

基于生物信息学的医学基础课程案例设计*

王哲¹, 祁冉², 刘媛媛¹, 高路^{1△}

1. 中国人民解放军海军军医大学基础医学院生理学教研室, 上海 200433;

2. 中国人民解放军海军航空大学青岛校区航空工程基础教研室, 山东青岛 264001

摘要:目的 结合特定案例,设计基于生物信息学的医学基础课程,以加深学生对所学知识的理解,提高学生的综合素质。方法 该文以医学生《生理学》授课过程中涉及的知识点作为出发点,设计授课方案,通过知识点引入、知识点延伸、前沿研究实例讲解、讲解与演示、学生分组讨论与现场实操环节进行授课。结果 该授课方案在介绍与该知识点相关的国际前沿最新进展的基础上,对文献中涉及的主要生物信息学分析方法进行示例分析、讲解与结果解读。让学生具备基础医学专业知识的同时培养其国际视野,将人才培养与社会需求相结合。通过这一系列学习不仅扩展了学生在相关问题上的知识储备,而且培养了学生发现问题、分析问题和解决问题的科研能力。通过课程学习,激发了学生对临床和科研工作的兴趣,让学生在未來工作中从实际出发,敬业奉献。结论 该文设计的授课方案从医学生的专业背景入手,以学生为中心,将基础理论知识、临床需求、科研进展与方法学分析相结合,加深学生对医学基础知识的理解,激发学生的探索欲和求知欲,提高学生的综合能力。

关键词:医学教育; 生物信息学; 生理学; 单细胞测序; 高通量测序**中图分类号:**G423**文献标志码:**B**文章编号:**1672-9455(2024)04-0568-05

随着高通量测序技术的不断发展,其在医学临床上的应用日益广泛,例如,高通量测序和生物信息学分析可为出生缺陷三级预防、遗传病辅助诊断和肿瘤靶向治疗等提供依据^[1]。医学院校的学生不论将来从事临床工作还是相关领域的基础研究工作,都需要掌握一定的生物信息学知识和数据分析技能,以配合临床检测报告解读和研发新的诊断、治疗方法等。这些生物信息学知识与医学基础主干课程具有天然的联系,因此,在医学基础课程教学中适当增加生物信息学分析示例教学对学生深入理解基础知识、提高综合素质具有重要意义。但是,生物信息学是宏观概念,其研究范围广、相关方法繁杂,涉及多种计算机语言编程和数学模型算法,与医学知识结构存在较大差异。因此,对医学生进行生物信息学教学需要从学生的专业背景、知识结构和目标需求等角度出发,科学设立教学目标、规划教学内容、探索教学方法,在促进学生综合素质的同时,不过多增加学生的课程负担。因此,本研究尝试将生物信息学的方法、实际操作与学生的专业学习相结合。在理论课程中寻找一个知识点,结合权威文献介绍相关研究的最新进展,对文献涉及的主要生物信息学方法进行理论讲解、分析示范和结果解读,让学生初步了解分析步骤,掌握分析结果的解读方法,有兴趣的学生可以进一步进行实战分析练习和分组讨论。以期让学生在扩展专业知识、

了解国际前沿进展的同时学习生物信息学分析方法。相对于知识灌输型的传统课堂,在课程设计过程中结合“以学生为中心”的教学理念^[2-4],最大程度发掘学生的自主学习能力和自我能动性。本研究结合《生理学》内分泌章节的相关知识点^[5],设计了一次课程的基本教学内容,现报道如下。

1 材料与方法

1.1 知识点引入 在《生理学》内分泌章节中,下丘脑-垂体系统是其重点内容。首先对该知识点进行简要回顾^[5]:(1)下丘脑神经细胞分泌促释放激素和释放抑制激素,腺垂体释放促靶腺激素和直接作用激素,包括生长素、催乳素、促黑激素、促甲状腺激素、促肾上腺皮质激素、促卵泡激素和促黄体激素等^[2]。(2)生长素在促进生长过程中具有重要作用,幼年时期生长素分泌不足,可导致侏儒症;幼年时期生长素分泌过多,可导致巨人症,若成年时期生长素分泌过多,可导致肢端肥大症,其中垂体瘤是造成生长素分泌过多的重要原因之一。

