# Supplementary Table S1: Database search strategy after retrieval of relevant SRs

Disease group(s)	Date range	Database
COPD, ILD and	1st January 2020 to 8 <sup>th</sup> July	MEDLINE, CENTRAL
bronchiectasis	2022	
COPD and bronchiectasis	1st January 2020 to 8 <sup>th</sup> July	EMBASE
	2022	
ILD	1 <sup>st</sup> January 2019 to 8 <sup>th</sup> July	EMBASE
	2022	
CF	1st of January 2019 to 8 <sup>th</sup> July	MEDLINE, EMBASE,
	2022	CENTRAL
PHT	1st January 2015 to 8 <sup>th</sup> July	MEDLINE, EMBASE,
	2022	CENTRAL

## Supplementary Table S2: Relevant prior systematic reviews

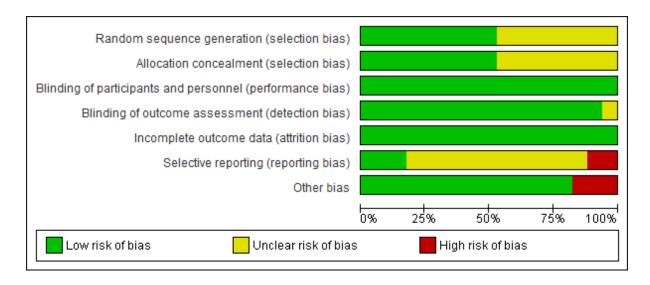
Prior Systematic Review	Relevant Studies
Bajwah, Colquitt [40]	None.
Barnes, McDonald [18]	Abernethy, Currow [51]
	Eiser, Denman [46]
	Johnson, Woodcock [47]
	Poole, Veale [54]
	Rice, Kronenberg [55]
	Woodcock, Gross [50]
Ekström, Nilsson [41]	Abernethy, Currow [51]
	Eiser, Denman [46]
	Johnson, Woodcock [47]
	Light, Stansbury [49]
	Poole, Veale [54]
	Woodcock, Gross [50]
	Woodcock, Johnson [56]
Jaiswal, Singh [39]	None.
Yamaguchi, Saif-Ur-Rahman [42]	Abdallah, Wilkinson-Maitland [45]
	Abernethy, Currow [51]
	Currow, Louw [31]
	Kronborg-White, Andersen [53]

	Eiser, Denman [46]
	Ferreira, Louw [52]
	Johnson, Woodcock [47]
	Kronborg-White, Andersen [53]
	Light, Muro [48]
	Light, Stansbury [49]
	Poole, Veale [54]
	Verberkt, Van Den Beuken-Van Everdingen [30]
	Woodcock, Gross [50]
	Woodcock, Johnson [56]
Total included relevant studies from	Abdallah, Wilkinson-Maitland [45]
prior systematic reviews	Abernethy, Currow [51]
	Currow, Louw [31]
	Eiser, Denman [46]
	Ferreira, Louw [52]
	Johnson, Woodcock [47]
	Kronborg-White, Andersen [53]
	Light, Muro [48]
	Light, Stansbury [49]
	1
	Poole, Veale [54]
	Poole, Veale [54]  Light, Stansbury [49]

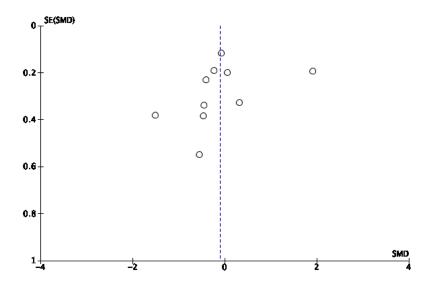
Woodcock, Gross [50]
Woodcock, Johnson [56]
Verberkt, Van Den Beuken-Van Everdingen [30]

	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other bias
Abdallah 2017	•	•	•	?	•	•	•
Abernethy 2003	•	•	•	•	•	?	•
Currow 2020	•	•	•	•	•	?	•
Eiser 1991(A)	?	?	•	•	•	?	•
Eiser 1991(B)	?	?	•	•	•	?	•
Ekstrom 2022	•	•	•	•	•	•	•
Ferreira 2018	•	•	•	•	•	?	
Ferreira 2020	•	•	•	•	•	?	
Johnson 1983	•	•	•	•	•	?	•
Kronborg-White 2020	•	•	•	•	•	•	•
Light 1989	?	?	•	•	•	?	•
Light 1996	?	?	•	•	•	?	•
Poole 1998	?	?	•	•	•	?	•
Rice 1987	?	?	•	•	•	?	
Verberkt 2020	•	•	•	•	•	•	•
Woodcock 1981	?	?	•	•	•	•	•
Woodcock 1982	?	?	•	•	•	?	•

