Table S1: Recommended use cases for the confirmatory setting (type 1 error rate less than standard test and power within 0.05 of the standard test) using the 0.2 cutoff for equal subgroup sample sizes.

| **Number  of Arms** | **Outcome  Type** | **Model** | **Variance** | **Sample Size** | **Effect Size** |
| --- | --- | --- | --- | --- | --- |
| 1 | binary | MEM | -- | 200 | ≥ 0.3 |
| 300 - 1000 | ≥ 0.2 |
| MEMr | -- | 200 - 300 | ≥ 0.3 |
| 400 - 1000 | ≥ 0.2 |
| continuous | MEM | 1 | 200 - 1000 | ≥ 1 |
| 10 | 800 - 1000 | ≥ 1 |
| MEMr | 1 | 200 - 1000 | ≥ 1 |
| 10 | 200 - 900 | ≥ 2 |
| 1000 | ≥ 1 |
| 100 | 600 - 1000 | ≥ 4 |
| 2 | binary | MEM | -- | 800 | ≥ 0.1 |
| 1000 | ≥ 0.1 |
| MEMr | -- | 1000 | ≥ 0.4 |
| continuous | MEMr | 1 | 200 - 300 | ≥ 2 |
| 400 - 1000 | ≥ 1 |
| 10 | 200 - 300 | ≥ 4 |
| 400 - 900 | ≥ 3 |
| 1000 | ≥ 2 |

Table S2: Recommended use cases for the confirmatory setting (type 1 error rate less than standard test and power within 0.05 of the standard test) using the 0.8 cutoff for equal subgroup sample sizes.

| **Number  of Arms** | **Outcome  Type** | **Model** | **Variance** | **Sample Size** | **Effect Size** |
| --- | --- | --- | --- | --- | --- |
| 1 | continuous | MEMr | 1 | 800 - 1000 | ≥ 1 |
| 10 | 800 - 1000 | ≥ 1 |
| 100 | 800 - 1000 | ≥ 1 |
| 2 | binary | MEMr | -- | 800 - 1000 | ≥ 0.1 |
| continuous | MEMr | 1 | 800 - 1000 | ≥ 1 |
| 10 | 800 - 1000 | ≥ 1 |
| 100 | 800 - 1000 | ≥ 1 |

Table S3: Recommended use cases for the confirmatory setting (type 1 error rate less than standard test and power within 0.05 of the standard test) using the calibrated cutoff for equal subgroup sample sizes.

| **Number  of Arms** | **Outcome  Type** | **Model** | **Variance** | **Sample Size** | **Effect Size** |
| --- | --- | --- | --- | --- | --- |
| 1 | continuous | MEM | 1 | 300 - 400 | ≥ 1 |
| 600 - 1000 | ≥ 1 |
| 10 | 300 - 400 | ≥ 1 |
| 600 - 1000 | ≥ 1 |
| 100 | 300 - 400 | ≥ 1 |
| 600 - 1000 | ≥ 1 |
| MEMr | 1 | 300 - 400 | ≥ 1 |
| 600 - 1000 | ≥ 1 |
| 10 | 300 - 400 | ≥ 1 |
| 600 - 1000 | ≥ 1 |
| 100 | 300 - 400 | ≥ 1 |
| 600 - 1000 | ≥ 1 |
| 2 | binary | MEM | -- | 200 - 800 | ≥ 0.1 |
| 1000 | ≥ 0.1 |
| MEMr | -- | 200 - 800 | ≥ 0.1 |
| 1000 | ≥ 0.1 |
| continuous | MEM | 1 | 200 - 600 | ≥ 1 |
| 900 - 1000 | ≥ 1 |
| 10 | 200 - 600 | ≥ 1 |
| 900 - 1000 | ≥ 1 |
| 100 | 200 - 600 | ≥ 1 |
| 900 - 1000 | ≥ 1 |
| MEMr | 1 | 200 - 600 | ≥ 1 |
| 900 - 1000 | ≥ 1 |
| 10 | 200 - 600 | ≥ 1 |
| 900 - 1000 | ≥ 1 |
| 100 | 200 - 600 | ≥ 1 |
| 900 - 1000 | ≥ 1 |

Table S4: Recommended use cases for the exploratory setting (type 1 error rate < 0.15 and power at least 0.1 higher than standard test) using the 0.2 cutoff for equal subgroup sample sizes.

| **Number  of Arms** | **Outcome  Type** | **Model** | **Variance** | **Sample Size** | **Effect Size** |
| --- | --- | --- | --- | --- | --- |
| 1 | continuous | MEM | 10 | 200 | 1 |
| 2 | binary | MEM | -- | 200 - 400 | 0.1 - 0.4 |
| 500 - 700 | 0.1 - 0.3 |
| 800 - 1000 | 0.1 - 0.2 |

