

МОРФОЛОГИЧЕСКАЯ ДИФФЕРЕНЦИАЛЬНАЯ ДИАГНОСТИКА ПЕРВИЧНОГО МИЕЛОФИБРОЗА И ЭССЕНЦИАЛЬНОЙ ТРОМБОЦИТЕМИИ С ИСПОЛЬЗОВАНИЕМ КОМПЬЮТЕРНОГО КЛАСТЕРНОГО АНАЛИЗА МЕГАКАРИОЦИТАРНОГО РОСТКА МИЕЛОИДНОЙ ТКАНИ

© 2020 г. З.П. Асауленко*, Л.Б. Полушкина**, А.И. Лепский***, Ю.А. Криволапов*

*Северо-Западный государственный медицинский университет имени И.И. Мечникова МЗ РФ, 191015, Санкт-Петербург, Кирочная ул., 41

**Российский научно-исследовательский институт гематологии и трансфузиологии ФМБА РФ, 191024, Санкт-Петербург, 2-я Советская ул., 16

***Санкт-Петербургский государственный университет», 199034, Санкт-Петербург, Университетская набережная, 7/9

E-mail: Zakhariy@list.ru

Поступила в редакцию 22.11.2019 г.

После доработки 27.01.2020 г.

Принята к публикации 03.02.2020 г.

Изучены особенности гистотопографических характеристик мегакариоцитарного ростка в биопсиях костного мозга больных эссенциальной тромбоцитемией и первичным миелофиброзом в префибротической стадии с мутацией JAK2 или CALR с помощью частного метода машинного обучения без учителя – алгоритма кластеризации DBSCAN. Исследовали 95 биопсий костного мозга больных эссенциальной тромбоцитемией и первичным миелофиброзом. Оценивали следующие характеристики гистотопографии мегакариоцитов: среднее количество мегакариоцитов в одном кластере, среднее количество кластеров и мегакариоцитов на 1 мм² среза. Получена статистически значимая модель логистической регрессии ($\chi^2 = 14.703$, $p = 0.023$, Nagelkerke $R^2 = 19.6\%$). Анализ гистотопографических характеристик мегакариоцитов позволил правильно дифференцировать эссенциальную тромбоцитемию и первичный миелофиброз в 71.6% случаев. Различия гистотопографических характеристик мегакариоцитов в биоптатах костного мозга у больных эссенциальной тромбоцитемией и первичным миелофиброзом с мутацией JAK2 или CALR, выявленные помощью алгоритма кластеризации DBSCAN, позволяют связать нозологическую форму заболевания и особенности строения мегакариоцитарного ростка, а также создать модель логистической регрессии, способную дифференцировать эти болезни.

Ключевые слова: эссенциальная тромбоцитемия, первичный миелофиброз, гистотопография мегакариоцитов, кластерный анализ.

DOI: 10.31857/S0006302920030225

Эссенциальная тромбоцитемия (ЭТ) и первичный миелофиброз (ПМФ) относятся к группе «Ph-негативных» хронических миелопролиферативных новообразований (МПН). Гистологическое исследование биопсий костного мозга относится к большим диагностическим критериям МПН [1]. Чувствительность этого метода исследования в диагностике МПН варьирует от 32,5 до 75%, специфичность – 92–98% в зависимости от

заболевания (считается, что тест имеет достаточную диагностическую точность, если его чувствительность и специфичность превышают 80%) [2].

Гистотопографические и морфологические аномалии мегакариоцитов важны для дифференциальной диагностики ЭТ и ПМФ. При ПМФ различимы компактные группы и скопления мегакариоцитов в виде тесно сомкнутых клеток (плотные кластеры), которые могут содержать более десяти мегакариоцитов. Характерно расположение мегакариоцитов и их скоплений вокруг расширенных кровеносных синусоидов и возле поверхности костных балок. Ядра мегакариоцитов часто имеют гиполобулярное строение (обла-

Сокращения: ЭТ – эссенциальная тромбоцитемия, ПМФ – первичный миелофиброз, МПН – миелопролиферативные новообразования, префибрПМФ – первичный миелофиброз в префибротической стадии.

ковидные ядра). При ЭТ скопления мегакариоцитов более разреженные, клетки не соприкасаются друг с другом (рыхлые кластеры), мегакариоциты распределены в костном мозге более однородно, их тропность к поверхностям костных балок и кровеносным синусам выражена слабее (рис. 1). Ядра при ЭТ гиперплоидные, дольчатые, напоминают венчик или, как их называют в англоязычных руководствах, — «staghorn»-like («ядра типа оленьих рогов»).

