

徐州地区不同血型单采血小板临床供应量分析预测研究

王梦京 裴培培

徐州市中心血站 江苏 徐州 221000

【摘要】：目的：尝试利用自回归移动平均模型（ARIMA）分析徐州地区各血型单采血小板临床供应量，建立预测模型，为血小板的招募、预约和采集计划提供科学依据。**方法：**收集徐州中心血站 2013-2019 年单采血小板临床供应量，采用 Eviews 8.0 统计分析软件对 2013 年至 2018 年数据建立 ARIMA 模型，并对 2019 年单采血小板临床供应量进行预测。**结果：**A、B、O、AB 模型预测结果平均绝对百分比误差分别 6.4%、7.0%、8.1%、21.4%，A、B、O 三型能较好拟合实测数据，AB 型拟合效果略差。**结论：**ARIMA 模型能较好的预测拟合单采血小板的临床供应量，为各型血小板的招募、预约和采集计划提供科学依据。

【关键词】：徐州地区；ARIMA；血小板供应量；预测

随着医疗技术的迅速发展，血小板的临床用量随之大幅增长，其供应缺口也进一步扩大^[1]。不同血型血小板临床使用的不可预测性，更增加了血小板采集的难度。本文以徐州地区 2013-2018 年单采血小板供应量为数据库，基于自回归移动平均模型（autoregressive integrated moving average model, ARIMA）对 2019 年的单采血小板需求进行预测，进而为中心血站制定单采血小板招募、预约和采集计划提供依据。

1 资料与方法

1.1 数据来源

通过穿越安全输血标准化软件（PAss Spring System）收集徐州中心血站 2013 年 1 月-2019 年 12 月期间每月单采血小板临床供应量，以 1 个治疗量为 1U。

1.2 方法

（1）对 2013 年 1 月至 2018 年 12 月数据做基础描述统计，包括计算统计量与时间趋势图；（2）使用 ADF 单位根检验（Augmented Dickey-Fuller test）、自相关图和偏自相关图，检验序列平稳性和自相关性；（3）通过建立可能的 ARIMA 模型阶数，考察 AIC 和 SC 以及拟合优度 R 方，综合确定 ARIMA 模型的最优阶数，完成最优模型的建立；（4）用最优 ARIMA 模型进行预测，并计算预测的相关统计量，与 2019 年 1 月至 2019 年 12 月数据对比，判断预测效果。

1.3 统计学方法

采用 Excel 汇总整理原始数据，Eviews8.0 统计分析软件建立 ARIMA 模型并进行预测。

2 结果

2.1 描述统计

将 A、B、O、AB 四种血型单采血小板供应量做对数处理为 LA、LB、LO、LAB，并做时间趋势图（图 1），初步识别数据存在季节性特征，具体表现为在一定的时期内（半年），呈现一定的规律变化。

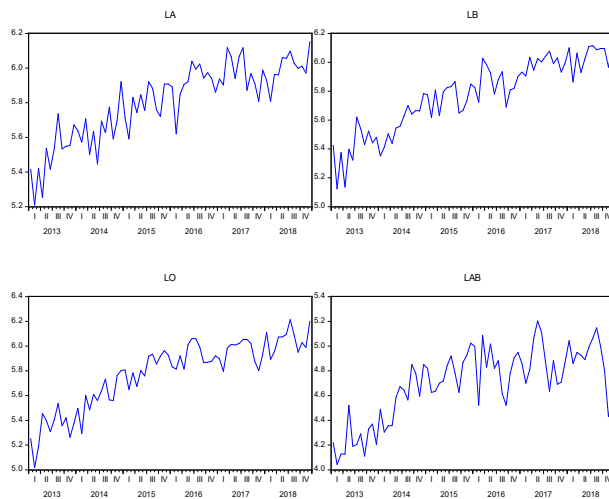


图 1 时间趋势图

2.2 平稳性检验

ADF 单位根检验结果（表 1）显示，一阶差分的 LA、LB、LO 和 LAB 的统计量小于临界值，LA、LB、LO 和 LAB 是一阶平稳的时间序列，因此应该建立 ARIMA(P,1,Q)模型。