Table S5: Recommended use cases for the exploratory setting (type 1 error rate < 0.15 and power at least 0.1 higher than standard test) using the 0.8 cutoff for equal subgroup sample sizes.

| **Number  of Arms** | **Outcome  Type** | **Model** | **Variance** | **Sample Size** | **Effect Size** |
| --- | --- | --- | --- | --- | --- |
| 1 | binary | MEM | -- | 300 - 400 | 0.1 |
| MEMr | -- | 200 - 300 | 0.1 |
| continuous | MEMr | 10 | 200 | 1 |
| 100 | 200 | 2 - 3 |
| 300 | 2 |
| 2 | binary | MEMr | -- | 200 | 0.2 - 0.4 |
| continuous | MEMr | 10 | 200 | 1 - 2 |
| 100 | 200 | 3 - 4 |
| 300 | 4 |

Table S6: Recommended use cases for the exploratory setting (type 1 error rate < 0.15 and power at least 0.1 higher than standard test) using the calibrated cutoff for equal subgroup sample sizes.

| **Number  of Arms** | **Outcome  Type** | **Model** | **Variance** | **Sample Size** | **Effect Size** |
| --- | --- | --- | --- | --- | --- |
| 1 | binary | MEM | -- | 200 | 0.1 |
| 2 | binary | MEM | -- | 200 - 300 | 0.2 - 0.4 |
| 400 | 0.1 - 0.4 |
| 500 - 700 | 0.1 - 0.3 |
| 800 - 1000 | 0.1 - 0.2 |

Table S7: Recommended use cases for the confirmatory setting (type 1 error rate less than standard test and power within 0.05 of the standard test) using the 0.2 cutoff for unequal subgroup sample sizes.

| **Sample  Size Dist** | **Number  of Arms** | **Outcome  Type** | **Model** | **Variance** | **Sample Size** | **Effect Size** |
| --- | --- | --- | --- | --- | --- | --- |
| 0.25 | 1 | binary | MEM | -- | 200 - 300 | ≥ 0.3 |
| 400 - 1000 | ≥ 0.2 |
| MEMr | -- | 200 - 400 | ≥ 0.3 |
| 500 - 1000 | ≥ 0.2 |
| continuous | MEMr | 10 | 200 | ≥ 3 |
| 300 - 1000 | ≥ 2 |
| 2 | continuous | MEMr | 10 | 300 - 500 | ≥ 4 |
| 600 - 1000 | ≥ 3 |
| 0.10 | 1 | binary | MEM | -- | 200 | ≥ 0.4 |
| 300 - 600 | ≥ 0.3 |
| 700 - 1000 | ≥ 0.2 |
| MEMr | -- | 300 - 400 | ≥ 0.4 |
| 500 - 800 | ≥ 0.3 |
| 900 - 1000 | ≥ 0.2 |
| continuous | MEMr | 10 | 200 | ≥ 4 |
| 300 - 600 | ≥ 3 |
| 700 - 1000 | ≥ 2 |
| 2 | continuous | MEMr | 10 | 600 - 1000 | ≥ 4 |

Table S8: Recommended use cases for the confirmatory setting (type 1 error rate less than standard test and power within 0.05 of the standard test) using the 0.8 cutoff for unequal subgroup sample sizes.

| **Sample  Size Dist** | **Number  of Arms** | **Outcome  Type** | **Model** | **Variance** | **Sample Size** | **Effect Size** |
| --- | --- | --- | --- | --- | --- | --- |
| 0.25 | 1 | continuous | MEMr | 10 | 800 - 1000 | ≥ 1 |
| 2 | binary | MEMr | -- | 800 - 1000 | ≥ 0.1 |
| continuous | MEMr | 10 | 800 - 1000 | ≥ 1 |
| 0.10 | 1 | continuous | MEMr | 10 | 800 - 1000 | ≥ 1 |
| 2 | binary | MEMr | -- | 800 - 1000 | ≥ 0.1 |
| continuous | MEMr | 10 | 800 - 1000 | ≥ 1 |

Table S9: Recommended use cases for the confirmatory setting (type 1 error rate less than standard test and power within 0.05 of the standard test) using the calibrated cutoff for unequal subgroup sample sizes.

