

Comparative Efficacy of Ibrutinib Versus Obinutuzumab + Chlorambucil in First-Line Treatment of Chronic Lymphocytic Leukemia: A Matching-Adjusted Indirect Comparison

Suzy Van Sanden · Simona Baculea · Joris Diels · Sarah Cote

Received: April 3, 2017 / Published online: June 1, 2017
© The Author(s) 2017. This article is an open access publication

ABSTRACT

Introduction: Ibrutinib (ibr) monotherapy and the combination of obinutuzumab plus chlorambucil (obi) are approved for previously untreated chronic lymphocytic leukemia (CLL). No trials directly comparing their efficacy are available. Therefore a matching-adjusted indirect comparison (MAIC) was performed to provide insight into their relative efficacy in terms of progression-free survival (PFS) and overall survival (OS). MAIC attempts to adjust for between-trial differences in factors known or suspected to influence treatment effects, to minimize bias.

Methods: A MAIC within a Bayesian framework was conducted using individual patient data from the RESONATE-2 study of ibr versus chlorambucil and published data from the CLL11

study of obi versus chlorambucil. Both studies were conducted in patients ineligible for full-dose fludarabine-based therapy. After matching, the reweighted adjusted relative efficacy measure of ibr versus chlorambucil from RESONATE-2 [hazard ratio (HR), 95% credible interval (CrI)] was compared with that of obi versus chlorambucil from CLL11 for each endpoint, using a Bayesian indirect comparison.

Results: Our results suggest that in a population with similar average baseline characteristics to CLL11, ibr would improve PFS and OS outcomes compared to obi. Before matching, the HRs for ibr versus obi were 0.48 [CrI = 0.22–1.02, $p(\text{HR} < 1) = 97\%$], 0.85 [CrI = 0.44–1.63, $p(\text{HR} < 1) = 69\%$], and 0.40 [CrI = 0.10–1.54, $p(\text{HR} < 1) = 91\%$] for PFS by investigator assessment, PFS by independent review committee, and OS, respectively. After matching on all available characteristics the HRs decreased to 0.12 [CrI = 0.02–0.97, $p(\text{HR} < 1) = 98\%$], 0.24 [CrI = 0.04–1.35, $p(\text{HR} < 1) = 95\%$], and 0.21 [CrI = <0.01–8.89, $p(\text{HR} < 1) = 79\%$], respectively. There was a large variance around the treatment effect for OS due to the low number of deaths.

Conclusion: Our analysis suggests that ibrutinib is highly likely to provide greater PFS benefit than obinutuzumab plus chlorambucil in older or less fit patients with previously untreated CLL. There is also an indication of improvement in OS, albeit with a higher uncertainty due to the low number of events.

Enhanced content To view enhanced content for this article go to <http://www.medengine.com/Redeem/7498F060452AF42B>.

Electronic supplementary material The online version of this article (doi:10.1007/s12325-017-0564-1) contains supplementary material, which is available to authorized users.

S. Van Sanden (✉) · J. Diels
Janssen-Cilag, Beerse, Belgium
e-mail: SVSANDE1@ITS.JNJ.COM

S. Baculea · S. Cote
Janssen-Cilag, High Wycombe, Buckinghamshire,
UK

Funding: Janssen-Cilag Ltd.

Keywords: Chlorambucil; Chronic lymphocytic leukemia; CLL; Hematology; Ibrutinib; Indirect comparison; Obinutuzumab; Oncology; Overall survival; Progression-free survival

INTRODUCTION

Chronic lymphocytic leukemia (CLL) has a median age at diagnosis of 72 years, and many patients are not fit enough to be eligible for chemoimmunotherapy regimens [1]. In recent years several novel agents have been approved for use in previously untreated CLL, but there is little direct evidence comparing their efficacy in this population. Head-to-head randomized controlled trials (RCTs) remain the most valid way to compare two agents. However, where such trials are not available, indirect comparisons can provide useful insights into the relative efficacy of treatments, both to clinicians and to inform reimbursement decisions.

Ibrutinib monotherapy was approved in 2014 by the US Food and Drug Administration (FDA) [2] and the European Medicines Agency (EMA) [3] for use in patients with previously treated CLL or a 17p deletion. In 2016 both authorities extended this indication to include adult patients with previously untreated CLL [4, 5], based on the results of the RESONATE-2 study (NCT01722487) [6].

Obinutuzumab was approved in 2013 by the FDA for use in combination with chlorambucil for the treatment of previously untreated CLL [7], and by the EMA in 2014 for adult patients with previously untreated CLL and comorbidities making them unsuitable for full-dose fludarabine-based therapy [8]. The pivotal trial for obinutuzumab in CLL was the CLL11 study (NCT01010061) [9].

The European Society of Medical Oncology (ESMO) guidelines [10, 11] recommend that patients with previously untreated CLL without del17p/TP53 mutation who are not sufficiently fit for fludarabine-based chemoimmunotherapy should receive chlorambucil + a

CD20 antibody, or ibrutinib. In the National Comprehensive Cancer Network (NCCN) guidelines from the USA [12], these are the only therapies given a category 1 recommendation in this population. Ibrutinib is also recommended as an option when del17p/TP53 mutation is present. Both ibrutinib monotherapy and the CD20 antibody obinutuzumab in combination with chlorambucil are associated with a survival improvement vs. chlorambucil alone, although obinutuzumab + chlorambucil is also associated with increased toxicity [9].