表 1 单位根检验

变量	T 统计量	临界值			P 值
		1%	5%	10%	
DLA	-9.153	-2.599	-1.946	-1.614	0.000
DLB	-16.395	-2.598	-1.946	-1.614	0.000
DLA	-11.696	-2.598	-1.946	-1.614	0.000
DLAB	-11.679	-2.598	-1.946	-1.614	0.000

(注: 显著性水平为 10%, D 表示一阶差分)

2.3 自相关检验

对 LA、LB、LO 和 LAB 的原始序列以及一阶差分序列进行自相关检验, 四种血型的自相关图均表现出明显的拖尾性, 偏相关函数则分别在滞后 2 期、2 期、1 期和 1 期截尾; 随后的一阶差分的自相关图和偏相关图检验表现出了一定的季节性特征, 即分别在滞后 6 期、12 期左右。

结合两种检验, 我们确定了对 LA、LB、LO、LAB 建立季节性的 ARIMA 模型。

2.4 季节 ARIMA 模型

对 A 型单采血小板的初始模型中引入季节 SAR(6)、SMA(6) 以及普通 AR 和 MA 因子, 将数据进行半年度季节差分消除季节性, 最终确定了最优的拟合模型 (表 2), 模型形式为 $(6,1,1) \times (0,1,1)_6$ 。Q 统计量对应的 p 值均大于 0.1, 确定了残差为白噪声序列。进而验证了 A 型单采血小板建立的季节 ARIMA 模型是合理的。

表 2 LA 的季节 ARIMA 模型结果

参数	β	T	P
C	-0.001	-1.407	0.165
AR(6)	-0.388	-3.262	0.002
MA(1)	-0.793	-9.183	0.000
SMA(6)	-0.868	-16.506	0.000

B 型、O 型、AB 型识别过程同 A 型一样, 最优模型分别为 $(6,1,2) \times (0,1,1)_6$ 、 $(6,1,2) \times (0,0,0)_6$ 、 $(6,1,2) \times (0,1,1)_6$ 。各模型结果 (表 3) 显示, 其对应参数均有统计学意义。Q 统计量对应的 p 值均大于 0.1, 无统计学意义, 说明模型残差序列为白噪声, 各模型建立合理。

表 3 B、O、AB 型单采血小板季节 ARIMA 模型结果

参数	B			O			AB		
	$(6,1,2) \times (0,1,1)_6$			$(6,1,2) \times (0,0,0)_6$			$(6,1,2) \times (0,1,1)_6$		
	β	T	P	β	T	P	β	T	P
C	-0.001	-5.23	0	-0.002	-2.565	0.013	-0.005	-2.985	0.004
AR(1)	-0.243	-2.313	0.025	-0.197	-2.109	0.04	-0.474	-3.699	0.001
AR(2)	-	-	-	-	-	-	0.347	2.167	0.035
AR(6)	-0.51	-6.591	0	-0.681	-8.275	0	-0.248	-2.258	0.028
MA(1)	-0.501	-3.064	0.003	-0.328	-2.76	0.008	-	-	-
MA(2)	-0.498	-3.102	0.003	-0.614	-5.234	0	-0.793	-5.977	0
SMA(6)	-0.898	-20.193	0	-	-	-	-0.802	-10.126	0

2.5 模型的预测

应用最优模型对 2019 年 A、B、O 和 AB 血型单采血小板供应量进行预测, 并与实际数据进行比对, 结果如图 2。同时选取平均绝对误差 (MAE)、平均绝对百分比误差 (MAPE)

2 个评价指标进行预测效果比较 (表 4)。比较结果显示 A、B 和 O 型血的预测结果与真实数据较为相符。相比 AB 血型的预测效果略微较差, 仅在 2019 年 1 月和 2 月有较高的拟合。

