

Additional file 7 to: Can it be checked? A cross-sectional meta-research audit of access to, and reporting in, machine learning studies of pharmaceutical manufacturing and quality control.
Phani Kumar Balagam.

Supplementary tables

Table S1. Each item by venue class: Y/(Y+N), Fisher-Freeman-Halton Monte Carlo p and Holm-adjusted p

item	pharmaceutical sciences	chemical and process engineering	analytical chemistry and spectroscopy	biotechnology and bioprocess	computing and engineering	multidisciplinary	p_montecarlo	p_holm
R1	13/53	3/9	1/6	4/15	2/9	25/88	0.9829	1.0
R2	5/54	0/9	1/6	1/15	0/9	12/92	0.7792	1.0
R3	50/53	9/9	6/6	15/15	9/9	84/92	0.9132	1.0
R4	15/54	3/9	3/6	5/15	4/9	35/92	0.7378	1.0
R5	7/54	1/9	2/6	3/15	1/9	21/89	0.5544	1.0
R6	30/54	3/9	4/6	10/15	3/9	48/91	0.5009	1.0
R7	12/18	2/2	0/0	1/1	4/5	10/17	0.8108	1.0
R8	17/54	2/9	3/6	7/15	1/9	28/92	0.4726	1.0
R9	13/54	3/9	2/6	4/15	3/9	29/92	0.9336	1.0
composite	0/52	0/9	0/6	0/15	0/9	2/88	0.7151	1.0

Table S2. Each item by year of first online publication: odds ratio per year (Wald 95% CI), p and Holm-adjusted p

item	n	OR_per_year	ci_lo	ci_hi	p	p_holm
R1	180	0.979	0.813	1.179	0.824	1.0
R2	185	0.993	0.762	1.293	0.9556	1.0
R3	184	1.009	0.719	1.416	0.9586	1.0
R4	185	1.297	1.078	1.561	0.0059	0.0529
R5	182	1.434	1.116	1.842	0.0049	0.0488
R6	184	0.94	0.799	1.105	0.4525	1.0
R7	43	0.934	0.647	1.346	0.7127	1.0
R8	185	1.136	0.949	1.361	0.165	1.0
R9	185	1.144	0.951	1.376	0.1545	1.0

item	n	OR_per_y ear	ci_lo	ci_hi	p	p_holm
composit e	179	0.851	0.404	1.794	0.6711	1.0

Table S3. Item counts (Y, N, NA, NASS) with Wilson 95% CIs

item	name	Y	N	NA	NASS	denom inator	prop	ci_lo	ci_hi
R1	Data availab ility	48	132	0	6	180	0.2667	0.2074	0.3357
R2	Code availab ility	19	166	0	1	185	0.1027	0.0667	0.1548
R3	Data split stated	173	11	0	2	184	0.9402	0.8961	0.9663
R4	Held- out evaluat ion	65	120	0	1	185	0.3514	0.2862	0.4225
R5	Uncert ainty on headli ne metric	35	147	0	4	182	0.1923	0.1416	0.2557
R6	Baselin e compa rison	98	86	0	2	184	0.5326	0.4606	0.6033
R7	Class distrib ution	29	14	139	4	43	0.6744	0.5252	0.7951
R8	Extern al validat ion	58	127	0	1	185	0.3135	0.251	0.3836
R9	Repeat ed evaluat ion	54	131	0	1	185	0.2919	0.2311	0.3611
compo	Compo	2	177	0	7	179	0.0112	0.0031	0.0398

item	name	Y	N	NA	NASS	denom inator	prop	ci_lo	ci_hi
site	site (R1-R4 all Y)								

Table S4. Agreement between two model passes on the 20% sample (37 papers, 333 items): percent agreement and Cohen's kappa between the first-pass scores and a second model pass that did not see them; this is agreement between two passes of the same model, not inter-rater reliability

item	papers	passes agree	agreement %	Cohen's kappa
R1 data availability	37	32	86.5	0.66
R2 code availability	37	35	94.6	0.84
R3 data split stated	37	34	91.9	0.37
R4 held-out evaluation	37	15	40.5	0.08
R5 uncertainty	37	25	67.6	0.31
R6 baseline comparison	37	31	83.8	0.65
R7 class distribution	37	34	91.9	0.66
R8 external validation	37	32	86.5	0.65
R9 repeated evaluation	37	25	67.6	0.26
All items	333	263	79.0	0.64

Table S5. Provenance of the item values in the 20% sample: adjudicated (the two passes disagreed; resolved against the verbatim rule with a recorded reason), reviewed (the passes agreed; in the random 25% re-read the same way), accepted (the passes agreed; not re-read). All resolution was model work

item	adjudicated	reviewed	accepted	changed when adjudicated	changed when reviewed
R1	5	10	22	3	0
R2	2	7	28	1	0

item	adjudicated	reviewed	accepted	changed when adjudicated	changed when reviewed
R3	3	11	23	0	0
R4	22	5	10	1	0
R5	12	5	20	1	0
R6	6	5	26	1	0
R7	3	8	26	0	0
R8	5	9	23	0	0
R9	12	6	19	0	0
all items	70	66	197	7	0

Table S6. Sensitivity analysis: item counts and Wilson 95% CIs on the first-pass (model-only) scores

item	name	Y	N	NA	NASS	denom inator	prop	ci_lo	ci_hi
R1	Data availab ility	46	131	0	9	177	0.2599	0.2009	0.3291
R2	Code availab ility	18	166	0	2	184	0.0978	0.0628	0.1493
R3	Data split stated	173	11	0	2	184	0.9402	0.8961	0.9663
R4	Held- out evaluat ion	64	121	0	1	185	0.3459	0.2812	0.417
R5	Uncert ainty on headli ne metric	34	147	0	5	181	0.1878	0.1377	0.251
R6	Baselin e compa rison	97	87	0	2	184	0.5272	0.4552	0.598
R7	Class distrib ution	29	14	139	4	43	0.6744	0.5252	0.7951

item	name	Y	N	NA	NASS	denom inator	prop	ci_lo	ci_hi
R8	External validation	58	127	0	1	185	0.3135	0.251	0.3836
R9	Repeated evaluation	54	131	0	1	185	0.2919	0.2311	0.3611
composite	Composite (R1-R4 all Y)	2	174	0	10	176	0.0114	0.0031	0.0405

Table S7. Papers with public data, public code and a stated split (R1, R2 and R3 all Y), among papers with all three items assessable, with Wilson 95% CIs

Scores	Papers	Assessable	% (95% CI)
Resolved (main results)	12	179	6.7 (3.9-11.4)
First pass (model-only sensitivity)	12	176	6.8 (3.9-11.5)