

ОПИСАНИЕ КЪМ СВИДЕТЕЛСТВО
ЗА РЕГИСТРАЦИЯ
НА ПОЛЕЗЕН МОДЕЛ

ПАТЕНТНО ВЕДОМСТВО

(21) Заявителски № 6266

(22) Заявено на 22.07.2024

(24) Начало на действие

на регистрацията от: 19.07.2024

Приоритетни данни

(45) Отпечатано на 15.11.2024

(46) Публикувано в

бюлетин № 202411.1 на 15.11.2024

(56) Информационни източници:

(62) Разделена заявка от рег. №

(66) Трансформирано от:

(67) Паралелно на: 113931, 19.07.2024

(73) Притежател(и):

**МЕДИЦИНСКИ УНИВЕРСИТЕТ – ПЛОВДИВ,
4002 ПЛОВДИВ, УЛ. "ВАСИЛ АПРИЛОВ" 15А**

(72) Изобретател(и):

**Александра Марио Балджиева
Христо Борисов Тасков
Хасан Алиев Бурнусов
Теодора Миткова Калфова
Марияна Атанасова Мурджева
Маргарита Любенова Генова**

(74) Представител по индустриална собственост:

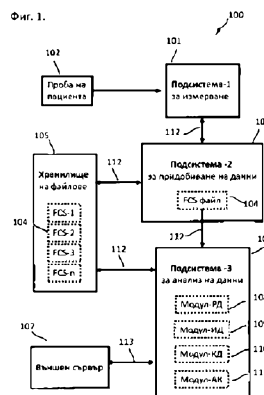
(86) № на РСТ заявка:

(87) № и дата на РСТ публикация:

**(54) СИСТЕМА ЗА ИЗМЕРВАНЕ И АНАЛИЗ
НА МИНИМАЛНА РЕЗИДУАЛНА БОЛЕСТ
ПРИ ДЕТСКА В-ПРЕКУРСОРНА ОСТРА
ЛИМФОБЛАСТНА ЛЕВКЕМИЯ ЧРЕЗ
МНОГОПАРАМЕТЪРНА
ФЛОУЦИТОМЕТРИЯ**

(57) Ползният модел се отнася до система за измерване на минимална резидуална болест (МРБ) при детска В-прекурсорна остра лимфобластна левкемия (дВ-ОЛЛ) чрез многопараметърна флоуцитометрия (МПФЦ), която ще намери приложение в здравеопазването и медицината за измерване на МРБ при дВ-ОЛЛ с висока чувствителност и специфичност, необходими за оценка на риска, мониториране на заболяването и определяне на режима на лечение. Системата за измерване на МРБ при дВ-ОЛЛ включва: подсистема-1 за измерване на пробата на пациента, подсистема-2 за приемане и обработване на данните и подсистема-3 за анализ на данните. Предимствата, които осигурява ползният модел, са високата чувствителност от 10-5 и висока специфичност,

и че за анализа не се изисква използването на референтни проби и/или машинно обучение, което го прави приложим в лаборатории с различна измервателна апаратура и използващи различни панели от антитела за идентифициране на левкемичните клетки.

2 претенции, 10 фигури

(54) СИСТЕМА ЗА ИЗМЕРВАНЕ И АНАЛИЗ НА МИНИМАЛНА РЕЗИДУАЛНА БОЛЕСТ ПРИ ДЕТСКА В-ПРЕКУРСОРНА ОСТРА ЛИМФОБЛАСТНА ЛЕВКЕМИЯ ЧРЕЗ МНОГОПАРАМЕТЪРНА ФЛОУЦИТОМЕТРИЯ

Област на техниката

Този полезен модел се отнася до система за измерване на минимална резидуална болест при детска В-прекурсорна остра лимфобластна левкемия. Този полезен модел разкрива система за установяване на малки количества аномални клетки в изследвана проба от клетки, използвайки многомерен автоматичен клъстерен анализ на клетъчни характеристики, измерени с помощта на многопараметърна флоуцитометрия. Системата може да намери приложение за оценка на риска, мониторинга и определяне на лечението при детска В-клетъчна остра лимфобластна левкемия, обект на изследване в педиатрията, хематологията и клиничната имунология. Но също така може да се използва и в други области на биологията и медицината за откриване на редки аномални клетки с помощта на многопараметърна флоуцитометрия.