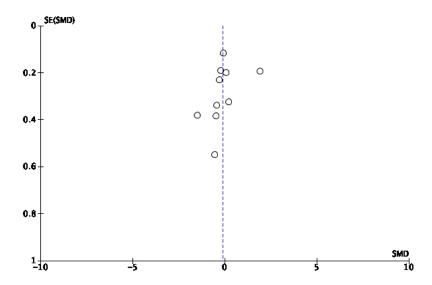
Supplementary Figure S1: Risk of Bias Summary



Supplementary Figure S2: Risk of bias domains



Supplementary Figure S3: Funnel plot for at home studies reporting breathlessness pragmatically during daily life



Supplementary Figure S4: Funnel plot for at home studies reporting breathlessness pragmatically during daily life using morning scores when available

			Opioids	Placebo		Std. Mean Difference	Std. Mean Difference	Risk of Bias
Study or Subgroup	Std. Mean Difference	SE	Total	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI	ABCDEFG
Abernethy 2003 (1)	-0.3007	0.2308	38	38	10.5%	-0.30 [-0.75, 0.15]	-	$\bullet \bullet \bullet \bullet \bullet ? \bullet$
Currow 2020 (2)	-0.0665	0.1187	145	139	11.2%	-0.07 [-0.30, 0.17]	+	$lackbox{}$
Ekstrom 2022 (3)	0.0619	0.1991	51	50	10.8%	0.06 [-0.33, 0.45]	+	•••••
Ferreira 2018 (4)	0.1996	0.3253	19	19	9.8%	0.20 [-0.44, 0.84]	+	••••• <del>•</del>
Ferreira 2020 (5)	1.9064	0.1947	74	81	10.8%	1.91 [1.52, 2.29]		••••• <del>•</del>
Johnson 1983 (6)	-0.444	0.3379	18	18	9.7%	-0.44 [-1.11, 0.22]	<del></del>	$lackbox{0.05}{\ }lackbox{0.05}{\ }lackbox{0.05}{\lackbox{0.05}{\ }lackbox{0.05}{\ }lackbox{0.05}{\ }lackbox{0.05}{\ }lackbox{0.05}{\ }lackbox{0.05}{\lackbox{0.05}{\ }lackbox{0.05}{\ }lackbox{0.05}{\lackbox{0.05}{\ }lackbox{0.05}{\lac$
Kronborg-White 2020 (7)	-1.4998	0.3824	18	18	9.3%	-1.50 [-2.25, -0.75]		•••••
Poole 1998 (8)	-0.457	0.3837	14	14	9.3%	-0.46 [-1.21, 0.30]	-+	??•••?•
Rice 1987 (9)	-0.5508	0.5484	7	7	7.8%	-0.55 [-1.63, 0.52]		??•••?
Verberkt 2020 (10)	-0.2335	0.1906	54	57	10.8%	-0.23 [-0.61, 0.14]	•	
Total (95% CI)			438	441	100.0%	-0.10 [-0.64, 0.43]	•	
Heterogeneity: Tau² = 0.66; C	hi² = 117.83, df = 9 (P	< 0.0000	1); I² = 92	:%			-10 -5 0 5 1	<u></u> 10
Test for overall effect: $Z = 0.3$	7 (P = 0.71)						Favours opioids Favours placeb	

- (1) 10cm VAS breathlessness intensity (final morning relative to baseline)
- (2) 10cm VAS breathlessness intensity NOW (days 5-7 average of mean morning/evening scores relative to...
- (3) NRS intensity of breathlessness (16mg/day dose, days 5 to 7 average scores relative to days -3 to -1 average...(C) Blinding of participants and personnel (performance...
- (4) 10cm VAS breathlessness intensity NOW (final evening score relative to baseline)
- (5) 10cm VAS breathlessness intensity NOW (days 5-7 average of mean morning/evening scores relative to...
- (6) 10cm VAS breathlessness (final early evening relative to baseline from alternating weeks period)
- (7) 10cm VAS breathlessness during LAST HOUR (change from baseline to followup)
- (8) CRQ Dyspnea subscale (change from baseline to 6-weeks)
- (9) 10cm VAS breathlessness during PAST 24 HOURS (change from baseline to followup)
- (10) NRS over PAST 24 HOURS (change from baseline to followup)