To date, the efficacy of ibrutinib and obinutuzumab + chlorambucil has not been directly compared in an RCT. Unadjusted or “naïve” indirect comparisons of treatment arms from different trials are prone to bias due to heterogeneity in patient populations. Such methods are not appropriate where there is a connected evidence network. Adjusted indirect treatment comparison methods, such as the Bucher method [13], involve comparing relative treatment effects across trials versus a common comparator. They preserve randomization and provide unbiased comparisons under the assumption that populations do not differ between studies with respect to treatment effect modifiers. However, if between-trial differences include factors that do influence the relative treatment effect, results from such indirect treatment comparisons can also be biased.

Matching-adjusted indirect comparison (MAIC) is a method originally developed by Signorovitch et al. [14] to adjust for potential biases related to treatment effect modifiers in traditional adjusted indirect comparisons. It attempts to make trial populations comparable by applying weightings to a trial population for which patient level data are available, so that it mimics the population of the other trial, for which only aggregate results are available [14, 15].

This paper attempts to compare the relative efficacy of ibrutinib with that of obinutuzumab + chlorambucil in patients with previously untreated CLL who are not eligible for full-dose fludarabine, using data from RESONATE-2 [6] and CLL11 [9]. Both trials were conducted in this population but differ with respect to some important baseline

characteristics. In CLL11, only patients with a Cumulative Illness Rating Scale (CIRS) score greater than 6 and/or creatinine clearance less than 70 mL/min were eligible, resulting in a lower proportion of patients with CIRS score of 6 or less compared to RESONATE-2 (26% vs. 64%, respectively). Patients with higher CIRS scores are considered to be less fit [11]. A retrospective study by Manda et al. suggested that higher CIRS score and age greater than 70 years are associated with worse overall survival outcomes [16]. In addition, higher baseline β 2-microglobulin and higher Binet stage are also recognized prognostic factors [11, 12, 17, 18] and are thus potential treatment effect modifiers. The RESONATE-2 trial includes fewer patients with age at least 75 years and more patients with β 2-microglobulin levels of at least 3.5 mg/L at baseline than the CLL11 study. Because of the potential bias arising from baseline differences, we conducted a MAIC using the methodology described by Signorovitch et al. [19], matching on all available, clinically relevant baseline characteristics.

METHODS

A MAIC within a Bayesian framework [20, 21] was conducted using individual patient data (IPD) for the RESONATE-2 study of ibrutinib vs. chlorambucil [6], and the data published by Goede et al. [9] for comparison of obinutuzumab + chlorambucil vs. chlorambucil in the CLL11 study.

Median follow-up for the progression-free survival (PFS) and overall survival (OS) analyses was not explicitly reported in CLL11, and it was estimated to range between 10 and 37 months [9]; a median follow-up of 18.4 months was reported for the final analyses of RESONATE-2 [6] and used in these analyses. Of note, longer follow-up survival data from RESONATE-2 were recently published [22, 23]; however, by this cutoff point there was significant crossover of patients from the chlorambucil arm to ibrutinib arm (41%), breaking randomization in the trial, and therefore biasing the OS estimate against ibrutinib. Therefore, the data cut as per final analyses reported in [6] was used for the

analysis, by which time a lower percentage of patients crossed over to ibrutinib. All analyses were conducted using SAS 9.2 (Cary, NC, USA).

Matching Process

The MAIC process was performed using the algorithm of Signorovitch et al. [14], which can be divided into two steps described below. Some of the technical details for the different steps are explained in the Supplementary Material.

Step 1: Updating HRs for RESONATE-2

In a first step, the treatment effect of ibrutinib vs. chlorambucil is estimated using RESONATE-2 patient level data, simulating the HR that would have been obtained if RESONATE-2 had included a patient population identical to the CLL11 population. This is implemented by first excluding patients from RESONATE-2 that did not satisfy the CLL11 inclusion/exclusion criteria: patients with a CIRS score of at most 6 and creatinine clearance of at least 70 mL/min, patients with creatinine clearance less than 30 mL/min, and patients with a diagnosis of small lymphocytic lymphoma (SLL) were excluded. To the remaining patients from RESONATE-2, weights were assigned in such a way that the reweighted population resembles (or “matches”) the trial population from CLL11 as close as possible. The approach is very similar to propensity score weighting, whereby patients from RESONATE-2 were weighted by the inverse odds of belonging to RESONATE-2 rather than to CLL11.

A weighted Cox regression model was used to estimate the HR for ibrutinib versus chlorambucil on the reweighted RESONATE-2 trial data for PFS by investigator assessment (the primary endpoint of CLL11), PFS by independent review committee (IRC; the primary endpoint of RESONATE-2), and OS.

