

A	B	C	D	E	F	G	H	I	K	L	M	N	O	P		
Totgevoerd door	Literatuur (titel, auteur, jaartal, journal)	Type literatuur (investerend, theoretisch, geprefereerd of gepubliceerd)	Qualiteit van evidence	Opt check	Toetsing Qof (Indirectness, Inconsistency, indirectness, Imprecision, effect size, right confounders, dose response)	Risk of bias (study limitations: design, inclusion and sample info, measurement, confounding, follow-up)	Toetsing bias	Type studie (zie werkblad hierboven)	Land	Verplichting of advies	Stekproef (grootte, populatie)	Recruitment (opvolgende studies in/excl cr., hoe geworven)	Representatief? (Is deze studie vergelijkbaar met NL situatie of populatie?)	Sleutelwoorden (gedrag/determinante/omstandigheden/redenen/maats van naleving/verschillen/interventies)	Doel studie	
1																
2	GLOBAL BEHAVIOR AND PERCEPTIONS AT THE ONSET OF THE COVID-19 PANDEMIC (2020) - At 2020	wet-ting paper	Ons een moderate kwalitatieve studie. Het is namelijk een observatieve en cross-sectorenkeerde vragenlijststudie, met een grote steekproef die verschillende robuustheidstests heeft. Er zijn geen info over de voorbereiding van de studie en de vragen zijn door via social media gevraagd. Er zijn verschillende verificaties die de vraaglijst niet kunnen bevestigen. De vragenlijst was erg goed ontworpen maar hoge precisie +1.	robustness of findings, precision, confounding, check, maar wel online vragenlijst. Wel heel groot.	Uitvoerende: wet-ting groepen via sociale media verspreiden de vragenlijst; was er dan wel een goede voorbereiding van de studie? Robustheid: Zij gebruikten snowball sampling, geen volledige representativiteit van de populatie over representatief van demografische groepen. Maar ze gaven wel diverse mogelijkheden om te reageren op de vragen. Ze wilden representatieve uitkomsten. De resultaten van de steekproef waren gelijk aan de resultaten van de populatie. De resultaten waren ook vergelijkbaar met andere studies die vergelijkbare vragenlijsten hadden. Methodische kwaliteit: De studie had een goede voorbereiding (+1), maar hoge precisie +1.	Uiter, selectie bias is wel gecontroleerd. To examine robustness of the findings to potential selection bias, we repeated the survey in April 2020. We recruited a representative sample from a representative online panel provided by Profilin. In total, 277 individuals participated in the survey. The results from the two surveys were similar to those obtained in the main article.	sampen resultaten van de 58 landen in welke er 200 mensen deelnamen, correspondeert met de resultaten van de oorspronkelijke studie.	online snowball sampling	interventie	interventie	interventie	interventie	interventie	interventie		
3																
4																
5	1	gepubliceerd paper	low tot moderate		■■■■■ 1 want online cross section en niet representatieve steekproef, indirectness (willingness to comply), inconsistent effect sizes, maar grote steekproef en confounders, gepreporter	selectie bias (welke niet veel verschilt in hoger inkomen), online vragenlijstobservaties -1, wel info gegeven over steekproef;			groot aantal mensen, gevonden online tussen 1 en 5 april 2020 (net hiervoor op 31 maart was er een verlening van nationale lockdown tot 31 april).	The final sample consisted of 14,433 users (59% female, 39% male, 0.4% diverse, 3.7% did not report their gender). Ages ranged between 10 and 80 years (M=44.2±13.6). Twenty-two percent of respondents were working in the medical/health care sector, 11% were working in the education sector, 10% were unemployed, 10% were retired, 13% were students, 23.1% had a secondary school certificate (Abitur), 18.1% had a vocational training certificate (Meister), 18% had a vocational training degree (Meister), 18% had a higher education degree (Bachelor), 17.6% had a master's degree and 10.4% had a postgraduate degree. The mean age of the study sample was comparable to the German population, but women and people with a higher education were slightly overrepresented. On average, 3% of all responses were missing. Given the large sample size and the low number of missing values, imputation was considered unnecessary here.	Deutsch					
6	1.1.7.1		moderate		sparsely written, representatieve respondenten, zelfverklaring maar wel wat denkt dat zou moeten dus lijkt teniet te doen dat gezondheidsgerelateerde schalen geen directe meetwaarden maar wel indirecte meetwaarden, percentage gemeld van antwoordvragen voor maatregelen.	steekproef representatief via bureau, online vragenlijst studie.			We do this via interview: Representative sample of 7,682 Italian adults between March 18th 2020 and March 20th 2020.							
7	1.1.7.2		low to moderate	minus 1 door ■■■■■ Gedrag indirect gemeten via intensities, geven verklaringen voor inconsistencie in resultaten, kleine sample mits confounders.	Study limit door design - 1. Mit opt. of de confounders hebben gevraagd en wel hun criteria was voor participant. Wel-voldende info gegeven over participanten.			During this period, metropolitan Melbourne was in stage 4 social distancing restrictions. During stage 4, residents were not permitted to leave their home for non-essential reasons for up to 12 hours a day. Residents were not permitted to leave their home for exercise or to visit their place of residence and could only leave their home for essential reasons such as shopping for groceries, exercise or to go to work.	Respondents consisted of 174 adults who were living in metropolitan Melbourne, Australia between 1 March 2020 and 15 September 2020. After providing informed consent, respondents were asked to provide information about their sex, age, ethnicity, their age, gender, residential postcode, marital status, education level and current working conditions.	For receiving Human Research Ethics approval, the survey was administered via Qualtrics online survey platform over a 3 weeks per-period from 28 August 2020 to 15 September 2020.						
8																
9																
10																
11																
12																
13																
14	1.1.7.3		moderat to high		Uitkomsten gemeten met gevuldeerde schalen,	longitudinal, grote steekproef, goede winning en beschrijving hiervan. Gevuldeerde schalen gebruikt, ééntesten bleek niet representatief maar hiervoor is gewogen.										
15			moderate		Directness, consistency, precision, het is een kleine consistentie dus geen correctie is gecontroleerd voor plausible continuïteit.											
16																
17																

