Supplementary material

Technical details of the MAIC process

- In order to match patients on the trial level, aggregate baseline characteristics for CLL11 were calculated as a weighted average across study arms (obinutuzumab + chlorambucil, rituximab + chlorambucil and chlorambucil). This was necessary as the baseline characteristics in the CLL11 trial were only reported by trial arm. However, the rituximab + chlorambucil arm was not considered further given that the interest lies in comparing ibrutinib to obinutuzumab + chlorambucil via the common comparator chlorambucil.
- A fixed effects model was used to perform the Bayesian indirect comparison as there is insufficient information in such a small network to consider a random effects model.
- ***Effective sample size***

  In general, matching larger numbers of baseline variables and adjusting for greater baseline differences between trials will require more extreme weights. Signorovitch et al. suggests the use of the effective sample size as a measure of the impact of reweighting on the available statistical information in the IPD.(14) It is estimated as the square of the summed weights divided by the sum of the squared weights.(19) The more extreme the weights, the smaller the effective sample size and the higher the uncertainty in the final statistical analysis of the weighted population.

Sensitivity analyses

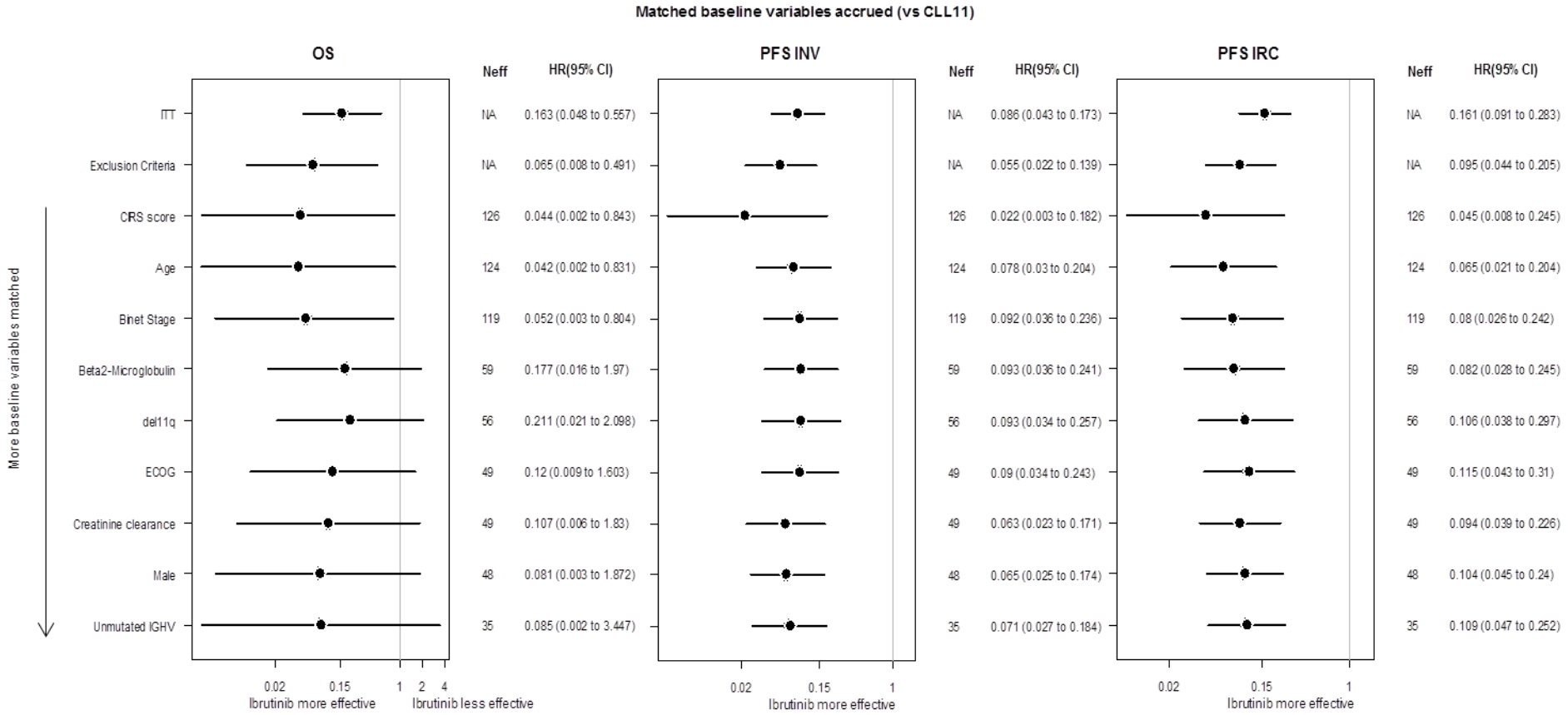
As a sensitivity analyses, the matching algorithm was repeated excluding the least important baseline variables from the matching process. This was performed in a stepwise manner, excluding more baseline characteristics from the MAIC process at each step.