#### Risk of bias legend

- (A) Random sequence generation (selection bias)
- (B) Allocation concealment (selection bias)
- (D) Blinding of outcome assessment (detection bias)
- (E) Incomplete outcome data (attrition bias)
- (F) Selective reporting (reporting bias)
- (G) Other bias

Supplementary Figure S5: At home studies reporting breathlessness pragmatically during daily life using morning scores when available

			Opioids	Placebo		Std. Mean Difference
Study or Subgroup	Std. Mean Difference	SE	Total	Total	Weight	IV, Random, 95% CI
Kronborg-White 2020 (1)	-2.7656	0.48	18	18	48.6%	-2.77 [-3.71, -1.82]
Verberkt 2020 (2)	-0.1389	0.206	44	51	51.4%	-0.14 [-0.54, 0.26]
Total (95% CI)			62	69	100.0%	-1.42 [-3.99, 1.16]

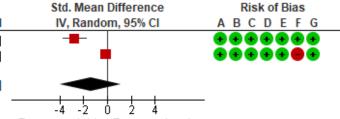
Heterogeneity: Tau $^2$  = 3.31; Chi $^2$  = 25.29, df = 1 (P < 0.00001);  $I^2$  = 96%

Test for overall effect: Z = 1.08 (P = 0.28)

#### Footnotes

- (1) Leicester Cough Score (change from baseline)
- (2) CAT Cough subscore (raw score at one month)

Supplementary Figure S6: At home studies reporting cough



Favours opioids Favours placebo

- (A) Random sequence generation (selection bias)
- (B) Allocation concealment (selection bias)
- (C) Blinding of participants and personnel (performance...
- (D) Blinding of outcome assessment (detection bias)
- (E) Incomplete outcome data (attrition bias)
- (F) Selective reporting (reporting bias)
- (G) Other bias

			Opioids	Placebo		Std. Mean Difference	Std. Mean Difference	Risk of Bias
Study or Subgroup	Std. Mean Difference	SE	Total	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI	ABCDEFG
Eiser 1991(B) (1)	1.013	0.5409	8	8	35.1%	1.01 [-0.05, 2.07]	-	??•••?•
Light 1989 (2)	0.4261	0.3974	13	13	64.9%	0.43 [-0.35, 1.20]	<b>-</b>  ■-	??•••?•
Total (95% CI)			21	21	100.0%	0.63 [0.00, 1.26]	•	
Heterogeneity: Tau² =	0.00; Chi² = 0.76, df = 1	(P = 0.38)	3); I² = 0%	ı			1 1	<del>.</del>
Test for overall effect:	Z=1.97 (P=0.05)						Favours opioids Favours placebo	•
<u>Footnotes</u>							Risk of bias legend	
(1) 1 hour post-dose							(A) Random sequence generation (s	selection bias)
(2) at isotime							(B) Allocation concealment (selectio	n bias)
							(C) Blinding of participants and pers	onnel (performance
							(D) Blinding of outcome assessmen	t (detection bias)
							(E) Incomplete outcome data (attritio	n bias)
							(F) Selective reporting (reporting bias	S)
							(G) Other bias	

Supplementary Figure S7: Laboratory-based exercise studies reporting arterial blood gases (partial pressure of carbon dioxide)

				Std. Mean Difference	Std. Mean Difference	Risk of Bias
Study or Subgroup	Std. Mean Difference	SE	Weight	IV, Random, 95% CI	IV, Random, 95% CI	ABCDEFG
Eiser 1991(B) (1)	-0.7839	0.5249	36.3%	-0.78 [-1.81, 0.24]		??+++?+
Light 1989 (2)	-0.3679	0.3961	63.7%	-0.37 [-1.14, 0.41]		??•••?•
Total (95% CI)			100.0%	-0.52 [-1.14, 0.10]	•	
Heterogeneity: Tau² =	0.00; Chi <sup>2</sup> = $0.40$ , df = $1$	(P = 0.53)	3); I <sup>z</sup> = 0%	)	4 -2 0 2	<del>,</del>
Test for overall effect:	Z = 1.64 (P = 0.10)				Favours placebo Favours opioid	ds
<u>Footnotes</u>					Risk of bias legend	
(1) 1 hour post-dose					(A) Random sequence generation	ı (selection bias)
(2) at isotime					(B) Allocation concealment (selec	tion bias)
					(C) Blinding of participants and pe	rsonnel (performance
					(D) Blinding of outcome assessm	ent (detection bias)
					(E) Incomplete outcome data (attri	tion bias)
					(F) Selective reporting (reporting b	ias)
					(G) Other bias	