For the base case analysis, all clinically relevant risk factors reported in [9] that may influence the relative treatment effect were included in the matching process: CIRS score, age, Binet stage, β 2-microglobulin, del11q, ECOG performance status, creatinine clearance, gender, and having unmutated IGHV.

However, the inclusion of more variables in the matching process increases the uncertainty in the end result. Sensitivity analyses were performed, whereby the baseline variables considered to be less important on the basis of clinical input were left out of the matching process in a stepwise manner. The effective sample size (Neff, see Supplementary Material), a measure used to quantify the impact of matching more variables on the uncertainty, will be reported for the main and sensitivity analyses (the smaller the Neff, the higher the uncertainty).

Step 2: Bayesian Indirect Comparison

In a second step, the updated HR for RESONATE-2 from step 1 was combined with the reported HR from the CLL11 study (obinutuzumab + chlorambucil vs. chlorambucil) to estimate the HR for ibrutinib (ibr) versus obinutuzumab + chlorambucil (obi), using the approach of Bayesian indirect comparison [21], with chlorambucil (chl) as the common comparator across both trials. The point estimate for the HR of ibr vs. obi generated by such indirect comparison is simply the ratio of the two original hazard ratios (HR ibr vs. chl divided by HR obi vs. chl).

The Bayesian approach generates a probability distribution around this point estimate. The probability that one treatment is more effective than the other can be derived from this distribution, which is a useful outcome measure in a decision-making context.

This article is based on previously conducted studies, and does not involve any new studies of human or animal subjects performed by any of the authors.

RESULTS

A total of 269 patients were randomized in RESONATE-2; 136 received ibrutinib and 133 received chlorambucil.

Matching Process

Baseline characteristics before and after adjustment of the RESONATE-2 population to match

the aggregate baseline characteristics of the CLL11 study are presented in Table 1. A total of 13 variables pertaining to nine baseline characteristics were matched. They are reported here in order of relevance to the outcome, except unmutated IGHV, which is placed last because of a larger number of missing values in RESONATE-2.

The del17p mutation was present in 6% of CLL11 patients but as patients with del17p were excluded from RESONATE-2 it was not possible to match on this characteristic. However, the impact of this difference is likely to be minimal because of the small number of patients with del17p in CLL11.

Excluding patients from RESONATE-2 who did not meet the inclusion criteria of CLL11 reduced the sample size used for matching to 191 patients. When matching on all available baseline variables the analysis population was reduced to 115 because some patients had missing values for certain characteristics.

After reweighting, the average characteristics from the reweighted RESONATE-2 population were virtually identical to the patient characteristics from CLL11, as would be expected when applying the MAIC algorithm. The effective sample size for the reweighted RESONATE-2 population was 35.

The variable “unmutated IGHV” had the most missing values in RESONATE-2. When this variable was removed from the matching analysis, 152 patients were included in the matching process and the effective sample size increased to 48 (see Table 1). Even though the matching algorithm did not include this variable, the percentage of patients with unmutated IGHV did not differ greatly between the two groups (61% in CLL11 versus 57% in the weighted RESONATE-2 population) (Table 1). An equivalent table showing the full set of stepwise sensitivity analyses is shown in the Supplementary Material.

Updated HRs for Reweighted RESONATE-2 Population

Results are shown in Table 2 for the following scenarios: before matching, based on randomized treatment arms (ITT); before matching,

Table 1 Baseline characteristics of the CLL11 and RESONATE-2 trial populations before and after matching for base case and sensitivity analysis for the removal of IGHV status from matching

	Trial data		MAIC results		
	CLL11 [9]	RESONATE-2	RESONATE-2		
	ITT	ITT	After exclusion ^a	Base case matched (<i>n</i> = 13 variables)	SA matched (<i>n</i> = 12 variables)
N (Neff)	589	269	191	115 (35)	152 (48)
CIRS score (median)	8	5	6	8	8
CIRS score ≤6 (%)	26	64	56	26	26
Age (median)	73	73	73	73	73
Age (≥75 years) (%)	43	35	40	43	43
Binet stage A (%)	22	19	19	22	22
Binet stage B (%)	42	43	40	42	42
Binet stage C (%)	36	38	41	36	36
β2-Microglobulin ≥3.5 mg/L (%)	35	71	77	35	35
del11q (%)	17	22	24	17	17
ECOG (median)	1	1	1	1	1
Creatinine clearance (median, mL/min)	62	61.21	55.35	62.02	62.02
Male (%)	62	63	59	62	62
Unmutated IGHV (%)	61	59	58	61	57 ^b

CIRS Cumulative Illness Rating Scale, ECOG Eastern Cooperative Oncology Group, IGHV immunoglobulin heavy variable cluster, ITT intention to treat population, *N* number of patients analyzed, *n* number of variables matched, *Neff* effective sample size, SA sensitivity analysis

^a Excludes patients from RESONATE-2 with CIRS ≤6 and creatinine clearance ≥70 mL/min, patients with creatinine clearance <30 mL/min, and SLL patients

^b Baseline characteristic that was not matched and hence value represents the average for the characteristic unadjusted

after exclusion of patients from RESONATE-2 who would not have been eligible for inclusion in CLL11; after matching on 13 baseline characteristic variables (base case analysis); and after matching on 12 variables (sensitivity analysis excluding unmutated IGHV from the matching).