1.2 知识点延伸

1.2.1 对垂体神经内分泌瘤进行延伸介绍 (1)垂体神经内分泌瘤是最常见的神经内分泌肿瘤,发病率在脑肿瘤中居第2位。(2)临床表现:占位效应和激素过度分泌,导致视力下降、停经、溢乳、肢端肥大等,累及全身多个重要脏器^[6]。(3)分类:根据激素分泌

* 基金项目:海军军医大学教学成果立项培育重点项目(JPY2020A19)。

△ 通信作者,E-mail:roadgao@163.com。

和免疫组化分为 10 种亚型^[7-8]。在临床治疗上,明确肿瘤起源和分化程度有助于预测患者预后情况。然而现有分类方式不能明确不同肿瘤亚型的起源。因此,需要基于肿瘤起源和分化程度进行垂体瘤分子分型,从而预测肿瘤远期预后。

1.2.2 介绍相关生物信息学技术 (1)传统技术的限制:传统转录组测序对组织直接进行 RNA 提取及建库测序,这样获得的转录组结果将各种细胞类型混合,无法区分细胞类型间的差异。如,在分析免疫微环境时,无法区分 T 淋巴细胞、B 淋巴细胞及巨噬细胞等不同类型免疫细胞;在分析肿瘤细胞时,无法精确分辨肿瘤细胞与癌旁细胞、肿瘤干细胞与高度分化细胞等,这不利于肿瘤分子分型,也不利于分析不同特征肿瘤细胞间的差异。(2)单细胞测序:单细胞转录组测序技术的出现实现了分别对每个细胞进行建库测序^[9]。通过分析单个细胞转录组特征,确定其细胞类型或细胞亚型,然后分析各个细胞类型或亚型的特征、类型间的基因表达差异及不同细胞类型间的相互作用关系等^[10-11]。这有助于检测并比较正常组织器官与各亚型肿瘤细胞的特征差异,发现稀有细胞类型,揭示肿瘤细胞来源,进行精确分子分型^[7]。

1.2.3 前沿研究实例讲解 这部分内容可以由学生分组进行文献讲解。以 ZHANG 等^[7]发表的论文为例。该团队通过单细胞转录组测序技术检测了 21 例垂体瘤组织和 3 例健康成人垂体前叶组织(对照)中合计 64 937 个细胞。解析了正常垂体的单细胞转录图谱,发现在 PIT1 谱系内,有一群内分泌细胞(PIT1_D)不表达经典的垂体激素(生长素、催乳素、促甲状腺激素),这群细胞最接近胚胎时期 PIT1 祖细胞的状态,命名为成人 PIT1 祖细胞。

为了探索垂体瘤的起源和分化程度,研究人员将正常垂体和垂体瘤的单细胞转录组表达谱进行整合分析;对垂体瘤各细胞亚群进行打分和重新分组,将各个谱系的垂体瘤划分为高分化组和低分化组;同时探索了每个分组特征性分子标志物、转录因子活性、富集通路等特征,筛选可能有助于预测肿瘤远期复发的标志物。最后研究人员在垂体瘤患者队列中,验证了候选标志物预测肿瘤远期复发的有效性^[8]。

1.3 讲解与演示 首先进行示教讲解,利用 GEO 数据库(<https://www.ncbi.nlm.nih.gov>)中 GSE205418 单细胞转录组数据进行基本的生物信息学分析示范。该数据是 10X Genomics 平台测序结果,并通过官方软件 Cell Ranger v3.1.0^[12]将原始数据处理为表达谱矩阵(包含 barcodes. tsv. gz, features. tsv. gz 和 matrix. mtx. gz),采用 R 语言进行下游分析^[13-14]。

