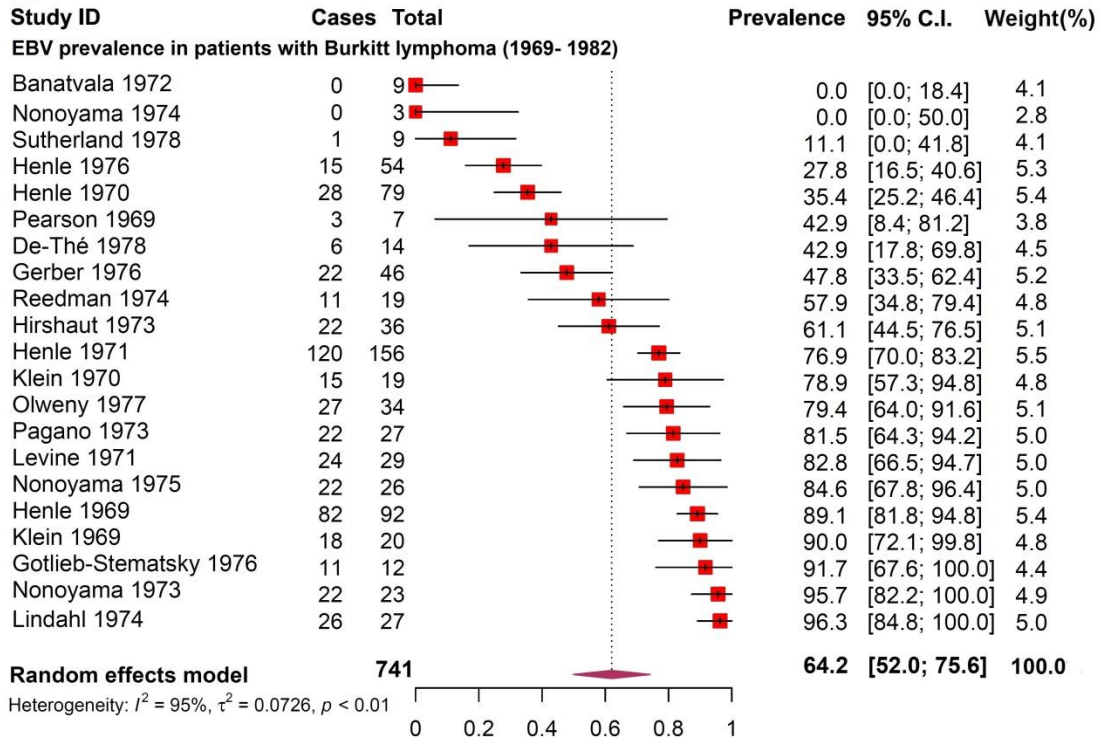
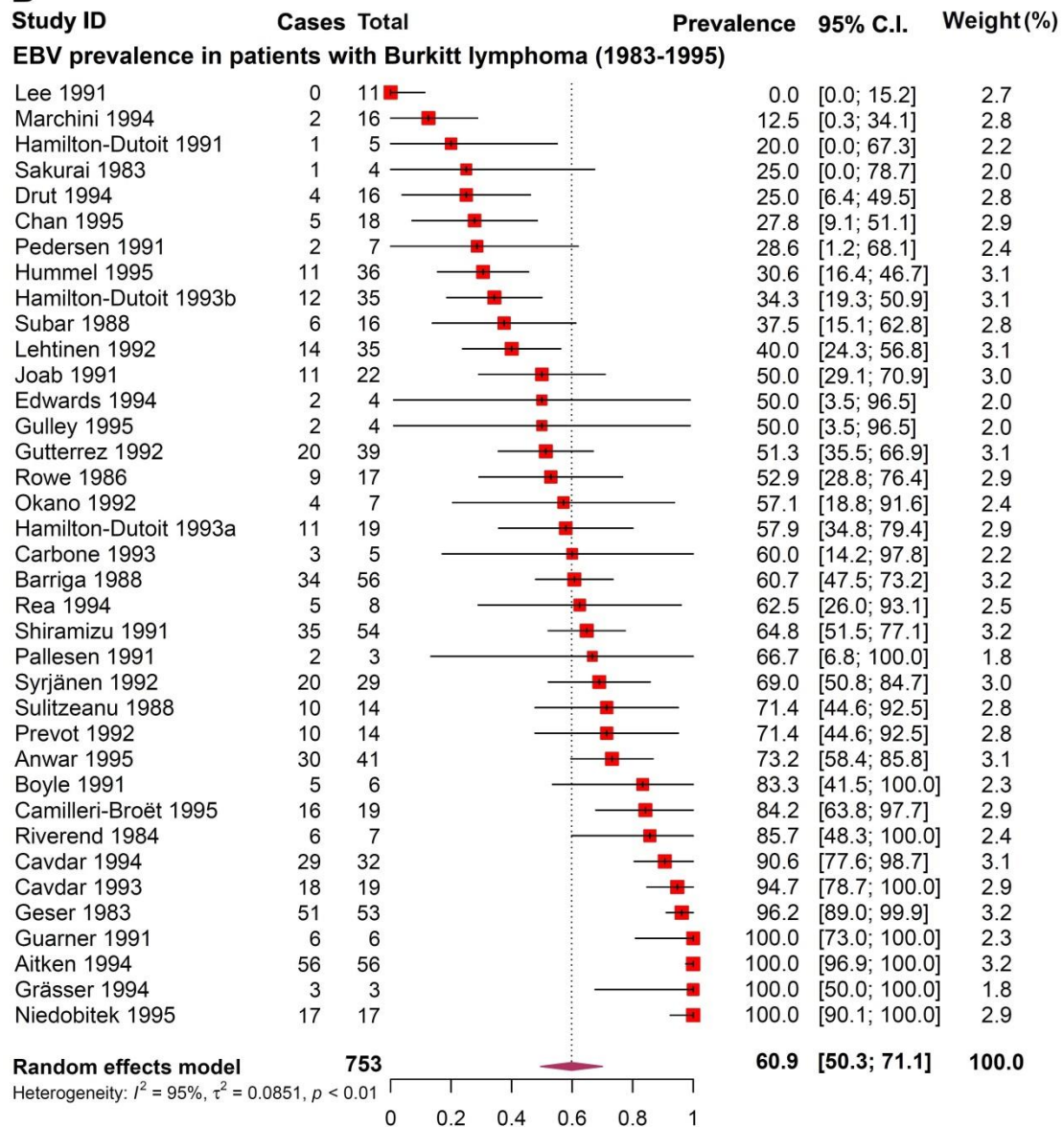


