Breast Cancer Research and Treatment (2019) 176:17–25
https://doi.org/10.1007/s10549-019-05227-7

REVIEW

Randomised clinical endpoint studies for trastuzumab biosimilars: a systematic review

Michael Coory[1] · **Kate Thornton**[1]

Received: 13 February 2019 / Accepted: 8 April 2019 / Published online: 12 April 2019

Abstract

Purpose Uptake of biosimilars depends on clinicians and patients having confidence in the evidential basis of marketing approval. The aim of this systematic review was to assess the evidential role of randomised clinical endpoint studies in the marketing approval of trastuzumab biosimilars.

Methods We searched PubMed for any published reports of randomised studies associated with the five trastuzumab biosimilars approved by the EMA, as on 31 January 2019. We also searched ClinicalTrials.gov for any ongoing studies for other trastuzumab biosimilars.

Results We identified eight published papers or abstracts for seven randomised clinical endpoint studies for five trastuzumab biosimilars approved by the EMA: four studies in the neoadjuvant setting and three in the first-line metastatic setting. Another six unpublished or ongoing studies for other trastuzumab biosimilars were identified via ClinicalTrials.gov. According to GRADE, and considered in isolation, the randomised studies would be categorised as low-quality evidence because of the use of surrogate endpoints and the small sample size. However, according to GRADE, the totality-of-evidence for each of the five approved trastuzumab biosimilars would be categorised as high quality in that further data would be unlikely to change the conclusion that each biosimilar was not different from Herceptin in any clinically important way.

Conclusion The pivotal data for each marketing approval was not the randomised clinical endpoint study, but the in vitro analytic characterisation. Regulatory confidence in in vitro analytic characterisation stems from years of experience with manufacturing changes for originator biological medicines. This emphasis on in vitro data, as the most sensitive way to detect clinically important differences, will be a new way of thinking for many oncologists.

Keywords Biosimilar · Trastuzumab · Marketing approval · In vitro characterisation · Pivotal evidence

Introduction

In 2015, trastuzumab was added to the World Health Organization's list of essential medicines [1]. Unfortunately, despite its extraordinary clinical achievements, trastuzumab is still not routinely available in many middle- and low-income countries [2]. Similar to small-molecule generics in the 1970s and 1980s, it is hoped that biosimilars will make essential biological medicines, like trastuzumab, more affordable and accessible [2, 3]. In high-income countries, where access to trastuzumab is already good, it is hoped that biosimilars will help to make room in constrained third-party-payer budgets for innovative new cancer medicines.

In various high-income countries, the originator trastuzumab medicine, intravenous Herceptin, started coming off patent in 2014 and will be off patent in almost all countries, including the United States, by the end of 2019 [4, 5]. Between December 2017 and January 2019, the European Medicines Agency (EMA) approved five trastuzumab biosimilars (Table 1) [6–10]. These approvals provide an opportunity to assess the role of randomised clinical endpoint studies, across five development programs against the same reference medicine (the intravenous formulation of Herceptin).

For smaller biosimilars, with an established pharmacodynamic endpoint, recent regulatory guidelines state that a randomised clinical study, against the reference (i.e. originator) biological medicine, might not be necessary. For example,

✉ Michael Coory
mcoory@unimelb.edu.au

[1] Faculty of Medicine, Dentistry and Health Sciences, University of Melbourne, Melbourne, Australia

Table 1 Randomised clinical endpoint studies for the five trastuzumab biosimilars approved by the EMA as of January 2019

Biosimilar	Neoadjuvant				First-line metastatic		
	ABP 980 Kanjinti [24]	CT-P6 Herzuma [23]	SB3 Ontruzant [20, 21]	PF-05280014 Trazimera [18]	CT-P6 [17]	MYL-1401O Ogivri [22]	PF-05280014 Trazimera [19]
Study number	NCT 01901146	NCT 02162667	NCT 02149524	NCT 02187744	NCT 01084876	NCT 02472964	NCT 01989676
Company Country	Actavis/ Amgen, US/ Netherlands	Celltrion South Korea	Samsung South Korea	Pfizer US	Celltrion South Korea	Mylan India	Pfizer US
Recruitment period	2013–2015	2014–2016	2014–2015	2014–2016	2010–2012	2012–2015	2014–2016
Trastuzumab regimen	Q3W, 4 cycles	Q3W, 8 cycles	Q3W, 8 cycles	Q3W, 6 cycles	Q3W	Q3W	Q1W
Associated chemotherapy	Paclitaxel Epirubicin Cyclo-phospha-mide	Docetaxel 5FU Epirubicin Cyclo-phospha-mide	Docetaxel 5FU Epirubicin Cyclo-phospha-mide	Docetaxel Carboplatin	Paclitaxel	Paclitaxel or docetaxel	Paclitaxel
Primary end-point	t.pCR	t.pCR	b.pCR	PK($Cmin_{ss}$)	ORR 6 months	ORR 6 months	ORR 6 months
Acceptance interval (RD)	± 13% 90% CI	± 15% 95% CI	± 13% 95% CI	Not specified	± 15% 95% CI	± 15% 95% CI	± 13% 95% CI
Primary analysis population	Per protocol	Per protocol	Per protocol	Per protocol	Intention to treat	Intention to treat	Intention to treat
Women ran-domised	725	549	875	226	475	458	707

for insulin biosimilars, standard clinical endpoints (e.g. HbA1c) are not considered as sensitive, for detecting clinically important differences, as pharmacodynamic endpoints (blood glucose levels) from insulin clamp studies [11]. Similar comments apply to pegylated and non-pegylated versions of filgrastim [12].