Supplementary Figure S8: Laboratory-based exercise studies reporting arterial blood gases (partial pressure of oxygen)

			Opioids	Placebo		Std. Mean Difference	Std. Mean Difference	Risk of Bias
Study or Subgroup	Std. Mean Difference	SE	Total	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI	ABCDEFG
Kronborg-White 2020 (1)	1.76	0.3993	18	18	25.2%	1.76 [0.98, 2.54]	-	
Rice 1987 (2)	0.9161	0.5722	7	7	20.4%	0.92 [-0.21, 2.04]	<del>  • -</del>	??•••?
Verberkt 2020 (3)	0.1131	0.1901	54	57	30.5%	0.11 [-0.26, 0.49]		$\bullet \bullet \bullet \bullet \bullet \bullet \bullet$
Woodcock 1982 (4)	0.8064	0.447	11	11	23.9%	0.81 [-0.07, 1.68]	-	??•••?•
Total (95% CI)			90	93	100.0%	0.86 [0.03, 1.69]	<b>•</b>	
Heterogeneity: Tau² = 0.55 Test for overall effect: Z = 2		: 0.002); f	²= 80%				-10 -5 0 5 Favours opioids Favours placeb	10 00

## <u>Footnotes</u>

- (1) Change from baseline
- (2) Raw score at follow up
- (3) change from baseline
- (4) Raw score at follow up

#### Risk of bias legend

- (A) Random sequence generation (selection bias)
- (B) Allocation concealment (selection bias)
- (C) Blinding of participants and personnel (performance...
- (D) Blinding of outcome assessment (detection bias)
- (E) Incomplete outcome data (attrition bias)
- (F) Selective reporting (reporting bias)
- (G) Other bias

Supplementary Figure S9: At home studies reporting arterial blood gases (partial pressure of carbon dioxide)

			Opioids	Placebo		Std. Mean Difference	Std. Mean Difference	Risk of Bias
Study or Subgroup	Std. Mean Difference	SE	Total	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI	ABCDEFG
Verberkt 2020 (1)	-0.2335	0.1906	54	57	83.4%	-0.23 [-0.61, 0.14]		
Woodcock 1982 (2)	-0.1431	0.4271	11	11	16.6%	-0.14 [-0.98, 0.69]	+	??•••?•
Total (95% CI)			65	68	100.0%	-0.22 [-0.56, 0.12]	•	
Heterogeneity: Tau² : Test for overall effect	= 0.00; Chi² = 0.04, df = 1 : Z = 1.26 (P = 0.21)	(P = 0.85	5); I² = 0%	ı			-4 -2 0 2 4 Favours placebo Favours opioids	_

- (1) change from baseline
- (2) Raw score at follow up

Supplementary Figure S10: At home studies reporting arterial blood gases (partial pressure of oxygen)

- (A) Random sequence generation (selection bias)
- (B) Allocation concealment (selection bias)
- (C) Blinding of participants and personnel (performance...
- (D) Blinding of outcome assessment (detection bias)
- (E) Incomplete outcome data (attrition bias)
- (F) Selective reporting (reporting bias)
- (G) Other bias

Study or Subgroup	log[Odds Ratio]	\$E	Opioids Total	Placebo Total	Weight	Odds Ratio IV, Random, 95% CI	Odds Ratio IV, Random, 95% CI	Risk of Bias ABCDEFG
Abdallah 2017 (1)	1.1486	1.6633	20	20	24.8%	3.15 [0.12, 82.16]		•••?••
Eiser 1991(A) (2)	1.1542	1.6667	18	18	24.7%	3.17 [0.12, 83.17]	<del></del>	?? • • • ? •
Eiser 1991(B) (3)	1.1987	1.6933	10	10	23.9%	3.32 [0.12, 91.61]		?? • • • ? •
Woodcock 1981 (4)	1.7838	1.6048	12	12	26.6%	5.95 [0.26, 138.26]	-	?? • • • •
Total (95% CI)			60	60	100.0%	3.79 [0.75, 19.18]	•	
Heterogeneity: Tau <sup>2</sup> =	0.00; Chi² = 0.11,	df = 3 (P	= 0.99); l²	= 0%			0.001 0.1 1 10 10	<del>   </del> 000
Test for overall effect:	Z=1.61 (P=0.11)	l					Favours opioids Favours place	
<u>Footnotes</u>							Risk of bias legend	
(1) count of self-repor	ted events						(A) Random sequence generatio	n (selection bias)
(2) count of self-repor	ted events						(B) Allocation concealment (selection)	ction bias)
(3) count of self-repor	ted events						(C) Blinding of participants and pe	ersonnel (performance
(4) count of self-repor	ted events						(D) Blinding of outcome assessm	nent (detection bias)
							(E) Incomplete outcome data (attr	ition bias)
							(F) Selective reporting (reporting b	bias)
							(G) Other bias	

