Supplementary Material for manuscript:

Combined conventional synthetic disease modifying therapy vs infliximab for rheumatoid arthritis: emulating a randomized trial in observational data

Andrei Barbulescu, MSc, Johan Askling, MD/PhD, Saedis Saevarsdottir, MD/PhD, Seoyoung C Kim, MD/ScD/MSCE, Thomas Frisell, PhD

Table of Contents

Supplementary Table S1 – Target trial and emulation protocol summaries

Protocol component	Target Trial (SWEFOT)	Observational Emulation	Comments
Eligibility criteria	1. Age ≥18 2. RA diagnosis according to ACR criteria with symptom duration of less than one year 3. No DMARD treatment before study entry 4. No oral glucocorticoids or stable treatment with 10 mg prednisone/day 5. After inclusion, all patients were treated with MTX for 3/4 months before being randomized to study treatments 6. DAS28-ESR>3.2 at both inclusion and randomization to study treatments	1. Age ≥18 2. RA as primary diagnosis in SRQ 3. No DMARD treatment other than MTX before study treatment initiation 4. No restriction to oral glucocorticoid use but adjustment for estimated average daily dose before baseline 5. First MTX prescription no more than 540 days after RA-debut and more than 30 days but less than 540 days before study treatment initiation 6. DAS28-ESR>3.2 at study treatment initiation	In SWEFOT, patients were treated with MTX for 3 to 4 months as part of the trial, before being randomly allocated to the receive additional infliximab or sulfasalazine + hydroxychloroquine. In the emulation, initiation of MTX at least 30 days before baseline (i.e. study treatment start) was required but continuation of MTX treatment to baseline was not imposed in order to increase sample size. Nonetheless, baseline co-medication with MTX was measured (the majority of patients receiving MTX at baseline) and was adjusted for.
Treatment strategies	A. Infliximab added to MTX (switch to etanercept allowed for safety reasons) B. Sulfasalazine and hydroxychloroquine added to MTX (stopping either drug or stopping both drugs and switching to ciclosporin A allowed for safety reasons)	A. Treatment start at first decision to initiate infliximab after RA-debut as identified in SRQ; treatment stop at decision to stop or at initiation of another DMARD; switch to etanercept allowed if within 90 days of infliximab stop B. Treatment start at first prescription of sulfasalazine or hydroxychloroquine identified in PDR no longer than 180 days after the first prescription of an ongoing treatment with the other drug; treatment stop at discontinuation of the last drug in combination or initiation of another DMARD; switch to ciclosporin A allowed if within 90 days of stopping last drug in combination	Sequential initiation of sulfasalazine and hydroxychloroquine was allowed in emulation in order to increase sample size in this arm (include close to 200 patients). The maximum of 180 days between the first prescription of the first drug and the first prescription of the second drug was decided based on the data, during the feasibility assessment phase

Treatment assignment	Random treatment assignment without blinding	Inverse probability of treatment weighting used to achieve a treatment assignment marginally independent of measured baseline covariates	
Outcome (primary)	Good EULAR response (DAS28-ESR ≤ 3.2 at evaluation and a decrease in DAS28-ESR larger than 1.2 units at evaluation compared to baseline)	Good EULAR response (DAS28-ESR ≤ 3.2 at evaluation and a decrease in DAS28-ESR larger than 1.2 units at evaluation compared to baseline)	
Follow-up	Patients were followed for 9 months after randomization, to the primary end-point. Patients who stopped (or never started) protocol treatment before the 9-month end-point were kept in the analysis as “non-responders”	Patients were followed for 9 months after study treatment start to the primary end-point. Patients who stopped protocol treatment before the 9-months end-point were kept in the analysis as “non-responders”	
Causal Contrast	Ratio between responder proportions	Ratio between responder proportions	
Statistical Analysis	Intention-to-treat (with “non-responder” imputation).	Intention-to-treat (with “non-responder” imputation).	The “intention-to-treat” was emulated as a decision to initiate treatment only in the infliximab arm. In the sulfasalazine + hydroxychloroquine arm treatment initiation was identified via prescription dispensation. However, the emulation analysis was considered an intention-to-treat analogue since all treatment initiators were analyzed, regardless of treatment discontinuation.