For large monoclonal antibodies, such as trastuzumab, randomised clinical endpoint studies remain the norm [13]. We were interested in their evidential role, given that the regulatory guidance of smaller biosimilars (e.g. insulin, filgrastim) has de-emphasised the role of clinical endpoint studies in identifying clinically important differences.

To help clarify this regulatory reasoning, the term sometimes used is "step-wise, totality-of evidence approach". Another term—"reverse pyramid"—is used to highlight that the reasoning for a biosimilar is the reverse of that for a new medicine. For a biosimilar, most of the evidence for marketing approval comes from the in vitro analytic characterisation. In contrast, for a new medicine, the pivotal evidence for marketing approval comes from the clinical endpoint studies.

An associated and important regulatory concept for biosimilars is bridging. In vitro analytic characterisation, along with evidence of pharmacokinetic equivalence, is often sufficient to allow bridging to all the efficacy and safety studies that were completed for the reference medicine. There is no need to independently establish the efficacy and safety of the biosimilar or to repeat all the Phase-3 regulatory studies for the reference medicine: the regulatory reasoning relies on bridging to the clinical evidence developed for the reference medicine [14, 15].

This systematic review only considers the evidence needed for marketing approval. We did not consider the evidence needed to allow switching between a biosimilar and the reference medicine, without reference to the treating doctor; that is, non-medical switching, as is allowed with small-molecule generics. In Australia, Canada and many European Union countries, decisions about whether to allow non-medical switching are the responsibility of third-party payers, not the regulator. In the US, the FDA has two-tiers of evidence: (1) that required for marketing approval. (2) that required for switching [15, 16]. This present paper only considers the first tier of evidence: that required for marketing approval.

Also, we did not compare decisions across agencies in high-income countries; nor ascertain which biosimilars had been submitted to which agencies. More specifically, we did not compare the timing of market access in the United States and the European Union because of the different dates of patent expiry. Our interest was in the evidential basis of the marketing approval, not the legal aspects of market access; although both are important. In particular, our aim with this systematic review is to consider the evidential role

of randomised clinical endpoint studies in the marketing approval of trastuzumab biosimilars.

Methods

We searched PubMed, using the term "trastuzumab biosimilar" in the title or abstract fields. No date range or language was specified. We excluded studies where the only endpoints were pharmacokinetic endpoints (e.g. area-under-the-curve, peak/trough levels at steady state), but included studies where the primary endpoint/s were pharmacokinetic and where there was a secondary clinical endpoint; for example, pathological complete response (pCR), objective response rate (ORR). We also cross checked the references of commentaries and editorials. The identified published articles, in peer-reviewed journals, were used in conjunction with information contained in the publicly available European Public Assessment Reports (EPARs).

To assess the role of any unpublished or ongoing randomised clinical endpoint studies, we also searched ClinicalTrials.gov. The search terms used were intervention = "trastuzumab"; study-type = "intervention"; study-phase = "Phase3".

All searches were current as on 31 January 2019.

Results

PubMed search

This search identified 31 articles; of which 24 were ineligible. Reasons for ineligibility were in vitro analytic characterisation (7); only pharmacokinetic endpoints (6), and commentaries or editorials (11). The remaining eight published articles were reports of seven studies that met the eligibility criteria (Table 1) [17–24].

Two of the articles were abstracts. One was for a 2010–2012 study, in the first-line metastatic setting, for CT-P6 (eventually marketed as Herzuma) [17]. It was not included in the registration dossier for the 2018 EMA approval of Herzuma, due to subsequent (relatively minor) changes in the manufacturing process [17, 25, 26]. The results from the randomised study for Trazimera in the first-line metastatic setting were also only available as an abstract [19], but that study was included in the registration dossier and complete results were available in the EPAR [9].

The inclusion and exclusion criteria were similar across the neoadjuvant studies; and as a result, the baseline characteristics were similar. For example, the average age was about 54 years; 80–90% were ECOG = 0; about 60% to 75% were ER+/PR+; about one-quarter were stage IIIa. For the two metastatic studies, average age was also 54 years; about one-half were ECOG = 0; two-thirds were less than 2 years from initial diagnosis; three-quarters had visceral metastases; 40% were ER+/PR+.