The results of the stepwise sensitivity analyses are provided in the Supplementary Material.

The within-trial hazard ratios (HRs) versus chlorambucil for each endpoint assessed are shown in Table 2. For ibrutinib vs.

chlorambucil, the HRs were obtained using the weighted Cox model.

The treatment effects on PFS with ibrutinib were statistically significantly greater than those with chlorambucil across all scenarios. Matching on 12 and 13 variables resulted in lower HRs (i.e., more favorable to ibrutinib) than those derived from the intention to treat population of RESONATE-2 before matching. The HRs for obinutuzumab + chlorambucil versus chlorambucil alone are taken directly from CLL11 as reported by Goede et al. [9].

Table 2 Within-trial comparison of progression-free and overall survival vs. chlorambucil for ibrutinib (RESONATE-2, derived using the weighted Cox model) and obinutuzumab + chlorambucil (CLL11)

		N (Neff)	Treatment vs. chlorambucil	HR			
				[95% CI]	PFS by IRC assessment	PFS by investigator assessment	Overall survival
					P value		
Trial data	CLL11	589	Obinutuzumab + chlorambucil	0.19	0.18	0.41	
				[0.14, 0.27]	[0.13, 0.24]	[0.23, 0.74]	
				<0.0001	<0.001	0.002	
	RESONATE-2 (ITT)	269	Ibrutinib	0.161	0.086	0.163	
				[0.091, 0.283]	[0.043, 0.173]	[0.048, 0.557]	
				<0.0001	<0.0001	0.0038	
MAIC results	RESONATE-2 (after exclusion) ^a	191	Ibrutinib	0.095	0.055	0.065	
				[0.044, 0.205]	[0.022, 0.139]	[0.008, 0.491]	
				<0.0001	<0.0001	0.0081	
	RESONATE-2 (matched n = 13)	115 (35)	Ibrutinib	0.045	0.022	0.085	
				[0.008, 0.245]	[0.003, 0.182]	[0.002, 3.447]	
				0.0003	0.0004	0.1922	
	RESONATE-2 (matched n = 12)	152 (48)	Ibrutinib	0.065	0.078	0.081	
				[0.021, 0.204]	[0.030, 0.204]	[0.003, 1.872]	
				<0.0001	<0.0001	0.1167	

CI confidence interval, HR hazard ratio, ITT intention to treat population, N number of patients analyzed, n number of variables matched, Neff effective sample size, PFS progression-free survival

^a Excludes patients from RESONATE-2 with CIRS ≤6 and creatinine clearance ≥70 mL/min, patients with creatinine clearance <30 mL/min, and SLL patients

For investigator assessment, the MAIC-adjusted results in the base case (after matching on 13 variables) and the sensitivity analysis adjusting for 12 variables suggest that in a population with similar average baseline characteristics to that of CLL11, ibrutinib would reduce the risk of progression or death by 98% and 92%, respectively, compared with chlorambucil. Using the PFS by IRC results the suggested reductions were similar, at 95% and 93% respectively. In the stepwise sensitivity analyses, where fewer variables were matched and effective sample sizes were consequently greater, HRs ranged from 0.080 to 0.115 (see Supplementary

Material), indicating that the findings were not markedly influenced by the number of variables used. In general, matching on fewer variables led to somewhat larger HRs.

A large decrease in the HR for OS, favoring ibrutinib, was observed after patients not eligible for inclusion in CLL11 were excluded: the HR before exclusion was 0.163 [95% CrI 0.048, 0.557] compared to 0.065 [95% CrI 0.008, 0.491] after exclusion; this pronounced effect on OS was maintained after matching on 13 variables (HR 0.085 [95% CrI 0.002, 3.447]).

The results for 13 and 12 variables matched suggest that ibrutinib reduces the risk of death

by approximately 92% in both analyses, compared with chlorambucil. The wide confidence intervals for OS indicated that the differences in treatment effects were not statistically significant for the analyses where between eight and 13 variables were matched (effective sample sizes 35–59; see Supplementary Material for the full set of analyses). This is likely due to the low number of deaths observed in the trial and, consequently, even fewer events in the effective sample after matching. Of note, when only seven variables were matched the effective sample size rose to 119, the HR became statistically significant (HR = 0.052; 95% CI 0.003, 0.804; $p = 0.0343$; see Supplementary Material).

Bayesian Indirect Comparison

The between-trial HRs for PFS and OS obtained from the MAIC are shown in Table 3. All the HRs in the Bayesian indirect comparison favor ibrutinib over obinutuzumab + chlorambucil, with a probability that ibrutinib is superior to obinutuzumab + chlorambucil in terms of improving PFS or OS (i.e., HR is less than 1) of at least 79% in all scenarios after matching.

