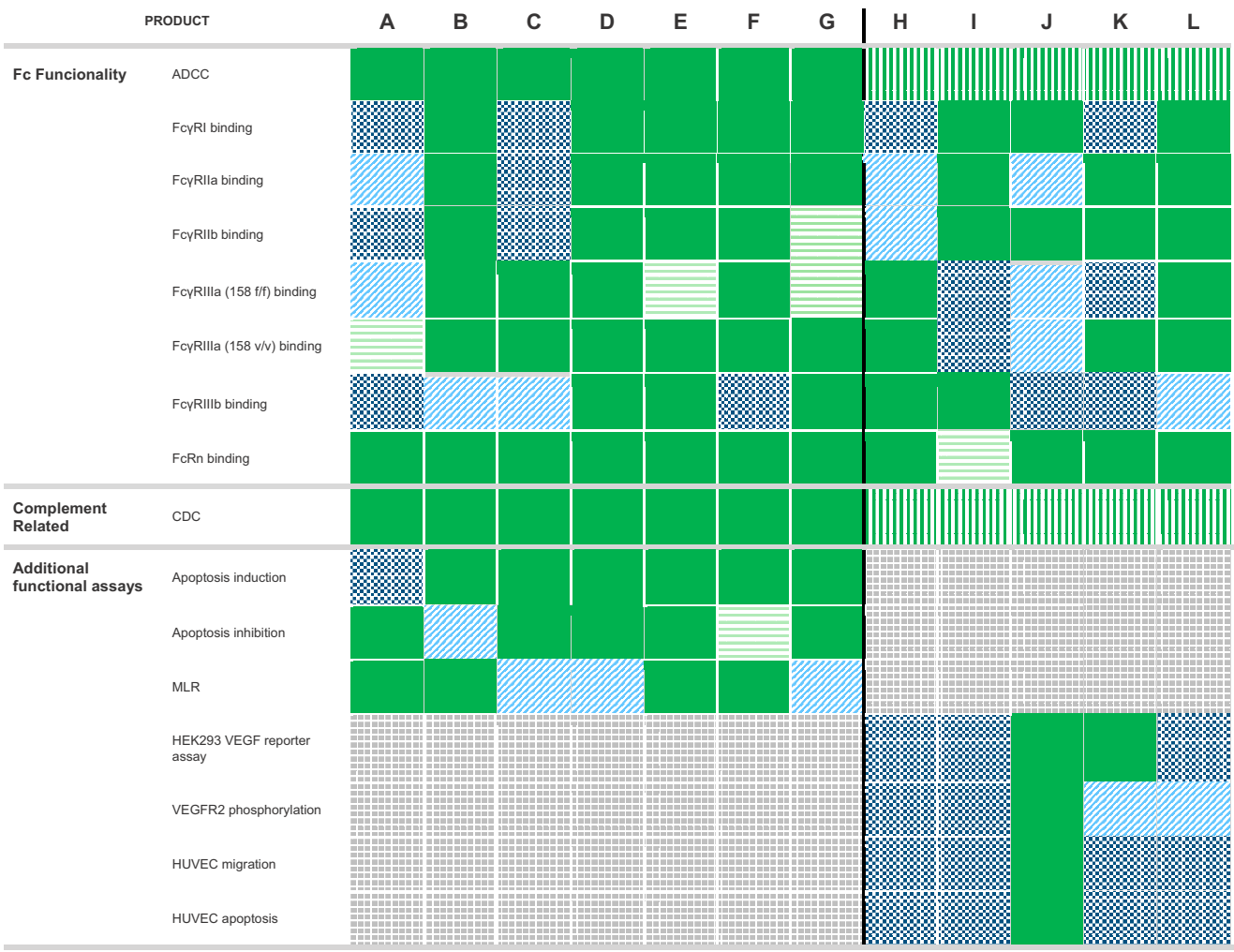
Additional QAs, for example, amino acid sequence, secondary and higher order structure, etc., are not included in this analysis because, in many cases, the data submitted were not entirely quantitative. For other QAs related to protein modifications, such as oxidation or deamidation, a quantitative comparison across products was also not possible due to different methodologies used by applicants. **Tables 2** and **3** summarize all the QAs tested, including also qualitative tests (information extracted from the EPARs).

Table 1 Similarity of QAs for all adalimumab (A–G) and bevacizumab (H–L) biosimilars. Color and patterns indicate the percentage of biosimilar batches within the similarity range derived from the EU reference product: solid dark green for 100%, horizontal light green stripes for 99–90%, diagonal light blue stripes for 89–50%, dark blue dots for < 50% and also when the QA was not assessed. Gray grid represents product specific QAs which reflect the mAbs main MoA. Green vertical stripes represent QAs that were tested but not found (in line with the mAbs MoA).



(Continued)

Table 1 (Continued)



ADCC, antibody-dependent cell-mediated cytotoxicity; CDC, complement dependent cytotoxicity; CE-SDS, capillary electrophoresis-sodium dodecyl sulfate; EU-RP, European reference product; HEK293, human embryonic kidney 293 cells; HMWS, high molecular weight species; HUVEC, Human Umbilical Vascular Endothelial Cell; Fc, fragment crystallizable region; Fab, fragment antigen-binding region; LMWS, low molecular weight species; mAb, monoclonal antibody; MLR, mixed lymphocyte reaction; MoA, mechanism of action; NR, non reduced; NGHC, non glycosylated heavy chain; QA, quality attribute; Red, reduced; SEC, size exclusion chromatography; VEGF, vascular endothelial growth factor; VEGFR2, vascular endothelial growth factor receptor 2.

Table 4 provides a summary of the instances where < 100% of batches were within the reference range (data on file at the EMA), and how the resulting uncertainty was resolved. In each case, the reason why these differences were accepted by the EMA is explained.

Comparison of clinical biosimilarity across products

Clinical data are presented as raw data in **Tables S1** and **S2** (the product columns are not in the same order as **Table 1** to maintain anonymity). Clinical data were extracted from EPARs, which contain public information that can be found on the EMA website and therefore anonymization is not necessary (European Medicines Agency. Find medicines. Available at: https://www.ema.europa.eu/en/medicines. Accessed April 2022).

For PK and efficacy parameters, acceptance ranges for comparability were defined before study start in the statistical analysis plan. Population PK (PopPK) analysis in patients was not model-based but descriptive. Safety and immunogenicity parameters are presented as raw data and were compared descriptively.

The few instances where uncertainties arose in the similarity of a specific clinical parameter are highlighted and discussed in context of other findings.

Table 5 provides a summary of all uncertainties stemming from clinical data and how they were resolved.

RESULTS

Table 1 provides a summary of the QAs considered for adalimumab (products A–G) and bevacizumab (H–L) biosimilars. For each adalimumab and bevacizumab biosimilar, the percentage of batches within the established similarity range for each individual QA is categorized in a color and pattern. The analytical similarity packages of the adalimumab and bevacizumab biosimilars comprised between 35 and 85 individual assays per product (for complete list see Tables 2 and 3). For most of the QAs, orthogonal analytical methods were used.