Supplementary Figure S11: Laboratory-based exercise studies reporting adverse events (nausea and/or vomiting)

			Opioids	Placebo		Odds Ratio	Odds Ratio	Risk of Bias
Study or Subgroup	log[Odds Ratio]	SE	Total	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI	ABCDEFG
Currow 2020 (1)	0.6699	0.2662	142	137	23.7%	1.95 [1.16, 3.29]		lacksquare
Ekstrom 2022 (2)	0.6568	0.5976	51	50	15.1%	1.93 [0.60, 6.22]	+-	$lackbox{}$
Ferreira 2018 (3)	1.8608	1.1519	19	19	6.7%	6.43 [0.67, 61.47]	+	lacksquare
Ferreira 2020 (4)	4.0771	1.0342	72	79	7.9%	58.97 [7.77, 447.69]		→ •••••?●
Kronborg-White 2020 (5)	1.0296	0.7392	18	18	12.1%	2.80 [0.66, 11.92]	+-	ullet
Poole 1998 (6)	1.6864	0.8028	16	14	11.0%	5.40 [1.12, 26.05]	-	??•••?•
Verberkt 2020 (7)	0.3542	0.4341	54	57	19.2%	1.43 [0.61, 3.34]	+-	$\bullet \bullet \bullet \bullet \bullet \bullet \bullet$
Woodcock 1982 (8)	2.7589	1.5262	16	16	4.3%	15.78 [0.79, 314.25]	<del>                                     </del>	<b>→ ??•••?•</b>
Total (95% CI)			388	390	100.0%	3.32 [1.70, 6.51]	•	
Heterogeneity: Tau² = 0.43;	; Chi² = 15.07, df =	7 (P = 0.0	$(4); I^2 = 54$	4%				<del></del>
Test for overall effect: Z = 3	.50 (P = 0.0005)	-					0.01 0.1 1 10 10 Favours opioids Favours placet	Ò0
	· · · · · · · · · · · · · · · · · · ·						ravours opioius ravours placei	JU

- (1) count of self-reported events
- (2) count of self-reported events
- (3) count of self-reported events
- (4) count of self-reported events
- (5) count of self-reported events
- (6) count of self-reported events
- (7) count of participants with increase of 2 or more points on symptom NRS relative to baseline
- (8) count of self-reported events

Supplementary Figure S12: At home studies reporting adverse events (nausea and/or vomiting)

- (A) Random sequence generation (selection bias)
- (B) Allocation concealment (selection bias)
- (C) Blinding of participants and personnel (performance...
- (D) Blinding of outcome assessment (detection bias)
- (E) Incomplete outcome data (attrition bias)
- (F) Selective reporting (reporting bias)
- (G) Other bias

		(	Opioids	Placebo		Odds Ratio	Odds Ratio	Risk of Bias
Study or Subgroup	log[Odds Ratio]	SE	Total	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI	ABCDEFG
Currow 2020 (1)	0.3558	0.2419	142	137	41.2%	1.43 [0.89, 2.29]	+-	$lackbox{0.05}{\ }lackbox{0.05}{\ }lackbox{0.05}{\lackbox{0.05}{\ }lackbox{0.05}{\ }lackbox{0.05}{\ }lackbox{0.05}{\ }lackbox{0.05}{\ }lackbox{0.05}{\ }lackbox{0.05}{\lackbox{0.05}{\ }lackbox{0.05}{\ }lackbox{0.05}{\lackbox{0.05}{\ }$
Ekstrom 2022 (2)	0.452	0.5695	51	50	7.4%	1.57 [0.51, 4.80]	<del></del>	$\bullet \bullet \bullet \bullet \bullet \bullet \bullet$
Ferreira 2018 (3)	0.4238	0.6535	19	19	5.6%	1.53 [0.42, 5.50]	<del></del>	⊕ ⊕ ⊕ ⊕ ? ●
Ferreira 2020 (4)	-0.2247	0.3275	72	79	22.5%	0.80 [0.42, 1.52]	<del></del>	lacksquare
Poole 1998 (5)	1.0986	0.7601	16	14	4.2%	3.00 [0.68, 13.31]	+	?? • • • ? •
Rice 1987 (6)	0.5754	1.0801	7	7	2.1%	1.78 [0.21, 14.77]	<del>-   -</del>	??•••?•
Verberkt 2020 (7)	0.539	0.3866	54	57	16.1%	1.71 [0.80, 3.66]	+•	$\bullet \bullet \bullet \bullet \bullet \bullet \bullet$
Woodcock 1982 (8)	1.8005	1.61	11	11	0.9%	6.05 [0.26, 142.03]	<del>-   ·  </del>	??
Total (95% CI)			372	374	100.0%	1.37 [1.01, 1.86]	<b>*</b>	
Heterogeneity: Tau <sup>2</sup> =	0.00; Chi² = 5.14,	df = 7 (P =	0.64); l²	= 0%			1 10 10	<del>(</del>
Test for overall effect:							0.01 0.1 1 10 100 Favours opioids Favours placebo	