For PFS by IRC assessment, the results before matching using the ITT populations suggest that ibrutinib is associated with a greater improvement in PFS compared with obinutuzumab + chlorambucil, with an HR of 0.85 [95% CrI 0.44, 1.63; $P(\text{HR} < 1)$ 69%]. After applying the exclusion criteria of the CLL11 study, the HR decreased to 0.5 with a higher probability of ibrutinib having a greater improvement in PFS (IRC) than obinutuzumab + chlorambucil (95%). After matching on all available baseline variables, the HR decreased further to 0.24 [95% CrI 0.04, 1.35; $P(\text{HR} < 1)$ 95%]. For the sensitivity analysis matching on 12 baseline variables, the HR was similar to the base case (HR 0.34; 95% CrI 0.11, 1.12; $P(\text{HR} < 1)$ 96%).

For PFS by investigator assessment, the results for the ITT population also suggest that ibrutinib is associated with a greater improvement in PFS compared with obinutuzumab + chlorambucil, with an HR of 0.48 [95% CrI 0.22, 1.02; $P(\text{HR} < 1)$ 97%]. The HR in

this case is lower than for PFS by IRC assessment. The HR decreased to 0.31 after exclusion of patients on the basis of the CLL11 eligibility criteria and further decreased (0.12; 95% CrI 0.02, 0.97; $P(\text{HR} < 1)$ 98%) after matching on all available baseline characteristics, as observed for the PFS IRC assessment. The analyses support the findings from the ITT analysis suggesting that ibrutinib reduces the risk of progression or death compared with obinutuzumab + chlorambucil, with an associated probability of 98% that it is superior to obinutuzumab + chlorambucil after matching on 13 baseline variables. For the sensitivity analysis matching on 12 baseline variables, the HR was 0.43 (95% CrI 0.16, 1.18), with a 95% probability that ibrutinib is superior to obinutuzumab + chlorambucil in terms of improving investigator-assessed PFS.

The results for OS based on the ITT population suggest that ibrutinib reduces the risk of death with an associated HR of 0.40 [95% CrI 0.10, 1.54; $P(\text{HR} < 1)$ 0.91]. After exclusion of patients and matching, the HRs for OS all favor ibrutinib, with probabilities of ibrutinib being superior to obinutuzumab + chlorambucil in improving OS ranging from 95% (after exclusion of patients not eligible for CLL11) to 79% after matching on 13 variables. When 13 baseline variables were matched the results suggest that ibrutinib improves OS compared to obinutuzumab + chlorambucil, with an HR of 0.21 (95% CrI 0.00, 8.89) and an associated probability of being superior to obinutuzumab + chlorambucil of 79%. For the sensitivity analysis (when matching on 12 variables), the results are similar, with the probability of ibrutinib being superior to obinutuzumab + chlorambucil increasing to 83% (see Supplementary Material for full set of results).

DISCUSSION

The matching-adjusted indirect comparison provides a means of comparing the effects of two novel agents for the first-line treatment of CLL, ibrutinib monotherapy and obinutuzumab + chlorambucil, in the absence of head-to-head trials comparing them directly. MAIC allows for indirect comparisons to be

Table 3 Between-trial comparison of progression-free and overall survival for ibrutinib vs. obinutuzumab + chlorambucil (Bayesian approach)

Population	HR [95% CrI] <i>P</i> (HR <1)		
	PFS by IRC assessment	PFS by investigator assessment	Overall survival
RESONATE-2 (ITT)	0.85 [0.44, 1.63]	0.48 [0.22, 1.02]	0.40 [0.10, 1.54]
	0.69	0.97	0.91
RESONATE-2 (after exclusion) ^a	0.50 [0.22, 1.15]	0.31 [0.12, 0.81]	0.16 [0.02, 1.34]
	0.95	0.99	0.95
RESONATE-2 (matched <i>n</i> = 13)	0.24 [0.04, 1.35]	0.12 [0.02, 0.97]	0.21 [0.00, 8.89]
	0.95	0.98	0.79
RESONATE-2 (matched <i>n</i> = 12)	0.34 [0.11, 1.12]	0.43 [0.16, 1.18]	0.20 [0.01, 5.16]
	0.96	0.95	0.83

CrI credible interval, *HR* hazard ratio, *ITT* intention to treat population, *n* number of variables matched, *P*(*HR* <1) probability that the HR is less than 1, i.e., ibrutinib is better than obinutuzumab + chlorambucil

^a Excludes patients from RESONATE-2 with CIRS ≤6 and creatinine clearance ≥70 mL/min, patients with creatinine clearance <30 mL/min, and SLL patients

made via a common comparator (in this case chlorambucil monotherapy) and reduces potential bias due to imbalances in treatment effect modifiers between trials by adjusting for differences in observed characteristics between trial populations.

This approach yielded HRs for ibrutinib versus obinutuzumab + chlorambucil of 0.12 [95% CrI 0.02, 0.97] for PFS by investigator assessment, 0.24 [0.04, 1.35] for PFS by IRC, and 0.21 [0.00, 8.89] for overall survival (HR <1 favors ibrutinib), with a 98%, 95%, and 79% probability, respectively, that ibrutinib provides a greater improvement in survival compared to obinutuzumab + chlorambucil. The greater benefit of ibrutinib was maintained across all the scenarios analyzed, although the magnitude of the differences in the treatment effect and the certainty associated with it varied. From the

results of the stepwise sensitivity analyses, a trend was observed whereby when matching on more variables, and therefore moving closer to the CLL11 population, the estimated HRs tended to show a greater improvement in survival outcomes for ibrutinib vs. obinutuzumab + chlorambucil. This indicates that the treatment effect of ibrutinib vs. chlorambucil is larger in a population more comparable to the CLL11 study.