(1)首先,打开 R 软件后,加载分析需要的 R 包:library(Seurat)library(harmony)library(dplyr)。

(2)以其中 1 个样本为例,读入数据:pr_name =

Shili" #项目名称,可以自己定义。input_dir = "~/Desktop/Shili_1" # barcodes. tsv. gz, features. tsv. gz 和 matrix. mtx. gz 3 个文件的位置,这里文件名不能修改。Data1 <- Read10X(data_dir = input_dir) # 读入 Shili_1 文件夹中的文件,导入数据。Or_ident = factor(c(rep("Sample1", ncol(Data1)))) # 定义样本名 "Sample1"。Data1 = CreateSeuratObject(counts = Data1, project = pr_name) # 创建 Seurat 对象,用于后续分析。Data1@meta.data\$orig.ident <- as.factor(Or_ident) # 见过定义的样本名加入 Seurat 对象。Data1["percent_mt"] <- PercentageFeatureSet(object = Data1, pattern = "^mt-") # 计算每个细胞表达线粒体基因的比例,一般认为表达线粒体基因比例过大的细胞,其细胞状态不好,可去掉。但是其具体标准无统一要求,可结合前人相似样本的过滤标准。P1 = VlnPlot(Data1, features = c("nFeature_RNA", "nCount_RNA", "percent_mt"), group.by = "orig.ident") # 对细胞表达线粒体基因比例、细胞表达基因数和捕获 UMI 数进行可视化,结果见图 1。

Data1 = subset(Data1, percent_mt < 20 & nFeature_RNA > 800 & nCount_RNA > 1000) # 过滤低质量细胞,这里仅保留表达线粒体基因比例小于 20%、表达基因数大于 800 且 UMI 数大于 1000 的细胞,可以根据实际情况调整。P2 = VlnPlot(Data1, features = c("nFeature_RNA", "nCount_RNA", "percent_mt"), group.by = "orig.ident") # 对过滤后细胞质量进行可视化,见图 2。

对于第 2 个样本,读入数据,方法同上。

(3)多个样本时,对样本 1(Data1)和样本 2(Data2)进行数据整合:All = merge(Data1, Data2) # 合并数据;All = NormalizeData(All) # 数据归一化处理,减少文库测序深度的影响;All = FindVariableFeatures(All, selection.method = "vst", nfeatures = 2000) # 鉴定细胞间表达量相差大的前 2000 个基因,用于后续鉴定细胞类型;All = ScaleData(All) # 数据标准化;All = RunPCA(All, pc.genes = All@var.genes, npcs = 50, verbose = FALSE) # PCA 降维并提取主成分,用于后续细胞降维;All = RunHarmony(All, group.by.vars = "orig.ident", plot.convergence = TRUE) # 按照"orig.ident"(样本名称)去除批次效应。

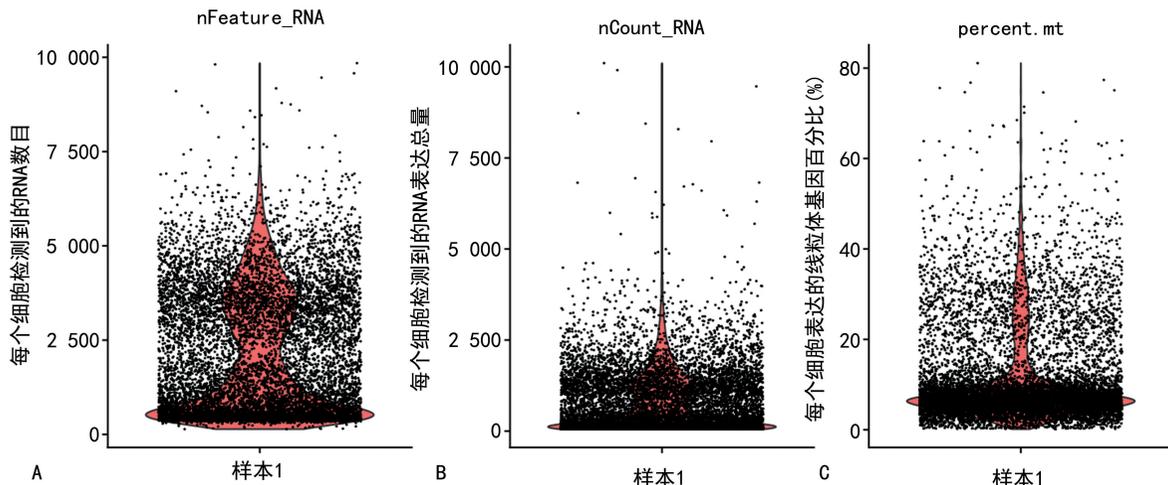
(4)对细胞进行降维、分群,见图 3:All = All %>% RunUMAP(reduction = "harmony", dims = 1:50) %>% # 降维算法 UMAP;RunTSNE(reduction = "harmony", dims = 1:50) %>% # 降维算法 tSNE, UMAP 和 TSNE 二选一即可;FindNeighbors(reduction = "harmony", dims = 1:50) %>%

```
FindClusters(resolution = 1) %>% # resolution
参数相当于放大镜,数值越大,分群越细致。identity
(P3) = DimPlot(All, reduction = "umap", group.
by = "seurat_clusters") # 按照无监督聚类的细胞群
进行可视化。P4 = DimPlot(All, reduction = "umap",
group. by = "orig. ident") # 按照样本进行可视化。
save(All, file = "All. Rdata") # 保存 Seurat 对象。
```