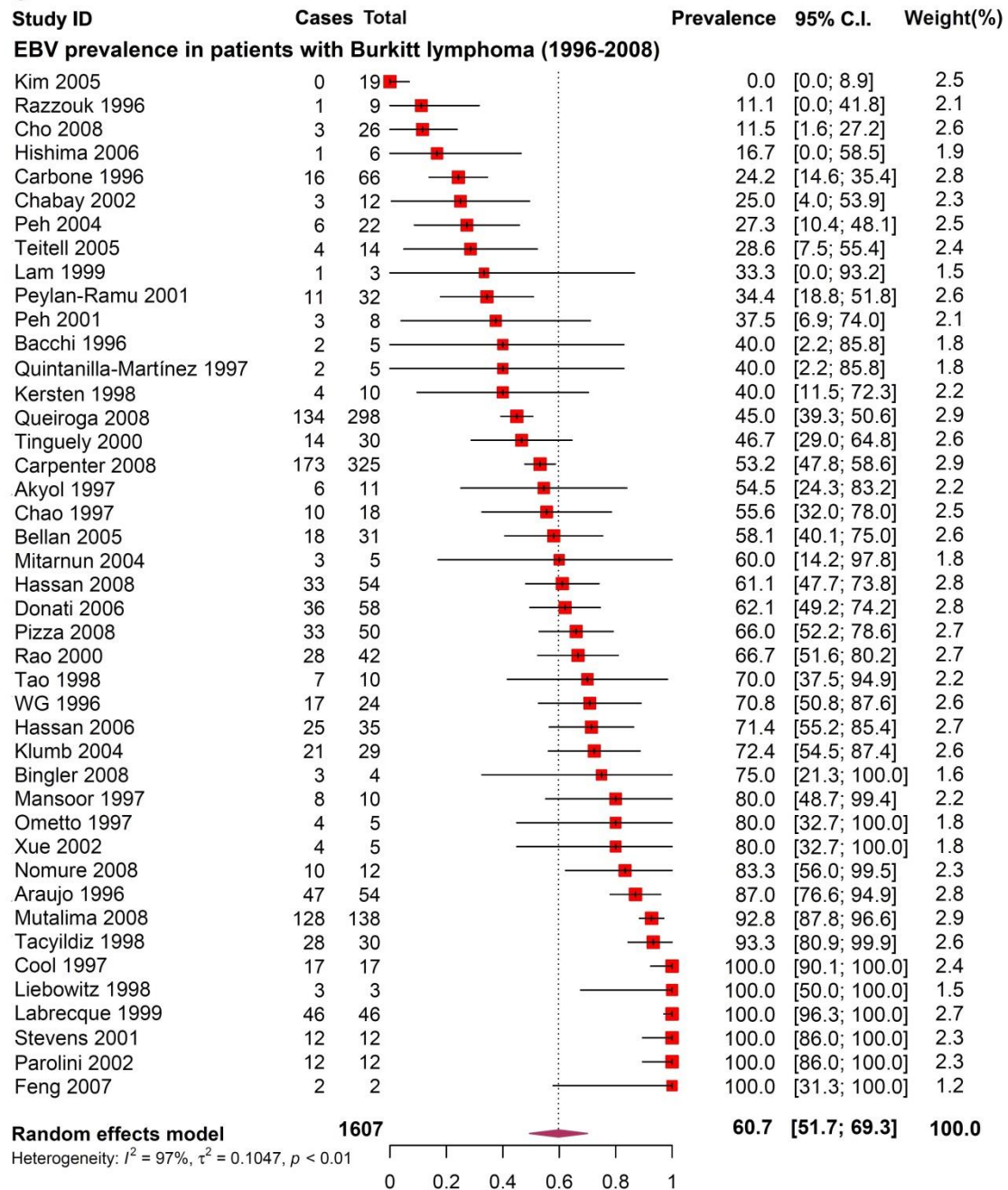
A



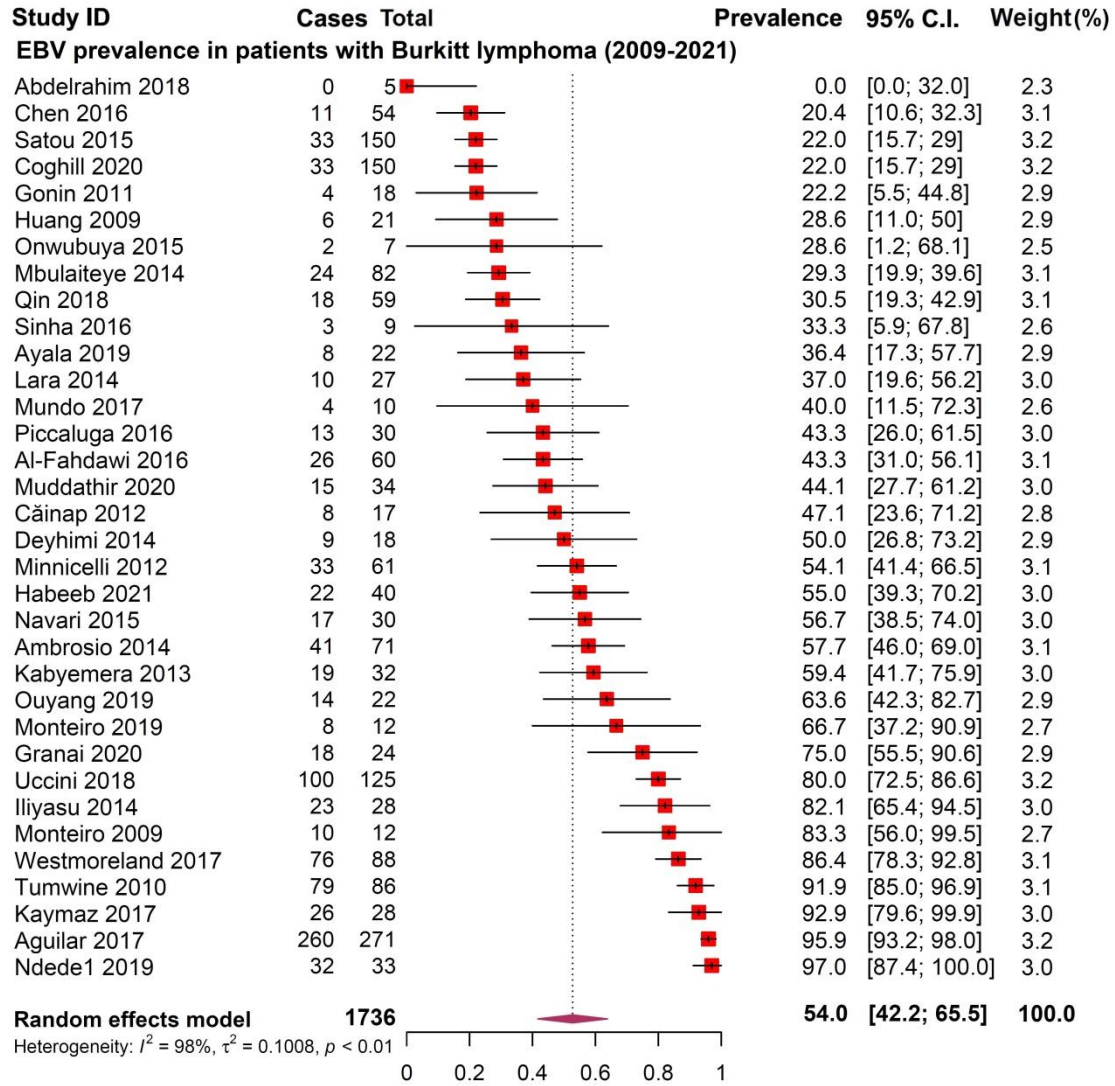
B



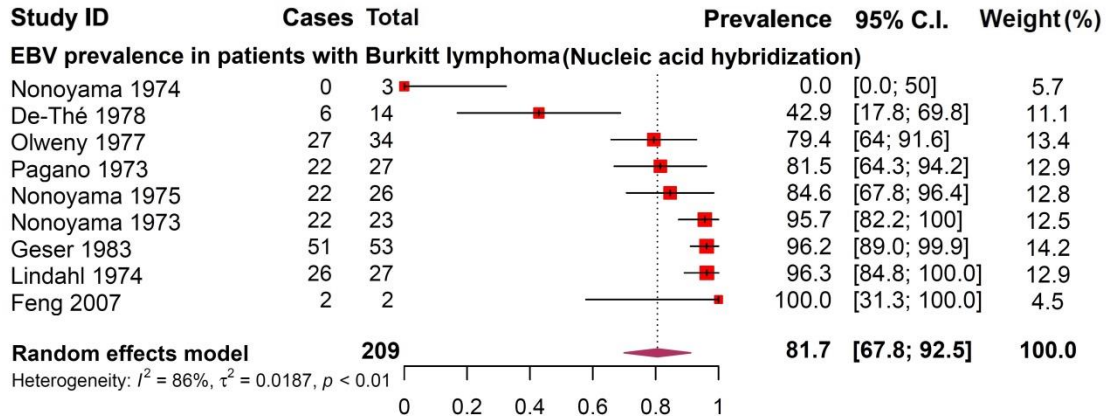
C



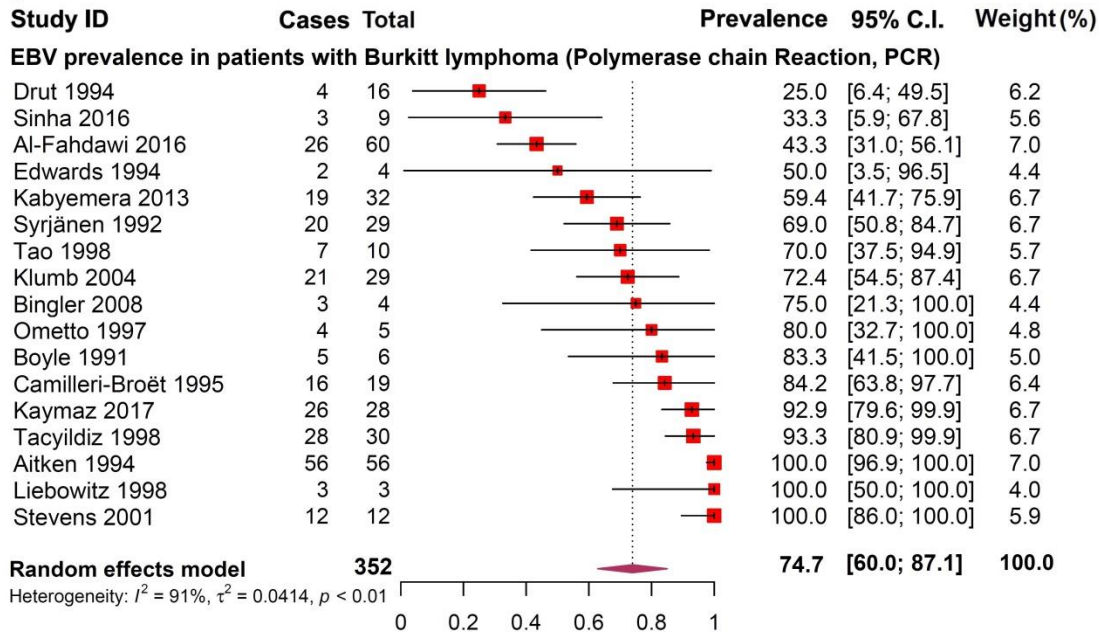