Protein content

Protein content is a highly critical QA which must be fully comparable between the biosimilar and the RP. For all biosimilars

Table 2 Summary of analytical assays performed for *adalimumab* biosimilars

Quality attribute (and analytical method/s) for comparative characterization	
Content	Protein content (UV-280)
Primary structure	Molecular weight/intact mass (RPLC-UV/MS)
	Amino acid sequence (peptide mapping)
	N-terminal sequencing (peptide mapping, Edman sequencing)
	C-terminal sequencing (peptide mapping)
	Peptide mapping
	Disulfide bond analyses (peptide mapping)
	Free thiols (Ellmans test)
Higher order structure	Secondary structure (FTIR)
	Secondary- and tertiary structure (far and near UV circular dichroism)
	Protein folding (Intrinsic and extrinsic fluorescence)
	Thermal stability (DSC)
	Tertiary structure (1D 1H NMR, 2D 1H-1H NOESY NMR, 2D-NMR, HDX, X-ray crystallography, antibody conformational array)
Protein modifications	N-term pyroglutamate (peptide mapping)
	C-terminal lysine (peptide mapping, CEX)
	Iso-aspartate (peptide mapping)
	Deamidation (peptide mapping)
	Oxidation (peptide mapping)
	Glycation (BAC)
	Succinimidation (peptide mapping)
	Isomerization (peptide mapping)
	Proline amide (peptide mapping)
	Thioether (peptide mapping)
	Cysteinylation (peptide mapping)
Glycosylation	N-glycan profile (LC-ESI-MS/MS, 2-AB labeling HILIC-UPLC)
	Afucosylation (LC-ESI-MS/MS, 2-AB labeling HILIC-UPLC)
	High mannose (LC-ESI-MS/MS, 2-AB labeling HILIC-UPLC)
	Sialylation (LC-ESI-MS/MS, 2-AB labeling HILIC-UPLC)
	G0F (LC-ESI-MS/MS, 2-AB labeling HILIC-UPLC)
	G1F (LC-ESI-MS/MS, 2-AB labeling HILIC-UPLC)
	G2F (LC-ESI-MS/MS, 2-AB labeling HILIC-UPLC)
	Galactosylation (LC-ESI-MS/MS, 2-AB labelling HILIC-UPLC)
Purity/ impurity profile and charged variants	Size heterogeneity (SEC, CE-SDS reducing and non-reducing, SV-AUC, SEC-MALS, DLS, FFF)
	Hydrophobic heterogeneity (HIC)
	N-linked glycosylation site (LC-ESI-MS/MS)
	Charge heterogeneity (CEX-HPLC, iCIEF, iCE, cIEF, IEC-HPLC)
Fab mediated	Soluble TNF-binding (ELISA, SPR, FRET)
	Membrane TNF-binding (cell-based assay)
	TNF-α neutralization (NF-kB reporter, viability/cell death)
Fc and complement mediated	ADCC *e.g., for one product, up to 20 assays were performed, including: • NK-PBMC ADCC using healthy and patient blood • Whole blood ADCC using healthy and patient blood • FcγRIIIa ADCC reporter • Addition of serum to these assays • Addition of IgG to these assays
	FcγRI binding (SPR)
	FcγRIIa (131H, 131R) binding (SPR)
	FcγRIIb binding (SPR)

(Continued)

Table 2 (Continued)

Quality attribute (and analytical method/s) for comparative characterization	
	FcγRIIIa (158F, 158V) binding (SPR, AlphaLISA, RGA)
	FcγRIIIb binding (SPR)
	FcRn binding (SPR)
	CDC (cell-based assay)
	C1q binding (ELISA)
Additional functional assays	Apoptosis induction, reverse signaling (cell-based assay)
	Apoptosis inhibition in intestinal epithelial cells
	MLR (T cell proliferation, regulatory macrophages (CD14/CD206))
	IL-8 release from HUVECs
	IL-8 release from PBMCs
	IL-8 release from keratinocytes
	IL-8 release from intestinal epithelial cells
	IL-6 release from synoviocytes
	sVCAM-1 release from HUVECs
	ICAM-1 expression on HUVECs
	ELAM-1 expression on HUVECs
	MIP-1β release from whole blood
	MCP-1 release from whole blood
	Lack of impact on Lymphotoxin α

ADCC, Antibody dependent cell-mediated cytotoxicity; AlphaLISA, Amplified Luminescent Proximity Homogeneous Assay; BAC, boronate affinity chromatography; CDC, complement dependent cytotoxicity; CE-SDS, capillary electrophoresis-sodium dodecyl sulfate; CEX, cation exchange chromatography; cIEF, capillary isoelectric focusing; DSC, differential scanning calorimetry; DLS, dynamic light scattering; ELAM-1, endothelial-leukocyte adhesion molecule 1; ELISA, enzyme-linked immunoassay; FcRn, neonatal Fc receptor; FcγR, fragment crystallizable gamma receptor; FFF, filed flow fractionation; FRET, Förster (flourescence) Resonance Energy Transfer; FT-IR, Fourier-transform infrared; HDX, hydrogen–deuterium exchange; HIC, hydrophobic interaction chromatography; HILIC-UPLC, hydrophilic interaction ultra performance liquid chromatography; HPLC, high-performance liquid chromatography; HUVECs, human umbilical vein endothelial cells; iCIEF, imaged capillary electrophoresis focusing; ICAM-1, intercellular adhesion molecule 1; IEC, ion exchange chromatography; IL-6, interleukin 6; IL-8, interleukin 8; LC-ESI-MS/MS, liquid chromatography electrospray ionization tandem mass spectrometric; MCP-1, monocyte chemoattractant protein-1; MIP-1-β, macrophage inflammatory protein 1-β; MLR, mixed lymphocyte reaction; NMR, nuclear magnetic resonance; NOESY, NOE correlated spectroscopy; PBMCs, peripheral blood mononuclear cells; SEC-MALS, multi-angle light scattering coupled with size exclusion chromatography; RGA, reporter gene assay; RPLC-UV/MS, reversed phase liquid chromatography-ultraviolet/mass spectrometry; SEC, size exclusion chromatography; SPR, surface plasmon resonance; SV-AUC, sedimentation velocity-analytical ultracentrifugation; sVCAM-1, soluble vascular cell adhesion molecule-1; TNF, tumor necrosis factor; UV-280, ultraviolet absorbance at 280 nm wavelength.

examined, 100% of biosimilar batches were within the reference range, except for one bevacizumab (product F) and one adalimumab (product L) biosimilar (≥ 90%).

Fragment antigen binding mediated functions

a. Binding to soluble tumor necrosis factor. Adalimumab is an IgG1 mAb that binds, via its fragment antigen binding (Fab) domain, to tumor necrosis factor α (TNFα) and prevents it from binding to its receptors TNFR1 and TNFR2, thereby blocking TNF-induced inflammation.[33,34] This is the primary mechanism of action (MoA) for adalimumab across all approved indications. The biological activity of adalimumab is determined by a combination of binding assays and a cell-based TNFα cytotoxicity inhibition assay. In addition, some applicants used a nuclear factor kappa B (NF-κB) reporter gene assay, which is viewed as supportive data (not included in **Table 1** but for those biosimilars where this assay was used, ≥ 90% of biosimilar batches were within the reference range). As shown in **Table 1**, for all 7 adalimumab biosimilars studied, 100% of the batches were within the similarity range for binding to soluble TNFα with the chosen assays.

b. Binding to transmembrane TNF and reverse signaling. In addition to binding to soluble TNFα, adalimumab can bind to membrane-associated TNFα (mTNFα) and mediate reverse (or outside-to-inside) signaling. Binding of adalimumab to mTNFα does not appear to be important for therapeutic efficacy in all indications, however, it may contribute to the clinical efficacy of adalimumab in inflammatory bowel diseases (IBDs). Several possible mechanisms explain the contribution of reverse signaling to the efficacy of adalimumab in IBD. For example, adalimumab-mediated apoptosis of lamina propria T cells may represent an additional key MoA of adalimumab in IBD indications and is thought to be mediated by reverse signaling, although it may also be mediated by binding to soluble TNF, which is in turn bound to its receptor.[35] Anti-TNF agents, such as adalimumab and infliximab, are also known to induce CD14+ CD206+ M2-type wound-healing macrophages (regulatory macrophages) which may contribute to mucosal healing in IBD.[36,37] Induction of regulatory macrophages can be assayed