Studies.org			
1	2	3	4
Targeted door	Methode (control group, etc)	Measures DV on IV (item/scale/qualidient/intentis/gdrag/self-report)	Confounders
1			
2	We use daily country level data on the extent of different government COVID-19 interventions [8]. To justify a causal interpretation, we exploit time variation in country-level COVID-19 responses, thus comparing across different countries at the same time. We find that, as a country imposes more stringent non-pharmaceutical public health measures, there is a significant increase in the probability of an individual mention by the public ($p < 0.001$) and the government ($p < 0.001$). (i) Increase trust that governments keep people safe and that they are doing what is best for them. (ii) Increase the sense of social control via vrag: 'in my opinion, there should be a general curfew in my country (with the exception of grocery shopping, necessary family trips, and the commute to work) because of the corona-virus right now'	In my opinion, there should be a general curfew in my country (with the exception of grocery shopping, necessary family trips, and the commute to work) because of the corona-virus right now. How many of 200 people in your country do you think believe there should be a general curfew in your country? (0-100) (with weight observations to make them representative at the country level, based on respondent gender, age, income, and education. Depending on the focus of the analysis, we also weight according to country population, or give all 100% weight)	
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			

A	B	C	D
Teugenseg door	Methode (controle groep, etc)	Measures en IV (item/schaal/gvaliditeit/intensie/gedrag/self-report)	Confounding
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			
25			
26			
27			
28			
29			
30			
31			
32			
33			
34			
35			
36			
37			
38			
39			
40			
41			
42			
43			
44			
45			
46			
47			
48			
49			
50			
51			
52			
53			
54			
55			
56			
57			
58			
59			
60			
61			
62			
63			
64			
65			
66			
67			
68			
69			
70			
71			
72			
73			
74			
75			
76			
77			
78			
79			
80			
81			
82			
83			
84			
85			
86			
87			
88			
89			
90			
91			
92			
93			
94			
95			
96			
97			
98			
99			
100			
101			
102			
103			
104			
105			
106			
107			
108			
109			
110			
111			
112			
113			
114			
115			
116			
117			
118			
119			
120			
121			
122			
123			
124			
125			
126			
127			
128			
129			
130			
131			
132			
133			
134			
135			
136			
137			
138			
139			
140			
141			
142			
143			
144			
145			
146			
147			
148			
149			
150			
151			
152			
153			
154			
155			
156			
157			
158			
159			
160			
161			
162			
163			
164			
165			
166			
167			
168			
169			
170			
171			
172			
173			
174			
175			
176			
177			
178			
179			
180			
181			
182			
183			
184			
185			
186			
187			
188			
189			
190			
191			
192			
193			
194			
195			
196			
197			
198			
199			
200			
201			
202			
203			
204			
205			
206			
207			
208			
209			
210			
211			
212			
213			
214			
215			
216			
217			
218			
219			
220			
221			
222			
223			
224			
225			
226			
227			
228			
229			
230			
231			
232			
233			
234			
235			
236			
237			
238			
239			
240			
241			
242			
243			
244			
245			
246			
247			
248			
249			
250			
251			
252			
253			
254			
255			
256			
257			
258			
259			
260			
261			
262			
263			
264			
265			
266			
267			
268			
269			
270			
271			
272			
273			
274			
275			
276			
277			
278			
279			
280			
281			
282			
283			
284			
285			
286			
287			
288			