All seven studies were predominantly conducted in middle- and low-income countries (e.g. China, India, South-East Asia, Central/South America, South Africa, Eastern Europe), which have different health systems and different approaches to the treatment of early breast cancer than high-income countries. For example, in the Ogivri study (metastatic setting), only 10% of women had previous (neoadjuvant or adjuvant) treatment with trastuzumab. This low percentage of (neo)adjuvant use of trastuzumab is not surprising, given that one of the main aims of biosimilar development is to improve access in low- and middle-income countries. Also, for high-risk women, pertuzumab was not a possible addition to the regimen.

All of the randomised clinical studies for trastuzumab biosimilars showed non-inferiority, according to the pre-specified non-inferiority margin on the pre-specified primary endpoint; pathological complete response (Table 2) or objective response rate (Table 3). However, at the other end of the acceptance interval, two of the biosimilars, Kanjinti and

Table 2 Total pathological complete response (t.pCR), neoadjuvant studies, per protocol populations

	Kanjinti		Herzuma		Ontruzant		Trazimera	
	Biosimilar	Herceptin	Biosimilar	Herceptin	Biosimilar	Herceptin	Biosimilar	Herceptin
N	358	338	248	256	382	380	100	86
n	172	137	116	129	175	136	47	43
t.pCR (%)	48	41	47	50	46	36	47	50
RD (95% CI)	7.3% (0.0%, 14.6%)		− 3.6% (− 12.4%, 5.2%)		11.1% (4.4%, 17.7%)		− 3.0% (− 17.4%, 11.4%)	
RR (95% CI)	1.19 (1.00, 1.40)		0.93 (0.78, 1.11)		1.32 (1.11, 1.57)		0.92 (− 0.69, 1.23)	

The pre-specified primary analysis populations were the per protocol populations. Analyses on the intention-to-treat populations were concordant (see text in the "Discussion" section)

Table 3 Overall response rate (ORR), first-line metastatic studies, intention-to-treat populations

	Ogivri		Trazimera		CT-P6	
	Biosimilar	Herceptin	Biosimilar	Herceptin	Biosimilar	Herceptin
N	230	228	352	355	244	231
n	160	146	220	236	139	143
ORR (%)	69.6	64.0	62.5	66.5	57.0	61.9
RR (95% CI)	1.09 (0.95, 1.24)		0.94 (0.84, 1.05)		0.92 (0.79, 1.07)	
RD (95% CI)	5.5% (− 3.1, 14.0)		− 4.0 (− 11.0, 3.1)		− 5.0% (− 14%, 4%)	

The pre-specified primary analysis populations were the intention-to-treat populations. Analyses on the per protocol populations were concordant (see text in the "Discussion" section)

Ontruzant, did not show non-superiority (see "Discussion" section).

There were no new or unexpected safety findings from any of the studies. Besides the neoadjuvant study for Trazimera [pharmacokinetic primary endpoint: area-under-curve (AUC)], all the other neoadjuvant studies have reported on safety in the subsequent adjuvant setting. Infusion-related reactions (IRR) were about 8–10% for adjuvant treatment, with similar incidence in the biosimilar and Herceptin groups. Similar percentages were reported for IRRs in the biosimilar studies in the first-line metastatic setting. There were a handful of women (< 1%) who developed antidrug antibodies (ADAs); percentages were similar across the biosimilar and Herceptin groups and not all the ADAs were neutralising. The average LVEF for the women recruited to the studies was 66% (women with heart disease were excluded). Across all the studies, between 1 and 4% of women had a decrease in LVEF of more than 10%, such that their LVEF was less than 50%. Decreases in LVEF seemed similar in the biosimilar and Herceptin groups; with the caveat that the studies were not powered to detect differences in that endpoint. In the LILAC study for Kanjinti, half the women, who received Herceptin in the neoadjuvant period, were switched to Kanjinti in the adjuvant period (171 of 338 women). This switching did not lead to any detectable safety signals.

ClinicalTrials.gov search

This search resulted in 139 hits. Of these, 133 were studies where the intervention was focussed on the various formulations of the reference medicine (i.e. intravenous or subcutaneous trastuzumab, intravenous trastuzumab emtansine). The remaining six studies were randomised clinical endpoint studies for intravenous trastuzumab biosimilars (Table 4).

The results for HERtiCAD (BCD-022) were available from ClinicalTrials.gov (but have not been published in a journal indexed by PubMed). The setting was first-line metastatic breast cancer; 225 women were randomised. The co-primary endpoints were area-under-the-curve: AUC (pharmacokinetic) and objective response rate: ORR (clinical) at 18 weeks. For ORR the results were HERtiCAD (56/113: 50%); Herceptin (48/100: 48%). This suggests equivalence (risk difference = 1.6%; 95% CI − 11.9%, 15.0%) although an acceptance interval (equivalence margin) was not pre-specified. HERtiCAD has marketing approval in Russia and other Eastern European countries. It has not been submitted for evaluation to a regulatory agency in a high-income country or region.