D



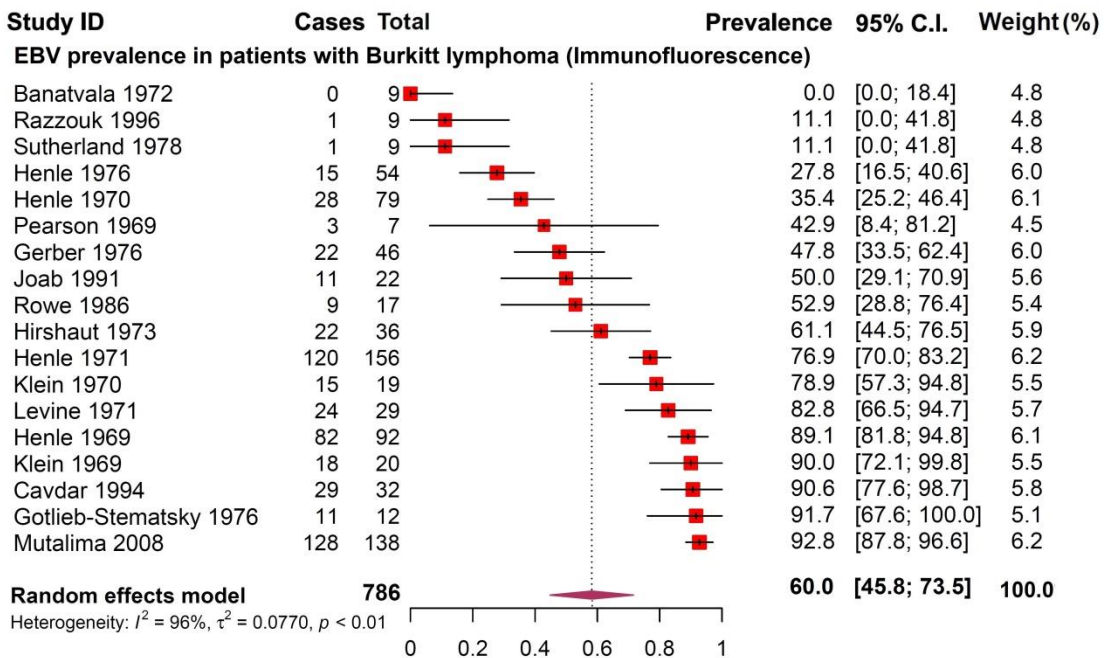
E



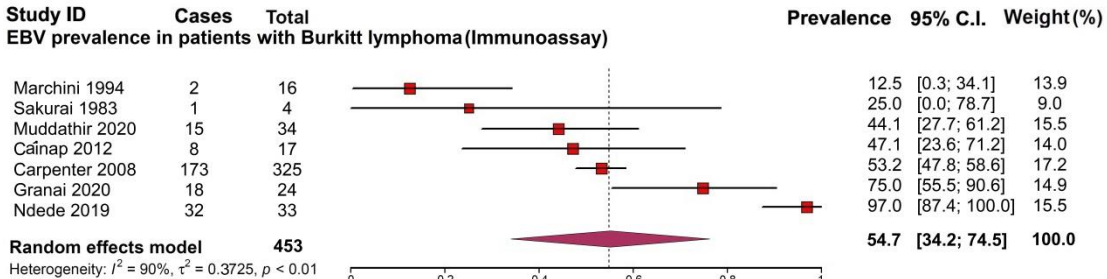
F

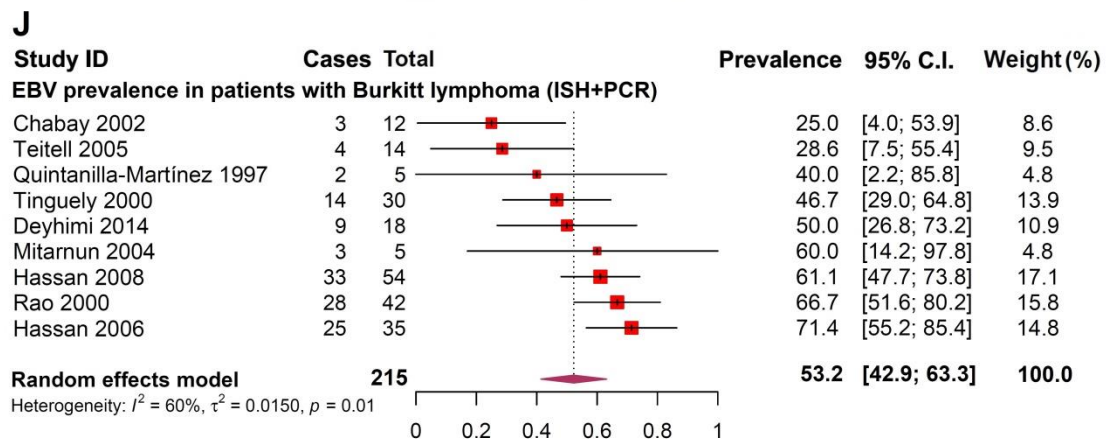
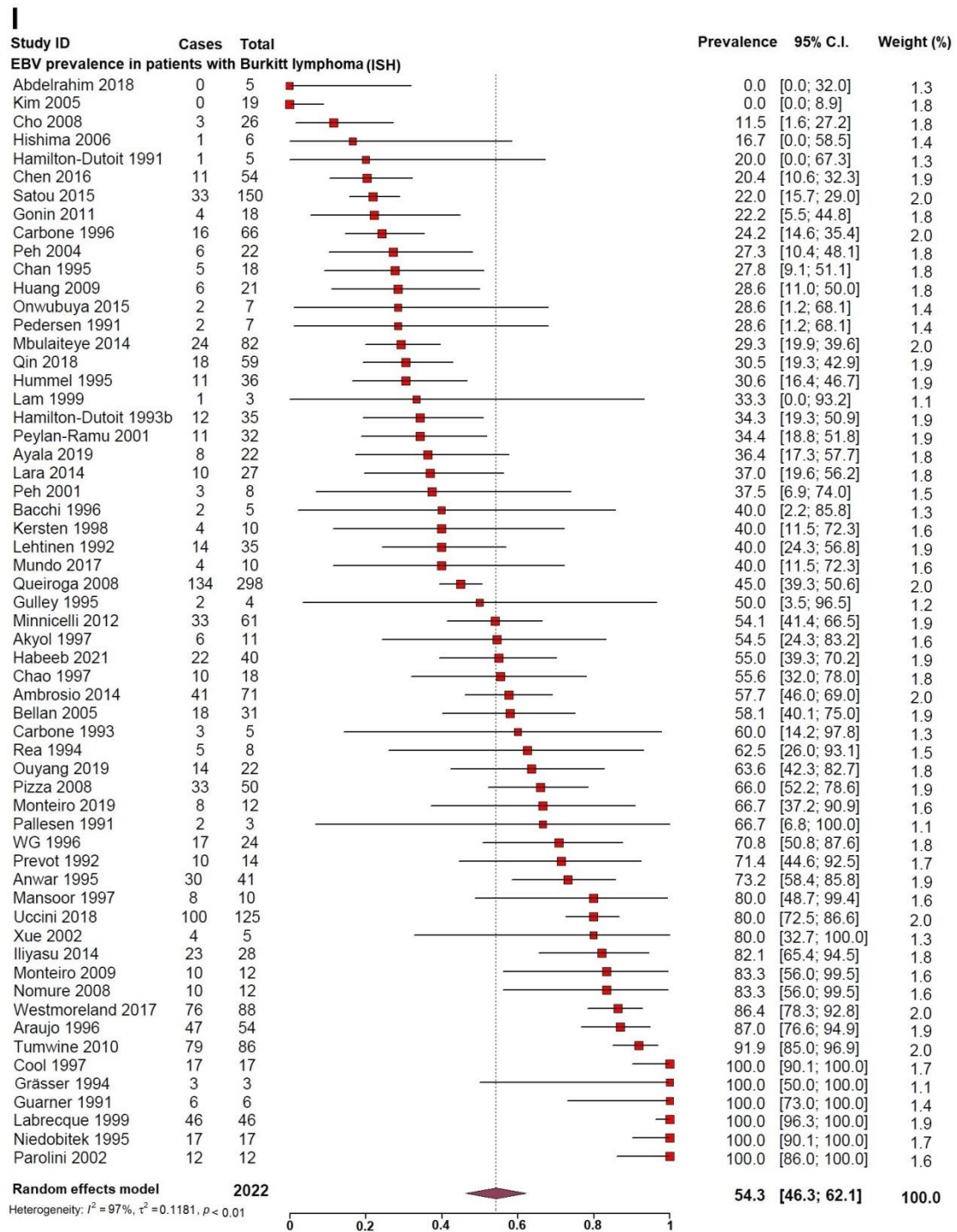