DAS28-ESR = Disease Activity Score (28 joints and Erythrocyte Sedimentation Rate); DMARD = Disease-Modifying Anti-Rheumatic Drug; MTX = Methotrexate; RA = Rheumatoid Arthritis;

Supplementary Table S2 – Baseline covariate definitions and measurement windows (in days relative to baseline)

Cancer	
Data Sources	Swedish Cancer Register
Definition	Any malignancy
Time Window	[-∞ to -1]
Serious (Hospitalized) **Infections**	
Data Sources	National Inpatient Register (hospital only)
Definition	Primary diagnosis using the following ICD10 codes: 'A' 'B' 'D733' 'E060' 'E321' 'G00' 'G01' 'G02' 'G042' 'G050' 'G051' 'G052' 'G06' 'G07' 'H000' 'H010' 'H030' 'H031' 'H061' 'H10' 'H130' 'H131' 'H160' 'H162' 'H190' 'H191' 'H192' 'H220' 'H30' 'H440' 'H600' 'H601' 'H603' 'H620' 'H621' 'H622' 'H623' 'H66' 'H670' 'H671' 'H68' 'H70' 'H730' 'H750' 'H940' 'I301' 'I330' 'I39' 'I400' 'I410' 'I411' 'I412' 'I430' 'I520' 'I521' 'I681' 'I980' 'I981' 'J00' 'J01' 'J02' 'J03' 'J04' 'J05' 'J06' 'J09' 'J10' 'J11' 'J12' 'J13' 'J14' 'J15' 'J16' 'J170' 'J171' 'J172' 'J173' 'J178' 'J18' 'J20' 'J21' 'J22' 'J32' 'J340' 'J36' 'J37' 'J383F' 'J383R' 'J390' 'J391' 'J440' 'J85' 'J86' 'K044' 'K045' 'K046' 'K047' 'K050' 'K052' 'K102' 'K113' 'K122' 'K140' 'K230' 'K35' 'K570' 'K572' 'K574' 'K578' 'K61' 'K630' 'K650' 'K67' 'L00' 'L01' 'L02' 'L03' 'L04' 'L05' 'L08' 'L303' 'M00' 'M01' 'M03' 'M462' 'M463' 'M465' 'M490' 'M491' 'M492' 'M493' 'M600' 'M630' 'M631' 'M632' 'M650' 'M651' 'M710' 'M711' 'M726' 'M86' 'M900' 'M901' 'M902' 'N080' 'N088' 'N10' 'N11' 'N136' 'N151' 'N159' 'N160' 'N290' 'N291' 'N300' 'N308' 'N33' 'N340' 'N341' 'N342' 'N390' 'N412' 'N431' 'N450' 'N481' 'N482' 'N49' 'N51' 'N61' 'N700' 'N710' 'N72' 'N730' 'N733' 'N74' 'N751' 'N760' 'N762' 'N770' 'N771' 'O070' 'O075' 'O080' 'O23' 'O353' 'O411' 'O753' 'O85' 'O86' 'O91' 'O98'
Time Window	[-365 to -1]
Diabetes	
Data Sources	National Patient Register (inpatient/outpatient)
Definition	Any diagnosis using the following ICD-10 codes: 'E10' 'E11' 'E12' 'E13' 'E14' 'O24'
Time Window	[-1825 to -1]
Acute Coronary Syndrome (ACS)	
Data Sources	National Inpatient Register (hospital only)
Definition	Any diagnosis using the following ICD-10 codes: 'I200' 'I21' 'I22'
Time Window	[-1825 to -1]
Congestive Heart Failure	
Data Sources	National Inpatient Register (hospital only)
Definition	Any diagnosis using the following ICD-10 codes: 'I110' 'I130' 'I132' 'I255' 'I420' 'I426' 'I427' 'I428' 'I429' 'I43' 'I50'
Time Window	[-1825 to -1]
Stroke (Cerebrovascular Disease)	
Data Sources	National Inpatient Register (hospital only)

Definition	Any diagnosis using the following ICD-10 codes: 'G45' 'I60' 'I61' 'I62' 'I63' 'I64' 'I67' 'I69'
Time Window	[-1825 to -1]

VTE (Venous Thromboembolism)	
Data Source	National Patient Register (inpatient/outpatient)
Definition	Any diagnosis using the following ICD-10 codes: 'I260' 'I269' 'I80' 'I822' 'I823' 'I828' 'I829' 'O222' 'O223' 'O870' 'O871'
Time Window	[-1825 to -1]