| **Sample  Size Dist** | **Number  of Arms** | **Outcome  Type** | **Model** | **Variance** | **Sample Size** | **Effect Size** |
| --- | --- | --- | --- | --- | --- | --- |
| 0.25 | 1 | continuous | MEM | 10 | 500 - 700 | ≥ 1 |
| 1000 | ≥ 1 |
| MEMr | 10 | 500 - 700 | ≥ 1 |
| 1000 | ≥ 1 |
| 2 | binary | MEM | -- | 200 - 1000 | ≥ 0.1 |
| MEMr | -- | 200 | ≥ 0.1 |
| 300 - 1000 | ≥ 0.1 |
| continuous | MEM | 10 | 200 | ≥ 1 |
| 400 - 1000 | ≥ 1 |
| MEMr | 10 | 200 | ≥ 1 |
| 400 - 1000 | ≥ 1 |
| 0.10 | 1 | continuous | MEM | 10 | 500 - 600 | ≥ 1 |
| 1000 | ≥ 1 |
| MEMr | 10 | 500 - 600 | ≥ 1 |
| 1000 | ≥ 1 |
| 2 | binary | MEM | -- | 200 - 1000 | ≥ 0.1 |
| MEMr | -- | 200 - 300 | ≥ 0.1 |
| 400 - 500 | ≥ 0.1 |
| 600 - 1000 | ≥ 0.1 |
| continuous | MEM | 10 | 200 - 300 | ≥ 1 |
| 500 | ≥ 1 |
| 800 | ≥ 1 |
| 1000 | ≥ 1 |
| MEMr | 10 | 200 - 300 | ≥ 1 |
| 500 | ≥ 1 |
| 800 | ≥ 1 |
| 1000 | ≥ 1 |

Table S10: Recommended use cases for the exploratory setting (type 1 error rate < 0.15 and power at least 0.1 higher than standard test) using the 0.2 cutoff for unequal subgroup sample sizes.

| **Sample  Size Dist** | **Number  of Arms** | **Outcome  Type** | **Model** | **Variance** | **Sample Size** | **Effect Size** |
| --- | --- | --- | --- | --- | --- | --- |
| 0.25 | 1 | continuous | MEM | 10 | 200 - 300 | 1 |
| 2 | binary | MEM | -- | 200 | 0.2 - 0.4 |
| 300 - 400 | 0.1 - 0.4 |
| 500 - 800 | 0.1 - 0.3 |
| 900 - 1000 | 0.1 - 0.2 |
| 0.10 | 1 | continuous | MEM | 10 | 400 - 600 | 1 |
| 2 | binary | MEM | -- | 200 | 0.1 - 0.4 |
| 300 - 600 | 0.2 - 0.4 |
| 700 - 800 | 0.1 - 0.4 |
| 900 - 1000 | 0.1 - 0.3 |

Table S11: Recommended use cases for the exploratory setting (type 1 error rate < 0.15 and power at least 0.1 higher than standard test) using the 0.8 cutoff for unequal subgroup sample sizes.

| **Sample  Size Dist** | **Number  of Arms** | **Outcome  Type** | **Model** | **Variance** | **Sample Size** | **Effect Size** |
| --- | --- | --- | --- | --- | --- | --- |
| 0.25 | 1 | binary | MEM | -- | 400 - 500 | 0.1 |
| MEMr | -- | 200 - 300 | 0.1 |
| continuous | MEMr | 10 | 200 | 1 |
| 2 | binary | MEMr | -- | 200 | 0.2 - 0.4 |
| continuous | MEMr | 10 | 200 | 1 - 2 |
| 0.10 | 1 | binary | MEM | -- | 800 - 1000 | 0.1 |
| continuous | MEMr | 10 | 200 | 1 - 2 |
| 2 | binary | MEMr | -- | 300 | 0.4 |
| continuous | MEMr | 10 | 200 | 2 - 3 |
| 300 | 2 |

Table S12: Recommended use cases for the exploratory setting (type 1 error rate < 0.15 and power at least 0.1 higher than standard test) using the calibrated cutoff for unequal subgroup sample sizes.

| **Sample  Size Dist** | **Number  of Arms** | **Outcome  Type** | **Model** | **Variance** | **Sample Size** | **Effect Size** |
| --- | --- | --- | --- | --- | --- | --- |
| 0.25 | 1 | binary | MEM | -- | 200 | 0.1 |
| 2 | binary | MEM | -- | 200 - 400 | 0.2 - 0.4 |
| 500 | 0.2 - 0.3 |
| 600 - 800 | 0.1 - 0.3 |
| 900 - 1000 | 0.1 - 0.2 |
| 0.10 | 2 | binary | MEM | -- | 200 | 0.4 |
| 300 - 400 | 0.3 - 0.4 |
| 500 - 800 | 0.2 - 0.4 |
| 900 - 1000 | 0.2 - 0.3 |