G

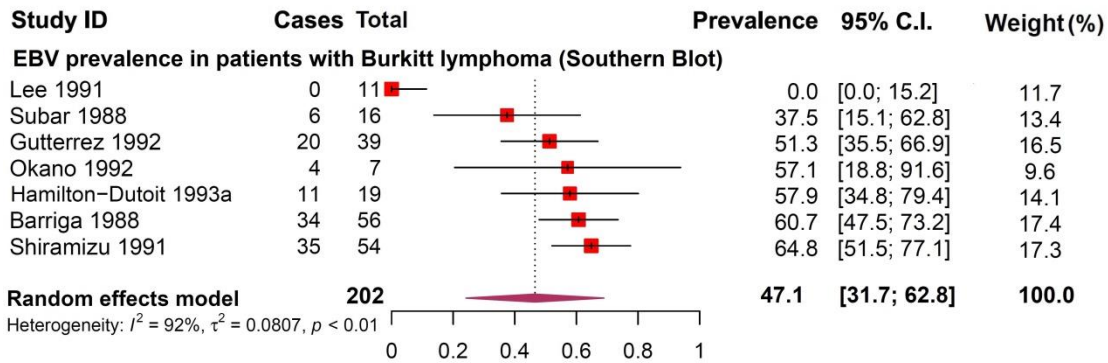


H

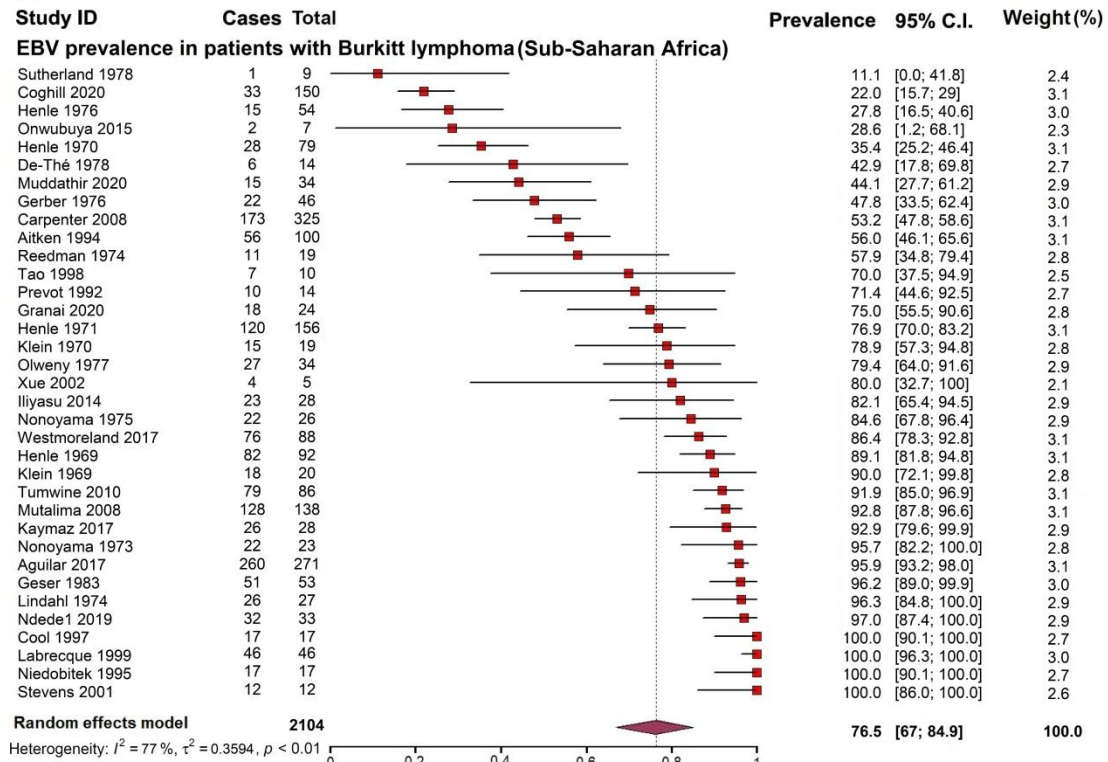




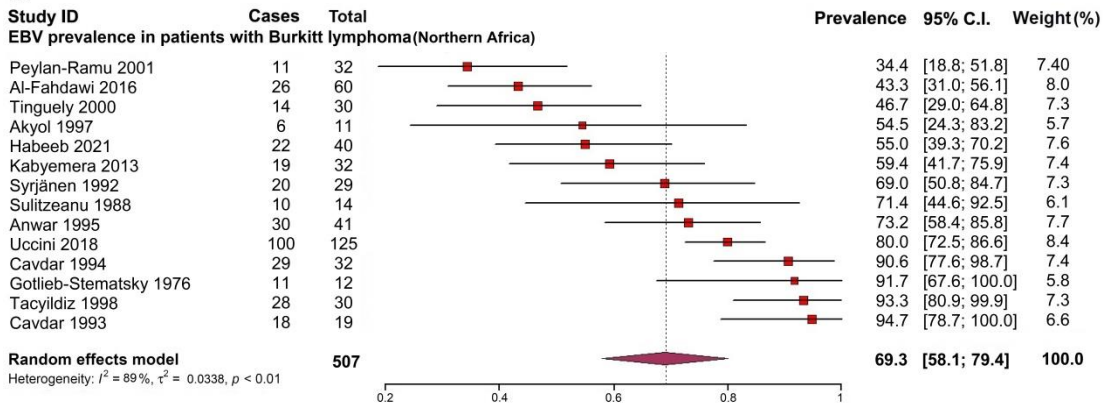
K



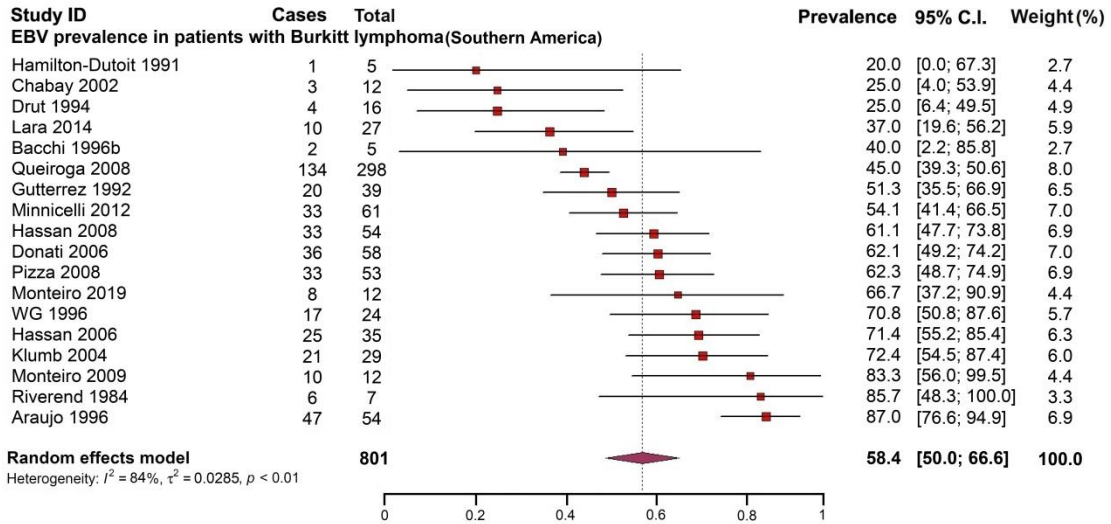
L



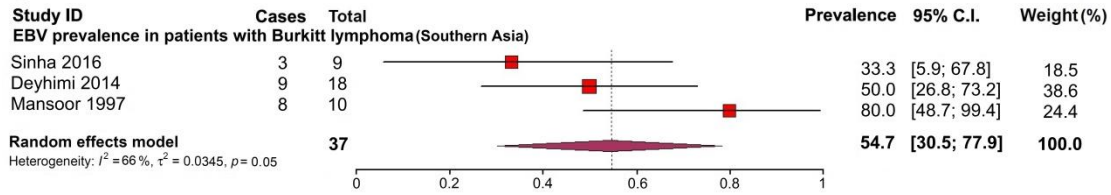
M