(5) 计算各个细胞群的特征基因, 见图 4: DEG = FindAllMarkers(All, only. pos = T) Head(DEG) avg_log2FC 表示该基因在组合间的差异倍数; cluster 表

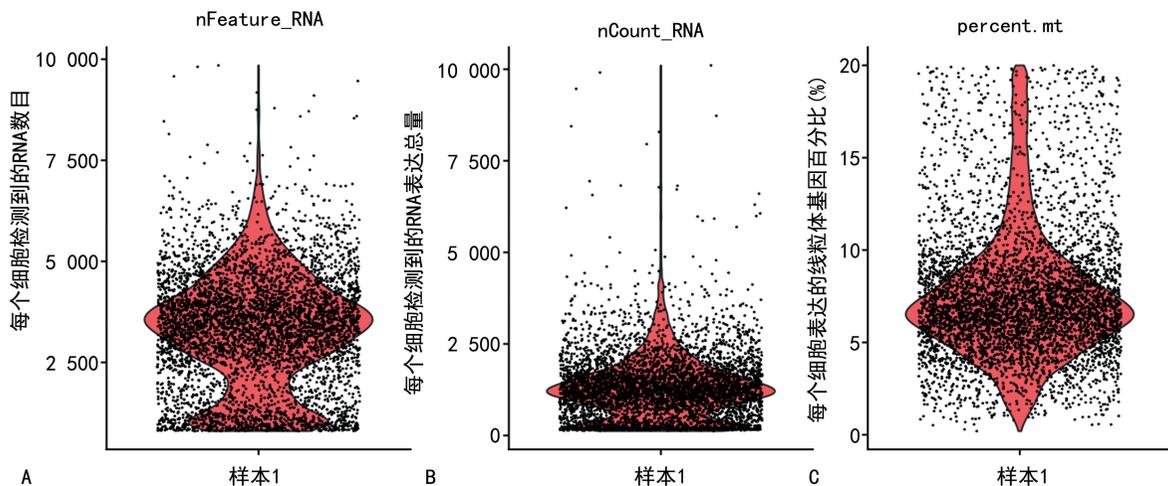
示该基因在对应的细胞群中高表达; pct. 1 表示在当前细胞群中检测到的表达该基因的细胞比例; pct. 2 表示在其他细胞群中检测到的表达该基因的细胞比例; p_val_adj 表示该基因的矫正 P 值; gene 表示基因名称。

(6) 细胞类型鉴定: 根据每个细胞群的特征基因鉴定该细胞群对应的细胞类型, 见表 1。除了这些常见类型, 其余细胞类型需要根据自己分析的样本来查阅文献和数据库, 如 CellMarker 数据库 (<http://xteam.xbio.top/CellMarker/>) 等^[15]。



注: A 为样本 1 中每个细胞检测到的 RNA 数目; B 为样本 1 中每个细胞检测到的 RNA 表达总量; C 为样本 1 中每个细胞表达的线粒体基因百分比。