Предшестващо състояние на техниката

Острата лимфобластна левкемия (ОЛЛ) е най-честият педиатричен рак, представляващ приблизително 25% от диагностицираните злокачествени заболявания. Около 60% от всички случаи се срещат при деца и юноши под 20 години, с годишна честота от >90 случая на 1 милион души. В детската възраст най-често срещана е В-клетъчната прекурсорна остра лимфобластна левкемия (дВКП-ОЛЛ). Настоящите протоколи за цитостатично или цитотоксично лечение не са в състояние да унищожат всички злокачествени клетки във всички случаи. При някои пациенти остава малко количество левкемични клетки в костния мозък, които водят до рецидив на заболяването. Тези клетки се наричат минимална *резидуална (остатъчна) болест* (МРБ). Понастоящем измерването на количеството на клетките на МРБ е най-мощният прогностичен индикатор на преживяемостта при ОЛЛ в детска възраст.

Идентифицирането на левкемичните клетки на МРБ става най-често по фенотипни белези, наречени "клетъчни маркери" или "CD-молекули". Обикновено фенотипът на една клетка се описва от съвкупността на 10-20 маркера. Доказването на тези маркери става чрез антитела, които се свързват специфично с тях. Чрез антителата, конюгирани с различни флуорохроми и по тяхната флуоресценция, тези маркери могат да бъдат открити.

За измерването на клетките на МРБ се изисква високо чувствителен метод, който да открива между 1 на 10,000 и 1 на 100,000 клетки (0.01% - 0.001%). Понастоящем такъв метод е многопараметърната флоуцитометрия. При флоуцитометрията една след друга клетките пресичат лазерен лъч типично със скорост 5,000 - 10,000 клетки в секунда. При осветяването клетката разсейва светлина напред (FSC) и настрани (SSC), както и излъчва флуоресценция от всеки неин маркер, свързан със съответното антитяло, което е конюгирано с различни флуорохроми. Светлината от FSC и SSC, както и всеки цвят на флуоресценциите, се улавят от отделни системи от филтри и детектори, които се наричат "канални" или

"параметри". Съвременните цитометри разполагат с между 12 и 50 параметъра. Процесът на събиране на данни от проби с помощта на проточен цитометър се нарича "придобиване".

Сигналят от клетката, уловен от всеки от детекторите, се преобразува в дигитално "събитие", което се записва във файл със стандартизиран формат FCS 3.0, Flow Cytometry format, разработен от Международно дружество за усъвършенстване на цитометрията (Flow Cytometry Standard на International Society for Advancement of Cytometry - ISAC), и който може да бъде прочетен от различни софтуери за анализ на флоуцитометрични данни. Измерените клетки се представят под форма на събитие и в този полезен модел понятията "клетка" и "събитие" са взаимнозаменяеми.

При **класическия анализ** на флоуцитометричните данни събитията се представят графично под формата на "хистограма" и двумерни диаграми, наречени "бивариатен плот" или "би-плот". На хистограмата се представят събитията от един параметър, разпределени по сила на флуоресценцията (X) и брой на събитията (Y). На би-плота събитията се представят най-често като отделни точки с координати - силата на флуоресценцията по два параметъра. Заради това би-плотът се нарича още "точков плот". На би-плота събитията (клетките) със сходни координати могат да се групират в "облаци" или "кълъстери", като в този полезен модел ще се употребява понятието "кълъстер", които представляват една клетъчна популация. Такава клетъчна популация може да се загради чрез електронен прозорец, който се нарича "гейт". Чрез този гейт може да се получи разнообразна количествена информация за клетъчната популация или последната да се селектира за последващ анализ чрез хистограми или бивариатни плотове, като параметрите се комбинират всеки срещу всеки. За откриване на МРБ е необходимо да се анализират множество плотове, за да се достигне до кълъстера на клетките на МРБ. Броят на бивариатни плотове, които могат да бъдат генерирани за цитометричен протокол с N флуорохроми, е $(N \times (N-1))/2$. Така например при 6 параметъра трябва да се анализират 15 плота. С увеличаване на броя на параметрите броят на плотовете за анализ расте в геометрична прогресия. Например при 16 параметъра ще трябва да се анализират 120 плота. Расте и вероятният брой на кълъстерите. Например при 10 параметъра възможният брой на кълъстерите е 1024, а при 16 параметъра - 65,536.

