Paper	Domain 1	Domain 2	Domain 3	Domain 4	Domain 5	Overall ROB
Efficacy and	1.1.: Stratified	2.1: Study was	3.1: Population for	4.1: Tools used for	5.1: Data produced	High risk
Safety of	randomization was	double-blind and	ITT analysis was	measurement are	and analyzed were	
Upadacitinib in	noted, but	quadruple masked.	randomized.	valid and appropriate	consistent with the	
Patients With	randomization	Placebo was used	Greater than 5%	tools to measure	pre-specified analysis	
Moderate to	technique used	(N).	and less than 20%	outcomes (N).	plan per the protocol	
Severe Atopic Dermatitis	was not discussed	2.2: Study was	of patients were	4.2: Same	(Y).	
Dermatitis	(NI).	double-blind, and	discontinued from	measurement	5.2: Primary and	
	1.2: Double-blind	carers were masked.	the study (N).	methods and	secondary outcome	
	was mentioned,	All placebo subjects	3.2: Multiple	thresholds were used	domains were	
	but blinding was	were rerandomized	imputation was	in all groups and	measured in multiple	
	not discussed	to investigational	conducted for	measurements were	ways, and data from	
	further. Clinical	treatment (PN).	both groups using	done at similar time	all measures was	
	trial registration	2.6: ITT analysis was	the Markov Chain	points (N).	appropriately	
	indicated	conducted, and is	Monte Carlo	4.3: Double-blind	included (N).	
	quadruple	considered	method. Multiple	experiment (N).	5.3: Eligible reported	
	masking, but no	appropriate (Y)	imputation is not	Overall: Low risk	results for the	
	further	Overall: Low risk	sufficient to		outcome	
	information (PY)		control for bias		measurement	
	1.3: No suspicious		due to missing		correspond to all	
	or excessive		data, and MCMC		intended analyses	
	baseline		is known to result		(N).	
	imbalances or		in bias (N).		Overall: Low risk	
	similarities noted		3.3: In general			
	(N)		reasons for			
	Overall: Some		missingness were			
	concerns		similar between			
			groups, but the			
			placebo group had			
			disproportionately			
			large missingness			
			of data compared			

			to trooting and			
			to treatment			
			groups (PY).			
			3.4:			
			Disproportionately			
			high non-			
			completion by			
			participants in			
			placebo group,			
			with withdrawal			
			by subject and			
			other being most			
			common reasons			
			(PY)			
			Overall: High risk			
Efficacy and	1.1.: Stratified	2.1: Although	3.1: Population for	4.1: Tools used for	5.1: Data produced	Overall:
Safety of	randomization was	participants were	ITT analysis was	measurement are	and analyzed were	Some
Upadacitinib	noted; a	blinded, those in the	randomized.	valid and appropriate	consistent with the	concerns
vs Dupilumab	randomization	upadacitinib group	Greater than 5%	tools to measure	pre-specified analysis	
in Adults With	number used to	disproportionally	and less than 20%	outcomes (N).	plan per the protocol	
Moderate-to-	encode each	experienced acne;	of patients were	4.2: Same	(Y).	
Severe Atopic	patient's	acne is a well-known	discontinued from	measurement	5.2: All outcome	
Dermatitis	placement based	side effect of	the study (N).	methods and	measurements and	
	on a	upadacitinib (PY).	3.2: Non-	thresholds were used	analyses were	
	randomization	2.2: See 2.1	responder	in all groups and	included in either the	
	schedule (PY).	2.3: No indication	imputation,	measurements were	main paper or the	
	1.2: Double-blind	that blinding was	multiple	done at similar time	supplementary	
	was mentioned,	compromised.	imputation, and	points (N).	appendix (N).	
	but blinding was	Rescue therapies	per-protocol	4.3: Double-blind,	5.3: Eligible reported	
	not discussed	were used as	analyses were	double-dummy	results for the	
	further. Clinical	outlined in protocol	used as sensitivity	experiment (N).	outcome	
	trial registration	(PN).	analyses and are	Overall: Low risk	measurement	
	indicated	2.6: ITT analysis was	insufficient to	2.2.3	correspond to all	
	quadruple	conducted, and is	correct for bias		intended analyses	
	masking, but no		(PN).		(N).	
	masking, but no		\' ' <b>\'</b>		('*/-	