Results of sensitivity analyses

Supp. Table 1 Baseline characteristics of the CLL11 and RESONATE-2 trial populations before and after matching, showing base case and stepwise sensitivity analyses

	CLL11 (Goede 2014)	RESONATE-2										
	ITT	ITT	After exclusion*	Base case matched (n=13)	SA matched (n=12)	SA matched (n=11)	SA matched (n=10)	SA matched (n=9)	SA matched (n=8)	SA matched (n=7)	SA matched (n=4)	SA matched (n=2)
N (Neff)	589	269	191	115 (35)	152 (48)	152 (49)	152 (49)	152 (56)	161 (59)	173 (119)	173 (124)	173 (126)
CIRS score (median)	8	5	6	8	8	8	8	8	8	8	8	8
CIRS score ≤ 6 (%)	26	64	56	26	26	26	26	26	26	26	26	26
Age (median)	73	73	73	73	73	73	73	73	73	73	73	73
Age (≥ 75 years) (%)	43	35	40	43	43	43	43	43	43	43	43	38
Binet Stage A (%)	22	19	19	22	22	22	22	22	22	22	17	17
Binet Stage B (%)	42	43	40	42	42	42	42	42	42	42	37	38
Binet Stage C (%)	36	38	41	36	36	36	36	36	36	36	46	45
β2-Microglobulin ≥ 3.5 mg/L (%)	35	71	77	35	35	35	35	35	35	73	75	74
del11q (%)	17	22	24	17	17	17	17	17	14	22	20	20
ECOG (median)	1	1	1	1	1	1	1	1	1	1	1	1
Creatinine clearance (median)	62	61.21	55.35	62.02	62.02	61.98	65.2	64.57	64.79	60	59.48	60.38
Male (%)	62	63	59	62	62	55	55	53	54	62	62	62
Unmutated IGHV (%)	61	59	58	61	57	58	58	56	57	59	59	59

*Excludes patients from RESONATE-2 with CIRS ≤ 6 and creatinine clearance ≥ 70 ml/min, patients with creatinine clearance < 30 ml/min, and SLL patients, in line with exclusion criteria in the CLL11 trial; cells shaded grey indicate the baseline characteristic that was not matched for the scenario and hence represent the average for the characteristic unadjusted in each case. Note: CLL11 has 6% del17p patients while RESONATE-2 excluded such patients, hence it was not possible to adjust for this difference. CIRS, Cumulative Illness Rating Scale; ECOG, Eastern Cooperative Oncology Group; IGHV, immunoglobulin heavy variable cluster; N, number of patients analyzed; n, number of variables matched; Neff, effective sample size; SA, sensitivity analysis



Suppl Figure 1: Forest plot of within-trial hazard ratios for ibrutinib vs chlorambucil for the stepwise sensitivity analyses. The covariate list used to determine the HRs is cumulative from CIRS score to the bottom of the figures.

Footnote: excludes patients from RESONATE-2 with CIRS ≤ 6 and creatinine clearance ≥ 70 ml/min, patients with creatinine clearance < 30 ml/min, and SLL patients, in line with exclusion criteria in the CLL11 trial

Suppl Table 2: Between-trial HRs for ibrutinib vs obinutuzumab + chlorambucil (Bayesian approach) for the stepwise sensitivity analyses

Population	HR [95% CrI] p(HR<1)		
	PFS by IRC assessment	PFS by investigator assessment	Overall survival
RESONATE-2 (ITT)	0.85 [0.44, 1.63] 0.69	0.48 [0.22, 1.02] 0.97	0.40 [0.10, 1.54] 0.91
RESONATE-2 (after exclusion)*	0.50 [0.22, 1.15] 0.95	0.31 [0.12, 0.81] 0.99	0.16 [0.02, 1.34] 0.95
RESONATE-2 (matched n=13)	0.24 [0.04, 1.35] 0.95	0.12 [0.02, 0.97] 0.98	0.21 [<0.01, 8.89] 0.79
RESONATE-2 (matched n=12)	0.34 [0.11, 1.12] 0.96	0.43 [0.16, 1.18] 0.95	0.20 [0.01, 5.16] 0.83
RESONATE-2 (matched n=11)	0.42 [0.13, 1.34] 0.93	0.51 [0.19, 1.37] 0.91	0.26 [0.01, 4.81] 0.82
RESONATE-2 (matched n=10)	0.43 [0.14, 1.34] 0.93	0.52 [0.19, 1.40] 0.90	0.29 [0.02, 4.15] 0.82
RESONATE-2 (matched n=9)	0.56 [0.19, 1.64] 0.86	0.52 [0.18, 1.48] 0.89	0.52 [0.05, 5.52] 0.71
RESONATE-2 (matched n=8)	0.61 [0.21, 1.71] 0.83	0.50 [0.18, 1.40] 0.91	0.43 [0.04, 5.11] 0.75
RESONATE-2 (matched n=7)	0.50 [0.19, 1.26] 0.93	0.35 [0.12, 1.00] 0.98	0.13 [0.01, 2.19] 0.92
RESONATE-2 (matched n=4)	0.55 [0.22, 1.34] 0.91	0.36 [0.13, 1.00] 0.98	0.10 [<0.01, 2.19] 0.93
RESONATE-2 (matched n=2)	0.57 [0.23, 1.41] 0.89	0.39 [0.14, 1.08] 0.97	0.11 [<0.01, 2.31] 0.92

* Excludes patients from RESONATE-2 with CIRS ≤ 6 and creatinine clearance ≥ 70 ml/min, patients with creatinine clearance < 30 ml/min, and SLL patients. 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