Подходът за анализ със серия от гейтове върху множество плотове, въпреки че е често ефективен, отнема много време с място за субективност при множеството стъпки, при които на базата на специфични комбинации от маркери се вземат решения за изваждане на клетки с нормална експресия и запазване на клетки с аберантна експресия. Второ, докато анализираме повече бивариатни плотове, констатациите стават по-фрагментирани и по-трудни за тълкуване. И накрая, което е по-важно, оценката на отделните клетъчни кълъстери не дава никаква представа за връзките между тях.

Друг недостатък на класическия метод за измерване на МРБ е свързан с необходимостта да се открият малък брой клетки (поне 40 клетки сред 4×10^6 клетки). На точковите плотове се представя ограничена извадка от всички клетки, което крие риска малки популации клетки (0.01% - 0.001%) да не бъдат представени и следователно да не могат да бъдат анализирани.

Поради изтъкнатите по-горе недостатъци, класическият анализ е трудно приложим при флоуцитометрични изследвания с повече от 8 параметъра.

За измерването на МРБ за откриване на миниатюрна бластна популация е необходимо определянето на много клетъчни маркери и за целта понастоящем се използва от 10 до 16-параметърна флоуцитометрия, като тенденцията е броят на параметрите да се увеличава. Това от своя страна затруднява анализа на данните по класическия подход. Ето защо се търсят други подходи и системи, които позволяват анализ на големи количества данни от многопараметърна флоуцитометрия (МПФЦ). Компютърно-подпомогнат автоматизиран анализ и техники за машинно обучение в появиха през последните години, предоставяйки нови начини за анализиране на все по-сложни данни от флоуцитометрията [1, 2]. В допълнение на това се прилагат подходите за неконтролирано и контролирано машинно обучение.

В патентен документ EP 2347352 B1 [3], публ. на 6.11.2019 г. авторите правят опит да преодолеят недостатъците на класическия подход за анализ на десетки плотове при данни от МПФЦ. Те разкриват система и метод за разпределянето на клетъчните популации в йерархично дърво и по този начин тяхното по-лесно идентифициране. Изобретението включват методи и системи, които позволяват на изследователя да изследва повече от 512 фенотипа (т.е. 9 или повече флуорохрома) с един интерактивна дървовидна графика. Системите позволяват на изследователя да използва множество графики, всяка със съответни маркери, при което графиките са свързани помежду си. Субпопулации от данни, изобразени в една графика, могат да се селектират и така да се използват като гейтове за друга графика, за да се видят нагоре до 1024 фенотипове.

Създаването на йерархични графики става чрез мануално поставяне на гейтове на хистограми на бивариатни плотове на базата на експерта оценка на клетъчните субпопулации от оператора, както е при конвенционалния анализ на флоуцитометрични данни. Броят на клъстерите зависи от броя на мануално поставените гейтове, което в значителна степен зависи от субективната оценка на оператора. Чрез интерактивни действия операторът може да борави с клъстерите. Предимството на системата е, че представя количествено клъстерите и съответните клетъчни субпопулации на една графика, което улеснява тяхното визуализиране. Липсва обаче автоматизиране на процеса на разделянето на отделни клъстери. Това изобретение цели да улесни и автоматизира класическия анализ, но като и при последния, то е подходящо за откриване и анализ на големи популации клетки над 1%, но е неподходящо за откриване на малки популации клетки под 0.01%. На хистограмите и плотовете, където мануално се поставят гейтовете, тези малки популации трудно могат да бъдат забелязани от оператора.

В нашия полезен модел процесът на клъстерифицирането е напълно автоматизиран и не зависи от субективната оценка на оператора, което позволява да се открият популации клетки, съставляващи 0.01% и дори 0.001% от всички измерени клетки.

В патентен документ US 2003078703 A1 [4], публ. на 24.04.2003 г. е разкрита цитометрична мрежа, включваща база данни, позволяваща анализ и преглед на големи количества цитометрични файлове от множество цитометрии. Данните от множество инструменти се качват в база данни, където комбинираните матрици за компенсация и калибриране се прилагат автоматично, за да коригират както спектралното припокриване между каналите за придобиване, така и зависещите от инструмента вариации. Върху данни автоматично се поставят гейтове с помощта на общи шаблони. Потребител на клиентски компютър може да преглежда и извършва статистически анализи на всяко желано подмножество от данни. Потребител от

клиентски компютър може да преглежда и извършва статистически анализ на получените и обработени данни.