- (1) count of self-reported events
- (2) count of self-reported events
- (3) count of self-reported events
- (4) count of self-reported events
- (5) count of self-reported events
- (6) count of self-reported events
- (7) count of participants with increase of 2 or more points on symptom NRS relative to baseline
- (8) count of self-reported events

#### Risk of bias legend

- (A) Random sequence generation (selection bias)
- (B) Allocation concealment (selection bias)
- (C) Blinding of participants and personnel (performance...
- (D) Blinding of outcome assessment (detection bias)
- (E) Incomplete outcome data (attrition bias)
- (F) Selective reporting (reporting bias)
- (G) Other bias

Supplementary Figure S13: At home studies reporting adverse events (drowsiness)

			Opioids	Placebo		Odds Ratio	Odds Ratio	Risk of Bias
Study or Subgroup	log[Odds Ratio]	SE	Total	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI	ABCDEFG
Abernethy 2003 (1)	2.3838	1.0761	48	48	6.1%	10.85 [1.32, 89.38]		$lackbox{0.05}$
Currow 2020 (2)	0.5055	0.2415	142	137	21.7%	1.66 [1.03, 2.66]	<del></del>	lacksquare
Ekstrom 2022 (3)	1.7492	0.6	51	50	12.8%	5.75 [1.77, 18.64]	<del></del>	$\bullet \bullet \bullet \bullet \bullet \bullet \bullet$
Ferreira 2018 (4)	3.5635	1.5006	19	19	3.6%	35.29 [1.86, 668.22]	<del></del>	→ •••••?•
Ferreira 2020 (5)	0.0809	0.3259	72	79	19.5%	1.08 [0.57, 2.05]	+	$lackbox{0.5}{\bullet} lackbox{0.5}{\bullet} lackbox{0.5}{\bullet} lackbox{0.5}{\bullet} lackbox{0.5}{\bullet}$
Kronborg-White 2020 (6)	1.6275	0.8923	18	18	8.0%	5.09 [0.89, 29.26]	<del>  • -</del>	$\bullet \bullet \bullet \bullet \bullet \bullet \bullet$
Poole 1998 (7)	2.5337	0.9394	16	14	7.5%	12.60 [2.00, 79.43]	<del></del>	??+++?+
Verberkt 2020 (8)	0.7072	0.3979	54	57	17.6%	2.03 [0.93, 4.42]	<del>  • -</del>	$lackbox{0.05}$
Woodcock 1982 (9)	1.8005	1.61	11	11	3.2%	6.05 [0.26, 142.03]	-	??•••?•
Total (95% CI)			431	433	100.0%	3.08 [1.69, 5.61]	•	
Heterogeneity: $Tau^2 = 0.37$ Test for overall effect: $Z = 3$		8 (P = 0.	02); I² = 57	7%			0.002 0.1 1 10 50 Favours opioids Favours placeb	<del>-</del>

- (1) count of self-reported events
- (2) count of self-reported events
- (3) count of self-reported events
- (4) count of self-reported events
- (5) count of self-reported events
- (6) count of self-reported events
- (7) count of self-reported events
- (8) count of participants with increase of 2 or more points on symptom NRS relative to baseline
- (9) count of self-reported events

Supplementary Figure S14: At home studies reporting adverse events (constipation)

- (A) Random sequence generation (selection bias)
- (B) Allocation concealment (selection bias)
- (C) Blinding of participants and personnel (performance...
- (D) Blinding of outcome assessment (detection bias)
- (E) Incomplete outcome data (attrition bias)
- (F) Selective reporting (reporting bias)
- (G) Other bias