289			
290			
291			
292			
293			
294			
295			
296			
297			
298			
299			
300			
301			
302			
303			
304			
305			
306			
307			
308			
309			
310			
311			
312			
313			
314			
315			
316			
317			
318			
319			
320			
321			
322			
323			
324			
325			
326			
327			
328			
329			
330			
331			
332			
333			
334			
335			
336			
337			
338			
339			
340			
341			
342			
343			
344			
345			
346			
347			
348			
349			
350			
351			
352			
353			
354			
355			
356			
357			
358			
359			
360			
361			
362			
363			
364			
365			
366			
367			
368			
369			
370			
371			
372			
373			
374			
375			
376			
377			
378			
379			
380			
381			
382			
383			
384			
385			
386			
387			
388			
389			
390			
391			
392			
393			
394			
395			
396			
397			
398			
399			
400			
401			
402			
403			
404			
405			
406			
407			
408			
409			
410			
411			
412			
413			
414			
415			
416			
417			
418			
419			
420			
421			
422			
423			
424			
425			
426			
427			
428			
429			
430			
431			
432			
433			
434			
435			
436			
437			
438			
439			
440			
441			
442			
443			
444			
445			
446			
447			
448			
449			
450			
451			
452			

	A	B	C	D
1	Quality of evidence is a continuum; any discrete categorisation involves some degree of arbitrariness.			
2	While factors influencing the quality of evidence are additive – such that the reduction or increase in each individual factor is added together with the other factors to reduce or increase the quality of evidence for an outcome – grading the quality of evidence involves judgements which are not exclusive. Therefore, GRADE is not a quantitative system for grading the quality of evidence. Each factor for downgrading or upgrading reflects not discrete categories but a continuum within each category and among the categories. When the body of evidence is intermediate with respect to a particular factor, the decision about whether a study falls above or below the threshold for up- or downgrading the quality (by one or more factors) depends on judgment.			
3				
4				
5				
6				
7	Table 5.1: Quality of Evidence Grades			
8	Grade	Definition		
9	High	We are very confident that the true effect lies close to that of the estimate of the effect.		
10	Moderate	We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different		
11	Low	Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect.		
12	Very Low	We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect		
13				
14				
15	Study Design			
16	Study design is critical to judgments about the quality of evidence.			
17	For recommendations regarding management strategies – as opposed to establishing prognosis or the accuracy of diagnostic tests –			
18	randomized trials provide, in general, far stronger evidence than observational studies, and rigorous observational studies provide stronger evidence than uncontrolled case series.			
19	In the GRADE approach to quality of evidence:			
20	randomized trials without important limitations provide high quality evidence			
21	observational studies without special strengths or important limitations provide low quality evidence			
22				
23	Limitations or special strengths can, however, modify the quality of the evidence of both randomized trials and observational studies.			
24	Note:			
25	Non-randomised experimental trials (quasi-RCT) without important limitations also provide high quality evidence, but will automatically be downgraded for limitations in design (risk of bias) – such as lack of concealment of allocation and tie with a provider (e.g. chart number).			