Несмотря на возросшую роль гистологического исследования биопсий костного мозга в постановке диагноза МПН, данные о воспроизводимости этого критерия диагностики противоречивы. В ряде публикаций представлены данные о совпадении диагнозов МПН в 76–88% случаев [3–6]. Другие работы свидетельствуют о более низкой воспроизводимости результатов гистологического исследования костного мозга — совпадение диагнозов в 62–70% случаев. Наибольшую трудность представляет дифференциальная диагностика между ЭТ и ПМФ в префибротической стадии (префибрПМФ) — совпадение морфологических диагнозов наблюдается лишь у 53% больных [7–10]. Авторы работы [11] показали, что при гистологическом исследовании биопсий костного мозга больных МПН мнение трех из четырех врачей-патологоанатомов о формировании плотных кластеров мегакариоцитов совпало в 88% случаев, согласованность мнения трех специалистов из четырех о наличии рыхлых кластеров мегакариоцитов или их отсутствии наблюдалось в 71 и 73% случаев соответственно. Оценка плотности кластеризации мегакариоцитов субъективна и варьирует даже у врачей, специализирующихся в гематопатологии: каппа Коэна = 0.21 (граница между плохой и удовлетворительной степенью согласованности) [2].

Метаанализ, проведенный авторами работы [12], продемонстрировал, что у больных ЭТ частота обнаружения мутаций JAK2 и CALR составляет 31.3–72.1% и 12.6–50% соответственно, у пациентов с ПМФ мутацию JAK2 можно обнаружить в 25.0–85.7% случаев, мутацию CALR — у 10–100% больных. Доказана роль драйверных мутаций JAK2 и CALR как прогностических и предиктивных биомаркеров у больных ЭТ и ПМФ, но влияние мутационного статуса на пространственные характеристики мегакариоцитарного роста в костном мозге изучено мало [13].

Для улучшения качества диагностики ЭТ и префибрПМФ при гистологическом исследовании биопсий костного мозга и изучении связи мутационного профиля с особенностями строения мегакариоцитарного роста может быть полезно внедрение алгоритмов машинного обучения, способных распознавать морфологические и

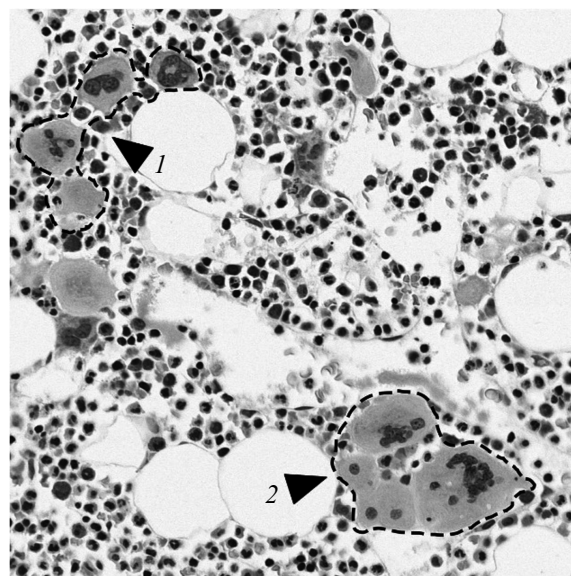


Рис. 1. Эссенциальная тромбоцитемия. Стрелками отмечены рыхлый (1) и плотный (2) кластеры мегакариоцитов.

гистотопографические признаки, недоступные человеческому глазу.

Цель исследования — найти отличия гистотопографических характеристик мегакариоцитарного роста в биопсиях костного мозга больных ЭТ и ПМФ в префибротической стадии с мутацией JAK2 или CALR с помощью частного метода машинного обучения без учителя — алгоритма кластеризации DBSCAN (Density-Based Spatial Clustering of Applications with Noise).

МАТЕРИАЛЫ И МЕТОДЫ

В патологоанатомическом отделении клинической молекулярной морфологии СЗГМУ им. И.И. Мечникова МЗ РФ в период 2012–2018 гг. были отобраны для исследования 95 биопсий костного мозга больных «Ph-» МПН с выявленной мутацией JAK2 или CALR. Распределение по группам было следующим: ЭТ с мутацией JAK2 — 30 больных, ЭТ с мутацией CALR — 30 больных, префибрПМФ с мутацией JAK2 — 25 больных и префибрПМФ с мутацией CALR — 10 больных. У всех пациентов определяли JAK2V617F-статус с помощью метода полиморфизма длин рестрикционных фрагментов. Наличие мутации гена CALR проводили методом прямого секвенирования по Сенгеру на автоматической капиллярной системе MegaBACE 1000 DNA Analysis System (Amersham Biosciences, Великобритания). Все молекулярно-генетические исследования были выполнены в лаборатории молекулярной генетики РосНИИГТ ФМБА России.