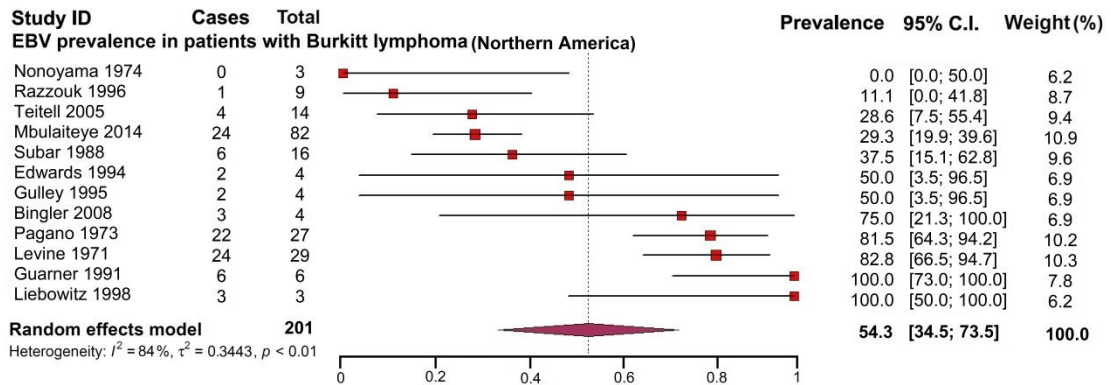
N



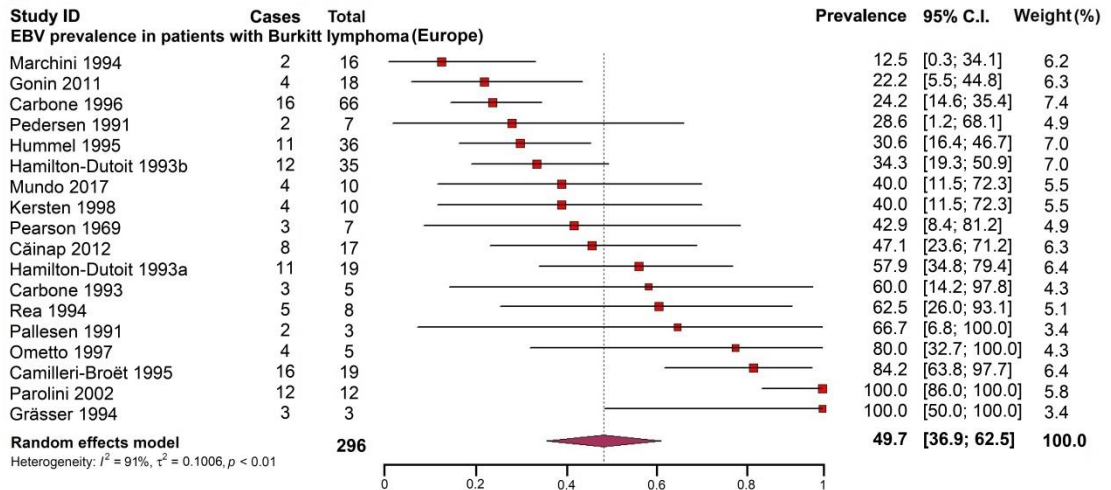
O



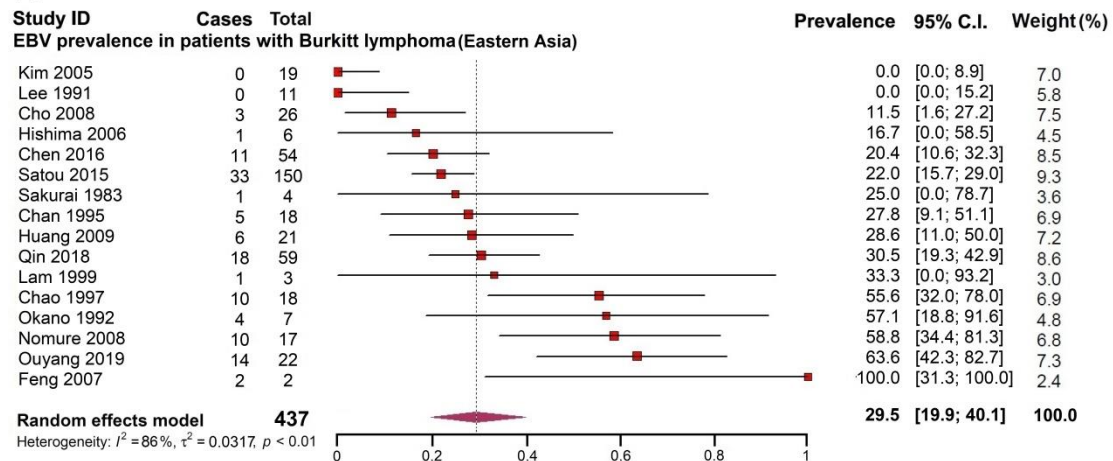
P



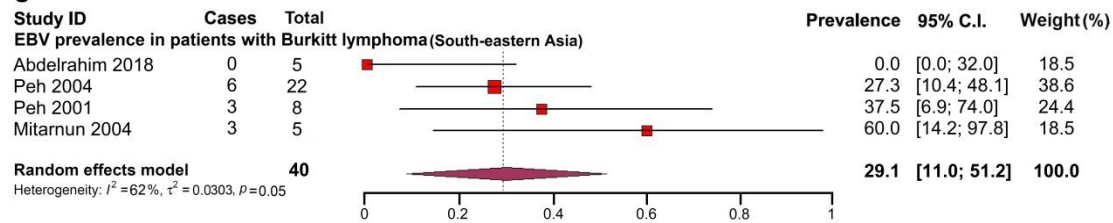
Q



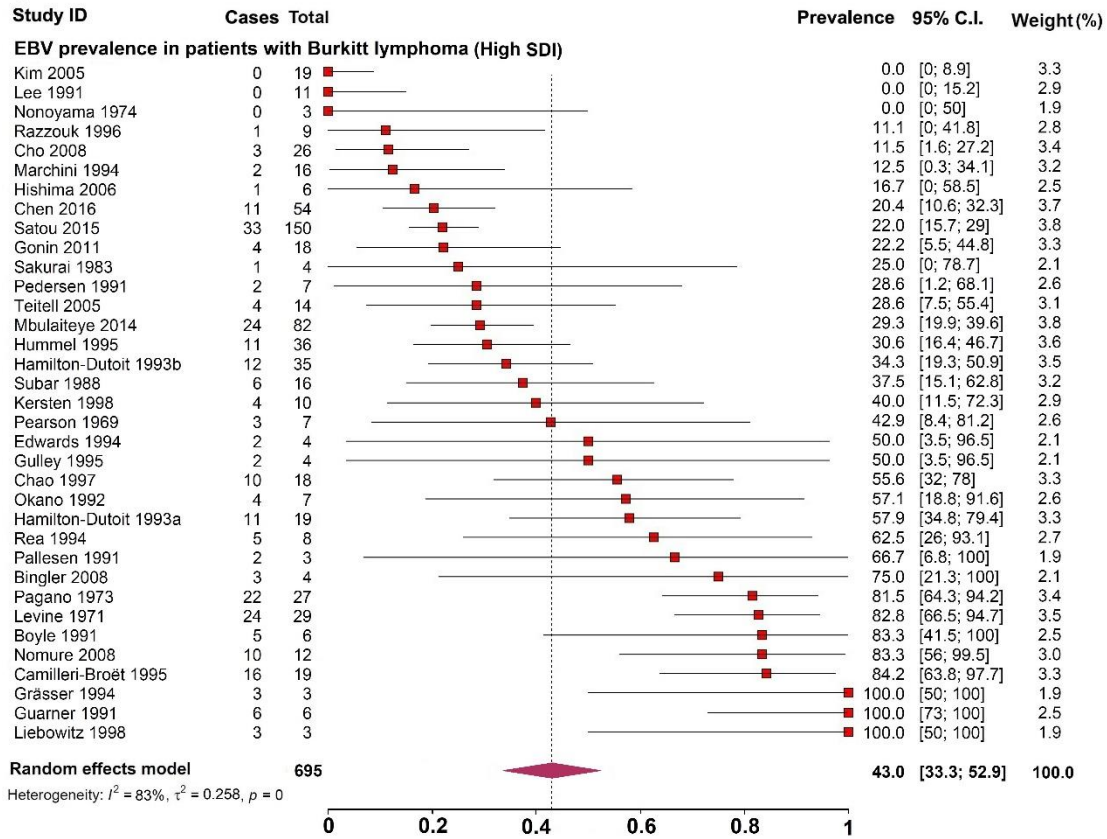
R



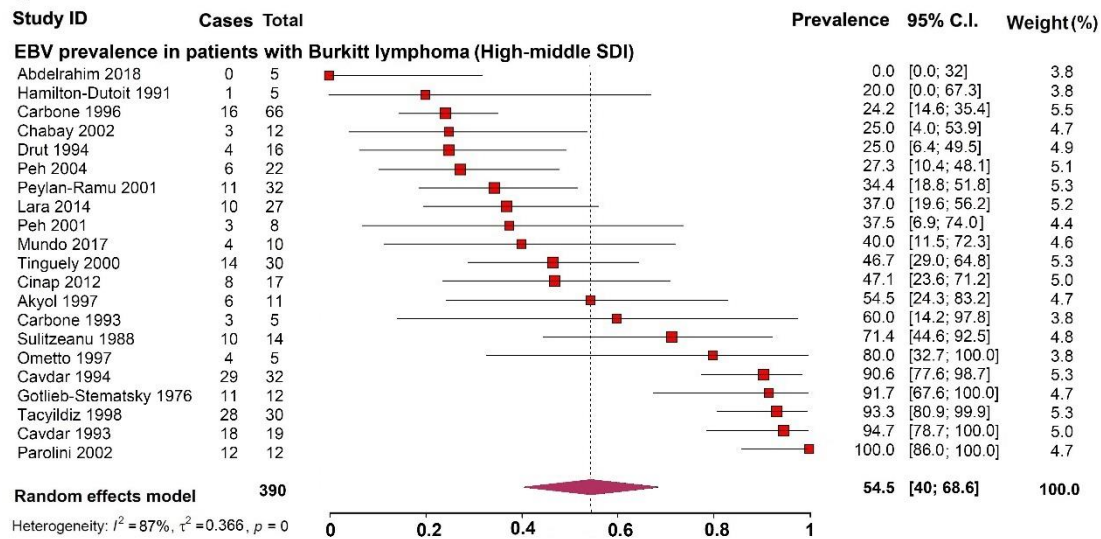