	A	B	C	D
1	Quality of evidence is a continuum; any discrete categorisation involves some degree of			
26	Case series and case reports are observational studies that investigate only patients exposed to the intervention. Source of control group results is implicit or unclear, thus, they will usually warrant downgrading from low to very low quality evidence.			
27	Expert opinion is not a category of quality of evidence. Expert opinion represents an interpretation of evidence in the context of experts' experiences and knowledge. Experts may have opinion about evidence that may be based on interpretation of studies ranging from uncontrolled case series (e.g. observations in expert's own practice) to randomized trials and systematic reviews known to the expert. It is important to describe what type of evidence (whether published or unpublished) is being used as the basis for interpretation.			

	E	F	G	H
1	Table 5.2: Factors that can reduce the quality of the evidence			VOOR REVIEWS
2	Factor	Consequence	toelichting	
3	Limitations in study design or execution (risk of bias)	↓ 1 or 2 levels	zie linksonder Study Design en volgend tabblad voor Risk of Bias.	5.1.2h
4	Inconsistency of results	↓ 1 or 2 levels	Niet toegelichte heterogeniteit van resultaten (vooral bij syst reviews, als er veel verschillende bevindingen zijn, gemengd bewijs).	
5	Indirectness of evidence	↓ 1 or 2 levels	Bijvoorbeeld gemeten met een surrogaat maat (niet gedrag, maar intentie of zelfgerapporteerde gedrag). Of nt andere interventie (niet thuisblijven bij klachten maar thuisblijven in het algemeen).	
6	Imprecision	↓ 1 or 2 levels	Kleine steekproef of kleine hoeveelheid events, dus wijd confidence interval	
7	Publication bias	↓ 1 or 2 levels	Lastig te achterhalen, gaat erom in hoeverre er studies met negatieve of andere resultaten niet zijn gepubliceerd en dus niet zijn opgenomen. Vooral voor syst reviews relevante factor. Bij losse studies gaat het om reporting bias (zijn er resultaten weggelaten die wel relevant zijn, nulbevindingen bijv)	
8	Table 5.3: Factors that can increase the quality of the evidence			
9	Factor	Consequence		
10	Large magnitude of effect	↑ 1 or 2 levels	Als er een groot effect wordt gevonden. For simple regression β is like R. Thus I would use R rules of thumb... I use the following with my Psychology students: $\beta < 0.1$ - Small effect size $\beta \in [0.1; 0.5]$ - Medium effect size $\beta \geq 0.5$ - Large effect size. For multiple regression these rules are not that straightforward, but for Social Sciences they seem to hold (also following Cohen's d suggestions).	
11	All plausible confounding would reduce the demonstrated effect or increase the effect if no effect was observed	↑ 1 level	Is er gecontroleerd voor plausibele confounders?	
12	Dose-response gradient	↑ 1 level		
13				
14				
15				
16				
17				
18				
19				
20				
21				
22				
23				
24				
25				

	E	F	G	H
1	Table 5.2: Factors that can reduce the quality of the evidence			VOOR REVIEWS
26				
27				