Таблица 1. Гистотопографические характеристики мегакариоцитов у больных эссенциальной тромбоцитемией и первичным миелофиброзом в префибротической стадии с мутациями JAK2 или CALR

Параметр	ЭТ JAK2	ЭТ CALR	ПМФ JAK2	ПМФ CALR
Количество мегакариоцитов в кластере	5.9 ± 2.4	7.2 ± 2.9	8.0 ± 4.3	7.4 ± 1.5
Мегакариоциты, мм ²	21.7 ± 12.7	20.2 ± 9.8	25.6 ± 16.0	17.9 ± 6.7
Кластеры, мм ²	2.66 ± 1.02	2.09 ± 0.96	2.34 ± 1.09	1.87 ± 0.72

Таблица 2. Переменные, включенные в модель логистической регрессии

Предиктор	<i>B</i>	Стандартная ошибка	Вальд	<i>p</i>	Exp (<i>B</i>) (95% ДИ)
Количество мегакариоцитов в кластере	0.350	0.212	2.723	0.099	1.420 (0.936–2.152)
Мегакариоциты, мм ²	–0.091	0.113	0.658	0.417	0.913 (0.732–1.138)
Кластеры, мм ²	0.365	1.056	0.120	0.729	1.44 (0.18–11.40)
Мутация	–1.588	0.583	7.420	0.006	0.204 (0.065–0.640)

Гистологические срезы костного мозга, окрашенные гематоксилином и эозином или азуром и эозином, с помощью сканера Pannoramic 250 Flash III (3DHISTECH, Венгрия) переводили в цифровой формат. В программе Pannoramic Viewer определяли координаты каждого мегакариоцита. Координаты всех мегакариоцитов каждого среза экспортировали в MS Excel и конвертировали в txt-формат. Обработка координат с целью анализа особенностей расположения мегакариоцитов в костном мозге была выполнена на языке программирования Python с помощью алгоритма кластеризации DBSCAN [15]. В биопсиях костного мозга оценивали следующие характеристики: среднее количество мегакариоцитов в одном кластере, среднее количество мегакариоцитарных кластеров и мегакариоцитов на 1 мм² среза. Для анализа использовали следующие параметры алгоритма: минимальное количество соседних мегакариоцитов, необходимых для образования кластера – 3, максимальное расстояние между ними – 100 мкм, расстояние между мегакариоцитами – евклидово. Для описания связи гистотопографических характеристик мегакариоцитов и наличия мутации JAK2 или CALR с диагнозом использовали многофакторный дисперсионный анализ и логистический регрессионный анализ. Проведение многофакторного дисперсионного анализа, построение логистической регрессии и кривой ошибок были выполнены с использованием пакета R.

РЕЗУЛЬТАТЫ

Морфометрические характеристики гистотопографии мегакариоцитов в биопсиях костного мозга у больных ЭТ и префибрПМФ с мутациями JAK2 или CALR представлены в табл. 1.

Выявлена статистически значимая связь между заболеванием, мутацией и гистотопографическими характеристиками мегакариоцитов, $F = 3,185$, $p = 0.001$; Wilks' $\Lambda = 0.739$.

Модель логистической регрессии на основе кластеризации и мутационного статуса была статистически значима ($\chi^2 = 14,703$, $p = 0.023$, Nagelkerke $R^2 = 19,6\%$). Коэффициент при свободном члене равнялся $-2,19$ (1,19). Включенные переменные представлены в табл. 2.

Качество дифференциальной диагностики ЭТ и префибрПМФ с использованием логистической регрессии отражено в табл. 3.

Как видно из табл. 3, процент корректных предсказаний диагноза ЭТ значительно выше, чем процент корректных предсказаний префибрПМФ. Вероятность правильного диагноза составляет 71,6%. Обоснованность модели проверена на тестовой и контрольной выборке путем построения модели логистической регрессии на 75 и 25% исходных данных.