	further information (PY) 1.3: No suspicious or excessive baseline imbalances or similarities noted (N) Overall: Low risk	considered appropriate (Y) Overall: Low risk	3.3: Substantial proportion of individuals withdrawing from the study or being lost to follow-up. No further information regarding their reasons was given. (PY) 3.4: Similar proportion of missing data per		Overall: Low risk	
			data. It is unlikely that missingness of outcome depends on true value (N). Overall: Some concerns			
Upadacitinib plus topical corticosteroids in atopic dermatitis: Week 52 AD Up study results	1.1: Stratified randomization occurred using an interactive response technology system (Y). 1.2: Double-blind and double-blind extension	2.1: The study was double-blinded and quadruple masked. Differences between treatments would not be readily obvious. Placebo was used (N). 2.2: See 2.1. 2.6: ITT analysis was	3.1: Greater than 5% and less than 20% of patients were discontinued from the study (N). 3.2: Mixed-effect model with repeated measures was	4.1: Tools used for measurement are valid and appropriate tools to measure outcomes (N). 4.2: Same measurement methods and thresholds were used in all groups and	5.1: Data produced and analyzed were consistent with the pre-specified analysis plan per the protocol. Any posthoc analyses were declared (Y). 5.2: Primary and secondary outcome	Some concerns
	indicated, but	conducted, and is	used, and was	measurements were	domains were	

	exact method used for blinding is unknown (PY). 1.3: No suspicious or excessive baseline imbalances or similarities noted (N) Overall: Low bias	considered appropriate (Y) Overall: Low risk	sufficient to account for bias due to missing information. Overall: Low risk	done at similar time points (N). 4.3: Double-blind experiment (N). Overall: Low risk	measured using multiple scales, and data from each measure was included (N). 5.3: No information regarding adjustment strategy indicated (NI). Overall: Some	
Efficacy and Safety of Oral Janus Kinase 1 Inhibitor Abrocitinib for Patients With Atopic Dermatitis: A Phase 2 Randomized Clinical Trial	1.1: Randomization conducted using a computer-generated randomization schedule and assignment was done through an Interactive Voice Response System (Y). 1.2: Study record detail and protocol indicated that study was double-blinded and triple masked (Y). 1.3: No suspicious or excessive baseline imbalances or similarities noted (N).	2.1: Participants were blinded throughout the study and placebo was used (N). 2.2: It was not stated whether those delivering the intervention were masked (NI). 2.3: There is no evidence of unintended changes to treatment and use of rescue medications were predetermined in study protocol (PN). 2.6: Modified ITT analysis was conducted, and is considered appropriate (Y).	3.1: Population for ITT analysis was randomized. Greater than 20% of patients were discontinued from the study (N). 3.2: Modified ITT and sensitivity analyses were performed for missing data. Models used for handling missing data include generalized linear mixed model and mixed-effects model; these models did not use imputation (PY). Overall: Low risk	4.1: Tools used for measurement are valid and appropriate tools to measure outcomes (N). 4.2: Same measurement methods and thresholds were used in all groups and measurements were done at similar time points (N). 4.3: Double-blind experiment, outcome assessors were masked (N). Overall: Low risk	5.1: Data produced and analyzed were consistent with the pre-specified analysis plan per the protocol (Y). 5.2: All outcome measurements and analyses were included in either the main paper, study record detail, or the supplementary appendix (N). 5.3: Eligible reported results for the outcome measurement correspond to all intended analyses (N). Overall: Low risk	Low risk