The other five studies identified on ClinicalTrials.gov (summarised in Table 4) are yet to post results or have not started recruiting. However, some of these trastuzumab biosimilars have already been approved for marketing in middle-income and low-income countries: AryoTrust (Iran); HLX02 (China) [27].

We did not identify any trastuzumab biosimilars for which the EMA had refused approval, once the manufacturing site had been cleared. Marketing approval has been refused for some other types of biosimilars (e.g. for interferon alpha-2a, insulin) [28, 29].

Discussion

We identified eight published papers or abstracts in PubMed, reporting on seven studies, for the five approved biosimilars: four studies in the neoadjuvant setting and three in the first-line metastatic setting. Taken on their own, the randomised clinical studies were not designed to provide convincing or definitive evidence of equivalent efficacy. The primary endpoints were surrogates (pCR or ORR) and the sample size was relatively small (100 to 500 in each arm). Also, the studies were not powered for safety endpoints; although, reassuringly, no cardiac, immunogenicity, or other safety signals were identified. The six additional studies, identified via ClinicalTrials.gov, were of similar design to the seven studies for the five approved biosimilars; and are intended to have the same role: supportive, not definitive nor pivotal.

Table 4 Additional randomised clinical endpoint studies identified via ClinicalTrials.gov, as at January 2019

Biosimilar	Neoadjuvant				First-line metastatic	
	AryoGen	HD201	EG12014	TX05	Biocad BCD-022	HLX02
Study number	NCT 03425656	NCT 03013504	NCT 03433313	NCT 03556358	NCT 01764022	NCT 03084237
Company Country	AryoGen Pharmed Iran	Prestige Biopharm S.Korea	EirGenix Taiwan	Tanvex BioPharma USA	Biocad Russia	Shanghai Henlius Biotech China
Recruitment	Complete	Complete	Not started	Continuing	Complete	Continuing
Results	Not posted	Not posted	Not posted	Not posted	Posted	Not posted
Start Recruitment	2016	2016	2018	2018	2012	2017
Trastuzumab regimen	Four cycles Q3W	Eight cycles Q3W	Q3W Cycles not posted	Four cycles Q3W	Q3W	Q3W
Associated chemotherapy	Doxorubicin cyclo-phosphamide	Epirubicin cyclo-phosphamide docetaxel	Epirubicin cyclo-phosphamide	Epirubicin cyclo-phosphamide paclitaxel	Paclitaxel	Docetaxel
Primary endpoint	t.pCR	t.pCR	t.pCR	t.pCR	ORR PK: AUC	ORR
Primary analysis population	Per protocol	Per protocol	Per protocol	Per protocol	intention to treat	Intention to treat
Women (to be) randomised	108	500	800	800	225	600

Setting

For two biosimilars (Herzuma, Trazimera), studies were conducted in both the neoadjuvant and first-line metastatic settings; and the efficacy of each biosimilar was equivalent to Herceptin in both settings. Based on these results, Rugo and co-workers concluded that both the neoadjuvant and first-line metastatic settings are appropriate for assessing efficacy [25].

Some commentators have stated a preference for the using the neoadjuvant setting for biosimilar trastuzumab development [26]. The reasoning is that women having neoadjuvant treatment are a more homogeneous population, given they have localised disease and are treatment naive. In contrast, women in the first-line metastatic setting might have varying tumour burden and metastatic sites; and varying prior treatments. Interestingly, for the two biosimilars studied in the first-line metastatic setting, for which baseline data were reported (Ogivri and Trazimera), only 10% of women had prior treatment with trastuzumab in the neoadjuvant or adjuvant setting. This reflects the lack of access to trastuzumab in the middle- and low-income countries, where these studies are conducted; and may mean that the women recruited to these studies are more homogeneous than originally thought.

Another theoretical concern with the metastatic setting is that these women might be more immune-suppressed than women with early breast cancer; so that the metastatic setting might have less sensitivity for detecting differences in immunogenicity [30, 31]. There was no evidence of this in the studies we identified, with the caveat that the studies were powered for a surrogate efficacy endpoint (ORR), not immunogenicity. It may be that first-line metastatic breast cancer is an appropriate setting, but that post-first-line (beyond progression) might not be an appropriate setting for assessing differences in immunogenicity.

Breast versus total pathological complete response (pCR) for the neoadjuvant studies

Pathological complete response (pCR) is broadly defined as the absence of histological evidence of tumour at the time of surgery. Three of the four studies in the neoadjuvant setting pre-specified total pCR (t.pCR)—absence of invasive cancer in both the breast and the axillary lymph nodes—as the primary endpoint. The other study (for Ontruzant) pre-specified breast pCR (b.pCR) as the primary endpoint; although it also reported t.pCR. In all cases, the definition was agnostic to the presence of ductal carcinoma in situ (DCIS).