图 1 样本 1 过滤前的细胞质量

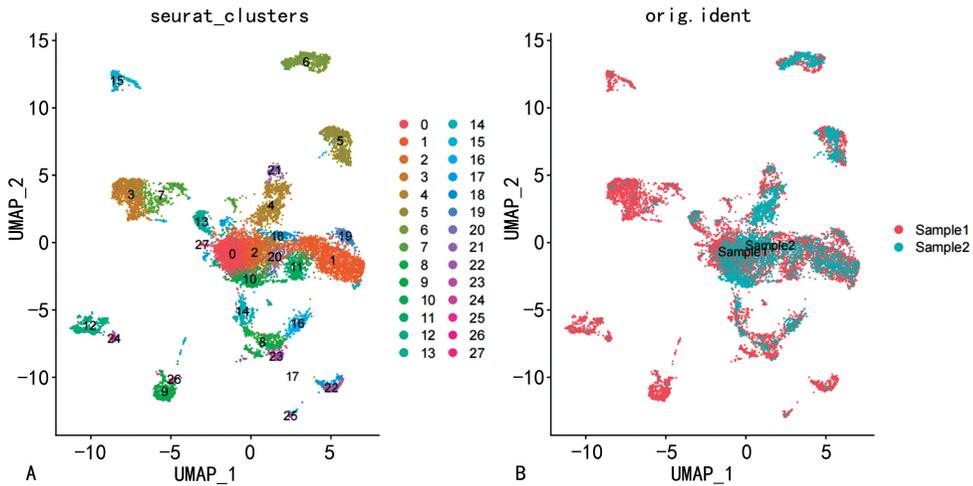


注: A 为样本 1 中每个细胞检测到的 RNA 数目; B 为样本 1 中每个细胞检测到的 RNA 表达总量; C 为样本 1 中每个细胞表达的线粒体基因百分比。

图 2 样本 1 过滤后的细胞质量

(7) 其他分析: 后续分析可以对比不同样本类型中各个细胞类型数目的差异, 寻找特有或稀有细胞类型; 鉴定同一细胞类型不同样本之间的基因表达差异, 如肿瘤组织和正常组织之间的差异特征基因、细胞类型之间的相关关系、同一细胞类型不同亚型之间的发育轨迹及发育过程中的重要基因等。

1.4 学生分组讨论与现场实操 后续分析可让学生分组学习、探索与现场实操, 将遇到的问题反馈给教师, 增强教学互动, 发挥学习者的能动性, 利用学生多元化的学习需求推动优化教学设计。这有利于后续教学中细化需要达成的教学目标, 分步骤实现, 不断提升课程内容和质量, 加强学生的综合能力培养。



注:A 为细胞降维聚类后的类群分布;B 为细胞降维聚类后的样本分布;UMAP 为均匀流形近似和投影,是一种非线性降维和可视化算法。

图 3 两个样本合并后的细胞群分布(左)和样本分布(右)

	p_val	avg_log2FC	pct.1	pct.2	p_val_adj	cluster	gene
Gm42418	0.000000e+00	1.2091845	1.000	0.999	0.000000e+00	0	Gm42418
Cga	1.349824e-232	0.9365945	0.994	0.991	4.191742e-228	0	Cga
Pomc	2.115853e-193	0.7830013	0.896	0.881	6.570570e-189	0	Pomc
AY036118	5.062477e-145	1.2822302	0.885	0.959	1.572102e-140	0	AY036118
Nop58	8.624485e-106	0.2609571	0.174	0.479	2.678247e-101	0	Nop58
Hnrnph1	8.234576e-101	0.2549790	0.314	0.668	2.557165e-96	0	Hnrnph1

图 4 特征基因计算结果示例

表 1 常见细胞类型和分子标记

细胞类型	分子标记
T 细胞	CD3D、CD3E、CD8A
NK 细胞	GNLY、NKG7、FGFBP2、FCG3RA、CX3CR1
B 细胞	CD19、CD79A、MS4A1
单核/巨噬细胞	CD68、CD163、CD14
成纤维细胞	FGF7、MME、DCN、LUM、GSN
内皮细胞	PECAM1、VWF
上皮细胞	EPCAM、KRT19、PROM1、ALDH1A1、CD24