Для демонстрации зависимости количества верно классифицированных диагнозов одного заболевания от количества неверно классифициро-

Таблица 3. Процент корректных предсказаний диагноза заболевания на основе переменных, включенных в модель логистической регрессии

Фактическое заболевание	Процент корректных предсказаний		Процент корректных предсказаний
	ЭТ	ПМФ	
ЭТ	54	6	90.0
ПМФ	21	14	40.0
Общий процент			71.6

ванных диагнозов другого заболевания построена кривая ошибок (операционная характеристика приемника). Площадь под кривой ошибок = 72,5 (95% ДИ 61,8–83,1) свидетельствует о хорошем качестве модели (рис. 2).

ОБСУЖДЕНИЕ

Полученные результаты подтверждают различия в расположении мегакариоцитов у пациентов с ЭТ и префибрПМФ с мутациями JAK2 или CALR, которые не воспринимаются глазом при микроскопическом исследовании и могут быть обнаружены лишь с применением алгоритмов машинного обучения. Выявленная статистически значимая связь между заболеванием, мутационным статусом и гистотопографическими характеристиками мегакариоцитов (табл. 1) демонстрирует влияние мутаций JAK2 или CALR на пространственные характеристики мегакариоцитарного ростка в костном мозге. Результаты логистической регрессии свидетельствуют о большей связи мутационного статуса и пространственных особенностей расположения мегакариоцитов у больных ЭТ, чем у пациентов с префибрПМФ (табл. 3). Использование алгоритма кластеризации DBSCAN для оценки гистотопографии мегакариоцитов позволяет оценить плотность кластеризации, размеры кластеров мегакариоцитов и выраженность гиперплазии мегакариоцитарного ростка. Методика кластеризации заключается в группировке объектов, расположенных близко друг к другу, и определении областей с малой плотностью расположения точечных объектов как шум. В методике DBSCAN используются следующие понятия [14, 15]:

- ϵ – окрестность объекта x ;
- V – множество всех наблюдений;
- $U(x, \epsilon) = \{y \in V: p(x, y) \leq \epsilon\}$.

Корневой объект степени M (для заданного ϵ) – объект, ϵ -окрестность которого содержит не менее M других объектов. При заданном значении M объект y плотно-достижим из объекта x , если $y \in U(x, \epsilon)$ и объект x является корневым.

Объект y плотно-достижим из объекта x , если существуют такие объекты x_1, \dots, x_n , где $x_1 = x$,

$x_n = y$, что при всех $i = 1, \dots, n-1$ объект x_{i+1} непосредственно плотно-достижим из x_i .

«Шум» – объекты, не достижимые ни из одного другого объекта.

Число кластеров K алгоритм DBSCAN определяет в процессе работы.

Принцип работы алгоритма DBSCAN:

1. Задаются значения параметров ϵ и M .
2. Если все объекты $x \in V$ уже просмотрены, завершить выполнение алгоритма. В противном случае выбирается любой из них и отмечается как просмотренный.
3. Если x – корневой объект, создается новый кластер ($K := K + 1$), следует переход к пункту 4; в противном случае точка x помечается как «шум», следует переход к шагу 1.
4. В созданный кластер включаются все объекты, которые являются плотно-достижимыми из (корневого) объекта x , следует переход к шагу 2.

В этой работе точечные объекты представляют собой экспортированные координаты мегакариоцитов, определенные при исследовании от-

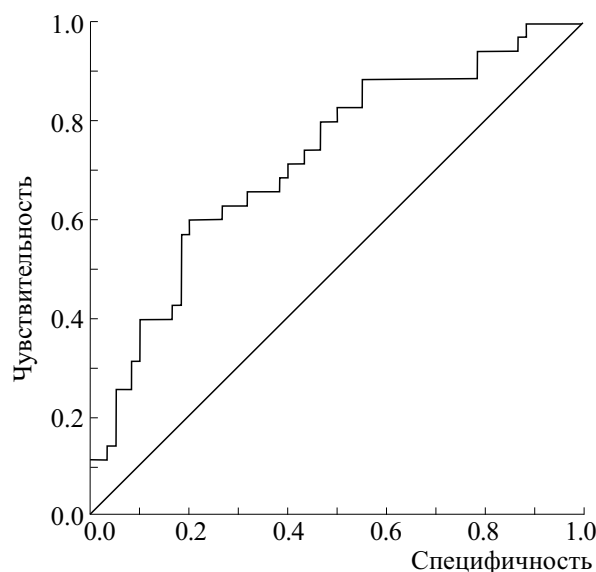


Рис. 2. Кривая ошибок (операционная характеристика приемника).