Peripheral Vascular Disease	
Data Source	National Patient Register (inpatient/outpatient)
Definition	Any diagnosis using the following ICD-10 codes: ‘I70’ ‘I71’ ‘I731’ ‘I739’ ‘I771’ ‘I790’ ‘I792’ ‘K55’
Time Window	[-1825 to -1]

Hepatic Disease	
Data Sources	National Patient Register (inpatient/outpatient)
Definition	Any diagnosis using the following ICD-10 codes: 'K70' 'K71' 'K72' 'K73' 'K74' 'K75' 'K76' 'K77' 'I85' 'I982' 'I983' ‘R18’
Time Window	[-1825 to -1]

Renal Disease	
Data Sources	National Patient Register (inpatient/outpatient)
Definition	Any diagnosis using the following ICD-10 codes: 'N032' 'N033' 'N034' 'N035' 'N036' 'N037' 'N11' 'N18' 'N19' 'N25' 'I120' 'I131' 'Q611' 'Q612' 'Q613' 'Q614' 'Z49' 'Z940' 'Z992'
Time Window	[-1825 to -1]

Obstructive Pulmonary Disease	
Data Sources	National Patient Register (inpatient/outpatient)
Definition	Any diagnosis using the following ICD-10 codes: 'J41' 'J42' 'J43' 'J44' 'J47'
Time Window	[-1825 to -1]

Interstitial Pulmonary Disease	
Data Sources	National Patient Register (inpatient/outpatient)
Definition	Any diagnosis using the following ICD-10 codes: 'J60' ’J61’ ’J62’ ’J63’ ’J64’ ’J66’ ’J67’ 'J82' 'J84'
Time Window	[-1825 to -1]

Anemias and **Hemorrhagic Conditions**	
Data Sources	National Patient Register (inpatient/outpatient)
Definition	Any diagnosis using the following ICD-10 codes:

	‘D50’ 'D51' 'D52' 'D53' 'D55' 'D56' 'D57' 'D58' 'D59' 'D600' 'D61' ‘D62’ 'D63' 'D64' ‘D66’ ‘D67’ ‘D68’ ‘D69’
Time Window	[-365 to -1]
Immunodeficiency Syndromes	
Data Sources	National Patient Register (inpatient/outpatient)
Definition	Any diagnosis using the following ICD-10 codes:
	‘D70’ ‘D730’ ‘D80’ ‘D81’ ‘D82’ ‘D83’ ‘D84’
Time Window	[-365 to -1]
Psoriasis	
Data Sources	National Patient Register (inpatient/outpatient)
Definition	Any diagnosis using the following ICD-10 code:
	‘L40’
Time Window	[-1825 to -1]
Neuropathies and Demyelinating Disease	
Data Sources	National Patient Register (inpatient/outpatient)
Definition	Any diagnosis using the following ICD-10 codes:
	'G35' 'G36' 'G37' 'G60' 'G61' 'G62' 'G63' 'G64'
Time Window	[-1825 to -1]
Pain Syndromes	
Data Sources	National Patient Register (inpatient/outpatient)
Definition	Any diagnosis using the following ICD-10 codes:
	‘F454’ ‘M791’ ‘M792’ ‘M797’ ‘R521’
Time Window	[-1825 to -1]
Mood Disorders (anxiety, depression)	
Data Sources	National Patient Register (inpatient/outpatient)
Definition	Any diagnosis using the following ICD-10 codes:
	‘F32’ ‘F33’ ‘F34’ ‘F40’ ‘F41’ ‘F43’ ‘F45’
Time Window	[-1825 to -1]
Osteoporosis	
Data Sources	National Patient Register (inpatient/outpatient)
Definition	Any diagnosis using the following ICD-10 codes:
	‘M80’ ‘M81’ ‘M82’
Time Window	[-1825 to -1]
Joint Surgeries	
Data Sources	National Patient Register (inpatient/outpatient)
Definition	Any of the following procedure codes:
	'NGB' '8423' '8424' '8426' 'NFB' '8400' '8401' '8402' '8403' '8404' '8405' '8406' '8407' '8408' '8409' '8410' '8411' '8412' '8413' '8414' '8415' '8419' 'NBB' '8437'

	'NHB' 'NHC' 'NHE' 'NHF' 'NHG' '8420' '8421' '8422' '8436' 'NDB' 'NDC' 'NDE' 'NDF' 'NDG'
Time Window	[-1825 to -1]