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1														
2														
3	RISK OF BIAS + Limitations in the study design and execution may bias the estimates of the treatment effect. Our confidence in the estimate of the effect and in the following recommendation decreases if studies suffer from major limitations. The more serious the limitations are, the more likely it is that the quality of evidence will be downgraded. Numerous tools exist to evaluate the risk of bias in randomized trials and observational studies.													
4	Risk of bias	Uitieg												
5	None													
6	Due to confounding	Baseline confounding occurs when one or more prognostic variables (factors that predict the outcome of interest) also predict the intervention received at baseline. ROBINS-I can also address time-varying confounding, which occurs when individuals switch between the interventions being compared and when post-baseline prognostic factors affect the intervention received after baseline.												
7	In participant selection	When selection of some eligible participants, or the initial follow-up of some participants, or later attrition even in relation to both intervention and outcome, there will be an association between interventions and outcome even if the effects of the interventions are identical. This form of selection bias is distinct from confounding—a specific example is bias due to the inclusion of prevalent users, rather than new users, of an intervention.												
8	Due to missing data	Bias that arises when later follow-up is missing for individuals initially included and followed (such as differential loss to follow-up that is affected by prognostic factors), bias due to exclusion of individuals with missing information about intervention status or other variables such as confounders.												
9	In measurement of predic/outcome	Bias introduced by either differential or non-differential errors in measurement of outcome data. Such bias can arise when outcome assessors are aware of intervention status, if different methods are used to assess outcomes in different intervention groups, or if measurement errors are related to intervention status.												
10	In selection of reported result	Selective reporting of results in a way that depends on the findings and prevents the estimate from being included in a meta-analysis (or other synthesis).												
11	In misclassification of intervention (randomization)	Bias introduced by either differential or non-differential misclassification of intervention status. Non-differential misclassification is unrelated to the outcome and will usually bias the estimated effect of intervention towards the null. Differential misclassification, when misclassification of intervention status is related to the outcome or the risk of the outcome, and is likely to lead to bias.												
12	Due to deviation from intended intervention	Bias that arises when there are systematic differences between experimental intervention and comparator groups in the care provided, which represent a deviation from the intended intervention(s).												
13	14													
15	Tussen studies	- Comparatie: bij interventie-studies, goed bekijken wat de comparatie conditie is en of studies vergelijkbaar zijn met elkaar. - Outcomes: zijn gebruikte uitkomstmaten vergelijkbaar? (gaat het om intentie van gedrag, zelfgerapporteerde naleving, daadwerkelijke naleving, etc.)												
16														
17														
18														
19														
20														
21														
22														
23														
24														
25														
26														
27														
28														
29														
30														
31														
32														
33														
34														
35														
36														

Table 5.4: Study limitations in randomized controlled trials

Explanation
Lack of allocation concealment Those enrolling participants are aware of the group (or period) in a crossover trial to which the next assigned patient will be allocated (a major problem in "pseudo" or "quasi" randomized trials with allocation by day of week, birth date, chart number, etc.)
Lack of blinding Patient, caregivers, those recording outcomes, those adjudicating outcomes, data analysts are aware of the group to which patients were allocated (or the medication currently being received in a crossover trial).
Incomplete accounting of patients and outcome events Loss to follow-up and failure to adhere to the intervention protocol are especially in superiority trials, or in noninferiority trials, loss to follow-up, and failure to conduct both analyses considering only those who adhered to treatment, and all patients for whom outcome data are available. The significance of particular rates of loss to follow-up, however, varies widely and is dependent on the intervention, the number of follow-up and number of events. The higher the proportion lost to follow-up in relation to interventions and control group event rates, and differences between intervention and control groups, the greater the threat of bias.
Selective outcome reporting Incomplete or absent reporting of some outcomes and not others on the basis of the results.
Other limitations Stopping trial early for safety or futility, overrecruitment are likely in trials with fewer than 500 events, but large overestimates are likely in trials with fewer than 200 events. Prognostic factors are often external stopping rules do not reduce this bias. Use of unvalidated outcome measures (e.g. patient-reported outcomes) Carryover effects in crossover trial Recruitment bias in cluster-randomized trials

Table 5.5: Study limitations in observational studies

	Explanation
Failure to develop and apply appropriate eligibility criteria (inclusion of control population)	<ul style="list-style-type: none"> Under- or over-recruiting in case-control studies Selection of exposed and unexposed in cohort studies from effect populations
Failure to measure both exposure and outcome	<ul style="list-style-type: none"> Differences in measurement of exposure (e.g. recall bias in case-control studies) Differential surveillance for outcome in exposed and unexposed in cohort studies
Failure to adequately control confounding	<ul style="list-style-type: none"> Failure of accurate measurement of all known prognostic factors Failure to match for prognostic factors and/or adjustment in statistical analysis
Incomplete or inadequately short follow-up	Especially within prospective cohort studies, both groups should be followed for the same amount of time.