S



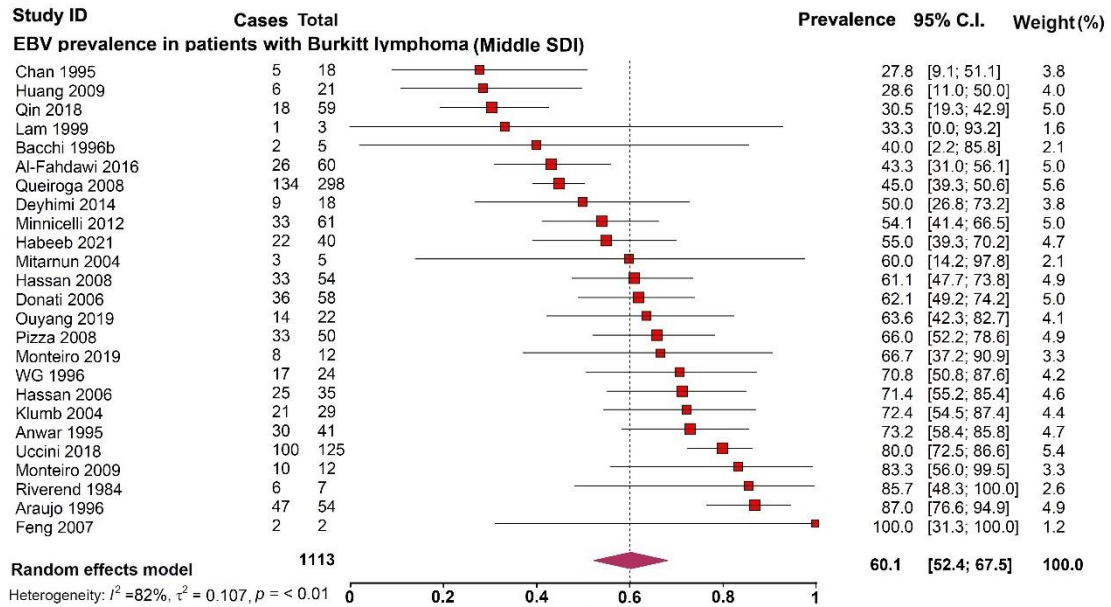
T



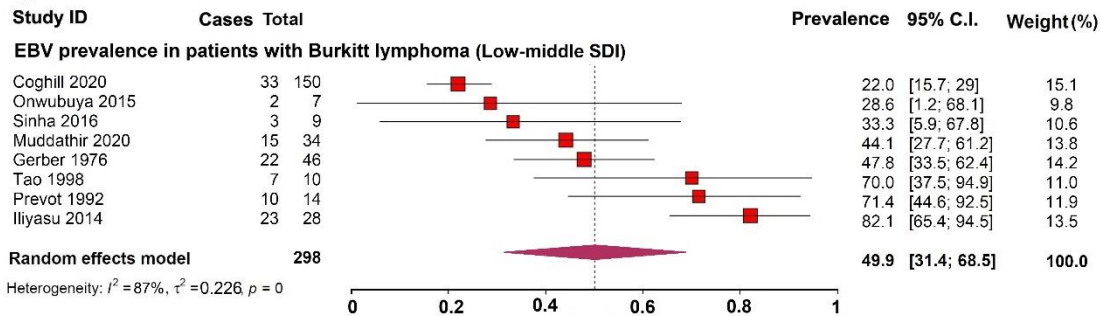
U



V



W



X

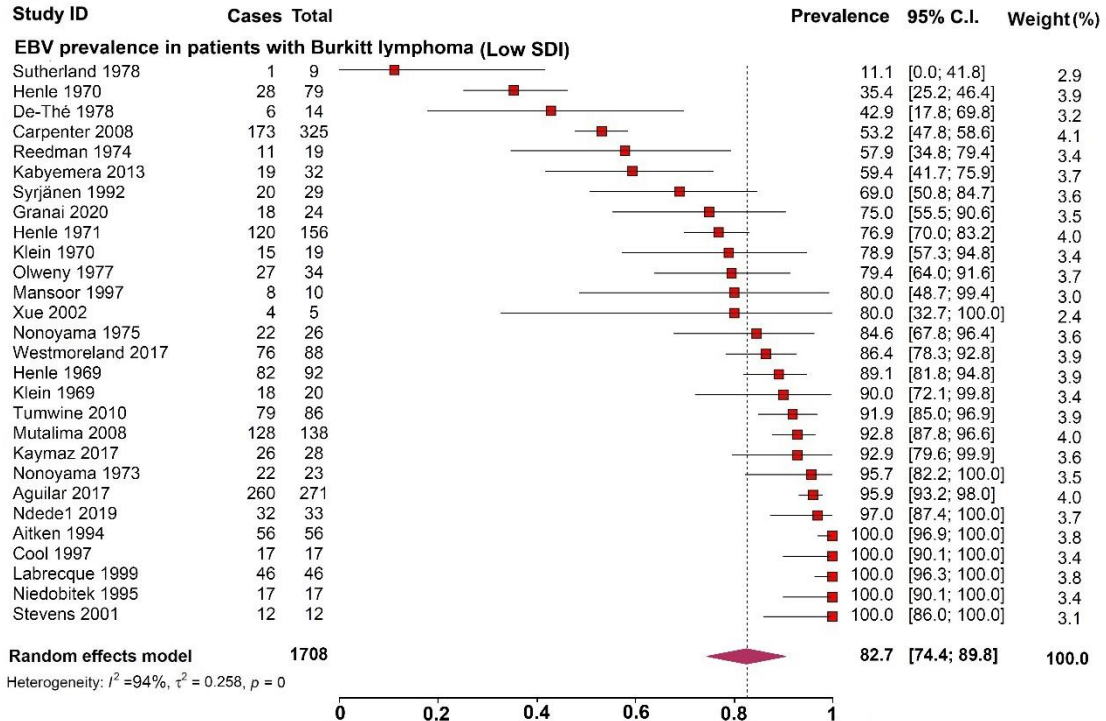
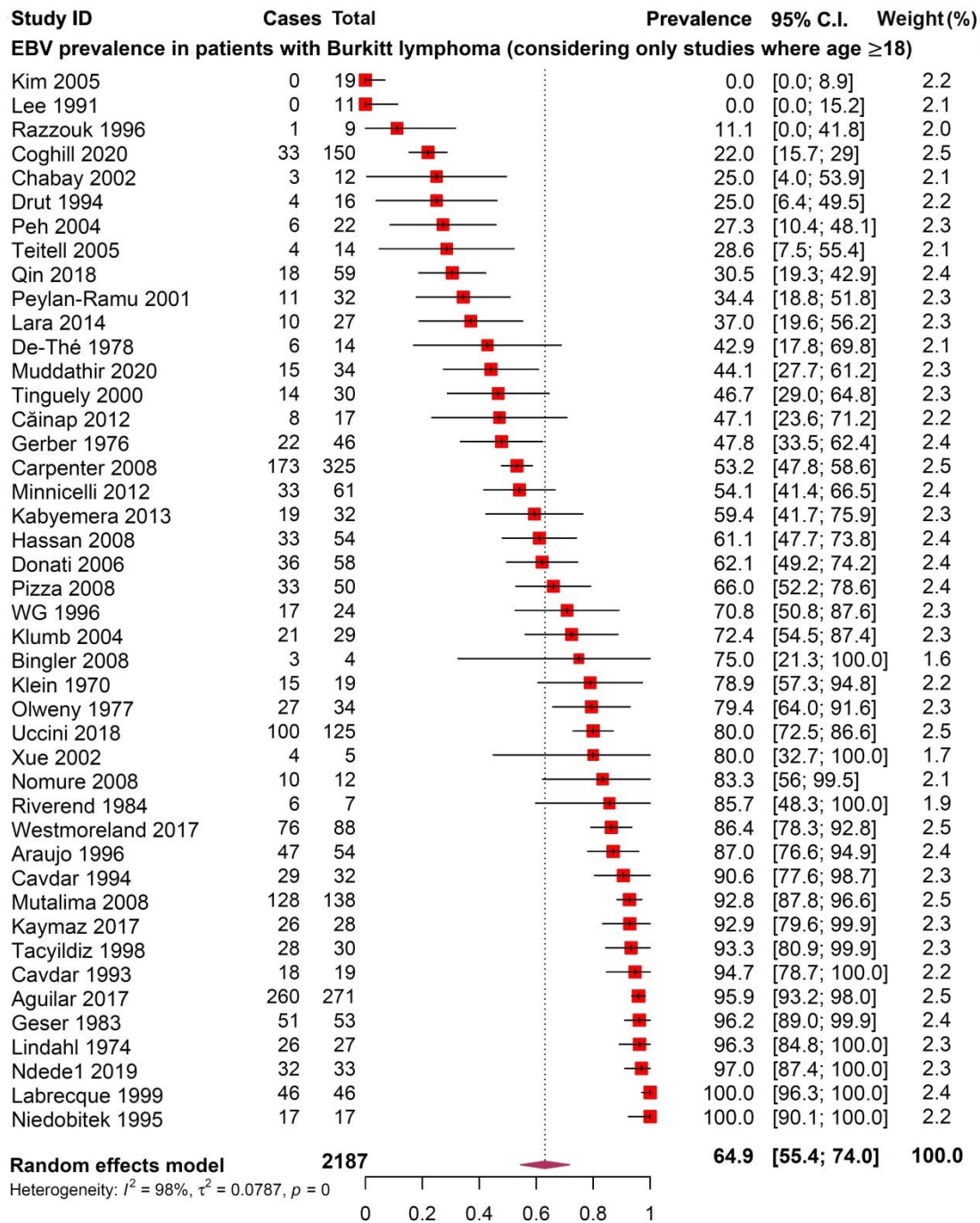
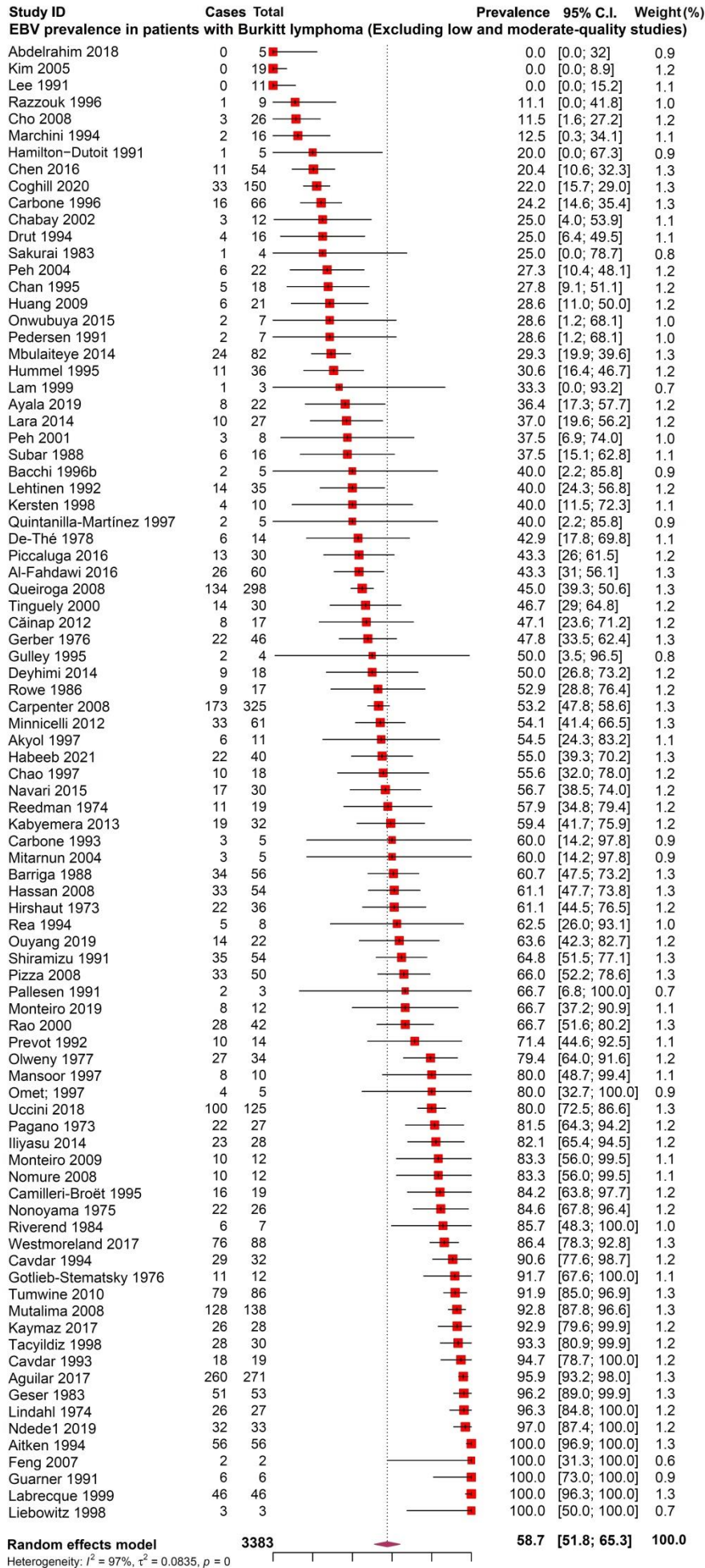


