

Supplementary Table 1. Descriptive statistics for 5- to 3-cluster solutions of sickness absences (SA) in 2021. The 4-cluster solution was selected

	5-Clusters	4-Clusters	3-Clusters
Cluster 1, n (%)	119 (0.3)	358 (1)	8399 (21)
Cluster 2, n (%)	31136 (78)	31320 (79)	30548 (77)
Cluster 3, n (%)	5108 (13)	5149 (13)	844 (2)
Cluster 4, n (%)	403 (1)	2965 (7)	
Cluster 5, n (%)	3025 (8)		
Minimum distance between initial seeds	11.65	12.93	17.26
Criterion based on final seeds	0.56	0.58	0.67
RSQ/(1-RSQ) ^a : SA days	4.89	4.16	2.47
RSQ/(1-RSQ) ^a : Short (1-9 days) SA episodes	1.36	1.35	0.84
RSQ/(1-RSQ) ^a : Long (10-365 days) SA episodes	1.97	1.49	0.99
RSQ/(1-RSQ) ^a : Overall	2.23	1.94	1.25
Pseudo F statistic	22186.37	25667.33	24915.89
Approximate exp. overall R-squared ^b	0.66	0.60	0.44

^a The ratio of between-cluster variance to within-cluster variance

^b Note the high correlation between SA days and long episodes making this statistic invalid for conclusions on overall coefficient of determination. However, it can be used for comparison between cluster solutions.

Supplementary Table 2. Descriptive statistics for 5- to 3-cluster solutions of sickness absence (SA) in 2019. The 4-cluster solution was selected

	5-Clusters	4-Clusters	3-Clusters
Cluster 1, n (%)	3749 (9)	29608 (74)	5800 (15)
Cluster 2, n (%)	29608 (74)	4931 (12)	33135 (83)
Cluster 3, n (%)	4948 (12)	4484 (11)	856 (2)
Cluster 4, n (%)	290 (0.7)	768 (2)	
Cluster 5, n (%)	1196 (3)		
Minimum distance between initial seeds	9.00	11.79	16.14
Criterion based on final seeds	0.49	0.53	0.66
RSQ/(1-RSQ) ^a : SA days	4.96	3.37	3.14
RSQ/(1-RSQ) ^a : Short (1-9 days) SA episodes	1.25	1.18	0.18
RSQ/(1-RSQ) ^a : Long (10-365 days) SA episodes	7.0	4.94	3.14
RSQ/(1-RSQ) ^a : Overall	3.16	2.56	1.33
Pseudo F statistic	31458.9	33913.4	26443.7
Approximate exp. overall R-squared ^b	0.66	0.60	0.44

^a The ratio of between-cluster variance to within-cluster variance

^b Note the high correlation between SA days and long episodes making this statistic invalid for conclusions on overall coefficient of determination. However, it can be used for comparison between cluster solutions.

Supplementary Table 3. Clustering variables of sickness absences (SA) in 2019 among 39791 Finnish public sector employees

	Sickness absence patterns 2019			
	Low SA (n = 29608)	Repeated short SA episodes (n = 4931)	Repeated long SA episodes (n = 4484)	Very high SA long episodes (n = 768)
Number of SA days in 2019				
Mean (SD) ^a	2.9 (3.5)	18.3 (9.8)	34.8 (17.2)	117.1 (48.5)
Range	0-22	1-102	0-91	48-365
Number of short (1-9 days) SA episodes in 2019				
Mean (SD) ^a	1.3(1.3)	7.1 (2.8)	2.8 (2.3)	3.9 (3.9)
Range	0-4	4-30	0-14	0-24
Number of long (10-365 days) SA episodes in 2019				
Mean (SD) ^a	0(0)	0.1 (0.3)	1.1 (0.4)	2.0 (0.9)
Range	0	0-2	1-3	1-6

^a SD, standard deviation

Supplementary Table 4. Employee and work characteristics in 2020 associated with the total number of sickness absence (SA) days and of short (1-9 days) and long (10-365 days) SA episodes in 2021 among Finnish public sector employees (n=39791). Negative binomial regression, RR=risk ratio; CI=confidence interval.

	Sickness absence					
	Total number of SA days		Short SA episodes		Long SA episodes	
	RR	95% CI	RR	95% CI	RR	95% CI
Employee characteristics						
<i>Sex</i>						
Men	1		1		1	
Women	1.06	0.98-1.14	1.16	1.10-1.22	1.13	1.07-1.19
<i>Age</i>	1.14	1.11-1.18	0.96	0.94-0.97	0.97	0.95-0.99
<i>Marital status</i>						
Married/cohabiting	1		1		1	
Single	1.12	1.04-1.21	1.07	1.02-1.12	1.01	0.96-1.06
Divorced/widowed	0.99	0.91-1.07	1.08	1.03-1.14	1.00	0.95-1.06
<i>Body mass index</i>	1.00	1.00-1.01	1.01	1.01-1.02	1.01	1.00-1.01
<i>Smoking</i>						
Never	1		1		1	
Ex-smoker	1.06	0.99-1.14	1.02	0.97-1.07	1.01	0.97-1.06
Current smoker	1.12	1.03-1.21	1.15	1.09-1.21	1.09	1.03-1.15
<i>Alcohol use</i>						
Moderate	1		1		1	
No	1.08	1.01-1.15	1.05	1.00-1.09	1.06	1.02-1.11
At-risk use	1.03	0.90-1.18	0.97	0.88-1.07	1.04	0.95-1.14
<i>First-wave COVID-19</i>						
No	1		1		1	