Table 3 Summary of analytical assays performed for *bevacizumab* biosimilars

Quality attribute (and analytical method/s) for comparative characterization	
Content	Protein content (UV-280)
Primary structure	Molecular weight (RPLC-UV/MS)
	Intact mass/reduced mass (LC-ESI-MS)
	Isoelectric point (cIEF)
	Amino acid sequence (peptide mapping)
	N-terminal sequencing (peptide mapping, Edman sequencing)
	C-terminal sequencing (peptide mapping)
	Amino acid sequence (peptide mapping)
	Disulfide bond analyses (peptide mapping)
	Free thiols (Ellmans test)
Higher order structure	Secondary structure (FTIR, far and near UV circular dichroism)
	Tertiary structure (far and near UV circular dichroism, FL)
	Protein folding (Intrinsic and extrinsic fluorescence)
	Thermal stability (DSC)
	Epitope mapping (HDX-MS)
	Di-sulfide bridging (RP-HPLC-ESI-MS, non-reduced peptide mapping)
Protein modifications	Deamidation (peptide mapping)
	Oxidation (peptide mapping)
	Glycation (BAC)
	Aspartate Isomerization (peptide mapping)
	Thioether (peptide mapping)
	Cysteinylation (peptide mapping)
Glycosylation	N-glycan profile (peptide mapping, LC-ESI-MS/MS, HILIC-UPLC)
	O-glycosylation (peptide mapping)
	Ng-HC (CE-SDS, reduced)
	Afucosylation (NP-HPLC)
	Fucosylation (NP-HPLC)
	High mannose (NP-HPLC)
	Sialylation (NP-HPLC, UHPLC-FLR)
	G0F (LC-ESI-MS/MS, 2-AB labeling HILIC-UPLC)
	G1F (LC-ESI-MS/MS, 2-AB labeling HILIC-UPLC)
	G2F (LC-ESI-MS/MS, 2-AB labeling HILIC-UPLC)
	Galactosylation (NP-HPLC)
Purity/ impurity profile and charge variants	Size heterogeneity (SEC, CE-SDS non-reduced and reduced, CGE non-reducing and reducing, SV-AUC, SEC-MALS, DLS, FFF)
	Particles (MFI)
	Charge heterogeneity (CEX-HPLC, iCIEF, cIEF)
	Hydrophobic heterogeneity (HIC)
Fab mediated	VEGF121 binding (HUVEC-cell based assay, SPR, ELISA)
	VEGF165 binding (HUVEC-cell based assay, SPR, ELISA)
	VEGF189 binding (HUVEC-cell based assay, SPR, ELISA)
	VEGF206 binding (HUVEC-cell based assay, SPR, ELISA)
	VEGF B, C, D binding (BLI)
	HUVEC neutralization assay (cell-based assay)
	VEGFR phosphorylation inhibition (cell-based assay)
	Cell signaling assay (HEK293 RGA)
	KDR/KDR dimerization assay (cell-based assay)

(Continued)

Table 3 (Continued)

Quality attribute (and analytical method/s) for comparative characterization	
Fc and complement mediated	ADCC (cell-based assay)
	FcγRI binding (SPR)
	FcγRIIa FcγRIIa (131H, 131R) binding (SPR)
	FcγRIIIa (158F, 158V) binding (SPR, AlphaLISA)
	FcγRIIIb binding (SPR)
	FcRn binding (SPR, ELISA)
	CDC (cell-based assay)
	C1q binding (ELISA)
Off-target binding	VEGF-B (SPR)
	VEGF-C (SPR)
	VEGF-D (SPR)
	PIGF-1 (SPR)
	PIGF-2 (SPR)

BAC, boronate affinity chromatography; BLI, bioluminescence imaging; CDC, complement dependent cytotoxicity; CE-SDS, capillary electrophoresis-sodium dodecyl sulfate; CEX, cation exchange chromatography; CGE, capillary gel electrophoresis; cIEF, capillary isoelectric focusing; DSC, differential scanning calorimetry; HDX-MS, hydrogen/deuterium exchange mass spectrometric; DLS, dynamic light scattering; ELISA, enzyme-linked immunoassay; FFF, filed flow fractionation; HEK293 RGA, reporter gene assay based on the HEK-293 cell; HIC, hydrophobic interaction chromatography; HILIC-UPLC, hydrophilic interaction ultra performance liquid chromatography; HPLC, high-performance liquid chromatography; HUVECs, human umbilical vein endothelial cells; iCIEF, imaged capillary electrophoresis focusing; KDR, kinase insert domain receptor; LC-ESI-MS/MS, liquid chromatography electrospray ionization tandem mass spectrometric; MFI, micro-flow imaging; PIGF, placental growth factor; SEC-MALS, multi-angle light scattering coupled with size exclusion chromatography; RPLC-UV/MS, reversed phase liquid chromatography-ultraviolet/mass spectrometry; SEC, size exclusion chromatography; SPR, surface plasmon resonance; SV-AUC, sedimentation velocity-analytical ultracentrifugation; TNF, tumor necrosis factor; UV-280, ultraviolet absorbance at 280 mm wavelength; VEGF, vascular endothelial growth factor.

by measuring antiproliferative effects in a mixed lymphocyte reaction (MLR).

In all cases, 100% of adalimumab biosimilar batches were within the reference range for binding to mTNF.

c. Binding to VEGF. Bevacizumab is an IgG1 mAb which binds to VEGF-A and prevents the signaling of VEGF receptors.[38] VEGF comprises at least 16 different isoforms due to alternate mRNA splicing. Inhibition of VEGF-A blocks the proliferation of vascular endothelial cells and angiogenesis. Although soluble VEGF isoforms ($VEGF_{121}$ and $VEGF_{165}$) are the most predominant isoforms in tumors, cell-associated VEGF (such as $VEGF_{189}$ and $VEGF_{206}$) is also expressed in a significant number of lung and colon cancers.[38]

For all bevacizumab biosimilars, 100% of batches were within the reference range for at least 2 of 3 VEGF isoforms. Binding to $VEGF_{165}$ could be demonstrated for all batches and binding to $VEGF_{121}$ and $VEGF_{189}$ was observed for all but one product (product L and product H, respectively). Binding to $VEGF_{206}$ was frequently not performed by applicants which was accepted because this isoform is seen as less important.[39]

Cell-based assays. Cell-based potency assays are considered highly important for determination of biosimilarity. In the absence of comparable biological activity, a product cannot be approved as a biosimilar. For adalimumab biosimilars, the functional cell-based assays were based on measuring adalimumab inhibition of TNFα mediated cell death. In 6 out of 7 adalimumab biosimilars, all 100% of batches were within the similarity range and for one product (product F), ≥ 90% of batches were within the similarity range. For bevacizumab biosimilars, an HUVEC-based antiproliferation assay was used, and, in all cases, 100% of batches were within the similarity range.

Fc-related assays. Adalimumab is known to induce antibody-dependent cellular cytotoxicity (ADCC) through the binding of the Fab region to mTNFα and the Fc region to FcγRIIIa, which is expressed on effector cells, such as NK cells (mainly via high affinity receptor genotype 158 v/v).[40] It is well-known that the binding of IgGs to FcγRIIIa is influenced by the glycan profile of the antibody. For example, levels of afucosylated glycans are generally correlated with ADCC activity. All seven adalimumab biosimilar applicants performed one or more comparative ADCC assays, which usually included peripheral blood mononuclear cells (PBMCs) or natural killer (NK) effector cells. As shown in **Table 1**, 100% of biosimilar batches were within the reference range for ADCC for all products. Although most applicants used 1 or 2 ADCC assay formats, for 1 biosimilar product, ADCC activity was measured using more than 20 different ADCC assay setups (see **Table 2**). This is an example of the large variety of assays that can be used to study a single QA.

Activation of complement-dependent cytotoxicity (CDC) is also viewed as a relevant MoA for adalimumab. For all adalimumab biosimilars, 100% of batches were within the reference range for CDC activity.

Bevacizumab is theoretically capable of mediating Fc-related effector functions. However, none of the authorized bevacizumab biosimilars displayed ADCC or CDC activity (represented as green vertical stripes in **Table 1**, meaning it was tested but not found), which is in line with previously published results for originator bevacizumab.

Fc binding assays. Neonatal Fc receptor (FcRn) has been shown to play a role in regulating IgG levels in the serum through recycling of bound antibodies, with an impact on the serum levels of therapeutic mAbs.[41] For this reason, binding to FcRn is considered as a critical QA.[42] In all cases except one bevacizumab biosimilar (product I), 100% of batches were found to be within the similarity range.

