Description of simulation results

Simulation 1: Type 1 error rate with maximum standard deviation[1] for three different nominal significance levels[2]			
Significance level = 0.1	**Type 1 error rate (95% CI)[3]**		
Maximum standard deviation	**Kruskal-Wallis**	**Fisher's ANOVA**	**Welch's ANOVA**
1	0.098 (0.092 - 0.103)	0.097 (0.091 - 0.103)	0.098 (0.092 - 0.104)
3	0.114 (0.108 - 0.12)	0.130 (0.123 - 0.136)	0.099 (0.094 - 0.105)
5	0.124 (0.117 - 0.13)	0.156 (0.149 - 0.163)	0.099 (0.093 - 0.105)
7	0.125 (0.119 - 0.132)	0.162 (0.155 - 0.17)	0.101 (0.095 - 0.107)
9	0.130 (0.124 - 0.137)	0.159 (0.152 - 0.166)	0.097 (0.091 - 0.103)
Significance level = 0.05	**Type 1 error rate (95% CI)**		
Maximum standard deviation	**Kruskal-Wallis**	**Fisher's ANOVA**	**Welch's ANOVA**
1	0.048 (0.044 - 0.052)	0.051 (0.046 - 0.055)	0.053 (0.048 - 0.057)
3	0.062 (0.058 - 0.067)	0.088 (0.082 - 0.093)	0.051 (0.047 - 0.055)
5	0.064 (0.06 - 0.069)	0.114 (0.108 - 0.12)	0.049 (0.045 - 0.053)
7	0.072 (0.067 - 0.077)	0.121 (0.114 - 0.127)	0.052 (0.047 - 0.056)
9	0.078 (0.073 - 0.083)	0.118 (0.112 - 0.125)	0.049 (0.045 - 0.053)
Significance level = 0.01	**Type 1 error rate (95% CI)**		
Maximum standard deviation	**Kruskal-Wallis**	**Fisher's ANOVA**	**Welch's ANOVA**
1	0.011 (0.009 - 0.013)	0.011 (0.009 - 0.013)	0.011 (0.009 - 0.013)
3	0.017 (0.015 - 0.02)	0.038 (0.035 - 0.042)	0.012 (0.01 - 0.014)
5	0.020 (0.017 - 0.023)	0.054 (0.049 - 0.058)	0.008 (0.006 - 0.01)
7	0.024 (0.021 - 0.027)	0.061 (0.056 - 0.065)	0.011 (0.009 - 0.013)
9	0.024 (0.021 - 0.027)	0.066 (0.061 - 0.071)	0.011 (0.009 - 0.013)
Simulation 2: Empirical probability of rejecting the null hypothesis with imbalanced sample sizes			
Significance level = 0.05	**Type 1 error rate (95% CI)**		
Sample size for each of two groups	**Kruskal-Wallis**	**Fisher's ANOVA**	**Welch's ANOVA**
20:40	0.032 (0.029 - 0.036)	0.018 (0.015 - 0.02)	0.051 (0.047 - 0.055)
30:30	0.059 (0.054 - 0.064)	0.053 (0.048 - 0.057)	0.051 (0.047 - 0.055)
40:20	0.090 (0.084 - 0.095)	0.110 (0.103 - 0.116)	0.050 (0.046 - 0.054)

1. We assumed the standard deviation of G – 1 groups is 1 and that of the rest of one group is $a(\geq 1)$, i.e. $\{\sigma_1, \sigma_2, \dots, \sigma_G\} = \{1,1, \dots, a\}$

2. The type 1 error rates for each section are expected to be the same as the significance level

3. The type 1 error rate was defined as the number of simulations with p-value ($p_i; i = 1,2, \dots, N$) below significance level (α): $\hat{p} = \sum_{i=1}^{N} I(p_i < \alpha)/N$. The 95% CI of type I error rates ($\hat{p}$) was computed as $\hat{p} \pm 1.96\sqrt{\hat{p}(1-\hat{p})/N}$.