Stebbing and co-workers have concluded that, based on the published meta-analyses, absence of invasive cancer in the breast and axillary lymph nodes—t.pCR, (ypT0/is ypN0)—is a stronger prognostic marker of overall survival (OS) than b.pCR [32]. DCIS had no effect on the ability of pCR to predict OS [33]. Briefly put, the available data seem to support the preference for t.pCR over b.pCR in randomised neoadjuvant studies for trastuzumab biosimilars.

Per protocol (PP) versus intention to treat (ITT)

For an equivalence (and non-inferiority) study, it is important to be clear about who was included in the analysis [34]. Traditionally, PP analyses have been preferred because they were thought to be more likely to detect differences. However, a systematic review of non-inferiority studies reported that ITT analyses did not lead to more conclusions of equivalence (or non-inferiority) than PP [35]. One potential problem with PP analyses is that the benefits of randomisation might be lost. Statistical and regulatory guidelines on equivalence and non-inferiority trials recommend that both ITT and PP analyses are reported [36].

For each of the four studies in the neoadjuvant setting, the pre-specified primary analysis was in the PP population. This was variously defined as those women who remained in the study at the time of surgery (the Kanjinti study) or those who had all pre-specified cycles of trastuzumab in the neoadjuvant setting and who remained in the study at the time of surgery (the other 3 neoadjuvant studies: Herzuma, Ontruzant and Trazimera). For all four studies, the ITT results were concordant with the PP results.

For each of the three studies in the metastatic setting, the pre-specified analysis was in the ITT population. No explicit reason was given, but the definition of a PP population can be somewhat arbitrary in the metastatic setting. Possible criteria, which could be used to define a PP population include no major protocol violations (e.g. that would preclude radiological tumour assessment), at least one post-baseline tumour assessment and at least two complete cycles of treatment (excluding patients who have already died or progressed). In any case, for Ogivri and Trazimera, the results of the ITT and PP analyses were concordant. In both cases, the PP population was defined as those women who had no major protocol violations. For CT-P6 (later marketed as Herzuma), only an abstract was available and only the ITT results were reported.

Equivalence margins

Five of the six randomised clinical studies associated with EMA approvals, used a synthesis of results, from various Herceptin studies, to justify the choice of equivalence margin. For the sixth study, the neoadjuvant study of Trazimera, the primary endpoint was area-under-the-curve (AUC); pathological complete response (pCR) was a secondary endpoint and was not associated with an equivalence margin or sample size calculation (Table 1).

As per general regulatory guidelines for non-inferiority and equivalence studies [37], a percentage of the Herceptin treatment effect (over chemotherapy alone) was pre-specified, which, variously, gave an interval for pCR on the arithmetic/absolute scale (risk difference) of ∓12.5% to ∓15%, depending on the trastuzumab biosimilar. The equivalence margins for the ORR in the metastatic setting were also based on retaining a percentage of the Herceptin treatment effect over placebo (Table 1).

Whether the pre-specified margins were appropriate is a matter of clinical judgement [37]. The HannaH study, which compared subcutaneous to intravenous Herceptin in the neoadjuvant setting, used a similar equivalence margin for pCR of ∓12.5% (95% CI). Similar to the trastuzumab biosimilars, a stepwise totality-of-evidence approach was used to justify marketing approval of this different route of administration (subcutaneous versus intravenous) of Herceptin [38].

Non-inferiority and non-superiority

One striking feature of the t.pCR (and b.pCR) results for the four neoadjuvant studies was their consistency in the biosimilar arm across the studies (t.pCR: 46% to 48%; Table 2). This was despite differences in the intervention; that is, number of cycles of trastuzumab (4 to 8) and the chemotherapy backbone (taxane ∓ anthracycline) (Table 1).

In contrast, t.pCR (and b.pCR) varied considerably in the Herceptin arm of the neoadjuvant studies: Ontruzant 36%; Kanjinti 41%; Herzuma 50%; Trazimera 50%. This meant that, although the non-inferiority margin was met, non-superiority was not met for Ontruzant and Kanjinti.

This raises the question of whether non-superiority is a reason for not allowing marketing approval of a biosimilar. Based on the EMA's publicly available EPARs, in both cases, the manufacturers put forward the explanation that shifts in antibody-dependent cellular cytotoxicity (ADCC) in some of the Herceptin lots had affected the efficacy of Herceptin [39]. For Kanjinti, an additional explanation was given: a lack of standardisation for the assessment of pCR (a post hoc, central re-assessment of pCR just met the pre-specified non-superiority limit).

These explanations were accepted by the EMA, suggesting that the Agency was not particularly concerned about non-superiority, especially in the absence of any safety signals. It is unknown whether regulators would have allowed a post hoc justification if non-inferiority had not been met.

Extrapolation to all indications for Herceptin

EMA and FDA guidelines for extrapolation discuss: mechanism of action; pharmacokinetics and distribution; dose; route-of-administration; immunogenicity; toxicity [15, 40].