Porphyria	
Data Sources	National Patient Register (inpatient/outpatient)
Definition	Any of the following procedure codes: 'E800' 'E801' 'E802'
Time Window	[-1825 to -1]

Glucose-6-Phosphate-Dehydrogenase Deficiency	
Data Sources	National Patient Register (inpatient/outpatient)
Definition	Any of the following procedure codes: 'D550'
Time Window	[-1825 to -1]

Folate Deficiency	
Data Sources	National Patient Register (inpatient/outpatient)
Definition	Any of the following procedure codes: 'E358' 'D52'
Time Window	[-1825 to -1]

Retinal Disorders	
Data Sources	National Patient Register (inpatient/outpatient)
Definition	Any of the following procedure codes: 'H30' 'H31' 'H32' 'H33' 'H34' 'H35' 'H36' 'H534' 'H535' 'H536' 'H540' 'H541' 'H542' 'H544' 'H545' 'H546'
Time Window	[-1825 to -1]

Days in Hospital	
Data Sources	National Inpatient Register (hospital only)
Definition	Days spent in hospital within the specified window
Time Window	[-1825 to -1]

DMARD use before baseline (same ATC codes used to identify next DMARD during FU)	
Data Sources	Swedish Rheumatology Quality Register Prescribed Drugs Register
Definition	Used for identifying eligibility for study entry and protocol treatment stop. Treatments identified using the following ATC codes: **biological or targeted synthetic DMARDs**: 'L01XC02' 'L01XC10' 'L04AA24' 'L04AA26' 'L04AB01' 'L04AB02' 'L04AB04' 'L04AB05' 'L04AB06' 'L04AC03' 'L04AC05' 'L04AC07' 'L04AC10' 'L04AC13' 'L04AC14' 'L04AA29' 'L04AA37' 'L04AA44' **conventional synthetic DMARDs**: 'A07EC01' 'L01AA01' 'L01AA02' 'L04AA06' 'L04AA10' 'L04AA13' 'L04AA32' 'L04AD01' 'L04AD02' 'L04AX01' 'M01CC01'

	'M01CB01' 'M01CB03' 'M04AA01' 'M04AA03' 'M04AC01' 'P01BA01' 'P01BA02' 'J06BA01' 'J06BA02'
	Methotrexate: 'L04AX03'
Time Window	[RA-debut to -1]
Co-medication at baseline from SRQ	
Data Sources	Swedish Rheumatology Quality Register
Definition	Used for identifying co-medication NSAID (non-steroidal anti-inflammatory drug use) and methotrexate. Any co-medication data recorded for the visit closest to baseline (from the specified time-window). If two visits were equally distant from baseline (before and after) and no closer visit was identified, visit before treatment start was used.
Time Window	[-90 to +30]
Co-medication at baseline from PDR	
Data Sources	Prescribed Drugs Register
Definition	Used for identifying co-medication with glucocorticoids, average dose of glucocorticoids and co-medication with methotrexate. Use = (at least one prescription collected in the specified time window) Average daily dose = the sum of dispensed systemic glucocorticoids quantities (in mg Prednisone Equivalent) divided by the duration between the dispensation date of the first prescription and the end date of the last prescription (within the time window). *The end of a prescription was estimated by adding to the dispensation date the number of *defined daily doses* dispensed. *The following coefficients were used to transform quantities in *prednisone equivalents*: betamethasone [6.7]; dexamethasone [6.7]; methylprednisolone [1.25]; prednisone/prednisolone [1]; hydrocortisone [0,25]
Time Window	[-180 to +30]
RA activity (and disability) proxies	
Data Sources	Swedish Rheumatology Quality Register
Definition	DAS28-ESR; HAQ-DI; CDAI; Pain (Visual Analogue Scale). Use RA-activity data from the visit closest to baseline within the specified time-window. If two visits are equally distant from baseline (before and after) and no closer visit is identified, use the average value.
Time Window	[-90 to +30]
RA duration	
Data Sources	Swedish Rheumatology Quality Register
Definition	Number of years between RA debut (first symptoms) and baseline (study treatment start)
Time Window	[RA-debut to baseline]
Demographic Characteristics	
Data Sources	Swedish Total Population Register