The results of binding assays for five other FcγR (FcγRIa, FcγRIIa, FcγRIIb, FcγRIIIa (158 f/f), and FcγRIIIb) showed a variable percentage of batches lying within the reference range. Only 3 of 7 adalimumab products had ≥ 90% or 100% of batches within the similarity range for all 5 FcR binding assays (products D, E, and G). Binding to the FcγRIIIa by therapeutic mAbs is known to enhance ADCC activity (relevant MoA for adalimumab). For the high affinity FcγRIIIa 158 v/v genotype, 6 out of 7 adalimumab biosimilars had 100% of batches within the similarity range and for one product ≥ 90% of batches were within the similarity range (product A). For bevacizumab biosimilars, the results were more variable, with between one (product L) and 4 (products J and K) FcγR binding assays having < 90% batches within similarity range. However, as discussed above, bevacizumab does not exhibit effector function, and therefore binding to FcγR is not considered a critical aspect of biosmiliarity.[30,31]

Glycosylation profile. Tests for glycosylation profile included as a minimum: G0F, G1F, G2F, afucosylation, sialylation, and high mannose content. It can affect the immunogenicity and, in some cases, (adalimumab) also the functionality of the mAb. In most cases, < 90% of batches tested were within the similarity range or the assay was not performed. This is not unexpected, because it is known that the glycoprofile is highly dependent on the cell line that is used as expression system, media, and several growth conditions.[43,44] Although differences in glycoprofile could impact the ADCC activity, in all cases for adalimumab, the ADCC activity was shown to be highly similar.

Purity testing. The purity/impurity profile is viewed as critical by the EMA as certain impurities may impact on safety and immunogenicity. Protein impurities can be measured by size exclusion chromatography (SEC) and capillary electrophoresis-sodium dodecyl sulfate (CE-SDS) under reduced and non-reduced conditions to detect several relevant impurities, such as fragmentation, truncation, and aggregation. For both the adalimumab and bevacizumab biosimilars, there were several instances where < 50% of the batches were inside of the similarity range. However, during the assessment process these differences were judged to be irrelevant in terms of safety and efficacy. In most cases, this was due to the fact that the absolute difference in impurity levels was so small as to not to be clinically meaningful. Furthermore, the purity is controlled by the release specification.

Charge variants. Due to the complex contribution of numerous QAs to the overall charge profile of mAbs, charge variations may quantitatively differ between biosimilars and their RP. As can be seen in **Table 1**, in most cases, the charge profile differed, with only one adalimumab biosimilar (product E) with 100% batches within similarity range. Differences in charge profile can generally be accepted, provided applicants justify why any observed charge differences would not have an adverse clinical impact.

Additional assays. Additional assays include, for example, inhibition of TNFα-induced apoptosis and inhibition of release of IL-8 or sVCAM-1 in cell culture (not included in **Table 1**) for adalimumab, and induction of HUVEC migration or apoptosis, site-specific phosphorylation of VEGFR2, and HEK293 VEGF reporter assay for bevacizumab biosimilars. These assays are not considered mandatory but can be useful to strengthen the claim of biosimilarity.

For the majority of adalimumab biosimilars, 100% of batches were within the reference range when measuring induction of apoptosis. For one biosimilar (product A), this function was not addressed or outside the similarity range. Data from MLR studies were provided for all adalimumab biosimilars, with 3 products (C, D, and G) showing < 90% batches within range. Only one bevacizumab biosimilar (product J) had all additional functions assessed and with 100% batches within range.

However, due to the inherent variability of these assays and the low numbers of batches tested, the evidence provided by these assays was considered supportive only.

Table 4 provides a summary of the instances where < 100% of batches were within the reference range. In each case, the reason why these differences were accepted by the EMA is explained.

RESULTS OF CLINICAL COMPARABILITY STUDIES

The clinical results obtained for each adalimumab biosimilar are provided in **Table S1**.

The clinical results obtained for each bevacizumab biosimilar are provided in **Table S2**.

In the following section, results of PK analyses (obtained in healthy volunteers (HVs) and patients), efficacy analyses of clinical trials in patients and safety, and immunogenicity evaluation (obtained from PK and efficacy studies) are presented.

PK studies

PK in healthy volunteers. In **Tables S1** and **S2**, the primary end points with prespecified margins and all secondary end points, including safety and immunogenicity, are presented.

Observation period. For adalimumab biosimilars, the length of follow-up ranged from 62 to 71 days, and for bevacizumab biosimilars from 85 to 100 days, which represents ~ 5 half-lives.

Primary end point. In all instances the primary end points (area under the curve to infinity (AUC_{inf}), maximum concentration (C_{max}), and AUC from time of administration up to the time of the last quantifiable concentration (AUC_{last})) were contained within the prespecified acceptance range of 0.8–1.25. For three adalimumab products (Hyrimoz/Halimatoz/Hefiya, Hulio, and Amsparity) and one bevacizumab (Alymsys/Oyavas) the end points were such that unity was not included in the 90%

Table 4 QAs with <100% of batches meeting similarity ranges and how the resulting uncertainty during MAA and how this was resolved

Adalimumab	QA	Percentage of batches within the similarity range	How resolved
Product F	Protein content	≥90% of batches	The small difference in protein content was concluded to be of no clinical relevance. Batch-to-batch variability of the biosimilar within the expected range.
Product F	Cell based potency assay	≥90% of batches	Minor difference not expected to affect the clinical performance of the product.
Products A,B,C,F	Binding to several FcγR receptors (FcγRIa, FcγRIIa, FcγRIIb, FcγRIIIa-158 f/f and FcγRIIIb)	Variable, see **Table 1**	Minor differences in binding results, similarity confirmed in cell-based functional assays.
Product A	Binding to FcγRIIIa 158v/v	≥90% of batches	Viewed as sufficient based on ADCC assay results
Product A – G (all)	Glycosylation (7 attributes)	Variable, often <90%, see **Table 1**	Similarity confirmed in cell-based functional assays. No clinically significant difference in PK profile.
Products A, B, C, D, and F (all except E)	Purity testing	Variable, often <90%, see **Table 1**	Based on regulatory experience, the small difference was seen as negligible. In most cases purity of biosimilar was marginally increased.
Products A, B, C, D, and F (all except E)	Charge variants	Variable, often <90%, see **Table 1**	Acceptable based on product understanding.
Product A	Apoptosis induction	<50% of batches or not done	The assay is not considered as highly critical, accepted based on high similarity for binding to transmembrane TNFα. Alternative assay used as a functional readout of transmembrane TNF binding, e.g., MLR.
Product B	Apoptosis inhibition	Variable, <90% in one case, see **Table 1**	Additional orthogonal assays supported biosimilarity. Accepted based on the totality of evidence.
Bevacizumab	**QA**	**Percentage of batches within the similarity range**	**How resolved**
Product L	Protein content	≥90% of batches	The small difference in protein content was concluded be of no clinical relevance. Batch-to-batch variability of the biosimilar within the expected range.
Product L	Binding to VEGF 121	<50% of batches or not done	High similarity for binding to other VEGF isoforms confirmed using orthogonal methods
Product H	Binding to VEGF 189	<50% of batches or not done	High similarity for binding to other VEGF isoforms confirmed using orthogonal methods
Products H, I, J, and L (all except K)	Binding to VEGF 206	Variable, often <90%, see **Table 1**	High similarity for binding to other VEGF isoforms confirmed using orthogonal methods VEGF 206 is a less frequent isoform in human tissues[39]
Product I	Binding to FcRn	≥90% of batches	Based on regulatory experience and the results from the comparative PK study, the minor difference was seen as negligible.
Products H–L (all)	Binding to several FcγR receptors	Variable, see **Table 1**	Binding to Fcγ receptors are not considered critical for the mode of action of bevacizumab.
Products H–L (all)	Glycosylation (7 attributes)	Variable, often <90%, see **Table 1**	Due to the lack of Fc functions for bevacizumab, glycosylation pattern is not critical for bevacizumab. The PK profiles demonstrated similarity.
Products H–L (all)	Purity testing	Variable, often <90%, see **Table 1**	Based on regulatory experience, the small difference was seen as negligible.
Products H–L (all)	Charge variants	Variable, often <90%, see **Table 1**	Acceptable based on product understanding.
Products H, I, K, and L (all except J)	Additional functional assays	Variable, often <90%, see **Table 1**	The assays are not considered as highly critical, differences accepted based on the totality of evidence presented for similarity.

