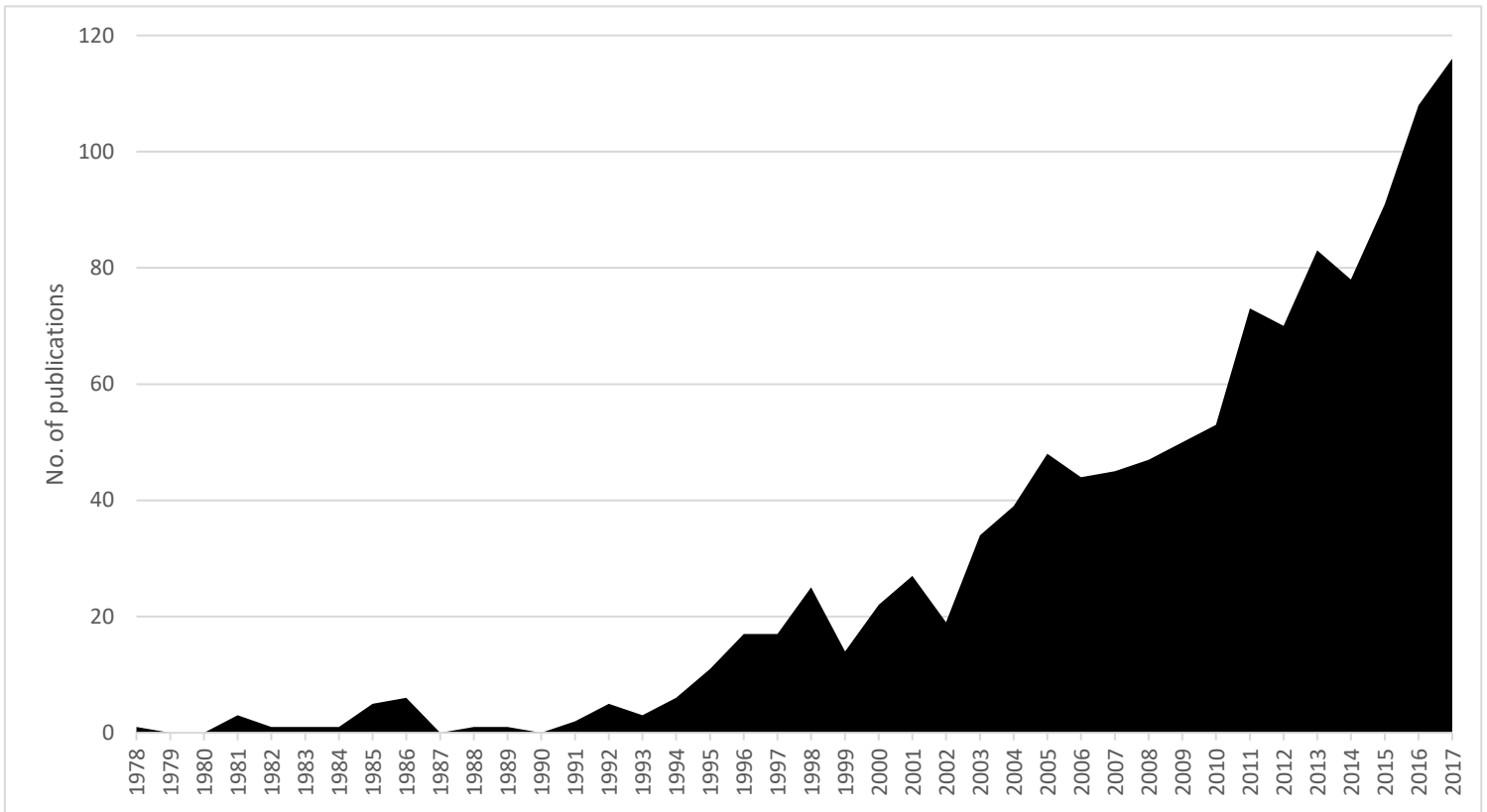


Supplementary Material File 2: Tables and figures

Supplementary Figure 1: Publications by year (n=1167)



Supplementary Table 1: Cohort Characteristics – Host studies

Articles could be coded to as many categories as relevant within each field.