Definition	Sex, Age (number of years from birth to index treatment start), Country of origin (categorized into: Sweden, other Scandinavian Country, other).
Time window	NA
Smoking status	
Data Sources	Swedish Rheumatology Quality Register
Definition	Data from the evaluation closest to index treatment start. Categorized into three levels: has never smoked, has smoked, current smoker.
Time window	[-∞ to +30]

CDAI=Clinical Disease Activity Index; DAS28-ESR = Disease Activity Score (28 joints and Erythrocyte Sedimentation Rate); DMARD = Disease-Modifying Anti-Rheumatic Drug; HAQ-DI=Health Assessment Questionnaire Disability Index; ICD= International Statistical Classification of Diseases and Related Health Problems

Supplementary Table S3 – Protocol analysis missing values

variable	***Infliximab*** *N (%)*	***SSZ+HCQ*** *N (%)*
Rheuma. Factor	4 (1.3)	8 (4.1)
Baseline VAS-pain	2 (0.6)	2 (1.0)
Baseline CDAI	7 (2.2)	9 (4.6)
Baseline DAS28-ESR	0 (0)	0 (0)
Baseline HAQ-DI	10 (3.2)	4 (2.0)
Baseline Smoking	190 (60.5)	133 (67.9)
EULAR response*	87 (27.7)	48 (25.0)

CDAI=Clinical Disease Activity Index; DAS28-ESR = Disease Activity Score (28 joints and Erythrocyte Sedimentation Rate); EULAR=European Alliance of Associations for Rheumatology; HAQ-DI=Health Assessment Questionnaire Disability Index; VAS=Visual Analogue Scale

*After “non-responder” imputation of discontinued treatments.

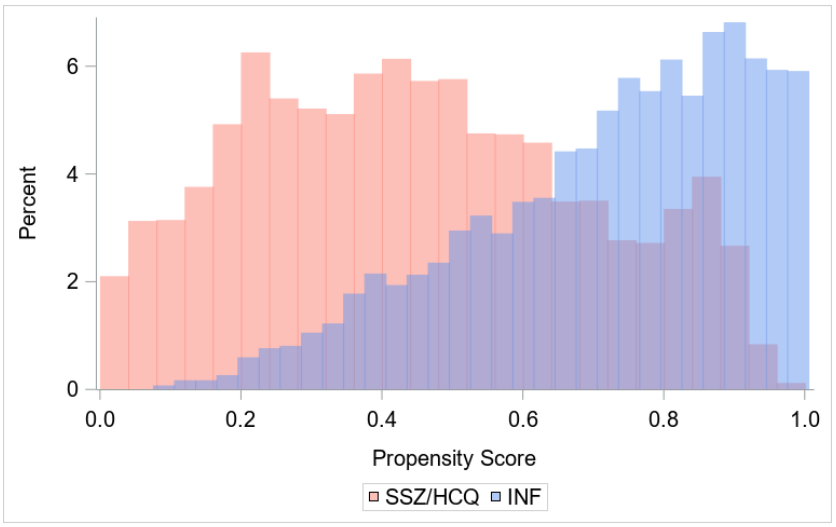
Supplementary Table S4 – Expectations (means or percentages) of baseline characteristics before and after inverse probability of treatment weighting (IPTW)