2 结 果

在本次授课中,学生回顾了生理学课程中重要的知识点,即下丘脑-垂体系统。通过文献扩展了解了垂体瘤的现行分类标准、现行分类标准的不足之处,以及国内外相关研究前沿提出的对这些不足的解决方案。在教学内容中引入学科领域的最新成果,更新和完善讲解内容。以国内外最新研究成果为牵引,让学生具备基础医学专业知识的同时培养其国际视野,将人才培养与社会需求相结合。通过这一系列学习不仅扩展了学生在相关问题上的知识储备,而且培养了学生发现问题、分析问题和解决问题的科研能力。

此外,对文献中所用的主要生物信息学方法进行讲解和分析示范,对分析结果进行数据解读。一方面告诉学生什么是生物信息学,生物信息学有什么用处;另一方面也鼓励学生在课堂示范的基础上自主探索和学习,为学生未来从事临床或基础研究工作奠定一定基础。同时,通过课程学习,激发了学生对临床

和科研工作的兴趣,让学生在未來工作中从实际出发,敬业奉献。

3 讨 论

随着生物信息学技术的发展,其与医学临床检验和基础研究的联系越来越紧密。因此,在医学生本科学习阶段开展生物信息学授课对学生综合能力具有重要帮助。但是,在教学设计和实施过程中应注意以下问题:(1)在授课教师的培训方面,该课程要求教师掌握计算机语言应用和生物信息学分析方法;同时,授课教师应具有良好的医学理论基础,关注医学国际前沿研究发展,能够结合学生的医学专业背景和专业知识进行教学实践,同时紧跟国际前沿研究进展、与时俱进。(2)在教学过程中应注意各个环节的时间分配,同时可设计翻转课堂,案例分析,小组讨论、现场实操等环节,例如,①在介绍基础课程知识点时以回顾为主,内容不宜过多;②讲解相关文献时,可由学生分组对教师指定文献进行解读。学生讲解之后,进行互动提问和教师总结;③数据分析讲解与示例部分主要由教师进行讲解和示范,与学生互动提问、现场实操和分组讨论,该部分是主要环节,可分配较长时长。(3)在课程设置上应充分考虑临床医学和基础研究的需要,设置实用的生物信息学课程内容。例如,概括介绍高通量测序的概念、常见技术、应用领域等。在后续具体教学实例中可结合文献进行常用的高通量方法的应用示范,例如,一代测序、全外显子测序、全基因组测序、转录组测序、甲基化测序等。(4)采用“以学生为中心”的教学模式^[2-4],应关注学生在大量互动和自主学习模式下的学习效果。考查学

生对知识的理解和熟悉情况,以及对教学模式的适应情况。与之配套的考核方式应将相当的比重放置在过程性考核上,关注学生在分组中的表现,注重检验学生能力提升情况。量化教学评价,培养德才兼备的高素质创新型人才。(5)在课程设置时应注重挖掘课程的思想政治元素,结合医学基础课程的特点,利用学习思维方法引领价值观,增强职业认同感,使学生更好地肩负起维护和促进人民群众身体健康的重要使命。

通过本文中这一系列课程设置,让学生回顾本专业知识点的同时进一步扩大知识面,了解国际前沿的新发现、新理论。同时培养学生发现问题、分析问题和解决问题的能力。再通过相关生物信息学方法的讲解、分析示范和结果解读,让学生在过程中了解生物信息学分析的一般方法、掌握对分析结果的解读,将临床与科研相结合,提高学生综合能力。