	Overall: Low risk	Overall: Low risk				
Efficacy and	Randomization	2.1: Study was	3.1:	4.1: Tools used for	5.1: Measures and	Low risk
Safety of	conducted using	double-blind and	Approximately	measurement are	statistical analyses	
Multiple	"predefined	triple masked,	11% of	valid and appropriate	conducted in study	
Dupilumab	random number	including participant	participants were	tools to measure	were consistent with	
Dose	sequence with	masking. Placebo was	discontinued from	outcomes (N).	the pre-specified	
Regimens	block size of 5	used (N).	the study (PN).	4.2: Same	analysis plan (Y).	
After Initial	within each	2.2: Although the	3.2: Multiple	measurement	5.2: Multiple scales	
Successful	combination of	research protocol	sensitivity	methods and	were used for both	
Treatment in	stratification	does not indicate	analyses were	thresholds were used	primary and	
Patients With	factors" (Y).	that care providers	conducted for	in all groups and	secondary outcomes,	
Atopic	1.2: Study record	were masked, the	placebo group and	measurements were	and data was	
Dermatitis: A	detail indicated	paper states that all	all intervention	done at similar time	provided for all	
Randomized	double-blinded	involved individuals	groups. Bias due	points (N).	measures (N).	
Clinical Trial	and triple masked	were blinded;	to missing	4.3: Double-blind	5.3: Eligible reported	
	(including	additionally, it is	outcome data is	experiment,	results for the	
	participant,	indicated that the	unlikely (PY).	outcome assessors	outcome	
	investigator, and	drug kits were	Overall: Low risk	were masked (N).	measurement	
	outcome assessor).	blinded and		Overall: Low risk	correspond to all	
	The paper further	interventions were			intended analyses	
	indicated that all	replaced with			(N).	
	individuals	identical placebos			Overall: Low risk	
	involved were	when intervention				
	blinded except for	was not administered				
	the statistician	(PN).				
	who conducted the	2.6: ITT analysis was				
	randomization; the	conducted, and is				
	statistician was not	considered				
	involved in the	appropriate (Y)				
	project in any	Overall: Low risk				
	other way. Study					
	drug kits were					
	blinded and coded					

	with a numbering system (PY). 1.3: No suspicious					
	or excessive					
	baseline					
	imbalances or					
	similarities noted					
	(N)					
	Overall: Low risk					
Dupilumab	1.1: Stratified	2.1: Study was	3.1: Population for	4.1: Tools used for	5.1: Data produced	High risk
does not	randomization	double-blind and	ITT analysis was	measurement are	and analyzed were	
affect	conducted using	triple masked,	randomized.	valid and appropriate	consistent with the	
correlates of	interactive	including participant	Greater than 5%	tools to measure	pre-specified analysis	
vaccine-	response	masking. Placebo was	and less than 20%	outcomes (N).	plan per the protocol	
induced	technology system	used (N).	of patients were	4.2: Same	(Y).	
immunity: A	(Y).	2.2: The study record	discontinued from	measurement	5.2: All outcome	
randomized,	1.2: Study record	detail did not specify	the study (N).	methods and	measurements and	
placebo-	detail indicated	whether care	3.2: Last observer	thresholds were used	analyses were	
controlled trial	that study was	providers were	carried forward	in all groups and	included in either the	
in adults with	double-blinded	masked (NI).	was used for	measurements were	main paper, study	
moderate-to-	and triple masked	2.3: No indication	missing	done at similar time	record detail, or the	
severe atopic	(including	was made as to	information, and	points (N).	supplementary	
dermatitis	participant,	whether there were	is not considered	4.3: Double-blind	appendix (N).	
	investigator, and	deviations due to	sufficient to	experiment,	5.3: Eligible reported	
	outcome assessor)	trial context (NI).	account for bias	outcome assessors	results for the	
	(Y).	2.6: ITT analysis was	due to missing	were masked (N).	outcome	
	1.3: No suspicious	not explicitly	information.	Overall: Low risk	measurement	
	or excessive	discussed; all	3.3:		correspond to all	
	baseline	individuals who were	Inconsistencies		intended analyses	
	imbalances or	discontinued from	between trial		(N).	
	similarities noted	the study were	groups related to		Overall: Low risk	
	(N)	included in efficacy	missing			
	Overall: Low risk	analyses (PY).	information may			
			indicate that			