	Before IPTW		After IPTW	
Characteristic	***Infliximab***	***SSZ+HCQ***	***Infliximab***	***SSZ+HCQ***
Mean start year	2012.0	2011.2	2011.8	2012.2
Mean age (years)	54.4	56.3	55.2	55.3
% Female	67.4	71.3	69.2	69.4
% Swedish (origin_d1)	83.1	86.2	84.4	88.3
% Other, Scandinavian (origin_d2)	3.8	7.2	4.4	5.1
% Other (origin_d3)	13.1	6.7	11.2	6.6
Mean RA-duration	1.1	1.0	1.1	1.1
% Rheumatoid Factor positive	72.9	65.8	72.4	68.2
Mean DAS28-ESR	5.2	4.6	5.0	5.0
Mean HAQ-DI	1.1	0.9	1.0	1.0
Mean Swollen Joint Counts	8.1	5.5	7.1	6.9
Mean Tender Joint Counts	8.4	6.2	7.6	7.4
Mean Global Health	58.6	46.0	55.0	54.7
Mean VAS-pain	57.5	45.4	54	52.4
Mean CDAI	27.9	20.7	25.4	24.9
Mean avg. daily GC dose (mg PEQ)	7.6	7.4	7.5	7.8
% MTX use	98.1	98.5	98.4	99.2
% NSAID use	33.2	36.9	34.2	33.3
% Joint Surgery	7.3	7.2	6.6	6.9
Mean days spent in hospital	5.5	4.2	4.5	3.8
% Cancer	2.6	7.2	4.1	4.0
% Diabetes	6.4	2.1	4.6	1.7
% Hosp. Infections	1.3	2.1	1.3	1.2
% Acute coronary syndrome	0.3	2.1	0.3	0.8
% Stroke	1.9	1.0	1.7	1.6
% Venous Thromboembolism	1.6	1.5	1.4	1.0
% Peripheral Vascular Dis.	1.3	1.0	1.0	0.6
% Obstruct Lung Dis.	1.3	2.6	1.3	1.3
% Neuropathies	1.0	1.0	0.9	0.7
% Pain Syndromes	4.2	3.1	3.6	3.3
% Mood Disorders	4.8	5.1	5.3	4.9
% Anemia	2.2	1.5	1.7	1.0
% Retinopathies	2.9	2.6	2.6	2.5
% Never Smoker (smoking_d0)	35.4	36.5	36.8	36.1
% Has Smoked (smoking_d1)	33.9	29.0	32.0	33.6
% Smoker (smoking_d2)	30.7	34.5	31.2	30.3

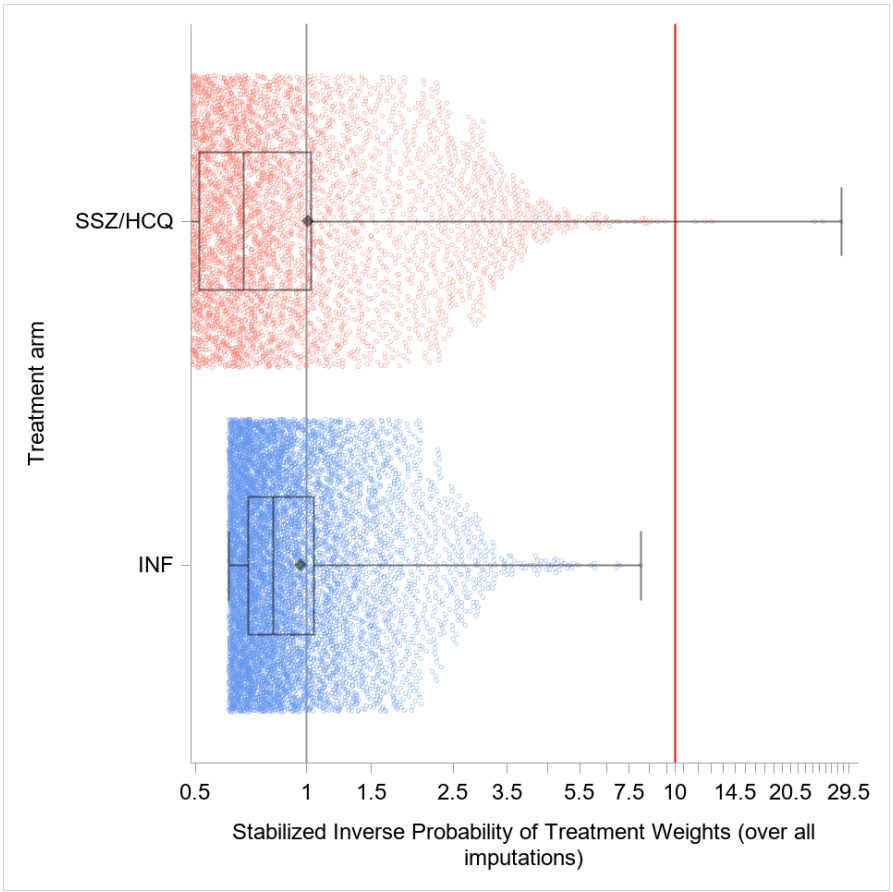
CDAI=Clinical Disease Activity Index; DAS28-ESR = Disease Activity Score (28 joints and Erythrocyte Sedimentation Rate); GC=Glucocorticoid; HAQ-DI=Health Assessment Questionnaire Disability Index; MTX=Methotrexate; NSAID=Non-Steroidal Anti-Inflammatory Drug; PEQ=Prednisone Equivalents; VAS=Visual Analogue Scale

* Inverse probability of treatment weights (IPTW) estimated separately in each imputation. Expectations estimated over all 30 imputations.