Yes	1.19	1.04-1.37	1.09	1.00-1.19	1.16	1.06-1.27
<i>Commute mode</i>						
Walking	0.98	0.96-1.00	1.01	0.99-1.02	1.00	0.99-1.02
Cycling	0.96	0.94-0.98	0.99	0.98-1.00	0.98	0.97-0.99
Public transport with active element	0.98	0.97-1.00	1.01	1.00-1.02	0.99	0.98-1.01
Public transport without active element	1.00	0.98-1.01	1.02	1.00-1.03	0.99	0.98-1.01
Car	0.98	0.96-1.00	1.00	0.98-1.01	0.99	0.98-1.00
Work characteristics						
<i>Socioeconomic status</i>						
High	1		1		1	
Intermediate	1.12	1.04-1.20	1.02	0.97-1.07	1.08	1.02-1.13
Low	1.13	1.05-1.22	1.11	1.06-1.17	1.06	1.01-1.12
<i>Job contract</i>						
Permanent	1		1		1	
Temporary	0.98	0.90-1.06	1.04	0.99-1.09	1.00	0.94-1.06
<i>Working time mode</i>						
No shift work	1		1		1	
Shift work	1.54	1.27-1.88	1.45	1.27-1.65	1.39	1.22-1.58
<i>Remote work due to COVID-19</i>						
No	1		1		1	
Yes	0.95	0.89-1.02	0.87	0.83-0.91	0.94	0.90-0.98
<i>Transfer into new tasks due to COVID-19</i>						
No	1		1		1	
Yes	0.99	0.90-1.09	1.00	0.94-1.07	1.01	0.94-1.08
<i>Team reorganisation due to COVID-19</i>						
No	1		1		1	
Yes	1.07	0.96-1.21	1.04	0.97-1.12	1.07	0.99-1.16

Supplementary Table 5. Descriptive statistics for 5- to 3-cluster solutions of sickness absence (SA) with two clustering (short and long episodes) variables

	5-Clusters	4-Clusters	3-Clusters
Cluster 1, n (%)	6165 (15)	1825 (5)	33098 (83)
Cluster 2, n (%)	28376 (71)	32439 (82)	6084 (15)
Cluster 3, n (%)	4630 (12)	5392 (14)	609 (2)
Cluster 4, n (%)	68 (0.2)	135 (0.3)	
Cluster 5, n (%)	552 (1)		
Minimum distance between initial seeds	8.47	11.51	15.09
Criterion based on final seeds	0.53	0.60	0.69
RSQ/(1-RSQ)*: Short (1-9 days) SA episodes	1.48	1.45	1.54
RSQ/(1-RSQ)*: Long (10-365 days) SA episodes	5.68	2.15	0.85
RSQ/(1-RSQ)*: Overall	2.62	1.76	1.14
Pseudo F statistic	26030.4	23305.2	22329.2
Approximate exp. overall R-squared	0.80	0.75	0.67

*The ratio of between-cluster variance to within-cluster variance

Supplementary Table 6. Patterns of sickness absence (SA) in 2021 among 39791 Finnish public sector employees, based on two clustering (short and long episodes) variables

	Sickness absence patterns		
	Low SA (n = 33098)	Repeated short SA episodes (n = 6084)	Repeated long SA episodes (n = 609)
Number of short (1-9 days) SA episodes in 2021			
Mean (SD) ^a	1.1 (1.2)	6.8 (2.9)	3.7 (3.2)
Range	0-4	3-35	0-16
Number of long (10-365 days) SA episodes in 2021			
Mean (SD) ^a	0.2 (0.4)	0.6 (0.7)	3.9 (1.6)
Range	0-2	0-3	3-16

^a SD, standard deviation

Supplementary Table 7. Self-rated health and work ability in 2020 among employees in different sickness absence (SA) patterns in 2021

	Sickness absence patterns			
	Low SA	Repeated short SA episodes	Repeated long SA episodes	Very high SA, long episodes
	(n = 31320)	(n = 5149)	(n = 2964)	(n = 358)
Self-rated health				
Mean (SD) ^a	4.2 (0.8)	3.9 (0.9)	3.7 (1.0)	3.4 (1.1)
Range	1-5	1-5	1-5	1-5
Self-rated work ability				
Mean (SD) ^a	8.2 (1.5)	7.7 (1.7)	7.3 (1.9)	6.7 (2.1)
Range	0-10	0-10	0-10	0-10

^a SD, standard deviation