## Supplementary Table S3: GRADE certainty of evidence in laboratory-based exercise studies

		(	Certainty assessr	nent			Nº of p	oatients		Effect	Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	opioids	no opioids	Relative (95% CI)	Absolute (95% CI)		
Breathless	sness after exercise (i	so-time & is	so-load)								•	
5	randomised trials – all crossover	Serious <sup>a</sup>	Not serious	Serious <sup>b</sup>	Serious <sup>c</sup>	Yes <sup>d</sup>	70	70	-	SMD -0.50 - 0.84 to -0.16 p=0.004 Favours opioids	⊕○○○ very low	CRITICAL
Breathless	sness after exercise (i	so-time onl	y)									
3	randomised trials – all crossover	Serious <sup>a</sup>	Not serious	Serious <sup>b</sup>	Serious <sup>c</sup>	Yes <sup>d</sup>	40	40	-	SMD -0.57 - 1.02 to -0.12 p=0.01 Favours opioids	⊕○○○ very low	CRITICAL
Breathless	sness after exercise (i	so-load only	y)									
2	randomised trials – all crossover	Serious <sup>a</sup>	Not serious	Not serious	Serious <sup>c,e</sup>	Yes <sup>d</sup>	30	30	-	SMD -0.41 -0.92 to 0.11 p=0.12	⊕○○○ very low	CRITICAL
ABG PaCC	)2 after exercise										l	
2	randomised trials – all crossover	Serious <sup>f</sup>	Not serious	Serious <sup>b</sup>	Serious <sup>c,g</sup>	Yes <sup>d a</sup>	21	21	-	SMD 0.63 0.0 to 1.26 p=0.05	⊕○○○ very low	IMPORTANT
ABG PaO2	2 after exercise			•			,				•	
2	randomised trials – all crossover	Serious <sup>f</sup>	Not serious	Serious <sup>b</sup>	Serious <sup>c,h</sup>	Yes <sup>d</sup>	21	21	-	SMD -0.52 -1.14 to 0.10 p=0.10	⊕○○○ very low	IMPORTANT
Adverse e	vents – Nausea or vo	miting	L	<u>I</u>			1		I		I	<u> </u>

4	randomised trials – all crossover	Serious <sup>a</sup>	Not serious	Serious <sup>b</sup>	Serious <sup>c+</sup>	Yes <sup>d</sup>	60	60	-	OR 3.79 0.75 to 19.18 p=0.11	⊕○○○ very low	IMPORTANT
---	-----------------------------------	----------------------	-------------	----------------------	-----------------------	------------------	----	----	---	------------------------------------	------------------	-----------

**CI:** confidence interval; **OR:** risk ratio; **SMD:** standardised mean difference

#### **Explanations:**

- a. >=50% of weighting comes from studies with uncertainty or high risk for selection bias and reporting bias
- b. All or most studies only include people with COPD
- c. Small numbers of patients in the included studies contributes to imprecision in the outcome estimate
- d. Most studies had risk of carryover effect due to inadequate washout period in crossover design. The carryover effect is not considered under risk of bias in the ROB1 tool, therefore this was considered separately as an additional consideration.
- e. The pooled estimate of the effect of opioids on breathlessness includes both small harm and large benefit
- f. All studies have uncertainty regarding for selection bias and reporting bias
- g. The pooled estimate of the effect of opioids on ABG PaCO2 includes both large harm and no benefit
- h. The pooled estimate of the effect of opioids on ABG PaO2 includes both large harm and small benefit
- i. The pooled estimate of the effect of opioids on nausea or vomiting includes both large harm and no harm