		Overall: Some	missingness is due			
			missingness is due			
		concerns	to true value (PN).			
			3.4: Differences			
			between groups			
			are substantial,			
			and likely indicate			
			that missingness is			
			due to the true			
			value.			
			Overall: High risk			
Abrocitinib	1.1: Randomization	2.1: Study was	3.1: Data for	4.1: Tools used for	5.1: Data produced,	Low risk
versus Placebo	conducted using	double-blind and	outcomes was	measurement are	analyzed, and	
or Dupilumab	center-based	quadruple masked.	available for	valid and appropriate	reported were	
for Atopic	permuted blocks	Placebo was used	nearly all	tools to measure	consistent with the	
Dermatitis	(Y).	(N).	participants	outcomes (N).	pre-specified analysis	
	1.2: Study record	Although the	randomized. The	4.2: Same	plan per the protocol	
	detail indicated	individuals providing	number of	measurement	(Y).	
	that study was	training to the carers	individuals that	methods and	5.2: All outcome	
	double-blinded	were unblinded, the	discontinued from	thresholds were used	measurements were	
	and quadruple	carers were blinded.	the study was less	in all groups and	included for each	
	masked (Y).	It is unlikely that	than 5% (Y).	measurements were	pre-specified	
	1.3: No suspicious	blinding was broken	Overall: Low risk	done at similar time	outcome	
	or excessive	during training (PN).	Overall. Low Hisk	points (N).	measurement used	
	baseline	2.6: A modified ITT		4.2: Outcome	(N).	
	imbalances or	analysis was		assessors were	5.3: Eligible reported	
	similarities noted	conducted, and is		masked (N).	results for the	
		considered		Overall: Low risk		
	(N)			Overall: Low risk	outcome	
	Overall: Low risk	appropriate (Y)			measurement	
		Overall: Low risk			correspond to all	
					intended analyses	
					(N).	
		,		,	Overall: Low risk	_
Dupilumab	n/a	n/a	n/a	n/a	n/a	High risk
Provides						due to

Favorable Safety and Sustained Efficacy for up to 3 Years in an Open-Label Study of Adults with Moderate-to- Severe Atopic Dermatitis						open label study.
Efficacy and safety of abrocitinib versus dupilumab in adults with moderate-to-severe atopic dermatitis: a randomised, double-blind, multicentre phase 3 trial	1.1: Block randomization was conducted using an interactive response system (Y). 1.2: Study record detail and protocol indicated that study was double- blinded and quadruple masked (Y). 1.3: No suspicious or excessive	2.1 and 2.2: Participants were blinded throughout the study. However, the two treatment groups were administered differently (oral and subcutaneous). Although, an interactive response technology was used to dispense tamper- free packaging, it is still possible that	3.1: Population for full analysis set was randomized. Greater than 5% and less than 20% of patients were discontinued from the study (N). 3.2: Multiple imputation was used as a sensitivity analyses, and is considered insufficient to	4.1: Tools used for measurement are valid and appropriate tools to measure outcomes (N). 4.2: Same measurement methods and thresholds were used in all groups and measurements were done at similar time points (N). 4.3: Outcome assessors were	5.1: Data produced and analyzed were consistent with the pre-specified analysis plan per the protocol (Y). 5.2: All outcome measurements and analyses were included in either the main paper, study record detail, or the supplementary appendix (N). 5.3: Eligible reported	Some
	baseline imbalances or similarities noted (N). Overall: Low risk	carers could determine the treatment during follow-up appointments, particularly in cases where subcutaneous injections were	correct for bias due to missingness of data (PN). 3.3: Due to the stratification by disease severity, the similarity in	masked (N). Overall: Low risk	results for the outcome measurement correspond to all intended analyses (N). Overall: Low risk	

		dolinonod are elte	haaalina	I		<u> </u>
		delivered on-site	baseline			
		rather than by the	characteristics			
		participant (PY).	between			
		2.3: Not enough	treatment groups,			
		information was	and the similarity			
		provided to	between groups			
		adequately assess	regarding			
		this domain (NI).	missingness of			
		2.6: A full analysis set	data, it is unlikely			
		was used and is	that missingness			
		considered	of data was due to			
		appropriate (Y).	the true value			
		Overall: Some	(PN)			
		concerns	Overall: Low risk			
Abrocitinib	1.1: Randomization	2.1: Participants were	3.1: Population for	4.1: Tools used for	5.1: Data produced	Some
versus Placebo	was conducted	blinded throughout	full analysis set	measurement are	and analyzed were	concerns
or Dupilumab	using an	the study and	was randomized.	valid and appropriate	consistent with the	
for Atopic	interactive	placebo was used	Greater than 5%	tools to measure	pre-specified analysis	
Dermatitis	response system	(N).	and less than 20%	outcomes (N).	plan per the protocol	
	and center-based	2.2: Both treatments	of patients were	4.2: Same	(Y).	
	randomly	and its matched	discontinued from	measurement	5.2: All outcome	
	permuted blocks	placebo was blinded	the study (N).	methods and	measurements and	
	(Y).	and administered	3.2: Multiple	thresholds were used	analyses were	
	1.2: Study record	similarly. Treatments	imputation was	in all groups and	included in either the	
	detail and protocol	and matched	used as a	measurements were	main paper, study	
	indicated that	placebos were	sensitivity	done at similar time	record detail, or the	
	study was double-	administered by an	analyses, and is	points (N).	supplementary	
	blinded and	unblinded	considered	4.3: Outcome	appendix (N).	
	quadruple masked	administrator who	insufficient to	assessors were	5.3: Eligible reported	
	(Y).	would not participate	correct for bias	masked (N).	results for the	
	1.3: No suspicious	in any other study-	due to	Overall: Low risk	outcome	
	or excessive	related procedures; it	missingness of		measurement	
	baseline	is unlikely that this	data (PN).		correspond to all	