ADCC, antibody-dependent cell-mediated cytotoxicity; QAs, quality attributes; MAA, marketing authorization assessment; MLR, mixed lymphocyte reaction; PK, pharmacokinetics; TNF, tumor necrosis factor; VEGF, vascular endothelial growth factor.

confidence intervals (CIs), which may be permissible.[45] Root cause analysis for not being included in the 90% CI was performed by additional supplementary analyses on the primary end point and scrutinizing relevant QAs (for example, high mannose and sialic acid) with no negative signals.

In addition, for two adalimumab products (Hyrimoz/Halimatoz/Hefiya, and Hulio), initially failed and subsequently successful PK studies were submitted. Root cause analysis was performed, without finding analytical dissimilarities that could have explained the initial failure to show bioequivalence. Further, in both instances, a second, more strictly standardized PK study was conducted with reduced intersubject variability, and PK similarity was shown.[6,14]

Population PK in patients. For some products, PopPK data were collected in a subset of patients as part of the clinical efficacy/safety study, with trough plasma concentration (C_{trough}) as an end point, as recommended in the mAb guideline.[10]

For adalimumab biosimilars, samples were typically collected at 5 timepoints (sparse sampling) for all patients in the initial 6 months study period and, in some instances, until week 50 or even 60. In all instances, PopPK results were considered comparable.

For bevacizumab biosimilars, C_{trough} was typically collected at time zero (baseline), and at weeks 4, 7, 13, and 19. Acceptable PopPK data were provided in four of six cases; for Alymsys, PopPK analysis was not carried out, and for Abevmy, the data set was viewed as insufficient after assessment. In both cases, the insufficiency of comparative PopPK data in patients was justified by proven PK similarity in HVs and totality of evidence from other parts of the similarity exercise.

Clinical efficacy studies. Adalimumab is currently approved in 13 autoimmune indications.[46] Four applicants chose to compare efficacy in subjects with rheumatoid arthritis (RA) as a model indication in the clinical trial. Three applicants chose chronic plaque-type psoriasis as the model indication. Both indications are viewed as sufficiently sensitive by the EMA to detect potential clinically relevant differences between the biosimilar and the RP due to the large treatment effect.

Bevacizumab is currently approved in six indications in the European Union.[47] All applicants chose newly diagnosed or recurrent stage (IIIB)/IV nonsquamous non-small cell lung cancer (NSCLC) as the most sensitive model indication due to the large treatment effect.

Observation period. The length of follow-up was typically 1 year for all adalimumab and bevacizumab biosimilars (for Hulio, follow-up was 24 weeks with an extension trial proceeding up to 1 year).

Primary end point. For adalimumab biosimilars, American College of Rheumatology Response (ACR 20) and Psoriasis Area and Severity Index (PASI) score were chosen as primary end points for RA and psoriasis, respectively. The equivalence margins for the risk difference varied between ± 10% and ± 15% for RA and ± 15% and ± 18% for psoriasis, depending on the number and nature of trials performed with the RP that were included in the meta-analysis to derive the equivalence margin. In all instances, the 95% CI for the primary end points were within the prespecified equivalence margins and all other secondary end points also supported similar clinical performance. Results obtained in both analysis sets (intention to treat (ITT); per protocol set (PPS)) were concordant in all instances. Secondary end points in the trials were ACR 50 and ACR 70 scores, and Disease Activity Score (DAS) 28 in RA, and PASI 50, 75, and 90 scores in psoriasis, as well as additional efficacy measurements at different timepoints.

For bevacizumab products, the predefined equivalence margin for the risk difference of overall response rate (ORR) varied between ± 12 and ± 13% depending on the chosen reference studies. The 95% CI for the ORR was fully contained within the prespecified acceptance range for all five substances. Results of secondary end points of progression-free survival (PFS) and duration of response (DOR) generally provided further support for biosimilarity. However, for Alymsys, PFS was seemingly worse for the biosimilar with a hazard ratio (HR) of 1.2 (0.98, 1.46); median (weeks): 36.0 (33.57–36.86) vs. 43.0 (36.14 to 45.14). Nevertheless, this finding was not viewed as critical as the primary end point was met and the study was not designed to demonstrate equivalence for PFS.

Secondary end points. Time-dependent end points were included as secondary end points, but are less sensitive and informative for conclusions on biosimilarity than end points reflecting the MoA, because they are likely influenced by patient-related factors, such as general health status. For bevacizumab biosimilars, median overall survival (OS) could not be estimated for either group in all instances due to limited observation time and due to the fact that > 50% of patients were still alive at the cutoff. In instances where the HR for OS was > 1.0 (e.g., Abevmy and Alymsys), suggesting higher mortality in the biosimilar group, the OS results were viewed with caution by the EMA because the studies were neither adequately powered to demonstrate equivalence, nor to detect differences in OS, and no type 1 error control was included.

Clinical safety data. Safety parameters, such as treatment emergent adverse events (TEAEs), adverse events of special interest (AESi), serious adverse events (SAEs), deaths, and TEAEs leading to discontinuation were comparable between groups in most cases, as seen in **Tables S1** and **S2**. Adverse events (AEs) were mild to moderate and the adverse episodes resolved in all instances with no deaths reported.

It should be noted that clinical trials are not powered for safety end points, because this is considered unnecessary and would usually require several thousand study participants.

Clinical immunogenicity data. Adalimumab is a highly immunogenic product and antidrug antibodies (ADAs) were detected in 30–88% of subjects across all trials. The variability may be explained by differences in study populations and sensitivity of antibody

Table 5 Discrepancies in clinical attributes and how they were resolved

Adalimumab	Clinical attribute	Observation	How resolved
Hyrimoz/Halimatoz/ Hefiya, Hulio, and Amsparity	PK	Unity was not included in the 90% CI	1. Permissable[44] 2. Relevant QAs (high mannose, sialic acid) showed close similarity
Hyrimoz/Halimatoz/ Hefiya and Hulio)	PK	Initial study failed to meet predefined acceptance range	1. Root cause analysis 2. Subsequently, successful PK studies were submitted
Bevacizumab	**Clinical attribute**	**Observation**	**How resolved**
Alymsys/Oyavas	PK	Unity was not included in the 90% CI	1. Permissable[44] 2. Relevant QAs (high mannose, sialic acid) showed close similarity
Alymsys/Oyavas	Pop PK	Not carried out	1. Pop PK only supportive 2. PK similarity proven in HV
Abevmy	Pop PK	Insufficient	1. Pop PK only supportive 2. PK similarity proven in HV
Alymsys/Oyavas	PFS	HR of 1.2 (0.98, 1.46); median (weeks): 36.0 (33.57–36.86) vs 43.0 (36.14 to 45.14).	1. Primary endpoint (ORR) met 2. Study not designed to demonstrate equivalence for PFS 3. Totality of evidence in overall biosimilarity assessment
Abevmy, Alymsys/ Oyavas	OS	HR for OS > 1.0; suggesting higher mortality in the biosimilar group	1. Primary endpoint (ORR) met 2. Study not designed to demonstrate equivalence for OS 3. Totality of evidence in overall biosimilarity assessment

CI, confidence interval; HV, healthy volunteers; HR, hazard ratio; ORR, overall response rate; OS, overall survival; PK, pharmacokinetic; PFS, progression-free survival; PopPK, population PK.

assays used. Importantly, patients with ADA-positive samples at any time were similar between treatment arms for all biosimilar products.

Bevacizumab is a low immunogenic product and bevacizumab ADAs and neutralizing antibodies were rarely detected, except for higher percentages observed for Alymsys/Oyavas and Aybintio/ Onbevzi; however, they were similar between treatments arms and almost all ADAs were transient and appeared not to have effects on PKs or safety.

A summary of all observations with clinically deviating results is provided in **Table 5**.

DISCUSSION

In the early years of biosimilar development, it was considered that even with a convincing quality and PK package, there would always be some "residual uncertainty" which in most cases could only be addressed by a sufficiently powered efficacy study in patients. However, since then, the discriminatory power of the analytical methods used has vastly increased. EU-regulators have gained a large body of knowledge on the quality profile of several mAbs, adalimumab and bevacizumab being just two examples. As shown in **Tables 2** and **3**, the panel of analytical testing for biosimilars is very comprehensive, with numerous orthogonal methods used to analyze dozens of QAs. Therefore, every relevant aspect of the mAb structure and activity is interrogated to ensure that it is sufficiently aligned with the RP in order to guarantee comparable clinical efficacy and safety.