参考文献

- [1] 杨建滨,尚世强.分子生物学技术在出生缺陷三级预防中的临床应用及思考[J].中华预防医学杂志,2021,55(9):1028-1032.
- [2] 杜函芮.以学生为中心构建高校科研育人网络的研究[J].高教探索,2022,38(6):34-39.
- [3] 卢威,梁凤霞,刘建民,等.以学生为中心的教学理念在《针灸学》教学中的应用[J].时珍国医国药,2022,33(10):2519-2520.
- [4] 龚广伟,黄春琳,许可,等.“以学生为中心”的融合式培育实践研究[J].工业和信息化教育,2022,10(11):1-16.
- [5] 王庭槐.生理学[M].9版.北京:人民卫生出版社,2018:78-81.
- [6] FERNANDEZ A, KARAVITAKI N, WASS J A H. Prev-

alence of pituitary adenomas: a community-based, cross-sectional study in Banbury (Oxfordshire, UK) [J]. Clin Endocrinol (Oxf), 2010, 72(3):377-382.

- [7] ZHANG Q, YAO B, LONG X, et al. Single-cell sequencing identifies differentiation-related markers for molecular classification and recurrence prediction of PitNET[J]. Cell Reports Medicine, 2023, 4(2):100934.
- [8] ASA S L, METE O, PERRY A, et al. Overview of the 2022 WHO classification of pituitary tumors[J]. Endocr Pathol, 2022, 33(1):6-26.
- [9] TANG F C, BARBACIORU C, WANG Y Z, et al. mRNA-Seq whole-transcriptome analysis of a single cell[J]. Nat Methods, 2009, 6(5):377-382.
- [10] POTTER S S. Single-cell RNA sequencing for the study of development, physiology and disease [J]. Nat Rev Nephrol, 2018, 14(8):479-492.
- [11] BASLAN T, HICKS J. Unravelling biology and shifting paradigms in cancer with single-cell sequencing[J]. Nat Rev Cancer, 2017, 17(9):557-569.
- [12] ZHENG G X Y, TERRY J M, BELGRADER P, et al. Massively parallel digital transcriptional profiling of single cells[J]. Nat Commun, 2017, 8:14049.
- [13] STUART T, BUTLER A, HOFFMAN P, et al. Comprehensive integration of Single-Cell data[J]. Cell, 2019, 177(7):1888-1902.
- [14] KORSUNSKY I, MILLARD N, FAN J, et al. Fast, sensitive and accurate integration of single-cell data with Harmony[J]. Nat Methods, 2019, 16(12):1289-1296.
- [15] ZHANG X X, LAN Y J, XU J Y, et al. CellMarker: a manually curated resource of cell markers in human and mouse[J]. Nucleic Acids Res, 2019, 47(D1):D721-D728.

(收稿日期:2023-07-10 修回日期:2023-12-12)

教学·管理 DOI:10.3969/j.issn.1672-9455.2024.04.033

多院区 POCT 同质化管理模式

阮 帅¹, 陈保德^{1△}, 金建敏², 徐 剑³

浙江大学医学院附属第一医院:1. 检验科;2. 质量管理部;3. 信息中心, 浙江杭州 311100

摘要:目的 探讨多院区的即时检测(POCT)质量管理体系,实现多院区之间同质化管理。方法 全面了解各院区的 POCT 仪器分布、厂家品牌、开展项目等情况,完善多院区 POCT 质量管理架构及职责,利用信息化手段搭建 POCT 平台打通各院区数据流,规范多院区 POCT 体系文件,统一室内质控、室间质评、院间比对,以及检测人员的培训考核、资格认证等内容。结果 在多院区同质化管理运行半年后,POCT 仪器室内质控执行率已达到 100%;室内质控(血气分析)不精密度(CV)明显降低,且符合目标 CV;院内比对(便携式血糖仪)通过率提高至 100%;室间质评(便携式血糖仪)偏倚结果逐年变小,确保检测结果的准确性。结论 利用信息化解决方案助力 POCT 管理的规范化,解决许多 POCT 质量管理的问题,实现多院区 POCT 统一协调的质量管理。

关键词:即时检测; 多院区; 同质化管理; 质量控制; 质量管理

中图分类号:R194

文献标志码:B

文章编号:1672-9455(2024)04-0572-05

近年来随着医疗体系深入变革,公立医院进入高速发展的时期,许多大型医院通过一院多区扩大规模

△ 通信作者, E-mail: baodeChen@zju.edu.cn.