数据分发工作总体可分为更新，筛选，下载，分发，回收与问题处理环节。核心是不断获取符合要求的非重复的单细胞转录组测序数据，并进行分发。本文档会对各个环节的重要节点进行质控说明和目前的工作办法（不涉及分发工作具体操作流程介绍）。

1. 更新环节
2. 需要记录每一次更新数据列表的输入，输出；
3. 每次更新设定的起始时间应比上一次更新时间早一天，以确保无缺漏；
4. 所有对数据分发记录应在工作记录表上，更新中使用的数据列表不得更改或添加额外信息；
5. 对于扩充期刊范围的情况下，应当尤其注意重复情况，以前没有直接检索这本期刊并不代表我们没做过上面的数据。
6. 筛选环节
7. 对于新增加的数据，进入筛选环节，需要依次根据文章查找数据集存储位置与编号，尤其注意对已经由脚本自动获得的GSE编号，也需要人工核验；
8. 筛选时的原则：单细胞，转录组，测序数据，且非公共数据重复使用的，非原始数据；
9. 下载环节
10. 对于一篇文章，可能有多个数据集，一个数据集下面可能还会有多个子集，而这些当中可能又有部分不是单细胞数据，不管采用怎样的方法，原则是，所有一篇文章下面的数据最终要合并到一起并分Part，且只要单细胞转录组测序数据；
11. 下载本身没有什么注意事项，只需按步骤运行脚本即可。对下载完的数据，有的需要合并、手动下载补充、判断数据格式等，这里重点说一下分part：首先，part本身代表不同的实验批次或实验设计，原则上同一实验设计下不需要分part；一般需要分part的情况有：不同子集，不同实验设计，同一种实验设计下完全不同的物种、用药、样本来源等。快捷方法：数据文件的命名会反应出一定的规律，如果有明显不同的批次，在文件名上会有体现，可以根据这一点并结合GEO上对数据实验的介绍，一般不需要细致阅读文章即可确定是否分part；
12. 对于下载失败的数据，需要重新试几次后，若一直失败，则记录下来，最后统一手动下载处理。
13. 分发

分发工作只需要严格记录，分发前根据分发记录和数据库已入库记录进行查重即可，此查重非常重要，不可缺少。

1. 回收与问题处理环节

以下两种情况需要进行数据回收：1.user主动申明不做；2.定期对长时间散布在外未收回的数据进行回收。回收的数据将原有分发记录清空，重新分发。

其他问题处理：对应相应问题，完善相应记录即可。