This is the first study that performs an in-depth analysis of all quality and clinical data for currently authorized biosimilars of two originator mAbs used in either oncologic (5 products) or in autoimmune indications (7 products).

Based on information provided to the EMA in the MA submissions, and following the scientific evaluation carried out by the Agency, we found that over 90% (and in most cases 100%) of the biosimilar batches met the EU-RP similarity range for critical QAs. A lower percentage of biosimilar batches were within the similarity range for QAs which may be considered less critical to safety and efficacy, such as glycosylation profile or charge variants (see **Tables 1, 4, 5**).

The most critical QAs for the determination of biosimilarity are those that could have an impact on the PK profile, on safety (including immunogenicity) and efficacy. Therefore, for high criticality QAs, a high degree of similarity to the RP is expected. Of the numerous QAs studied, the ones which as part of the assessment process were considered by the EMA to be of high criticality for determination of adalimumab biosimilarity, and which have demonstrated high concordance (marked solid green) are: protein content, soluble TNFα binding, transmembrane TNFα binding, FcγRIIIa binding, FcRn, biological activity (as measured in cell-based TNFα neutralization assay), ADCC, CDC, and a functional read out of reverse signaling (e.g., apoptosis induction). C1q binding is also considered critical, but these data were not presented in our analyses due to problems with anonymity. Similarly, high criticality for determination

of bevacizumab biosimilarity (as considered by the EMA and marked solid green) are: protein content, biological activity (as measured in cell-based antiproliferation assay), binding to main VEGF isoforms, and FcRn binding.

If < 100% of batches were within the similarity range for these highly critical QAs, data from relevant additional analytical and functional assays were reviewed in order to establish that the variations will not lead to differences in clinical performance of the biosimilar (see **Table 4**).

For all adalimumab products (A–G in **Table 1**) > 90% of batches were within the similarity range for highly critical QAs: protein content, potency, ADCC, CDC, and binding to soluble TNFa, mTNFa, FcγRIIIa (158 v/v), and FcRn. Regarding less critical QAs, < 90% of biosimilar batches were within the similarity range (for instance, products A and E, indicated as dark blue dots or light blue diagonal stripes). This was considered acceptable and the PK trial demonstrated no impact on safety/immunogenicity. In addition, the clinical data were supportive of biosimilarity.

Regarding the differences found in glycosylation, although high similarity was observed for all adalimumabs compared with the RP Humira with regard to FcγRIIIa (158 v/v) binding, as well as for ADCC and CDC activity, none of the adalimumab biosimilars had a fully comparable glycoprofile to the RP Humira. The specific glycoprofile is highly dependent on the manufacturing process, including the cell line and growth conditions used, therefore manufacturing an mAb with a highly similar glycoprofile is challenging. In all cases, the minor differences in glycoprofile were justified not to have a functional impact through orthogonal methods, for example, leading to a difference in ADCC activity. Therefore, any observed differences in the glycoprofile between the biosimilar and the originator were justified not to affect the clinical performance of the biosimilar.

For bevacizumab products (H–L), again, for highly critical QAs, ≥ 90% of batches were within the similarity range (i.e., protein content, HUVEC antiproliferation assay, and binding to FcRn and VEGF165). For products H and L, < 90% of biosimilar batches were within the similarity range for binding to VEGF189 and VEGF121, but this was accepted given as binding to other VEGF isoforms was highly similar. For products I and K, differences were apparent in several purity and glycosylation attributes, but again this was accepted by the EMA during the MA evaluation as Fc functionality (glycosylation profile) is not critical for bevacizumab.

Assessing the PK trial demonstrated no impact on safety/immunogenicity of patients and further clinical data were also supportive of biosimilarity.

Regarding the differences found in purity, these need to be justified or appropriately clinically qualified as they may affect efficacy and safety, including immunogenicity. In some instances, there were minor differences in impurity levels between the biosimilars and RP, with some biosimilars showing slightly higher purity levels and some slightly lower. However, in all cases, differences were considered minimal in absolute numbers and were justified to have no impact on safety or efficacy. Further, no immunogenicity signals were observed in the clinical PK or efficacy trials.

As a general comment, the list of critical QAs known to be of high importance for determination of biosimilarity should not be interpreted in such a way that these are the only QAs of interest and other QAs do not need to be studied. Rather, if differences between the biosimilar and the RP are detected, the biosimilar applicant needs to justify the impact of the difference. Moreover, although we did not include them in our analysis because the data could not be categorized in a quantitative way, the amino acid sequence secondary and higher order structure of a biosimilar is expected to be the same or highly similar to the RP and therefore these are also considered critical QAs. Information presented on the number and type of conducted assays for all products should not be leveraged by future developers as the presented results are reviewed by regulators on a case-by-case basis.

Several scientists[6,14,16,48,49,50] pointed out limitations of indiscriminative efficacy and safety studies in light of technical advances in analytical methods which provide more discriminative research tools. Even large molecules can currently be thoroughly characterized using state-of-the-art analytical and *in vitro* functional testing. This thorough characterization is also routinely applied as part of comparability studies conducted for biological medicinal products following introduction of manufacturing process changes.[51–53] The recently established EMA tailored scientific advice pathway for biosimilars acknowledges these scientific advances.[54]

For all adalimumab and bevacizumab biosimilars studied, a comprehensive clinical program was submitted consisting of a PK trial and a clinical efficacy study, which confirmed biosimilarity.

Secondary efficacy end points, safety, PopPK, and immunogenicity end points were always descriptive in nature and results generally concordant with those of the primary end points. In those cases where a trend toward a possible difference was observed, it was judged to be negligible and/or likely due to immaturity of the data. In some cases, certain data were not obtained or incomplete. These deviating or missing results did not preclude approval, as similarity was shown for the relevant QAs and in dedicated PK studies and confirmed by clinical data in the efficacy study. As this paper analysed already approved biosimilar products, and subsequent experience with these products to date has not resulted in any safety or efficacy issues following their approval,[16,55,56] it can be concluded that the regulatory decisions taken were correct.

Our study supports previous observations[48] that adequately powered PK trials, provide sufficient clinical safety and immunogenicity data, especially when close similarity in analytical and functional parameters together with comparable PK and impurity profiles can already largely predict similar safety and immunogenicity of the biosimilar and the RP.

As stated by Kurki *et al.*[48] the intrinsic immunogenicity observed for each RP was also observed for the respective biosimilars. In no instance did RPs with high immunogenicity have a biosimilar with low immunogenicity, or vice versa. Observations regarding comparability of immunogenicity made in the PK trial were, in all instances, confirmed in the efficacy/safety study. Similar observations were made with regard to safety parameters.

Furthermore, the EU pharmacovigilance systems and risk management planning are sufficiently robust[16,55,56] to detect safety signals in postmarketing use. However, safety signals (including reports on reduced efficacy) are not anticipated, because more than

a decade of clinical experience indicates that a new safety signal solely identified with a biosimilar is extremely unlikely.[55,56]

For the biosimilars included in this study, differences in several QAs were found. As part of the EMA approval process, applicants were challenged to justify that the observed differences would have no impact on the clinical performance of the biosimilar. Importantly, in all cases, these questions were answered by applicants based either on quality data alone or on a combination of quality, PK, and immunogenicity data (**Table 4**). In no instance were data from the clinical efficacy, safety, and immunogenicity study required to justify the differences at the quality level. On this basis, we argue that for the adalimumab and bevacizumab biosimilars, clinical efficacy, safety, and immunogenicity data were not needed to address residual uncertainty remaining from the quality and PK studies.

This analysis adds to the ongoing debate about the role of clinical studies for biosimilars.

In the authors' opinion, the usefulness of clinical efficacy, safety, and immunogenicity data for the purposes of regulatory decision could be questioned. Where the quality, PK, and immunogenicity data are sufficiently robust and convincing for regulatory decision making, as in the case of the adalimumab and bevacizumab examples cited in this paper, then it is our contention that the current expectations for clinical efficacy, safety, and immunogenicity could be re-examined. Therefore, we encourage a regulatory review of the standards for clinical data requirements for biosimilars, and propose that clinical data requirements should be further tailored.