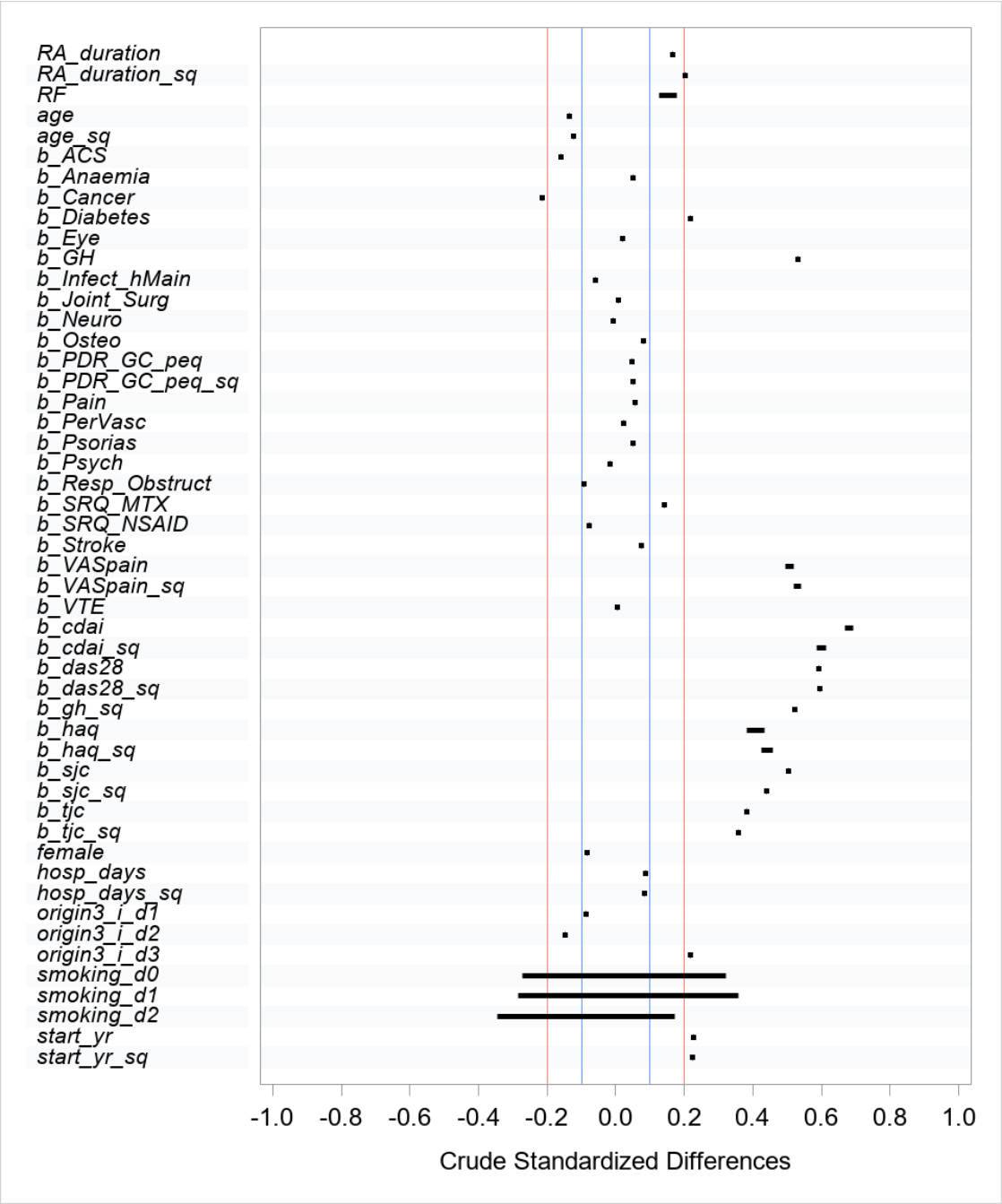
Supplementary Figure S1 – Protocol analysis propensity score distribution over all imputed data-sets



Supplementary Figure S2 – Protocol analysis stabilized inverse probability of treatment weights (IPTW) distribution over all imputed data-sets