The mechanism of action of trastuzumab is the same in all the indications (i.e. to inhibit the proliferation of human tumour cells that overexpress HER2). The target receptor involved in the mechanism of action (i.e. HER2) is the same across all indications. Studies on the reference medicine (Herceptin) show that the pharmacokinetics is similar for

all indications. The dosage is the same for all the indications and it is administered by the same route (IV) in all indications. (Currently, no biosimilars have been developed for the subcutaneous formulation of trastuzumab, which is still under patent.) Data for Herceptin show that trastuzumab has low immunogenicity across all indications. The extensive safety information for Herceptin does not indicate that there are any significant differences in expected toxicities for different indications.

Consequently, all five of the trastuzumab biosimilars, with marketing approval in the EU, had all the indications for Herceptin approved: neoadjuvant, adjuvant and metastatic breast cancer; metastatic gastric cancer.

GRADE categories

The Grading of Recommendations Assessment, Development and Evaluation (GRADE) working group has developed a common and transparent approach to grading quality (or certainty) of evidence and strength of recommendations [41]. According to GRADE, the totality-of-evidence for each of the five trastuzumab biosimilars would be categorised as high quality: Further data would be unlikely to change the conclusion that each biosimilar is highly similar to Herceptin, such that there are no clinically meaningful differences.

This confidence is based on the pivotal in vitro analytic characterisation, the scientific foundations of which have been built on the extensive regulatory experience with manufacturing changes for reference biological medicines. For example, Vezer et al. [42] reported that the intravenous formulation of Herceptin, has had 25 moderate- or high-risk manufacturing changes since it was first approved. Regulatory approval of these subsequent changes was based on rigorous in vitro analytic characterisation, which followed the same evidential reasoning as that used for trastuzumab biosimilars. Randomised clinical studies were not carried out for the manufacturing changes for Herceptin.

In contrast, according to GRADE [41], the randomised clinical studies, for each of the trastuzumab biosimilars, if taken in isolation, would be categorised as "low-quality" evidence. More specifically, because they were randomised studies, they would start as "high-quality". They would be downgraded one category to "medium-quality", because of indirectness: the primary endpoint was a surrogate. They would be downgraded a further category to "low-quality" because of imprecision: the sample size was relatively small. The design (i.e. choice of endpoint and sample size) only makes sense in light of the stepwise, totality-of-evident approach.

Use of established patient-relevant efficacy endpoints, such as disease-free survival, progression-free survival and overall survival would have required a larger sample size (in the thousands) and longer follow-up. For example, the PERSEPHONE study, which compared 6 months versus 12 months of adjuvant Herceptin, randomised more than 4000 women and followed them up for more than 4 years (primary endpoint: disease-free survival) [43].

Conclusion

Randomised clinical endpoint studies for trastuzumab biosimilars have been published in major medical journals. Oncologists will be very familiar with evaluating these types of studies. However, unlike the development program for new medicines or new indications, these randomised studies were the not the pivotal evidence for marketing approval. Instead, the pivotal evidence was the in vitro analytic characterisation, the details of which can be found in regulatory evaluations (e.g. European Public Assessment Reports: EPARs); rather than in medical journals. In spite of its unfamiliarity, the regulatory reasoning for biosimilars, which places the main emphasis on in vitro analytic characterisation, is not new. It has been successfully used for manufacturing changes for reference biological medicines for many years [44]. The approach is also used in other regulatory contexts; for example, it was used for marketing approval of the subcutaneous formulation of Herceptin [38]. Clinicians and patients can have confidence in the regulatory approach.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants performed by any of the authors.

Informed consent All the studies included in this systematic review were listed in ClinicalTrials.gov, which confirmed that written informed consent was obtained from all individual participants included in the studies.

References

1. WHO Essential Medicines List 20th Edition. http://www.who.int/medicines/publications/essentialmedicines/en/
2. Blackwell K, Gligorov J, Jacobs I, Twelves C (2018) The global need for a trastuzumab biosimilar for patients with HER2-positive breast cancer. Clin Breast Cancer 18(2):95–113
3. Gottlieb S (2018) Capturing the benefits of competition for patients. In: Keynote address by Commissioner Gottlieb to the 2018 FDLI annual conference 2018
4. Camacho LH (2017) Current status of biosimilars in oncology. Drugs 77(9):985–997
5. Nelson KM, Gallagher PC (2014) Biosimilars lining up to compete with Herceptin–opportunity knocks. Expert Opin Ther Pat 24(11):1149–1153