Given the 10 years of regulatory experience in assessing and approving biosimilars, and the positive performance of approved biosimilar mAbs on the market,[6,14,48,57] the authors suggest to move to a concept of "tailored evidence," depending on the nature of the product and the available orthogonal assays for quality similarity. For example, this could include removing the standard requirement for equivalence trials, accepting wider equivalence margins, omitting PopPK studies, and/or reducing secondary clinical end points.

Such tailored approaches may prove particularly useful in the case of biosimilars for orphan medicines or other treatments where there is a small patient population or products with a narrow treatment effect where a comparative efficacy study may not be feasible due to the inability to recruit a sufficient number of subjects for any meaningful statistical analysis.

In the authors' opinion, if the efficacy study is omitted, sponsors may consider expanding their PK studies with regard to study size or observation period to gather additional safety/immunogenicity data. In other instances, a clinical study generating some limited safety and immunogenicity data in patients may be beneficial.

Where the quality package or the PK data are not sufficiently convincing, a root cause analysis would be necessary, potentially requiring changes to the manufacturing process of the biosimilar candidate or an improved design/power of the PK study, as was, for example, observed during biosimilar adalimumab development.[14] Alternatively, a stand-alone application could be pursued. Whereas the analysis in our study is based on adalimumab and bevacizumab as representative examples, the principle could be generalized to mAbs as a class.

In conclusion, in the author's opinion, a tailored evidence approach for all biosimilars including mAbs and fusion proteins, where a robust and convincing analytical biosimilarity package is available in conjunction with an appropriately powered PK study that also provides safety and immunogenicity data, the extent of the clinical trial requirements can be further reduced, or such trials even omitted. This would allow for more rational use of clinical resources, reduce the type of clinical data analyzed or number of clinical trials, and streamline the development of biosimilar mAbs and fusion proteins to the benefit of patients and healthcare stakeholders which is also in line with the strategic priorities of the EMA.[58]

SUPPORTING INFORMATION

Supplementary information accompanies this paper on the *Clinical Pharmacology & Therapeutics* website (www.cpt-journal.com).

ACKNOWLEDGMENT

The authors thank Francesco Pignatti and Rosa Gonzalez-Quevedo (EMA) and Alexandre Moreau (French Agency for the Safety of Health Products (ANSM)) for their valuable suggestions and comments. Open Access funding enabled and organized by Projekt DEAL.

FUNDING

No funding was received for this work.

CONFLICT OF INTEREST

The authors declared no competing interests for this work.

AUTHOR CONTRIBUTIONS

E.G., N.E., S.B., M.W., and E.W.-H. wrote the manuscript. N.E., S.B., and E.W.-H. designed the research. E.G., N.E., S.B., and E.W.-H. performed the research. N.E., S.B., E.G., and E.W.-H. analyzed the data.

1. Guideline on similar biological medicinal products containing biotechnology-derived proteins as active substance: non-clinical and clinical issues. EMEA/CHMP/BMWP/42832/2005_Rev.1 <https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-similar-biological-medicinal-products-containing-biotechnology-derived-proteins-active_en-2.pdf> (2015). Accessed April 2022.
2. McClellan, J. *et al*. The 'totality-of-the-evidence' approach in the development of PF-06438179/GP1111, an infliximab biosimilar, and in support of its use in all indications of the reference product. *Therap Adv. Gastroenterol.* **12**, 1756284819852535 (2019).
3. Weise, M., Kurki, P., Wolff-Holz, E., Bielsky, M. & Schneider, C. Biosimilars: the science of extrapolation. *Blood* **124**, 3191–3196 (2014).
4. Schellekens, H. Biosimilar therapeutics, what do we need to consider? *NDT Plus* **2**, i27–i36 (2009).
5. O'Callaghan, J., Barry, S.P., Bermingham, M., Morris, J.M. & Griffin, B.T. Regulation of biosimilar medicines and current perspectives on interchangeability and policy. *Eur. J. Clin. Pharmacol.* **75**, 1–11 (2019).
6. Schiestl, M. *et al*. The path towards a tailored clinical biosimilar development. *BioDrugs* **34**, 297–306 (2020).
7. Guideline on non-clinical and clinical development of similar biological medicinal products containing recombinant human insulin and insulin analogues; EMEA/CHMP/BMWP/32775/2005_Rev.1

<https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-non-clinical-clinical-development-similar-biological-medicinal-products-containing_en-0.pdf> (2016). Accessed April 2022.