Several differences between the study populations of CLL11 and RESONATE-2 exist. Higher CIRS scores have been shown to predict poorer survival outcomes and are associated with “less fit” patients in CLL [16]. The eligibility criteria for CLL11 excluded patients with CIRS score of 6 or less and creatinine clearance of at least 70 mL/min, resulting in a lower proportion of fit patients (as indicated by CIRS score of 6 or less)

compared with RESONATE-2 (26% vs. 64%, respectively). After matching the proportion with CIRS score 6 or less in the CLL11 study, the estimated treatment effects from the indirect comparisons consistently showed a benefit for ibrutinib compared to obinutuzumab + chlorambucil in this less fit population. RESONATE-2 also included fewer patients aged 75 years or older than CLL11 (35% vs. 43%, respectively), though the eligibility criteria differed between the studies with only patients aged 65 years or older being eligible in RESONATE-2 and patients aged 18 years or older being eligible in CLL11. The median ages, however, were the same in both trials (73 years). Within the CLL11 trial, the del17p mutation was present in 6% of patients but these patients were not enrolled in the RESONATE-2 trial. Given the small proportion of del17p patients included within the CLL11 trial this difference was expected to have minimal impact on the results.

Not all baseline variables were associated with a poorer prognosis for the CLL11 population. RESONATE-2 included a higher proportion of patients with baseline β 2-microglobulin levels of at least 3.5 mg/L than CLL11 (71% vs. 35%), which would suggest poorer prognosis for patients from RESONATE-2 compared to CLL11. However, subgroup analyses within the RESONATE-2 trial showed a better treatment effect on PFS for ibrutinib vs. chlorambucil in patients with β 2-microglobulin levels greater than 3.5 mg/L compared to 3.5 mg/L or less [HR 0.15 (95% CI 0.08, 0.29) vs. HR 0.29 (95% CI 0.09, 0.92), respectively]. RESONATE-2 included a higher proportion of patients with del11q than CLL11 (22% vs. 17%, respectively) which has been shown to be associated with shorter PFS [11, 12, 24]. More advanced disease stages as captured by the Binet staging system have shorter OS [11], but the proportions of patients in each stage group were similar between the trials before matching. The distribution of ECOG performance status scores was also similar.

After matching to the CLL11 trial, and therefore balancing the populations with regard to a number of important prognostic characteristics, the relative treatment effect estimated

using the MAIC consistently indicated an improvement in PFS for ibrutinib compared with obinutuzumab + chlorambucil. The results of the stepwise sensitivity analysis suggested that the estimated within-trial HRs were not markedly influenced by matching on fewer variables and therefore increasing the effective sample size, as the results were consistent in showing an improvement in the relative treatment effect after matching (see Supplementary Material).

A similar observation was made for OS. However, the credible intervals of the indirect comparison for this endpoint tended to be quite wide. The low number of deaths in the experimental arms of either trial (only three deaths, all unrelated to study drug, were recorded in the ibrutinib arm of RESONATE-2) resulted in statistical uncertainty in the matched analyses for OS. In addition, statistical uncertainty is a consequence of performing a MAIC. The loss of statistical power is indicated by the relatively smaller effective sample sizes (35 if matched on 13 variables and 48 if matched on 12) and the wide 95% confidence intervals around the point estimates in the matched RESONATE-2 population, compared to the original population. Lastly, performing an adjusted indirect comparison introduces additional variance in the between-trial HR estimates, as the variances of the within-trial effects are effectively summed. Therefore, a large degree of uncertainty in the end result is inherent to the use of a MAIC.

The HR for the base case analysis of investigator-assessed PFS was estimated as 0.12 (95% CrI 0.02, 0.97), which is notably lower than the HR estimated for the sensitivity analysis matching on 12 variables [0.43 (95% CrI 0.16, 1.18)]. This difference is potentially influenced by the large number of missing data for unmutated IGHV (37 patients had missing baseline data for unmutated IGHV from RESONATE-2 across the arms). However, it is also related to the within-trial treatment effects from RESONATE-2 being close to 0 (HR for matching on 13 variables is 0.022, HR for matching on 12 variables is 0.078); this means that seemingly small differences in these HRs can lead to larger differences in the HR estimates from the adjusted indirect comparison.

MAIC is a relatively new technique that builds on conventional approaches to indirect comparison by attempting to adjust for differences between trial populations. Its use is becoming more widespread, and MAICs have been accepted in several health technology appraisals by England's National Institute for Health and Care Excellence (NICE), such as the appraisal of bortezomib as induction therapy for multiple myeloma [25].