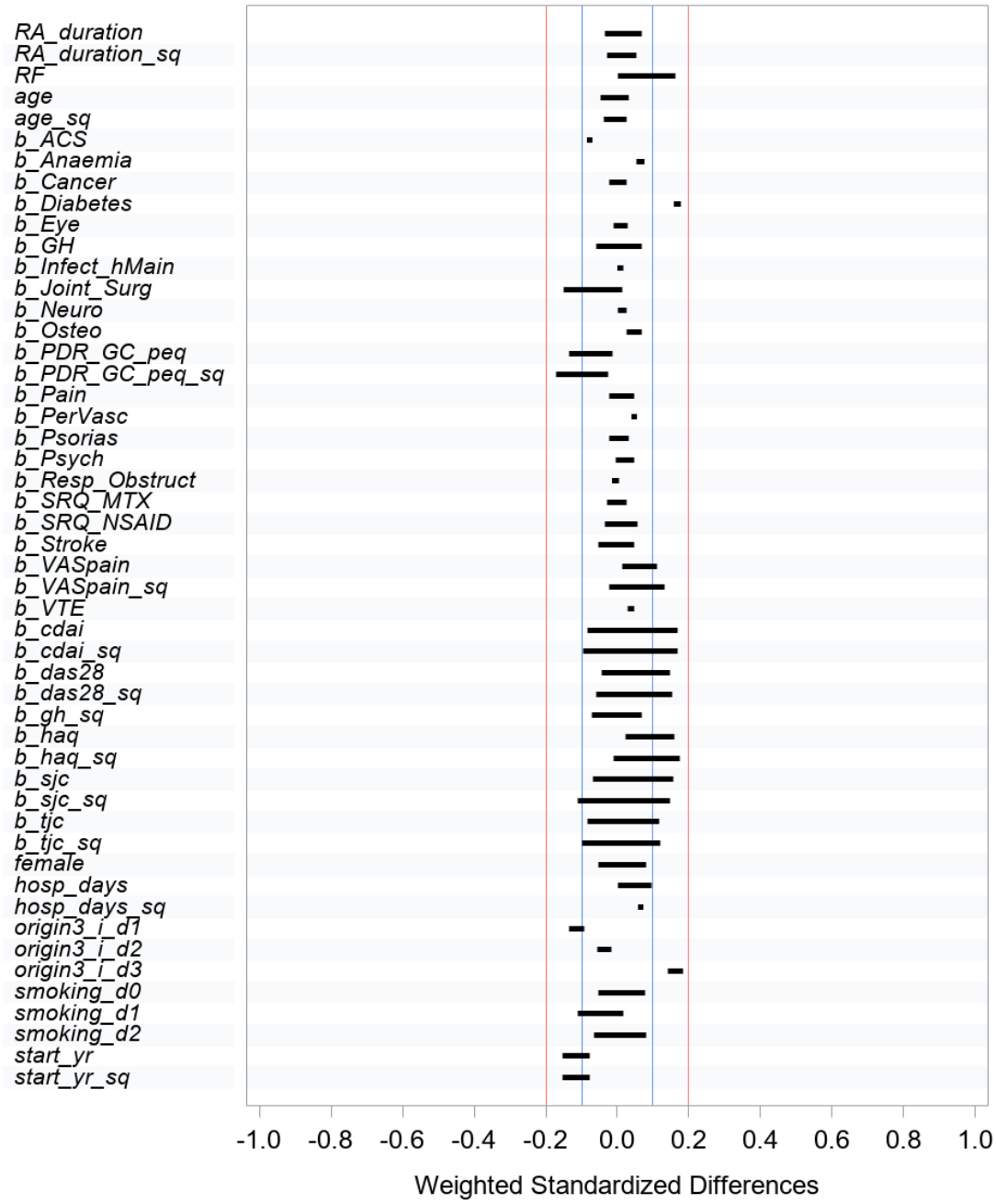


Supplementary Figure S3 – Protocol analysis standardized differences in baseline characteristics before inverse probability of treatment weighting (IPTW)



* The bars represent the range of standardized difference values over imputed data-sets. Only variables with missing data can produce variable standardized differences.

Supplementary Figure S4 – Protocol analysis standardized differences in baseline characteristics after inverse probability of treatment weighting (IPTW)



* The bars represent the range of standardized difference values over imputed data-sets. Due to inverse probability weights depending on covariates with missing data (thus imputed), the distribution of all covariates, and consequently of all standardized differences, will vary between imputations after weighting.