8. Guideline on non-clinical and clinical development of similar biological medicinal products containing low molecular-weight-heparins; EMEA/CHMP/BMWP/118264/2007 Rev. 1 <https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-non-clinical-clinical-development-similar-biological-medicinal-products-containing-low_en.pdf> (2016). Accessed April 2022.
9. Guideline on similar biological medicinal products containing recombinant granulocyte colony stimulating factor (rG-CSF), EMEA/CHMP/BMWP/31329/2005 Rev 1 <https://www.ema.europa.eu/en/documents/scientific-guideline/draft-guideline-similar-biological-medicinal-products-containing-recombinant-granulocyte-colony_en.pdf> (2018). Accessed April 2022.
10. EMA. Guideline on similar biological medicinal products containing monoclonal antibodies –non-clinical and clinical issues; EMA/CHMP/BMWP/403543/2010 <https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-similar-biological-medicinal-products-containing-monoclonal-antibodies-non-clinical_en.pdf> (2012). Accessed April 2022.
11. FDA. Guidance for industry bioavailability and bioequivalence studies submitted in NDAs or INDs — general considerations <https://www.fda.gov/files/drugs/published/Bioavailability-and-Bioequivalence-Studies-Submitted-in-NDAs-or-INDs-%E2%80%94-General-Considerations.pdf> (2014). Accessed April 2022.
12. MHRA. Guidance on the licensing of biosimilar products <https://www.gov.uk/government/publications/guidance-on-the-licensing-of-biosimilar-products/guidance-on-the-licensing-of-biosimilar-products> (2021). Accessed April 2022.
13. WHO. Guidelines on evaluation of biosimilars-Proposed revision of Annex 2 of WHO Technical Report Series, No. 977 <https://cdn.who.int/media/docs/default/source/biologicals/who-guidelines-on-evaluation-of-biosimilars---4-nov2021.pdf?sfvrsn=f17799ae_5> (2021). Accessed April 2022.
14. Wolff-Holz, E., Tiitso, K., Vleminckx, C. & Weise, M. Evolution of the EU biosimilar framework: past and future. *BioDrugs* **33**, 621–634 (2019).
15. EPAR Remsima EMA/CHMP/589317/2013 <https://www.ema.europa.eu/en/documents/assessment-report/remsima-epar-public-assessment-report_en-0.pdf> (2013). Accessed April 2022.
16. Bielsky, M.C. *et al*. Streamlined approval of biosimilars: moving on from the confirmatory efficacy trial. *Drug Discov. Today* **25**, 1910–1918 (2020).
17. Nepru, N., Joshi, S., Gulliarme, D. & Rathore, A.S. Analytical similarity assessment of biosimilars: global regulatory landscape, recent studies and major advancements in orthogonal platforms. *Front. Bioeng. Biotechnol.* **10**, 832059 (2022).
18. Druedahl, L.C. *et al*. A qualitative study of biosimilar manufacturer and regulator perceptions on intelectual property and abbreviated approval pathways. *Nat. Biotechnol.* **38**, 1253–1256 (2020).
19. McCamish, M. & Woollett, G. The state of the art in the development of biosimilars. *Clin. Pharmacol. Ther.* **91**, 405–417 (2012).
20. EPAR Amgevita EMA/106922/2017 <https://www.ema.europa.eu/en/documents/assessment-report/terrosa-epar-public-assessment-report_en.pdf> (2017). Accessed April 2022.
21. EPAR Imraldi EMA/CHMP/559383/2017 <https://www.ema.europa.eu/en/documents/assessment-report/imraldi-epar-public-assessment-report_en.pdf> (2017). Accessed April 2022.
22. EPAR Hyrimoz EMA/CHMP/559383/2017 <https://www.ema.europa.eu/en/documents/assessment-report/imraldi-epar-public-assessment-report_en.pdf> (2017). Accessed April 2022.
23. EPAR Hulio EMA/541826/2018 <https://www.ema.europa.eu/en/documents/assessment-report/hulio-epar-public-assessment-report_en.pdf> (2018). Accessed April 2022.
24. EPAR Idacio EMA/124342/2019 <https://www.ema.europa.eu/en/documents/assessment-report/idacio-epar-public-assessment-report_en.pdf> (2019). Accessed April 2022.
25. EPAR Amsparity EMA/2756/2020 <https://www.ema.europa.eu/en/documents/assessment-report/amsparity-epar-public-assessment-report_en.pdf> (2020). Accessed April 2022.
26. EPAR Yuflyma EMA/47907/2021 <https://www.ema.europa.eu/en/documents/assessment-report/yuflyma-epar-public-assessment-report_en.pdf> (2021). Accessed April 2022.
27. EPAR Mvasi EMA/CHMP/559383/2017 <https://www.ema.europa.eu/en/documents/assessment-report/imraldi-epar-public-assessment-report_en.pdf> (2017). Accessed April 2022.
28. EPAR Zirabev EMA/798844/2017 <https://www.ema.europa.eu/en/documents/assessment-report/mvasi-epar-public-assessment-report_en.pdf> (2017). Accessed April 2022.
29. EPAR Aybintio EMA/380645/2020 <https://www.ema.europa.eu/en/documents/assessment-report/aybintio-epar-public-assessment-report_en.pdf> (2020). Accessed April 2022.
30. EPAR Alymsys EMA/112177/2021 <https://www.ema.europa.eu/en/documents/assessment-report/alymsys-epar-public-assessment-report_en.pdf> (2021). Accessed April 2022.
31. EPAR Abevmy EEMA/176217/2021 <https://www.ema.europa.eu/en/documents/assessment-report/abevmy-epar-public-assessment-report_en.pdf> (2021). Accessed April 2022.
32. EMA. Reflection paper on statistical methodology for the comparative assessment of quality attributes in drug development; EMA/CHMP/138502/2017 <https://www.ema.europa.eu/en/documents/scientific-guideline/reflection-paper-statistical-methodology-comparative-assessment-quality-attributes-drug-development_en.pdf> (2017). Accessed April 2022.
33. Vena, G.A. & Cassano, N. Drug focus: adalimumab in the treatment of moderate to severe psoriasis. *Biol. Theory* **1**, 93–103 (2007).
34. Tracey, D., Klareskog, L., Sasso, E.H., Salfeld, J.G. & Tak, P.P. Tumor necrosis factor antagonist mechanisms of action: a comprehensive review. *Pharmacol. Ther.* **117**, 244–279 (2008).
35. Levin, W.M.E. & Van den Brink, G.R. Mechanism of action of anti-TNF therapy in inflammatory bowel disease. *J. Crohn's Colitis.* **10**, 989–997 (2016).
36. Vos, A.C. *et al*. Regulatory macrophages induced by infliximab are involved in healing in vivo and in vitro. *Inflamm. Bowel Dis.* **18**, 401–408 (2012).
37. Vos, A.C., Wildenberg, M.E., Duijvestein, M., Verhaar, A.P., Van den Brink, G.R. & Hommes, D.W. Anti-tumor necrosis factor-α antibodies induce regulatory macrophages in an Fc region-dependent manner. *Gastroenterology* **140**, 221–230 (2011).
38. Kazazi-Hyseni, F., Beijnen, J.H. & Schellens, J.H. Bevacizumab. *Oncologist* **15**, 819–825 (2010).
39. Burchardt, T. *et al*. Expression of VEGF splice variants 144/145 and 205/206 in adult male tissues. *IUBMB Life* **48**, 405–408 (1999).
40. de Taeye, S.W. *et al*. FcγR binding and ADCC activity of human IgG allotypes. *Front. Immunol.* **6**(11), 740 (2020).
41. Kim, J., Luo, H., White, W., Rees, W., Venkat, R. & Albarghouthi, M. Impact of Fc N-linked glycans on in vivo clearance of an immunoglobulin G1 antibody produced by NS0 cell line. *MAbs* **12**, 1844928 (2020).
42. Suzuki, T. *et al*. Importance of neonatal FcR in regulating the serum half-life of therapeutic proteins containing the Fc domain of human IgG1: a comparative study of the affinity of monoclonal antibodies and Fc-fusion proteins to human neonatal FcR. *J. Immunol.* **184**, 1968–1976 (2010).
43. Ramanan, S. & Grampp, G. Drift, evolution, and divergence in biologics and biosimilars manufacturing. *BioDrugs* **28**, 363–372 (2014).
44. Spearman, M., Dionne, B. & Butler, M. *The Role of Glycosylation in Therapeutic Antibodies* (2011).
45. CHMP request on biosimilarity to Pharmacokinetics Working Party (PKWP); EMA/CHMP/230419/2017 (2018).
46. EPAR Humira <https://www.ema.europa.eu/en/documents/variation-report/humira-h-c-481-ii-0082-epar-assessment-report-variation_en.pdf> (2021). Accessed April 2022.

47. EPAR Avastin <https://www.ema.europa.eu/en/documents/variation-report/avastin-h-c-582-a20-0038-epar-assessment-report-article-20_en.pdf> (2016). Accessed April 2022.
48. Kurki, P., Barry, S., Bourges, I., Tsantili, P. & Wolff-Holz, E. Safety, Immunogenicity and interchangeability of biosimilar monoclonal antibodies and fusion proteins: a regulatory perspective. *Drugs* **81**, 1881–1896 (2021).
49. McCamish, M. & Woollett, G. The continuum of comparability extends to biosimilarity: how much is enough and what clinical data are necessary? *Clin. Pharmacol. Ther.* **93**, 315–317 (2013).
50. Webster, C.J., Wong, A.C. & Woollett, G.R. An efficient development paradigm for biosimilars. *BioDrugs* **33**, 603–611 (2019).
51. Liu, J., Eris, T., Li, C., Cao, S. & Kuhns, S. Assessing analytical similarity of proposed Amgen biosimilar ABP 501 to Adalimumab. *BioDrugs* **30**, 321–338 (2016).
52. Saleem, R. *et al*. Analytical and functional similarity assessment of ABP 710, a biosimilar to infliximab reference product. *Pharm. Res.* **37**, 114 (2020).
53. Tailored Scientific Advice for biosimilar development Report on the experience from the pilot (2017–2020). EMA/297462/2021 <https://www.ema.europa.eu/en/documents/report/tailored-scientific-advice-biosimilar-development-report-experience-pilot-2017-2020_en.pdf> (2021). Accessed April 2022.
54. IHC Harmonized Tripartite Guideline Q5E. Comparability of biotechnological/biological products subject to changes in their manufacturing process. IHC Q5E <https://www.ema.europa.eu/en/documents/scientific-guideline/ich-q-5-e-comparability-biotechnological/biological-products-step-5_en.pdf> (2005). Accessed April 2022.
55. Vermeer, N.S., Giezen, T.J., Zastavnik, S., Wolff-Holz, E. & Hidalgo-Simon, A. Identifiability of biologicals in adverse drug reaction reports received from European clinical practice. *Clin. Pharmacol Ther* **105**, 962–969 (2019).
56. Pinheiro, L., Giezen, T.J., Wolff-Holz, E., Weise, M., Laslop, A. & Hidalgo-Simon, A. Identifiability of biologicals: an analysis using the EMA adverse drug reaction database, EudraVigilance. *Clin. Pharmacol Ther* **110**, 1311–1317 (2021).
57. Kurki, P., van Aerts, L., Wolff-Holz, E., Giezen, T., Skibeli, V. & Weise, M. Interchangeability of biosimilars: a European perspective. *BioDrugs* **31**, 83–91 (2017).
58. EMA. Regulatory science to 2025-strategic reflection <https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/ema-regulatory-science-2025-strategic-reflection_en.pdf> (2020).