Field	Category	Frequency of articles	% (n=1167)
Host Study Design	RCT	671	57%
	Cohort	320	27%
	Unknown	71	6%
	Cluster RCT	44	4%
	Survey	29	2%
	Factorial RCT	22	2%
	Non-Randomised Pilot/ Feasibility Study	17	1%
	Crossover RCT	16	1%
	Registry	16	1%
	Case Control	15	1%
	Early Phase Trials (Not Randomised)	10	1%
	Other	10	1%
	Patient Preference	7	1%
	Biobank	4	<1%
	Non-Randomised Pilot	2	<1%
	N-Of-1	1	<1%
Location	N. America	501	43%
	Europe	303	26%
	Unknown	266	23%
	Australasia	101	9%
	Asia	49	4%
	Africa	46	4%
	S. America	23	2%
	N/A	5	<1%
Recruitment approach	Unknown	547	47%
	Direct	533	46%
	Indirect	153	13%
Recruitment setting	Unknown	469	40%
	Secondary or Tertiary Care	340	29%
	Other	263	22%
	Primary Care	120	10%
	School	39	3%
	Emergency Care	30	3%
	Intensive Care	10	1%
	Dentist	6	1%
Health Area (HRCS)	Mental Health	292	25%
	Infection	136	12%
	Cancer	135	12%

	Generic	135	12%
	Unknown	109	9%
	Reproduction	87	7%
	Cardiovascular	80	7%
	Neurological	50	4%
	Respiratory	49	4%
	Metabolic and Endocrine	49	4%
	Musculoskeletal	40	3%
	Injuries	36	3%
	Inflammatory and Immune	27	2%
	Stroke	21	2%
	Renal	19	2%
	Oral And Gastro	16	1%
	Eye	8	1%
	Blood	6	1%
	Ear	6	1%
	Other	5	<1%
	Skin	5	<1%
	Congenital	4	<1%
Host intervention ¹ (n=866)	Behavioural	321	37%
	Drug	221	26%
	Unknown	165	19%
	Physical	54	6%
	Other	50	6%
	Care Pathway	46	5%
	Surgery	32	4%
	Vaccine	22	3%
	Medical Device	18	2%
	Screening	12	1%
	Complimentary Therapy	10	1%
	Diagnostic	3	<1%
Host intervention aim ¹ (n=866)	Treatment	458	53%
	Unknown	225	26%
	Prevention	174	20%
	Palliative	17	2%
	Other	11	1%
	Diagnostic	2	<1%
Host intervention setting ¹ (n=866)	Unknown	494	57%
	Secondary or Tertiary Care	185	21%
	Other	63	7%
	Online	59	7%

	Primary Care	36	4%
	School	19	2%
	Emergency Care	10	1%
	Dentist	4	<1%
	Intensive Care	4	<1%
Host data collection location	Unknown	510	44%
	Home	373	32%
	Clinic Visit	276	24%
	Online	88	8%
	EHR Data	45	4%
	Other	19	2%
	Inpatient	11	1%
	School	10	1%
Host data collection type	Questionnaire	541	46%
	Unknown	371	32%
	Clinical Measure	214	18%
	Sample	152	13%
	Interview	103	9%
	PROM	83	7%
	Diary	49	4%
	Actiograph	14	1%
	Other	9	1%
Host study participant age	Older Than 18	626	54%
	Older Than 60	341	29%
	Unknown	330	28%
	Younger Than 16	210	18%
	16-18 Yrs	149	13%
Host study participant gender	Mixed	947	81%
	Female Only	191	16%
	Male Only	58	5%
Host study blinding ² (n=826)	Unknown	488	59%
	Open	206	25%
	Patient and Healthcare Provider	98	12%
	Patient Only	56	7%
	Health Care Provider Only	2	<1%

¹ 301 studies were considered to be non-interventional. ²Blinding was considered not applicable in 341 studies due to study design.

Supplementary Table 2: Cohort Characteristics – Retention research
Articles could be coded to as many categories as relevant within each field.

Field	Category	Frequency	% (n=1167)
Retention methods	Case Report	519	44%
	Secondary Analysis Of A Case Report	233	20%
	Systematic Review And Reviews	150	13%
	Nested RCT	131	11%
	Survey	56	5%
	Qualitative Interviews	39	3%
	Pre/ Post-Test	37	3%
	Focus Groups	22	2%
	Workshop Proceedings	16	1%
	Other	16	1%
	Randomised Study	14	1%
	Nested Case Control	12	1%
	Vignettes	1	<1%
Retention Outcomes	Number of participants retained	979	84%
	Representativeness	317	27%
	Questionnaire Response Rate	233	20%
	Reasons for Withdrawal	120	10%
	No Evaluation	87	7%
	Attrition Causes	55	5%
	Other	51	4%
	Retention Cost	44	4%
Timing of retention research within host study	During Study	840	72%
	N/A	175	15%
	During Feasibility	156	13%
	Unknown	7	1%
	Other	1	<1%
Funder	Government	622	53%
	Unknown	351	30%
	Non-Commercial	226	19%
	Commercial	59	5%
	Not directly funded	20	2%