	imbalances or similarities noted (N). Overall: Low risk	would break the blinding of carers (PN).  2.6: A full analysis set was used and is considered appropriate (Y).  Overall: Low risk	3.3: The participants were not stratified by disease severity. Not enough information to properly assess this domain (NI). 3.4: Due to the similarities between reasons for noncompletion and unlikeliness that the circumstances of the trial contributed to the missingness of the data, it is unlikely that missingness depended on the true value (PN). Overall: Some		intended analyses (N). Overall: Low risk	
			concerns			
Efficacy and Safety of Abrocitinib in Patients With Moderate-to- Severe Atopic Dermatitis A Randomized Clinical Trial	1.1: Randomization was conducted using an interactive response system (PY). 1.2: Study record detail and protocol indicated that study was double-	2.1: Participants were blinded throughout the study and placebo was used (N). 2.2: Individuals delivering the treatment were blinded and both the	3.1: Population for full analysis set was randomized. Greater than 5% and less than 20% of patients were discontinued from the study (N). 3.2: Although imputation was	4.1: Tools used for measurement are valid and appropriate tools to measure outcomes (N). 4.2: Same measurement methods and thresholds were used in all groups and	5.1: Data produced and analyzed were consistent with the pre-specified analysis plan per the protocol (Y). 5.2: All outcome measurements and analyses were included in either the	Low risk

	blinded and quadruple masked (Y). 1.3: No suspicious or excessive baseline imbalances or similarities noted (N). Overall: Low risk	treatment and study kits were blinded (N). 2.6: A full analysis set was used and is considered appropriate (Y). Overall: Low risk	used as one type of sensitivity analysis, linear mixed-effect models and tipping point analysis was also used, and is considered sufficient (PY).  Overall: Low risk	measurements were done at similar time points (N). 4.3: Outcome assessors were masked (N). Overall: Low risk	main paper, study record detail, or the supplementary appendix (N). 5.3: Eligible reported results for the outcome measurement correspond to all intended analyses (N). Overall: Low risk	
Safety and Efficacy of Upadacitinib for Atopic Dermatitis in Japan: 2-Year Interim Results from the Phase 3 Rising Up Study	1.1: Randomization was mentioned but no further details were provided (NI). 1.2: Study record detail indicated that study was double-blinded and quadruple masked (PY). 1.3: No suspicious or excessive baseline imbalances or similarities noted (N). Overall: Low risk	2.1: Participants were blinded throughout the study. Placebo was used and participants were rerandomized into treatment groups (N). 2.2: Study record detail indicated that carer was masked (PN). 2.6: Not enough information was provided (NI). 2.7: Not enough information was provided (NI). Overall: High risk	3.1: Population was randomized. Greater than 5% and less than 20% of patients were discontinued from the study (N). 3.2: No sensitivity analyses were performed and were reported as is (N). 3.3: Not enough information (NI). 3.4: Not enough information (NI). Overall: High risk	4.1: Tools used for measurement are valid and appropriate tools to measure outcomes (N). 4.2: Same measurement methods and thresholds were used in all groups and measurements were done at similar time points (N). 4.3: Outcome assessors were masked (N). Overall: Low risk	5.1, 5.2, 5.3: Not enough information was provided to sufficiently assess this domain. Overall: High risk	High risk