6. European Medicines Agency (EMA) (2018) European Public Assessment Report. Ontruzant (trastuzumab) Procedure No. EMEA/H/C/004323/0000
7. European Medicines Agency (EMA) (2018) European Public Assessment Report. Kanjinti (trastuzumab) Procedure No. EMEA/H/C/004361/0000
8. European Medicines Agency (EMA) (2018) European Public Assessment Report. Herzuma (trastuzumab) Procedure No.EMEA/H/C/002575/0000
9. European Medicines Agency (EMA) (2018) European Public Assessment Report. Trazimera (trastuzumab) Procedure No.EMEA/H/C/004463/0000
10. European Medicines Agency (EMA) (2019) European Public Assessment Report. Ogivri (trastuzumab) Procedure No. EMEA/H/C/004916/0000
11. European Medicines Agency (EMA) (2015) 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 Committee for Medicinal products for Human Use (CHMP)
12. European Medicines Agency (EMA) (2018) Guideline on similar biological medicinal products 4 containing recombinant granulocyte-colony stimulating 5 factor (rG-CSF) DRAFT: EMEA/CHMP/BMWP/31329/2005 Rev 1.2 Committee for Medicinal Product for Human Use (CHMP)
13. European Medicines Agency (EMA) (2012) Guideline on similar biological medicinal products containing monoclonal antibodies—non-clinical and clinical issues: EMA/CHMP/BMWP/403543/2010 Committee for Medicinal Products for Human Use (CHMP)
14. Weise M, Bielsky MC, De Smet K, Ehmann F, Ekman N, Giezen TJ, Gravanis I, Heim HK, Heinonen E, Ho K et al (2012) Biosimilars: what clinicians should know. Blood 120(26):5111–5117
15. Lemery SJ, Ricci MS, Keegan P, McKee AE, Pazdur R (2017) FDA's approach to regulating biosimilars. Clin Cancer Res 23(8):1882–1885
16. Food and Drug Administration (FDA) (2017) Considerations in demonstrating interchangeability with a reference product, Guidance for Industry, Draft. In: U.S. Department of Health and Human Services, Food and Drug Administration (FDA), Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER)
17. Im Y-H, Odarchenko P, Grecea D, Komov D, Anatoliy CV, Gupta S (2013) Double-blind, randomized, parallel group, phase III study to demonstrate equivalent efficacy and comparable safety of CT-P6 and trastuzumab, both in combination with paclitaxel, in patients with metastatic breast cancer (MBC) as first-line treatment. J Clin Oncol 31(5):629
18. Lammers PE, Dank M, Masetti R, Abbas R, Hilton F, Coppola J, Jacobs I (2018) Neoadjuvant PF-05280014 (a potential trastuzumab biosimilar) versus trastuzumab for operable HER2+ breast cancer. Br J Cancer 119(3):266–273
19. Pegram M, Tan-Chiu E, Freyman A, Vana A, Hilton F, Zacharchuk C, Ewesuedo R (2017) A randomized, double-blind study of PF-05280014 (a potential trastuzumab biosimilar) vs trastuzumab, both in combination with paclitaxel, as first-line treatment for HER2-positive metastatic breast cancer (Abstract 238PD). Ann Oncol 28:v74. https://doi.org/10.1093/annonc/mdx1365
20. Pivot X, Bondarenko I, Nowecki Z, Dvorkin M, Trishkina E, Ahn JH, Im SA, Sarosiek T, Chatterjee S, Wojtukiewicz MZ et al (2018) A phase III study comparing SB3 (a proposed trastuzumab biosimilar) and trastuzumab reference product in HER2-positive early breast cancer treated with neoadjuvant-adjuvant treatment: final safety, immunogenicity and survival results. Eur J Cancer 93:19–27
21. Pivot X, Bondarenko I, Nowecki Z, Dvorkin M, Trishkina E, Ahn JH, Vinnyk Y, Im SA, Sarosiek T, Chatterjee S et al (2018) Phase III, randomized, double-blind study comparing the efficacy, safety, and immunogenicity of SB3 (trastuzumab biosimilar) and reference trastuzumab in patients treated with neoadjuvant therapy for human epidermal growth factor receptor 2-positive early breast cancer. J Clin Oncol 36(10):968–974
22. Rugo HS, Barve A, Waller CF, Hernandez-Bronchud M, Herson J, Yuan J, Sharma R, Baczkowski M, Kothekar M, Loganathan S et al (2017) Effect of a proposed trastuzumab biosimilar compared with trastuzumab on overall response rate in patients with ERBB2 (HER2)-positive metastatic breast cancer: a randomized clinical trial. JAMA 317(1):37–47
23. Stebbing J, Baranau Y, Baryash V, Manikhas A, Moiseyenko V, Dzagnidze G, Zhavrid E, Boliukh D, Stroyakovskii D, Pikiel J et al (2017) CT-P6 compared with reference trastuzumab for HER2-positive breast cancer: a randomised, double-blind, active-controlled, phase 3 equivalence trial. Lancet Oncol 18(7):917–928
24. von Minckwitz G, Colleoni M, Kolberg HC, Morales S, Santi P, Tomasevic Z, Zhang N, Hanes V (2018) Efficacy and safety of ABP 980 compared with reference trastuzumab in women with HER2-positive early breast cancer (LILAC study): a randomised, double-blind, phase 3 trial. Lancet Oncol 19(7):987–998
25. Rugo HS, Curigliano G, Cardoso F, Gradishar WJ, Pegram M, Barrios CHH, CortesCastan J, Pennella E, Muniz R (2018) 324P settings-based efficacy comparison of trastuzumab biosimilars in breast cancer: a systematic literature review. Ann Oncol 29(suppl_8):mdy272.314–mdy272.314
26. Pivot X, Petit T (2018) Can we establish a hierarchy among trastuzumab biosimilar candidates? Br J Cancer 119(3):263–265
27. Biosimilars of trastuzumab. http://www.gabionline.net/Biosimilars/General/Biosimilars-of-trastuzumab
28. European Medicines Agency (EMA) (2006) Refusal Assessment Report Alpheon (recombinant human interferon-alfa-2a) Procedure No. EMEA/H/C/000585
29. European Medicines Agency (EMA) (2015) Refusal Assessment Report. Solumarv (human insulin). Procedure No. EMEA/H/C/003858/0000
30. Jackisch C, Scappaticci FA, Heinzmann D, Bisordi F, Schreitmuller T, Minckwitz G, Cortes J (2015) Neoadjuvant breast cancer treatment as a sensitive setting for trastuzumab biosimilar development and extrapolation. Fut Oncol 11(1):61–71
31. Cortes J, Curigliano G, Dieras V (2014) Expert perspectives on biosimilar monoclonal antibodies in breast cancer. Breast Cancer Res Treat 144(2):233–239
32. Stebbing J, Baranau Y, Manikhas A, Lee SJ, Thiruchelvam P, Leff D, Esteva FJ (2018) Total pathological complete response versus breast pathological complete response in clinical trials of reference and biosimilar trastuzumab in the neoadjuvant treatment of breast cancer. Expert Rev Anticancer Ther 18(6):531–541
33. von Minckwitz G, Untch M, Blohmer JU, Costa SD, Eidtmann H, Fasching PA, Gerber B, Eiermann W, Hilfrich J, Huober J et al (2012) Definition and impact of pathologic complete response on prognosis after neoadjuvant chemotherapy in various intrinsic breast cancer subtypes. J Clin Oncol 30(15):1796–1804
34. DeMets DL, Cook T (2018) Challenges of non-intention-to-treat analyses. JAMA 321:145
35. Wangge G, Klungel OH, Roes KC, de Boer A, Hoes AW, Knol MJ (2010) Room for improvement in conducting and reporting non-inferiority randomized controlled trials on drugs: a systematic review. PLoS ONE 5(10):e13550
36. Rehal S, Morris TP, Fielding K, Carpenter JR, Phillips PP (2016) Non-inferiority trials: are they inferior? A systematic review of reporting in major medical journals. BMJ Open 6(10):e012594