В нашия полезен модел флоуцитометърът е свързан с база данни, която служи само за съхранение на файловете, и която е спомагателен елемент от полезния модел. Същността на нашия полезен модел се състои в система, способна да анализира многомерни данни с цел да се открият малки популации клетки на МРБ.

В патентен документ CN 103942415 A [4], публ. на 23.07.2014 г. е разкрита система и метод за автоматичен анализ на данни от поточен цитометър. Системата за автоматичен анализ на данни на поточния цитометър включва следните стъпки: автоматична идентификация на клъстери, след получаване броя на клъстерите с данни, данните се групират автоматично. Тя позволява извършването на автоматичен и бърз анализ на данните от поточния цитометър чрез компютърен софтуер.

Изобретението разкрива метод за автоматичен анализ на данни на поточен цитометър. Методът за автоматичен анализ на данни на поточния цитометър включва следните две стъпки: (1) автоматичната идентификация на клъстери се извършва върху данни чрез VIC метод; (2) след като се получи броят на клъстерите, съдържащи се в данните, се извършва автоматично групиране на данните чрез t модел на смесване.

Едно от съществените ограничения на описаното изобретение е, че не е демонстрирано приложението му за идентифициране на малки клетъчни субпопулации, както и приложението му за измерване на МРБ. В нашия полезен модел клъстерифицирането е само една стъпка от функционирането на системата. Последната включва и други три стъпки, които са необходими за идентифициране и измерване на клетките на МРБ при дВКП-ОЛЛ.

В патентен документ US 5605805 A2 [5], публ. на 25.02.1997 г. е разкрита система и метод за автоматично определяне на линията на остра левкемия. Методът използва осем епруветки със състав: 1. неоцветен; 2. изотипни контроли; 3. CD10 FITC, CD19 PE; 4. CD20 FITC, CD5 PE; 5. CD3 FITC, CD22 PE; 6. CD7 FITC, CD33 PE; 7. HLADR FITC, CD13 PE и 8. CD34 FITC, CD38 PE. Осемте файла с данни по четири параметъра се придобиват с поточния цитометър. При анализа на данните клъстерите от клетки се групират използвайки специален алгоритъм. След това клъстерите се съединяват във файлове с данни, за да образуват клетъчни популации. Средните позиции на всяка клетъчна популация се сравняват с дърво на решенията, което идентифицира нормалните клетъчни популации. За идентифициране на левкемични клетъчни популации, алгоритъмът елиминира нормалните клетъчни популации от пространството на данните и останалите популации се класифицират като В-линия ALL, Т-линия ALL, AML, AUL, В-CLL или неизвестни.

Това изобретение има следните недостатъци: 1. Алгоритъмът за клъстерифициране е приложим само за данните от горепосочените 8 епруветки. Те обаче не включват маркери на МРБ при дВКП-ОЛЛ; и 2. Системата и методът са подходящи за класифициране на основните линии В-линия ALL, Т-линия ALL, AML, AUL, В-CLL и не са подходящи за измерване на МРБ при дВКП-ОЛЛ.

В нашия полезен модел системата позволява да се измерят и анализират произволен брой маркери, които са известни понастоящем, както и нови маркери в бъдеще. Системата позволява да се измери МРБ при дВКП-ОЛЛ с висока специфичност и чувствителност.

В патентен документ EP 1 785 899 A2 [6], публ. на 27.08.2008 г. се разкрива изчислителна система за идентифициране на популации от събития в многоизмерен набор от данни, получен от флоуцитометър. Данните се обработват чрез: а) модела на крайно смесване (finite mixture model - FMM), като моделът включва претеглена сума от многомерни гаусови функции на плътност на вероятността; б) алгоритъм за максимизиране на очакванията, конфигуриран да работи върху подмножества от набора от данни и да актуализира параметри, свързани с функциите на плътност на FMM, чрез които става клъстерифициране на многомерните данни; в) набор от експертни знания, включващ една или повече трансформации на данни и едно или повече логически твърдения.

Използването на набора от експертни знания в комбинация с FMM позволява по-стабилни и точни процеси за автоматично класифициране на данни в една или повече популации. В контекста на флоуцитометрия, експерт-хематолог в резултат на предишна информация, има идея къде многомерните данни попадат в една или повече двуизмерни проекции. Това включва позиция на клъстер (напр. в двуизмерна проекция или графика на подмножество от данни), геометрична форма на клъстер в някакво двуизмерно прогнози и позиция на клъстера спрямо други клъстери. Разкритата система осигурява автоматизирана класификация на клетките.