One of the main limitations of the MAIC approach is that only imbalances in observed characteristics can be adjusted for and therefore residual confounding for unobserved patient characteristics may still be present. However, from our consideration of the literature, potential treatment effect modifiers that we have identified have mostly been observed in both trials and therefore accounted for in the matching. However, some differences between trial designs could not be adjusted for.

The dose of chlorambucil differed between the two studies in our analysis, as in RESONATE-2 a dose increase from 0.5 mg/kg up to 0.8 mg/kg was permitted, whereas the maximum dose allowed in CLL11 was 0.5 mg/kg. In addition, the maximum number of 28-day cycles of chlorambucil received was 12 in RESONATE-2 and six in CLL11. The higher dose and greater number of cycles of chlorambucil in RESONATE-2 compared with CLL11 suggests that the relative treatment effect for ibrutinib versus obinutuzumab + chlorambucil from this analysis may be underestimated, as efficacy outcomes are expected to be affected by the cumulative dose per patient received (which was 700 mg [26] and 400 mg [9], respectively).

PFS was assessed by slightly different criteria in each trial. RESONATE-2 used the modified IWCLL criteria (2012) whereas CLL11 states that it used the IWCLL criteria (2008) but does not mention or reference the modified criteria. Under the modified criteria, which were developed to take into account the pharmacodynamic characteristics of novel agents, isolated treatment-related lymphocytosis (in the absence of other evidence of disease progression) was not considered to indicate progressive disease [6, 27]. However, the three study drugs in CLL11 are not known to have lymphocytosis issues and so the

use of the modified IWCLL criteria is not expected to lead to different results; therefore this was not considered to be a limitation in the comparability of the two studies.

CONCLUSIONS

Although direct comparison in a randomized controlled trial provides the most reliable evidence on the relative treatment effects of two agents, such evidence is not always available. Adjusted indirect comparison methods may provide biased results where there are imbalances in treatment effect modifiers between trials. In these cases a matching-adjusted indirect comparison can be used to account for between-trial imbalances in observed characteristics. Our analysis showed that, when matching on patient baseline characteristics known to be potential treatment effect modifiers, the treatment effect of ibrutinib vs. chlorambucil is larger in a population more comparable to the CLL11 study. Taking this into account in the indirect comparison indicates a greater improvement in PFS for ibrutinib vs. obinutuzumab + chlorambucil compared to a traditional indirect comparison. The results of the matching-adjusted indirect comparison presented here indicate that ibrutinib is likely to provide greater progression-free benefit than the combination of obinutuzumab + chlorambucil in older or less fit patients with previously untreated CLL. While similar results are observed for OS, there is more uncertainty in the analysis related to this endpoint, mainly due to the low event rate observed with ibrutinib in the trial.

ACKNOWLEDGEMENTS

The study and article processing charges were funded by Janssen-Cilag Ltd. All authors had full access to all of the data in this study and take complete responsibility for the integrity of the data and accuracy of the data analysis. All named authors meet the International Committee of Medical Journal Editors (ICMJE) criteria for authorship for this manuscript, take responsibility for the integrity of the work as a

whole, and have given final approval for the version to be published. Editorial assistance was provided by Jo Whelan and Amie Padhiar from ICON plc.

Disclosures. Suzy Van Sanden is an employee of Janssen and holds stock in Johnson & Johnson. Simona Baculea is an employee of Janssen and holds stock in Johnson & Johnson. Joris Diels is an employee of Janssen and holds stock in Johnson & Johnson. Sarah Cote is an employee of Janssen and holds stock in Johnson & Johnson.

Compliance with Ethics Guidelines. This article is based on previously conducted studies, and does not involve any new studies of human or animal subjects performed by any of the authors.

Open Access. This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any non-commercial use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