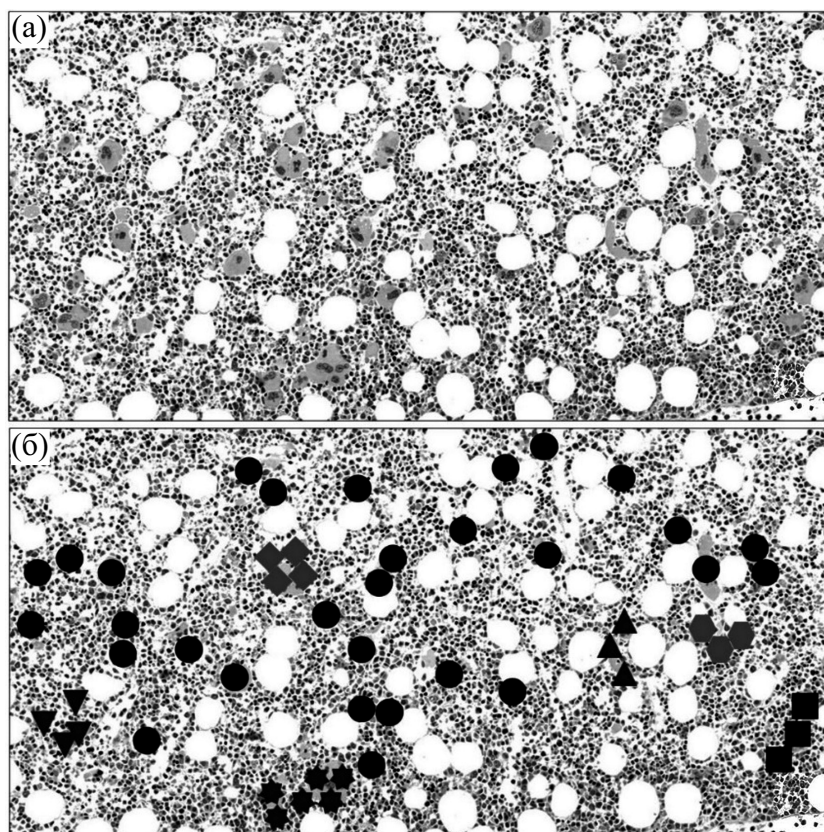


Рис. 3. Анализ кластеризации мегакариоцитов в биопсиях костного мозга. (а) – Фрагмент отсканированного гистологического препарата костного мозга больного эссенциальной тромбоцитемией. (б) – Мегакариоциты формируют отдельные кластеры, отмеченные разными геометрическими фигурами. Мегакариоциты, классифицированные как «шум», отмечены круглыми аннотациями.

сканированных гистологических препаратов костного мозга (рис. 3).

ЗАКЛЮЧЕНИЕ

Пространственные отличия мегакариоцитарного роста в биопсиях костного мозга больных эссенциальной тромбоцитемией и первичным миелофиброзом в префибротической стадии с подтвержденной мутацией JAK2 или CALR, установленные с помощью частного метода машинного обучения без учителя – алгоритма кластеризации DBSCAN, показывают целесообразность применения элементов искусственного интеллекта в научной и практической деятельности патолого-анатомических отделений медицинских организаций.

УЧАСТИЕ АВТОРОВ

Концепция и дизайн исследования – З.П.А. и Ю.А.К.; сбор и обработка материала – З.П.А., Л.Б.П. и А.И.Л.; написание текста – З.П.А. и Ю.А.К., редактирование – Ю.А.К.

КОНФЛИКТ ИНТЕРЕСОВ

Авторы заявляют об отсутствии конфликта интересов.

СОБЛЮДЕНИЕ ЭТИЧЕСКИХ СТАНДАРТОВ

Все процедуры, выполненные в исследовании с участием людей, соответствовали этическим стандартам Хельсинкской декларации 1964 г. и ее последующим изменениям. От участников исследования было получено информированное добровольное согласие.