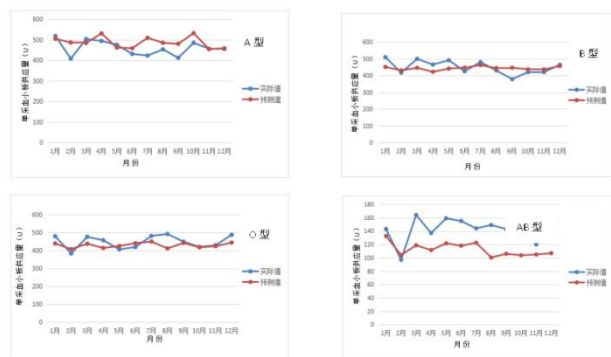


图2 2019年四种血型血小板供应量预测值与实际值对比

表4 预测评价指标

	MAE	MAPE
A 型	29.684	6.4%
B 型	31.637	7.0%
O 型	35.467	8.1%
AB 型	31.847	21.4%

3 讨论

ARIMA 模型是由 Box 和 Jenkins 提出的时间序列预测方法^[2], 广泛应用于经济、建筑、传染病等众多领域。余宇奇等^[3]采用回顾性研究方法, 建立 ARIMA 模型预测临床血液需

求量; 胡燕嫣等^[4]基于该模型对 RhD 阴性红细胞需求量进行预测; 马春会等^[5]等基于该模型对单采血小板临床需求进行预测。文献结果表明, ARIMA 模型能较好地预测短期内血液制品临床需求量。但是尚未见对不同血型单采血小板供应量的预测模型研究, 因此本文借鉴这一方法, 以本血站 2013-2018 年 4 个血型单采血小板每月临床供应量建立 ARIMA 模型, 对 2019 年的临床需求进行预测, 并与实际数据进行对比。2020 年受传染病影响, 该年度数据波动较大, 不稳定, 不具有代表性, 故不纳入统计分析范围。本文收集分析结果显示, A、B、O 型单采血小板的预测结果理想, 其 ARIMA 模型可用于本血站未来半年各型单采血小板的临床供应量预测。但 AB 型预测结果没有表现出很好的时间特性, 预测数值略低于实际数据, 与文章中的预测模型结果不同。分析原因是 AB 型单采血小板每月需求量较少, 对环境因素的影响较敏感, 预测难度较大。可以再积累一段时间的数据, 对 AB 型单采血小板供应量再次进行预测模型建立。

时间序列模型能分析趋势走向、季节周期, 方便我们在招募、预约不同血型的小血小板献血者时, 有所依据地开展工作。但是, 其没有考虑到政府相关政策、突发公共卫生事件等因素影响^[6], 在实际运用当中, 应尽可能多的收集、更新数据, 才能保障预测准确性, 充分发挥预测模型对实际工作的指导作用。

参考文献:

- [1] 孙俊,耿鸿武,吕杭军,等.中国输血行业发展报告[R],2018.
- [2] 叶柱江,刘赴平.时间序列自回归移动平均模型在临床红细胞用量预测中的应用[J].中国输血杂志,2013,26(02):131-134.
- [3] 余宇奇,李登清,邱明,等.2006~2012 年长沙地区用血情况分析 & 趋势预测[J].中国输血杂志,2013,26(05):466-468.
- [4] 胡燕嫣,章怿,范文明.用时间序列模型预测 Rh 阴性红细胞库存需求[J].中国输血杂志,2020,33(03):248-250.
- [5] 马春会,余晋林,温丽玲,郭如华,余卓丽,罗益红.ARIMA 模型在单采血小板临床用量预测中的应用[J].中国输血杂志,2016,29(12):1394-1396.
- [6] 张喜红,李慧,曹文君,等.SARIMA 模型在长治市肺结核预测中的应用[J].中国医科大学学报,2018,47(07):585-588.