Figure S1. Subgroup analyses estimate the prevalence of the EBV virus at (A-D) different times, using (E-K) different EBV virus detection methods, in (L-S) different regions, and based on the (T-X) Socio-demographic Index (SDI).

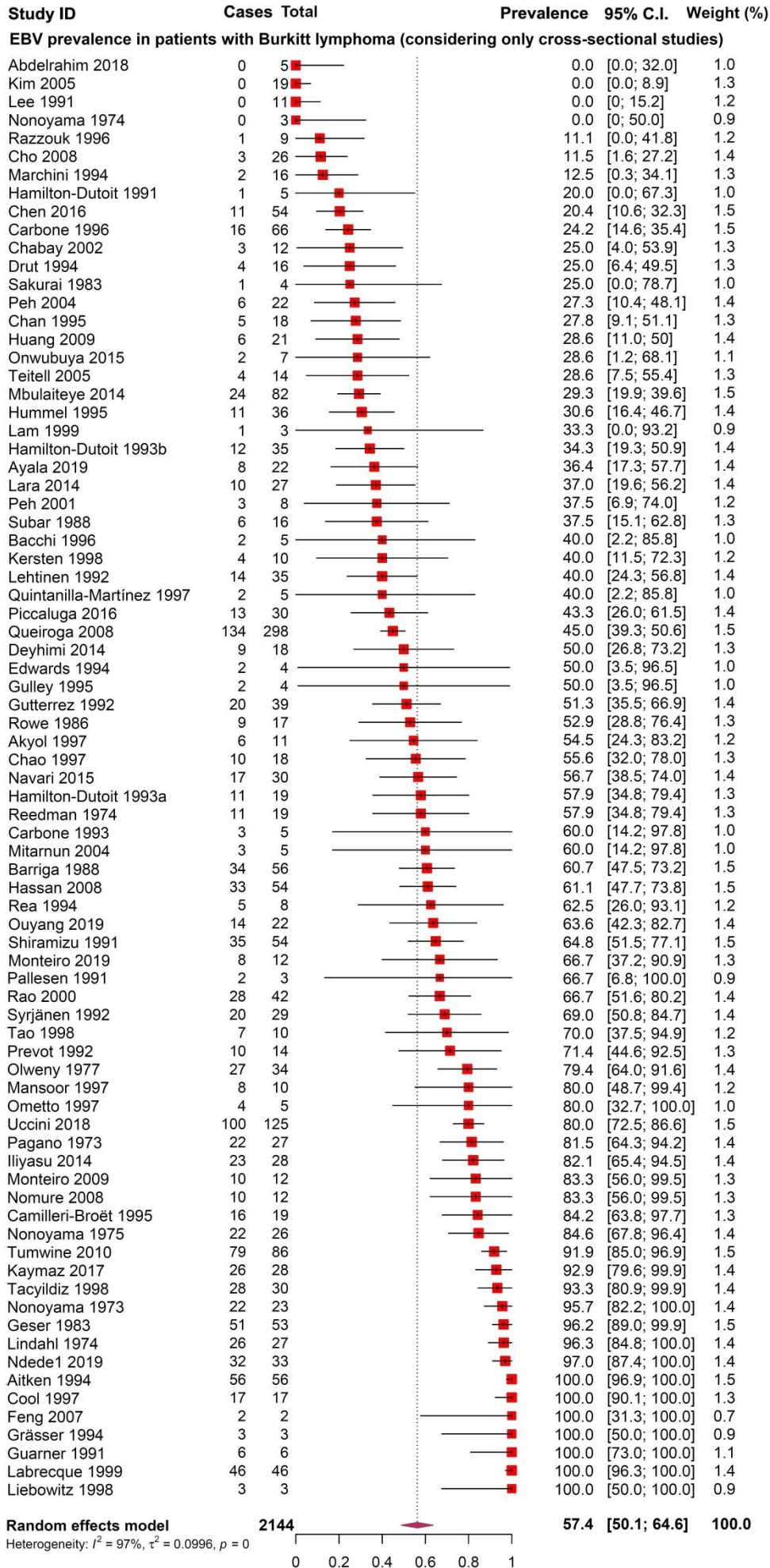
A



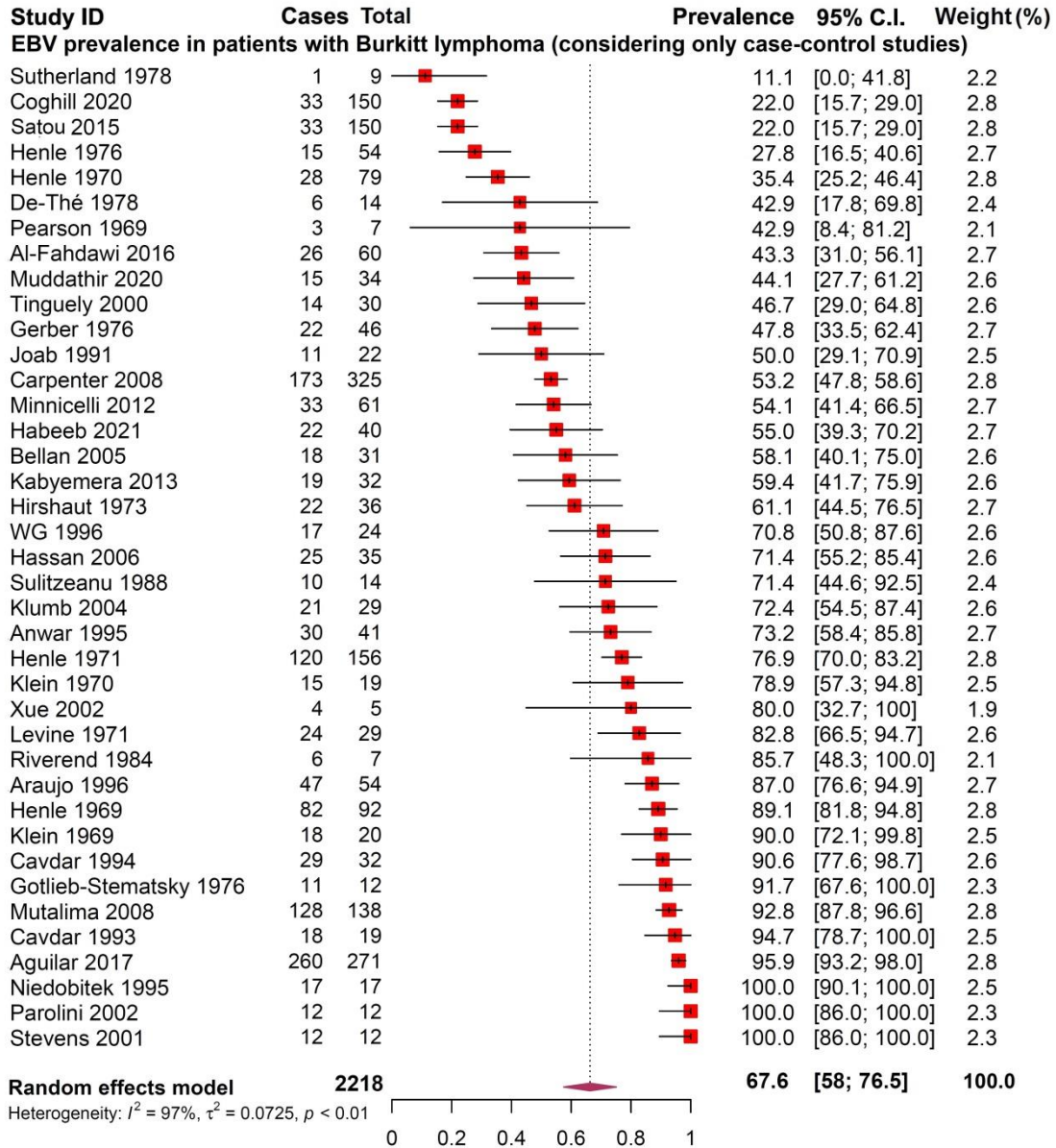
B



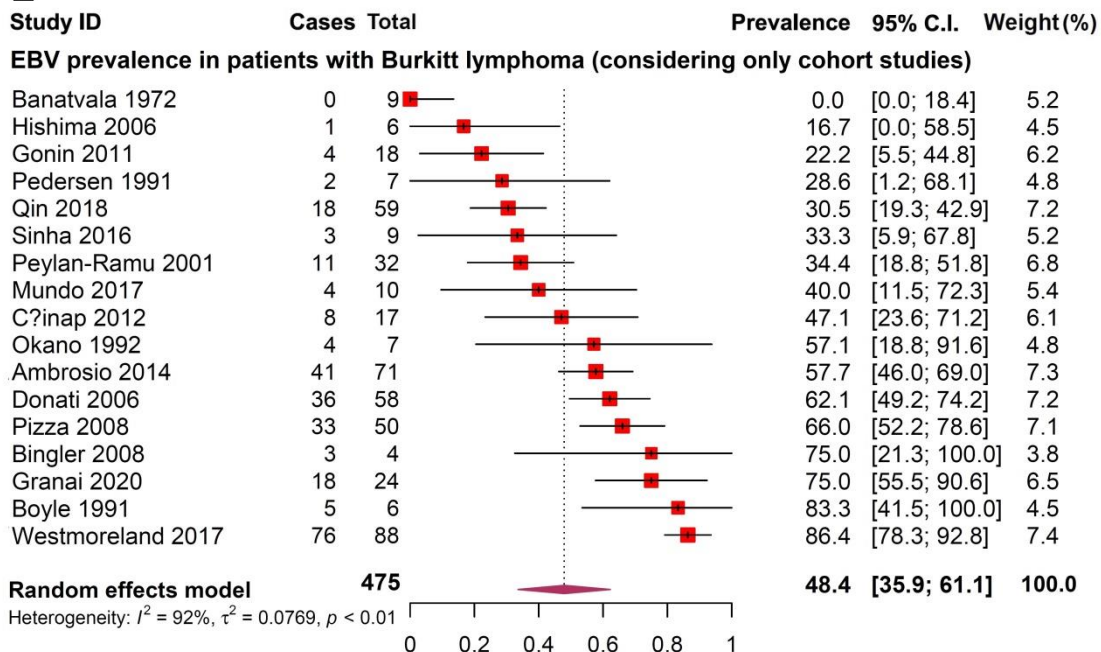
C



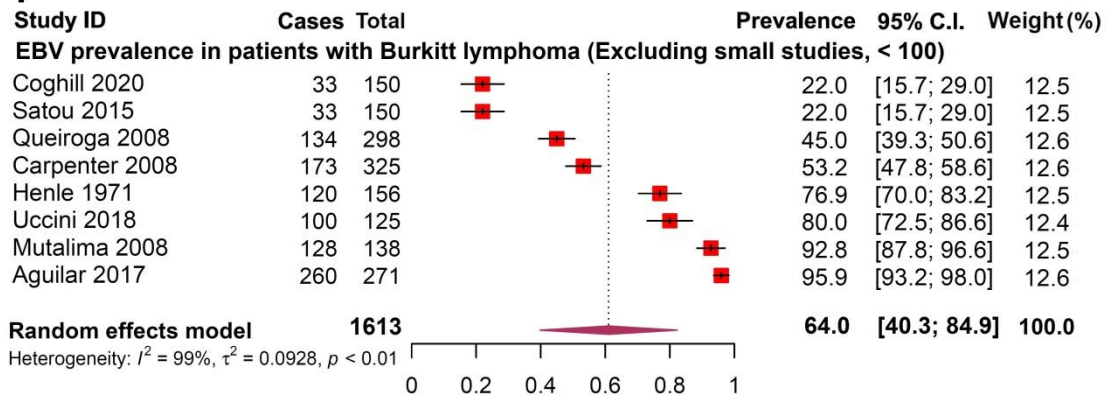
D



E



F



G

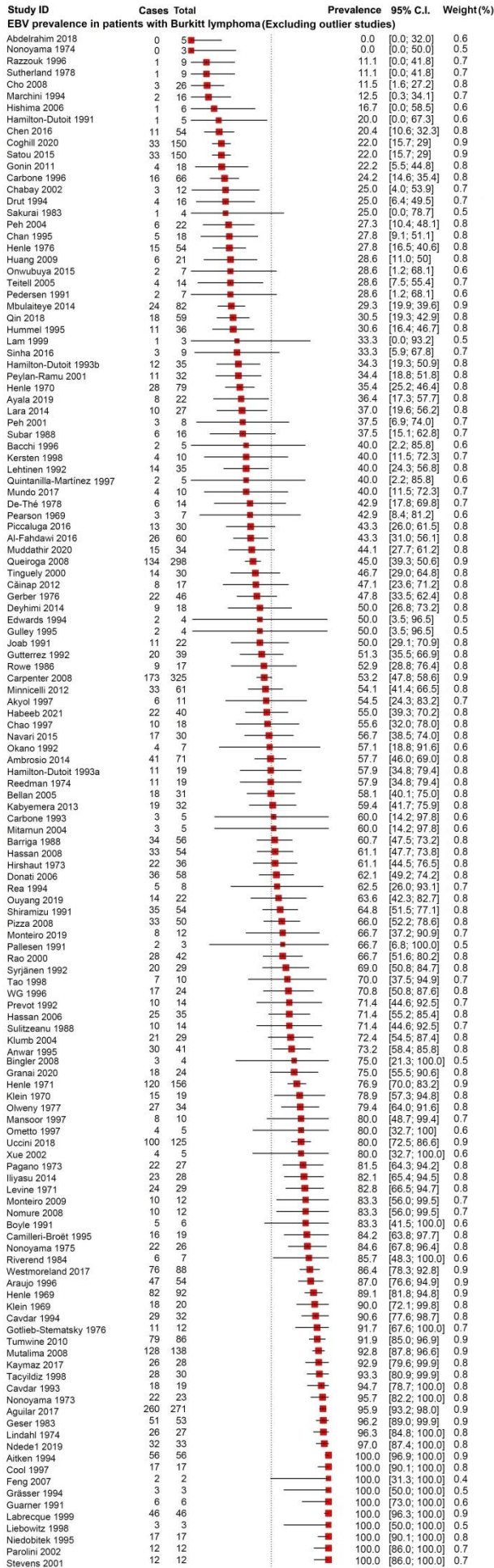


Figure S2. Sensitivity analyses (A) excluding small studies, (B) excluding low- and moderate-quality studies, (C) considering only cross-sectional studies, (D) considering only case-control studies, (E) considering only cohort, (F) considering only studies where the age was less than 18 years old, (G) excluding outlier studies estimating the prevalence of EBV virus in patients with BL.

Table S1. Search strategies

Databases	Search strategies
PubMed	(Burkitt[Title/Abstract] OR Burkitts[Title/Abstract] OR Burkitt's[Title/Abstract] OR "African Lymphoma"[Title/Abstract]) AND (Epstein-Barr[Title] OR EBV[Title] OR "Human Herpesvirus 4"[Title] OR HHV4[Title] OR HHV-4[Title] OR "EB virus"[Title])
Google Scholar	allintitle:(Burkitt OR Burkitts OR Burkitt's OR "African Lymphoma") (Epstein-Barr OR EBV OR "Human Herpesvirus 4" OR HHV4 OR HHV-4 OR "EB virus")
Scopus	TITLE(Burkitt OR Burkitts OR Burkitt's OR "African Lymphoma") AND TITLE(Epstein-Barr OR EBV OR "Human Herpesvirus 4" OR HHV4 OR HHV-4 OR "EB virus")
Web of Science	TI=(Burkitt OR Burkitts OR Burkitt's OR "African Lymphoma") TI=(Epstein-Barr OR EBV OR "Human Herpesvirus 4" OR HHV4 OR HHV-4 OR "EB virus")

Table S2. Quality assessment of the included cross-sectional studies

No.	Study ID	Questions assessing included cross-sectional studies								Yes (%)
		1	2	3	4	5	6	7	8	
1	Abdelrahim 2018	Y	Y	Y	Y	U	U	Y	Y	75
2	Aitken 1994	Y	Y	Y	Y	U	U	Y	Y	75
3	Akyol 1997	Y	Y	Y	Y	U	U	Y	Y	75
4	Ayala 2019	Y	Y	Y	Y	U	U	Y	Y	75
5	Bacchi 1996b	Y	Y	Y	Y	Y	Y	Y	Y	100
6	Barriga 1988	Y	Y	Y	Y	U	U	Y	Y	75
7	Camilleri-Bro ã 1995	Y	Y	Y	Y	Y	Y	Y	Y	100
8	Carbone 1993	Y	Y	Y	Y	Y	N	Y	Y	87.5
9	Carbone 1996	Y	U	Y	Y	Y	Y	Y	Y	87.5
10	Chabay 2002	Y	Y	Y	Y	Y	Y	Y	Y	100
11	Chan 1995	Y	Y	Y	Y	U	U	Y	Y	75
12	Chao 1997	Y	Y	Y	Y	U	U	Y	Y	75
13	Chen 2016	Y	Y	Y	Y	U	U	Y	Y	75
14	Cho 2008	Y	Y	Y	Y	U	U	Y	Y	75
15	Cool 1997	U	Y	Y	Y	U	U	Y	Y	62.5
16	Deyhimi 2014	Y	Y	Y	Y	U	U	Y	Y	75
17	Drut 1994	Y	Y	Y	Y	U	U	Y	Y	75
18	Edwards 1994	U	N	Y	Y	U	U	Y	Y	50
19	Feng 2007	Y	Y	Y	Y	U	U	Y	Y	75
20	Geser 1983	Y	Y	Y	Y	Y	Y	Y	Y	100
21	Gr ässer 1994	U	Y	Y	Y	U	U	Y	Y	62.5
22	Guarner 1991	Y	Y	Y	Y	U	U	Y	Y	75
23	Gulley 1995	Y	Y	Y	Y	Y	Y	Y	Y	100
24	Gutierrez 1992	U	Y	Y	Y	U	U	Y	Y	62.5
25	Hamilton-Dutoit 1991	Y	Y	Y	Y	U	U	Y	Y	75
26	Hamilton-Dutoit 1993a	Y	N	Y	Y	U	U	Y	Y	62.5
27	Hamilton-Dutoit 1993b	Y	N	Y	Y	U	U	Y	Y	62.5
28	Hassan 2008	Y	Y	Y	Y	Y	Y	Y	Y	100
29	Huang 2009	Y	Y	Y	Y	U	U	Y	Y	75
30	Hummel 1995	Y	Y	Y	Y	U	U	Y	Y	75
31	Iliyasu 2014	Y	Y	Y	Y	U	U	Y	Y	75
32	Kaymaz 2017	Y	Y	Y	Y	U	U	Y	Y	75
33	Kersten 1998	Y	Y	Y	Y	U	U	Y	Y	75
34	Kim 2005	Y	Y	Y	Y	U	U	Y	Y	75
35	Labrecque 1999	Y	Y	Y	Y	U	U	Y	Y	75
36	Lam 1999	Y	Y	Y	Y	U	U	Y	Y	75
37	Lara 2014	Y	Y	Y	Y	Y	Y	Y	Y	100
38	Lee 1991	Y	Y	Y	Y	U	U	Y	Y	75
39	Lehtinen 1992	Y	Y	Y	Y	Y	Y	Y	Y	100

1. Were the criteria for inclusion in the sample clearly defined? 2. Were the study subjects and the setting described in detail? 3. Was the exposure measured in a valid and reliable way? 4. Were objective, standard criteria used for measurement of the condition? 5. Were confounding factors identified? 6. Were strategies to deal with confounding factors stated? 7. Were the outcomes measured in a valid and reliable way? 8. Was appropriate statistical analysis used? Y=Yes; N=No; U=Unclear.