СПИСОК ЛИТЕРАТУРЫ

1. D. A. Arber, A. Orazi, R. Hasserjian, et al., *Blood* **127** (20), 2391 (2016). DOI: 10.1182/blood-2016-03-643544
2. A. Alvarez-Larrán, A. Ancochea, M. García, et al., *Brit. J. Haematol.* **166** (6), 911 (2014). DOI: 10.1111/bjh.12990
3. U. Gianelli, A. Bossi, I. Cortinovis, et al., *Modern Pathol.* **27**, 814 (2014). DOI: 10.1038/mod-pathol.2013.196

4. T. Barbui, J. Thiele, F. Passamonti, et al., *J. Clin. Oncol.* **29**, 3179 (2011). DOI: 10.1200/JCO.2010.34.5298
5. J. Thiele, H. M. Kvasnicka, L. Müllauer, et al., *Blood* **117**, 5710 (2011). DOI: 10.1182/blood-2010-07-293761
6. A. B. Madelung, H. Bondo, I. Stamp, et al., *Am. J. Hematol.* **88**, 1012 (2013). DOI: 10.1002/ajh.23554
7. B. S. Wilkins, W. N. Erber, D. Bareford, et al., *Blood* **111**, 60 (2008). DOI: 10.1182/blood-2007-05-091850
8. M. Brousseau, E. Parot-Schinkel, M. P. Moles, et al., *Histopathology* **56**, 758 (2010). DOI: 10.1111/j.1365-2559.2010.03545.x
9. T. Buhr, K. Hebeda, V. Kaloutsi, et al., *Haematologica* **97**, 360 (2012). DOI: 10.3324/haematol.2011.047811
10. B. S. Wilkins, W. N. Erber, D. Bareford, et al., *Blood* **111** (1), 60 (2007). DOI: 10.1182/blood-2007-05-091850
11. S. M. Koopmans, F. J. Bot, K. H. Lam, et al., *Am. J. Clin. Pathol.* **136** (4), 618 (2011). DOI: 10.1309/ajcp2ug9sggwahua
12. M. Mejia-Ochoa, P. A. Acevedo Toro, and J. A. Cardona-Arias, *BMC Cancer* **19** (1), 590 (2019). DOI: 10.1186/s12885-019-5764-4
13. T. R. Merlinsky, R. L. Levine, and E. Pronier, *Clin. Cancer Res.* **25** (10), 2956 (2019). DOI: 10.1158/1078-0432.CCR-18-3777
14. J. Sander, M. Ester, H. P. Kriegel, and X. Xu, *Data Mining Knowledge Discov.* **2**, 169 (1998). DOI: 10.1023/A:1009745219419
15. I. V. Konnov, O. A. Kashina, and E. I. Gilmanova, *Uchenye Zapiski Kazanskogo Universiteta. Seriya Fiziko-Matematicheskie Nauki*, **161** (3), 423 (2019). DOI: 10.26907/2541-7746.2019.3.423-437 (In Russian)

Morphological Differential Diagnosis of Primary Myelofibrosis and Essential Thrombocythemia with Computer Cluster Analysis of Megakaryocytic Lineage in Myeloid Tissue

Z.P. Asaulenko*, L.B. Polushkina**, A.I. Lepsky***, and Yu.A. Krivolapov*

*North-Western State Medical University named after I.I. Mechnikov, Kirochnaya ul. 41, St. Petersburg, 191015 Russia

**Russian Research Institute of Hematology and Transfusiology, 2-ya Sovetskaya ul. 16, St. Petersburg, 191024 Russia

***St. Petersburg State University, Universitetskaya nab. 7/9, St. Petersburg, 199034 Russia

DBSCAN clustering algorithm – unsupervised machine learning algorithm was used to explore particularities of the histotopographic features of the megakaryocytic lineage in the bone marrow biopsies of patients with JAK2+ or CALR-mutated essential thrombocythemia and primary myelofibrosis in prefibrotic stage. We examined 95 bone marrow biopsies of patients with essential thrombocythemia and primary myelofibrosis. We evaluated the following histotopographical features of megakaryocytes: the mean quantity of megakaryocytes in one cluster, the mean number of clusters and megakaryocytes per 1 mm² area in section. The logistic regression model was statistically significant, ($\chi^2 = 14.703$, $p = 0.023$, Nagelkerke $R^2 = 19.6\%$). Analysis of histotopographical features of megakaryocytes helped to differentiate correctly between essential thrombocythemia and primary myelofibrosis in 71.6% cases. The differences revealed by DBSCAN clustering algorithm in histotopographical features of megakaryocytes in biopsies of the bone marrow derived from patients with JAK2+ or CALR-mutated essential thrombocythemia and primary myelofibrosis suggest the relationship between the nozologic form of the disease and particularities of the development of megakaryocytic lineage and are also useful to create a logistic regression model for differentiating these diseases.

Keywords: essential thrombocythemia, primary myelofibrosis, histotopography of megakaryocytes, cluster analysis