37. Mauri L, D'Agostino RB Sr (2017) Challenges in the design and interpretation of noninferiority trials. N Engl J Med 377(14):1357–1367
38. Ismael G, Hegg R, Muehlbauer S, Heinzmann D, Lum B, Kim SB, Pienkowski T, Lichinitser M, Semiglazov V, Melichar B et al (2012) Subcutaneous versus intravenous administration of (neo) adjuvant trastuzumab in patients with HER2-positive, clinical stage I-III breast cancer (HannaH study): a phase 3, open-label, multicentre, randomised trial. Lancet Oncol 13(9):869–878
39. Kim S, Song J, Park S, Ham S, Paek K, Kang M, Chae Y, Seo H, Kim HC, Flores M (2017) Drifts in ADCC-related quality attributes of Herceptin(R): impact on development of a trastuzumab biosimilar. MAbs 9(4):704–714
40. Weise M, Kurki P, Wolff-Holz E, Bielsky MC, Schneider CK (2014) Biosimilars: the science of extrapolation. Blood 124(22):3191–3196
41. Guyatt G, Vist G, Falck-Ytter Y, Kunz R, Magrini N, Schunemann H (2006) An emerging consensus on grading recommendations? ACP J Club 144(1):A8–9
42. Vezer B, Buzas Z, Sebeszta M, Zrubka Z (2016) Authorized manufacturing changes for therapeutic monoclonal antibodies (mAbs) in European Public Assessment Report (EPAR) documents. Curr Med Res Opin 32(5):829–834
43. Hiller L, Dunn JA, Loi S, Vallier AL, Howe DL, Cameron DA, Miles D, Wardley AM, Earl HM (2018) Adjuvant trastuzumab duration trials in HER2 positive breast cancer—what results would be practice-changing? Persephone investigator questionnaire prior to primary endpoint results. BMC Cancer 18(1):391
44. Lemery SJ (2017) When one is a hammer, everything looks like a nail. J Oncol Pract 13:10s–11s