Table S2. Quality assessment of the included cross-sectional studies, continued

No.	Study ID	Questions assessing included cross-sectional studies								Yes (%)
		1	2	3	4	5	6	7	8	
40	Liebowitz 1998	Y	Y	Y	Y	Y	N	Y	N	75
41	Lindahl 1974	Y	Y	Y	Y	U	U	Y	Y	75
42	Mansoor 1997	Y	Y	Y	Y	U	U	Y	Y	75
43	Marchini 1994	Y	Y	Y	Y	U	U	Y	Y	75
44	Mbulaiteye 2014	Y	Y	Y	Y	Y	Y	Y	Y	100
45	Mitarnun 2004	Y	Y	Y	Y	U	U	Y	Y	75
46	Monteiro 2009	Y	Y	Y	Y	U	U	Y	Y	75
47	Monteiro 2019	Y	Y	Y	Y	U	U	Y	Y	75
48	Navari 2015	Y	Y	Y	Y	U	U	Y	Y	75
49	Ndede1 2019	Y	Y	Y	Y	U	U	Y	Y	75
50	Nomure 2008	Y	Y	Y	Y	U	U	Y	Y	75
51	Nonoyama 1973	N	Y	Y	Y	U	U	Y	Y	62.5
52	Nonoyama 1974	N	Y	N	Y	U	U	N	Y	50
53	Nonoyama 1975	Y	Y	Y	Y	Y	Y	Y	Y	100
54	Olweny 1977	Y	Y	Y	Y	U	U	Y	Y	75
55	Ometto 1997	Y	Y	Y	Y	U	U	Y	Y	75
56	Onwubuya 2015	Y	Y	Y	Y	U	U	Y	Y	75
57	Ouyang 2019	Y	Y	Y	Y	U	U	Y	Y	75
58	Pagano 1973	Y	Y	Y	Y	U	U	Y	Y	75
59	Pallesen 1991	Y	Y	Y	Y	Y	N	Y	Y	87.5
60	Peh 2001	Y	Y	Y	Y	U	U	Y	Y	75
61	Peh 2004	Y	Y	Y	Y	U	U	Y	Y	75
62	Piccaluga 2016	Y	Y	Y	Y	Y	Y	Y	Y	100
63	Prevot 1992	Y	Y	Y	Y	U	U	Y	Y	75
64	Queiroga 2008	Y	Y	Y	Y	U	U	Y	Y	75
65	Quintanilla-Mart ínez 1997	Y	Y	Y	Y	U	U	Y	Y	75
66	Rao 2000	Y	Y	Y	Y	U	U	Y	Y	75
67	Razzouk 1996	Y	Y	Y	Y	U	U	Y	Y	75
68	Rea 1994	Y	N	Y	Y	U	U	N	Y	75
69	Reedman 1974	Y	Y	Y	Y	U	U	Y	Y	75
70	Rowe 1986	Y	Y	N	Y	U	U	Y	Y	75
71	Sakurai 1983	Y	Y	Y	Y	U	U	Y	Y	75
72	Shiramizu 1991	Y	Y	Y	Y	Y	Y	Y	Y	100
73	Subar 1988	Y	Y	Y	Y	U	U	Y	Y	75
74	Syrj änen 1992	U	Y	Y	Y	U	U	Y	Y	62.5
75	Tacyildiz 1998	Y	Y	Y	Y	U	U	N	Y	75
76	Tao 1998	Y	Y	Y	N	U	U	Y	Y	62.5
77	Teitell 2005	Y	N	Y	Y	U	U	Y	Y	62.5
78	Tumwine 2010	Y	Y	Y	Y	U	U	Y	Y	75
79	Uccini 2018	Y	Y	Y	Y	U	U	Y	Y	75

1. Were the criteria for inclusion in the sample clearly defined? 2. Were the study subjects and the setting described in detail? 3. Was the exposure measured in a valid and reliable way? 4. Were objective, standard criteria used for measurement of the condition? 5. Were confounding factors identified? 6. Were strategies to deal with confounding factors stated? 7. Were the outcomes measured in a valid and reliable way? 8. Was appropriate statistical analysis used? Y=Yes; N=No; U=Unclear.

Table S3. Quality assessment of the included case-control studies

No.	Study ID	Questions assessing included case-control studies										Yes (%)
		1	2	3	4	5	6	7	8	9	10	
1	Aguilar 2017	Y	Y	Y	Y	Y	Y	Y	Y	U	Y	90
2	Al-Fahdawi 2016	Y	Y	Y	Y	Y	N	N	Y	U	Y	70
3	Anwar 1995	N	Y	Y	Y	Y	N	N	Y	U	Y	60
4	Araujo 1996	N	Y	Y	Y	Y	U	U	Y	U	Y	60
5	Bellan 2005	N	Y	Y	Y	Y	U	U	Y	U	Y	60
6	Carpenter 2008	Y	Y	Y	Y	Y	U	U	Y	U	Y	70
7	Cavdar 1993	Y	Y	Y	Y	Y	U	U	Y	U	Y	70
8	Cavdar 1994	Y	Y	Y	Y	Y	U	U	Y	U	Y	70
9	Coghill 2020	Y	Y	Y	Y	Y	U	U	Y	U	Y	70
10	De-Th é 1978	Y	Y	Y	Y	Y	U	U	Y	Y	Y	80
11	Gerber 1976	Y	Y	Y	Y	Y	U	U	Y	U	Y	70
12	Gotlieb-Stematsky 1976	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100
13	Habeeb 2021	Y	Y	Y	Y	Y	U	U	Y	U	Y	70
14	Hassan 2006	Y	U	Y	Y	Y	U	U	Y	U	Y	60
15	Henle 1969	N	N	Y	Y	Y	U	U	Y	U	Y	50
16	Henle 1970	U	U	Y	Y	Y	U	U	Y	U	N	40
17	Henle 1971	U	N	Y	Y	Y	U	U	Y	U	Y	50
18	Henle 1976	N	N	Y	Y	Y	U	U	Y	U	Y	50
19	Hirshaut 1973	Y	Y	Y	Y	Y	U	U	Y	U	Y	70
20	Joab 1991	U	U	U	Y	Y	U	U	Y	U	Y	40
21	Kabyemera 2013	Y	Y	Y	Y	Y	U	U	Y	U	Y	70
22	Klein 1969	U	N	Y	Y	Y	U	U	Y	U	Y	50
23	Klein 1970	U	N	Y	Y	Y	U	U	Y	U	Y	50
24	Klumb 2004	U	U	Y	Y	Y	U	U	Y	U	Y	50
25	Levine 1971	N	N	Y	Y	Y	U	U	Y	U	Y	50

1. Were the groups comparable other than the presence of disease in cases or the absence of disease in controls? **2.** Were cases and controls matched appropriately? **3.** Were the same criteria used for identification of cases and controls? **4.** Was exposure measured in a standard, valid and reliable way? **5.** Was exposure measured in the same way for cases and controls? **6.** Were confounding factors identified? **7.** Were strategies to deal with confounding factors stated? **8.** Were outcomes assessed in a standard, valid and reliable way for cases and controls? **9.** Was the exposure period of interest long enough to be meaningful? **10.** Was appropriate statistical analysis used? Y=Yes; N=No; U=Unclear.

Table S3. Quality assessment of the included case-control studies, continued

No.	Study ID	Questions assessing included case-control studies										Yes (%)
		1	2	3	4	5	6	7	8	9	10	
26	Minnicelli 2012	Y	Y	Y	Y	Y	Y	Y	Y	U	Y	90
27	Muddathir 2020	U	U	Y	Y	Y	U	U	Y	U	Y	50
28	Mutalima 2008	Y	U	Y	Y	Y	Y	Y	Y	U	Y	80
29	Niedobitek 1995	U	N	Y	Y	Y	U	U	Y	U	Y	50
30	Parolini 2002	U	U	Y	Y	Y	U	U	Y	U	Y	50
31	Pearson 1969	U	U	Y	Y	Y	U	U	Y	U	Y	50
32	Riverend 1984	Y	Y	Y	Y	Y	U	U	Y	U	Y	70
33	Satou 2015	Y	U	Y	Y	Y	U	U	Y	U	Y	60
34	Stevens 2001	N	U	Y	Y	Y	U	U	Y	U	Y	50
35	Sulitzeanu 1988	N	U	Y	Y	Y	U	U	Y	U	Y	50
36	Sutherland 1978	N	N	Y	Y	Y	U	U	Y	U	U	40
37	Tinguely 2000	Y	Y	Y	Y	Y	U	U	Y	U	Y	70
38	WG 1996	Y	U	Y	Y	Y	U	U	Y	U	Y	60
39	Xue 2002	Y	U	Y	Y	Y	U	U	Y	U	Y	60

1. Were the groups comparable other than the presence of disease in cases or the absence of disease in controls? **2.** Were cases and controls matched appropriately? **3.** Were the same criteria used for identification of cases and controls? **4.** Was exposure measured in a standard, valid and reliable way? **5.** Was exposure measured in the same way for cases and controls? **6.** Were confounding factors identified? **7.** Were strategies to deal with confounding factors stated? **8.** Were outcomes assessed in a standard, valid and reliable way for cases and controls? **9.** Was the exposure period of interest long enough to be meaningful? **10.** Was appropriate statistical analysis used? Y=Yes; N=No; U=Unclear.

Table S4. Quality assessment of the included cohort studies

No.	Study ID	Questions assessing included cohort studies											Yes (%)
		1	2	3	4	5	6	7	8	9	10	11	
1	Ambrosio 2014	N	Y	Y	U	U	N	Y	U	N	U	Y	36.36
2	Banatvala 1972	Y	Y	U	U	U	U	Y	U	Y	U	Y	45.45
3	Bingler 2008	Y	Y	Y	U	U	U	Y	U	N	Y	Y	54.54
4	Boyle 1991	Y	Y	Y	U	U	U	Y	Y	Y	U	Y	63.63
5	Căinap 2012	Y	Y	Y	U	U	U	Y	Y	Y	Y	Y	72.72
6	Donati 2006	Y	Y	Y	U	U	U	Y	Y	Y	N	Y	63.63
7	Gonin 2011	U	Y	Y	U	U	U	Y	N	N	Y	Y	45.45
8	Granai 2020	N	N	Y	U	U	U	Y	U	Y	U	Y	36.36
9	Hishima 2006	Y	Y	Y	U	U	U	Y	Y	Y	U	Y	63.63
10	Mundo 2017	Y	Y	Y	U	U	U	Y	Y	Y	N	Y	63.63
11	Okano 1992	Y	Y	Y	U	U	U	Y	Y	Y	U	Y	63.63
12	Pedersen 1991	Y	Y	Y	U	U	U	Y	Y	Y	Y	Y	72.72
13	Peylan-Ramu 2001	Y	Y	Y	U	U	U	Y	Y	N	U	Y	54.54
14	Pizza 2008	Y	Y	Y	U	U	Y	Y	Y	Y	Y	Y	81.81
15	Qin 2018	Y	Y	Y	U	U	Y	Y	U	Y	U	Y	63.63
16	Sinha 2016	Y	Y	Y	U	U	U	Y	Y	Y	N	Y	63.63
17	Westmoreland 2017	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	100

1. Were the two groups similar and recruited from the same population? **2.** Were the exposures measured similarly to assign people to both exposed and unexposed groups? **3.** Was the exposure measured in a valid and reliable way? **4.** Were confounding factors identified? **5.** Were strategies to deal with confounding factors stated? **6.** Were the groups/participants free of the outcome at the start of the study (or at the moment of exposure)? **7.** Were the outcomes measured in a valid and reliable way? **8.** Was the follow up time reported and sufficient to be long enough for outcomes to occur? **9.** Was follow up complete, and if not, were the reasons to loss to follow up described and explored? **10.** Were strategies to address incomplete follow up utilized? **11.** Was appropriate statistical analysis used? Y=Yes; N=No; U=Unclear; NA: Not applicable.