REFERENCES

- Eichhorst B, Dreyling M, Robak T, Montserrat E, Hallek M, ESMO Guidelines Working Group. Chronic lymphocytic leukemia: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol.* 2011;22(Suppl 6):vi50–4.
- FDA. <http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm406916.htm>. Accessed 5 Aug 2016.
- EMA. CHMP Assessment report. Imbruvica. International non-proprietary name: ibrutinib. Procedure No.: EMEA/H/C/003791/0000EMA/CHMP/645137/2014. EMA/CHMP/645137/2014. Accessed 5 Aug 2016.
- EMA. Committee for Medicinal Products for Human Use (CHMP). Assessment report. Invented name: Imbruvica. International non-proprietary name: ibrutinib. Procedure No. EMEA/H/C/003791/II/0016. EMA/373867/2016. Accessed 5 Aug 2016.
- Abbvie. US FDA expands IMBRUVICA® (ibrutinib) label to include overall survival data in previously untreated chronic lymphocytic leukemia (CLL). 2016. <https://news.abbvie.com/news/us-fda-expands-imbruvica-ibrutinib-label-to-include-overall-survival-data-in-previously-untreated-chronic-lymphocytic-leukemia-ctl-and-new-indication-for-small-lymphocytic-lymphoma-sll-patients.htm>. Accessed 5 Aug 2016.
- Burger JA, Tedeschi A, Barr PM, et al. Ibrutinib as initial therapy for patients with chronic lymphocytic leukemia. *N Engl J Med.* 2015;373(25):2425–37.
- FDA. Gazyva (obinutuzumab). <http://www.fda.gov/drugs/informationondrugs/approveddrugs/ucm373263.htm>. Accessed 5 Aug 2016.
- EMA. CHMP Assessment Report. Gazyvaro. International non-proprietary name: OBINUTUZUMAB. Procedure No.: EMEA/H/C/002799/0000. EMA/CHMP/231450/2014. Accessed 5 Aug 2016.
- Goede V, Fischer K, Busch R, et al. Obinutuzumab plus chlorambucil in patients with CLL and coexisting conditions. *N Engl J Med.* 2014;370(12):1101–10.
- Eichhorst B, Robak T, Montserrat E, et al. Appendix 6: chronic lymphocytic leukaemia: eUpdate published online September 2016 (<http://www.esmo.org/Guidelines/Haematological-Malignancies>). *Ann Oncol.* 2016;27(suppl 5):v143–v4.
- Eichhorst B, Robak T, Montserrat E, et al. Chronic lymphocytic leukaemia: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol.* 2015;26(Suppl 5):v78–84.
- National Comprehensive Cancer Network. Chronic lymphocytic leukemia/small lymphocytic leukemia version 1.2017. 2016.
- Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol.* 1997;50(6):683–91.
- Signorovitch JE, Sikirica V, Erder MH, et al. Matching-adjusted indirect comparisons: a new tool for timely comparative effectiveness research. *Value Health.* 2012;15(6):940–7.
- Malangone E, Sherman S. Matching-adjusted indirect comparison analysis using common SAS® 9.2: procedures. <http://support.sas.com/resources/>

- [papers/proceedings11/228-2011.pdf](#). Accessed 5 Aug 2016.
16. Manda S, James S, Wang R, Krishnan R, Danilov AV. Impact of comorbidities on treatment outcomes in chronic lymphocytic leukemia: a retrospective analysis. *Blood*. 2014;124(21):1312.
 17. Hallek M. Prognostic factors in chronic lymphocytic leukemia. *Ann Oncol*. 2008;19(Suppl 4):iv51–3.
 18. Hallek M, Wanders L, Ostwald M, et al. Serum beta(2)-microglobulin and serum thymidine kinase are independent predictors of progression-free survival in chronic lymphocytic leukemia and immunocytoma. *Leuk Lymphoma*. 1996;22(5–6):439–47.
 19. Signorovitch JE, Wu EQ, Yu AP, et al. Comparative effectiveness without head-to-head trials: a method for matching-adjusted indirect comparisons applied to psoriasis treatment with adalimumab or etanercept. *Pharmacoeconomics*. 2010;28(10):935–45.
 20. Phillippo DM, Ades AE, Dias S, Palmber S, Abrams KR, Welton NJ. NICE DSU technical support document 18: methods for population-adjusted indirect comparisons in submissions to NICE. 2016. <http://scharr.dept.shef.ac.uk/nicedsu/wp-content/uploads/sites/7/2017/05/Population-adjustment-TSD-FINAL.docx>. Accessed 5 Aug 2016.
 21. Dias S, Sutton AJ, Ades AE, Welton NJ. Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. *Med Decis Making*. 2013;33(5):607–17.
 22. Coutre S, Byrd JC, Hillmen P, et al. Integrated and long-term safety analysis of ibrutinib in patients with chronic lymphocytic leukemia (CLL)/small lymphocytic lymphoma (SLL). In: 58th Annual Meeting and Exposition of the American Society of Hematology; San Diego, CA, USA. 2016.
 23. Barr P, Robak T, Owen C, et al. Updated efficacy and safety from the phase 3 RESONATE-2 study: ibrutinib as first-line treatment option in patients 65 years and older with chronic lymphocytic leukemia. In: 58th Annual Meeting and Exposition of the American Society of Hematology; 3–6 December; San Diego, CA, USA. 2016.
 24. Gentile M, Zirlik K, Ciolli S, et al. Combination of bendamustine and rituximab as front-line therapy for patients with chronic lymphocytic leukaemia: multicenter, retrospective clinical practice experience with 279 cases outside of controlled clinical trials. *Eur J Cancer*. 2016;60:154–65.
 25. National Institute for Health and Care Excellence (NICE). Technology appraisal 311: bortezomib for induction therapy in multiple myeloma before high-dose chemotherapy and autologous stem cell transplantation. 2014. <https://www.nice.org.uk/guidance/ta311/chapter/1-guidance>. Accessed 5 Aug 2016.
 26. Pharmacyclics Inc. Data on file. A randomized, multicenter, open-label, phase 3 study of the Bruton's tyrosine kinase inhibitor PCI-32765 versus chlorambucil in patients 65 years or older with treatment-naïve chronic lymphocytic leukemia or small lymphocytic lymphoma. PCYC-1115-CA Clinical Study Report. 2015.
 27. Cheson BD, Byrd JC, Rai KR, et al. Novel targeted agents and the need to refine clinical end points in chronic lymphocytic leukemia. *J Clin Oncol*. 2012;30(23):2820–2.