**Table S3: Dataset scenarios**

Ten datasets that cover five different scenarios

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Dataset no. | Description | Scenario | | | | |
| 1. Identical cell types different technologies | 2. Non-identical cell types | 3. Multiple batches | 4. Big data | 5. Simulation |
| 1 | Human Dendritic Cells |  | √ |  |  |  |
| 2 | Murine Atlas | √ |  |  |  |  |
| 3 | Simulations |  |  |  |  | √ |
| 4 | Human Pancreas |  | √ | √ |  |  |
| 5 | Human Peripheral Blood Mononuclear Cell | √ |  |  |  |  |
| 6 | Cell line |  | √ | √ |  |  |
| 7 | Mouse Retina |  | √ |  | √ |  |
| 8 | Mouse Brain |  | √ |  | √ |  |
| 9 | Human Cell Atlas |  |  |  | √ |  |
| 10 | Mouse Haematopoietic Stem and Progenitor Cells |  | √ |  |  |  |