## Supplementary Table S4: GRADE certainty of evidence in at home studies

			Certainty assess	sment			Nº of p	oatients		Effect	Certainty	Importance
Nº of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	opioids	no opioids	Relative (95% CI)	Absolute (95% CI)		
Breathles	sness intensity (eve	ening measu	res or time not s	tated)								
10	randomised trials – 5 crossover	Not serious <sup>a</sup>	Serious <sup>b</sup>	Not serious	Serious <sup>c</sup>	Yes <sup>d</sup>	438	441	-	SMD -0.10 -0.64 to 0.44 p=0.71	⊕○○○ very low	CRITICAL
Breathles	sness intensity (mo	rning measu	ures or time not	stated)							•	
10	randomised trials – 5 crossover	Not serious <sup>a</sup>	Serious <sup>b</sup>	Not serious	Serious <sup>c</sup>	Yes <sup>d</sup>	438	441	-	SMD -0.10 -0.64 to 0.43 p=0.71	⊕○○○ very low	CRITICAL
Health-Re	elated Quality of Lif	e										
6	randomised trials – 1 crossover	Not serious <sup>a</sup>	Serious <sup>b</sup>	Not serious	Serious <sup>e</sup>	No	356	359	-	SMD -0.42 -0.98 to 0.13 p=0.13	⊕⊕○○ low	IMPORTANT
Cough			1	l	I		.1		l		1	1
2	randomised trials – 0 crossover	Not serious	Serious <sup>b</sup>	Not serious	Serious <sup>f</sup>	No	62	69	-	SMD -1.42 -3.99 to 1.16 p=0.28	⊕⊕○○ low	IMPORTANT
ABG PaCC	)2	•					•		•		1	
4	randomised trials – 2 crossover	Serious <sup>g</sup>	Serious <sup>b</sup>	Serious <sup>h</sup>	Serious <sup>f</sup>	Yes <sup>i</sup>	90	93	-	SMD 0.86 0.03 to 1.69 p=0.04 favours placebo	⊕○○○ very low	IMPORTANT

ABG PaO	2											
2	randomised trials – 1 crossover	Serious <sup>j</sup>	Not serious	Serious <sup>k</sup>	Serious <sup>1</sup>	Yes <sup>i</sup>	65	68	-	SMD -0.22 -0.56 to 0.12 p=0.21	⊕○○○ very low	IMPORTANT
Adverse 6	events - Drowsiness											
8	randomised trials – 4 crossover	Serious m	Not serious	Not serious	Serious <sup>n</sup>	No	372	374	-	OR 1.37 1.01 to 1.86 p=0.04	⊕⊕○○ low	IMPORTANT
Adverse (	events - Constipatio	n					•	1	•			
9	randomised trials – 4 crossover	Serious °	Serious <sup>p</sup>	Not serious	Not serious	No	431	433	-	OR 3.08 1.69 to 5.61 p=0.0002 favours placebo	⊕⊕○○ low	IMPORTANT
Adverse o	events – Nausea or v	omiting		I						I		1
8	randomised trials – 3 crossover	Serious <sup>q</sup>	Serious <sup>p</sup>	Not serious	Not serious	No	388	390	-	OR 3.32 1.70 to 6.51 p=0.0005 favours placebo	⊕⊕⊖⊖ low	IMPORTANT

CI: confidence interval; OR: risk ratio; SMD: standardised mean difference

#### **Explanations:**

- a. Nearly all studies had an unclear risk regarding selective reporting as most did not publish study protocols before publishing trial outcomes. One study was considered at high risk of bias for selective reporting. However, overall the risk of bias from all domains was considered not serious
- b. Significant heterogeneity identified I<sup>2</sup> >80%, which is not explained by differences in study design or study populations
- c. The pooled estimate of the effect of opioids on breathlessness includes both strong benefit and moderate harm
- d. 3 of the 5 crossover studies had a risk of carryover effect due to inadequate washout period in crossover design. The carryover effect is not considered under risk of bias in the ROB1 tool, therefore this was considered separately as an additional consideration.
- e. The pooled estimate of the effect of opioids on QOL includes both small harm and large benefit
- f. Small numbers of patients in the included studies contributes to imprecision in the outcome estimate
- g. 45% of weighting comes from studies with uncertainty regarding selection bias and 75% of weighting comes from studies with uncertainty or high risk regarding reporting bias
- h. 75% of weighting comes from studies with only people with COPD
- i. 1 of the 2 crossover studies had a risk of carryover effect due to inadequate washout period in crossover design. The carryover effect is not considered under risk of bias in the ROB1 tool, therefore this was considered separately as an additional consideration.

- j. 17% of weighting comes from studies with uncertainty regarding selection bias and 100% of weighting comes from studies with uncertainty or high risk regarding reporting bias
- k. Only patients with COPD included in the studies for this outcome
- I. The pooled estimate of the effect of opioids on PaO2 includes both moderate harm and small benefit
- m. All but 1 study had uncertainty or high risk regarding reporting bias and 30% of the weighting comes from studies with high risk for other bias
- n. The pooled estimate of the effect of opioids on drowsiness includes both negligible and large harm
- o. 79% of the weighting comes from studies with uncertainty or high risk regarding reporting bias and 23% of the weighting comes from studies with high risk for other bias
- p. Significant heterogeneity identified I<sup>2</sup> >50%, which is not explained by differences in study design or study populations
- q. 73% of the weighting comes from studies with uncertainty or high risk regarding reporting bias and 15% of the weighting comes from studies with high risk for other bias