	O	P
		5.1.2h
1.	Were the criteria for inclusion in the sample clearly defined?	The authors should provide clear inclusion and exclusion criteria that they developed prior to recruitment of the study participants.
2.	Were the study subjects and the setting described in detail?	The study sample should be described in sufficient detail so that other researchers can determine if it is comparable to their own interests. The authors should provide a clear description of the persons from whom the study participants were selected or recruited, including: demographics, location, and time period.
3.	Was the exposure measured in a valid and reliable way?	The study should clearly describe the method of measurement of exposure. Assessing validity requires that a 'gold standard' is available to which the measure can be compared. The validity of exposure measurement usually refers to whether the measure measures what it purports to measure, or whether a measure of past exposure is needed. Reliability refers to the processes included in an epidemiological study to check repeatability of measurements of the exposure. These include inter-rater reliability and inter-observer reliability.
4.	Were objective, standard criteria used for measurement of the condition?	It is useful to determine if patients were included in the study based on either a specified diagnosis or definition. This is more likely to decrease the risk of bias. Characteristics are another useful approach to matching groups, and studies that did not use specified diagnostic methods often show no evidence on matching by key characteristics.
5.	Were confounding factors identified?	Typical confounders include baseline characteristics, prognostic factors, or concomitant exposures (e.g., smoking). A confounder is a difference between the comparison group and the study group that may affect the study results. A high quality study at the level of cohort design will identify potential confounders and measure them (where possible). This is difficult for studies where behavioral, attitudinal or lifestyle factors may impact on the results.
6.	Were strategies to deal with confounding factors stated?	Strategies to deal with effects of confounding factors may be dealt within the study design or in data analysis. By matching or stratifying sampling of cases, effects of confounding factors can be adjusted for. When dealing with adjustment in analysis, assess the statistics used in the study. Most will be some form of multivariate regression analysis to account for the confounding factors measured.
7.	Were the outcomes measured in a valid and reliable way?	Importantly, determine if the measurement tools used were validated instruments as this has a significant impact on outcome assessment validity. Having established the objectives of the study, the measurement tools (or cancer instrument, if it's important to establish how the measurement was conducted. Were those involved in collecting data trained for each tool? Are the instruments the same (in particular, reporters or informants) and how specific confounders were measured. For studies utilizing regression analysis, it is useful to identify if the study took into account whether variables were related to how they related to the outcome, if stratification was the analytical approach used, were the strata of analysis defined by the specified variables? Additionally, it is also important to consider the statistical approach and analytical strategy in terms of the assumptions associated with the approach as differing methods of analysis are based on differing assumptions about the data and how it will respond.
8.	Was appropriate statistical analysis used?	As with any consideration of statistical analyses, consideration should be given to whether there was a more appropriate alternate statistical method that could have been used. The methods section should be detailed enough for readers to identify which analytical techniques were used (in particular, regression or stratification) and how specific confounders were measured. For studies utilizing regression analysis, it is useful to identify if the study took into account whether variables were related to how they related to the outcome, if stratification was the analytical approach used, were the strata of analysis defined by the specified variables? Additionally, it is also important to consider the statistical approach and analytical strategy in terms of the assumptions associated with the approach as differing methods of analysis are based on differing assumptions about the data and how it will respond.
9.		
10.		
11.		
12.		
13.		
14.		
15.		
16.		
17.		
18.		
19.		
20.		
21.		
22.		
23.		
24.		
25.		
26.		
27.		
28.		
29.		
30.		
31.		
32.		
33.		
34.		
35.		
36.		