Идентифицирането на клъстерите (популациите) в многоизмерния набор от данни може да бъде преобразувано в количествени или качествени данни, представени във възприемана от човека форма, като например графика или диаграма на данните с цветно кодиране за идентифициране на дискретни популации в набора от данни.

В сравнение с нашия полезен модел, това изобретение има следните недостатъци: 1) Изпълнението на системата е само за класифициране на основните клетъчни популации в кръвта - гранулоцити, моноцити и лимфоцити. Това са големи популации и не е ясно, доколко разкритата система е приложима за малки популации, каквито са клетките на МРБ; 2) В изобретението измерването на параметрите на клетките е само по физически параметри - разсейване на светлината. В този смисъл не е известна приложимостта му за клъстерифициране на клетките по флуоресценция; 3) Изпълнението на системата е за 7 параметърни флоуцитометрични данни и не е известна неговата приложимост за многопараметърни (14-16 параметъра) данни.

Нашият полезен модел предлага решение, при което: а) се откриват редки популации (10^5) в костния мозък; б) използват се както физическите, така и флуоресцентните параметри на флоуцитометъра и в) могат да се анализират данни от МПФЦ (16 параметъра).

В патентен документ на САЩ US 008214157 B2 [7], публ. на 03.06.2012 г. и **US 2013124522 A1**, публ. на 16.05.2013 г. с едно и също заглавие и едни и същи автори са разкрити методи и устройство за пренасяне на множество измерени данни. Проблемите с анализа на множеството данни от флоуцитометрия се решават чрез метод и апарат за представяне на разпределения на многомерни данни чрез набор от сегменти. За целта би-плота се разделя на сегменти по метод, наречен йерархично групиране с многомерна

минимална вариация с еднаква вероятност. Сегментите се разпределят в 5 йерархични нива, като броят на сегментите се удвоява във всяко следващо ниво, като по този начин се достига до общо 32 сегмента. Обединените данни от многомерната плътност на вероятността могат да се преобразуват в линейна последователност от числа като "пръстов отпечатък". Той е едномерно представяне на информацията, съдържаща се в многомерните данни. В допълнение, описаните тук апарати и методи включват алгоритми за използване на сегментите от един набор от данни, за обработване и представянето на сегменти на втори набор от данни и получаване на нов пръстов отпечатък. Чрез двата пръстови отпечатъка може да се представи степента на сходство или несходство на две или повече флоуцитометрични измервания. Не могат обаче да се открият редки популации.

Това решение има следните недостатъци по отношение на измерването на МРБ чрез МПФЦ: 1) Сегментирането се извършва върху бивариатни плотове, които представляват извадка от всички данни и вероятността е малка в даден сегмент да попадне малка популация клетки, каквито са клетките на МРБ; 2) На второ място според претенциите на авторите, тази система може да се използва за сравнение на две и повече измервания, но не и за откриване на малки популации и 3) И на трето място - изобретението е изпитано върху данни от 4 параметъра флоуцитометрия за анализ на основните клетъчни субпопулации на лимфоцитите в кръвта. Не е ясно, доколко то е приложимо за данни, съдържащи десетки и стотици плотове при многопараметърната флоуцитометрия.

В нашия полезен модел всички многомерни данни се групират в множество клъстери, сред които може да се открие клъстер с редки клетки на МРБ. Моделът е приложим за МПФЦ.

В патентен документ WO 2006089190 A2 [8], публ. на 28.12.2006 г. е разкрита система, която улеснява откриването на ниски нива на туморни клетки въз основа на техните фенотипни разлики спрямо нормалните им двойници, оценени чрез анализ на комплексни данни от нормални и аномални клетъчни популации. Изобретението описва система, включваща компютър и алгоритми, чрез която се създават центроид и радиуси, дефиниращи набор от клъстери в n -измерно пространство на флоуцитометричните данни, съответстващо на нормално съзряване за клетъчна линия в нормалния набор от клетки. Радиусите изхождат от центроид и съответстват на всички параметри на флоуцитометричните данни. Те се изчисляват на базата на силата на флуоресценцията, обаче тяхната посока и размер се настройват ръчно. Всяка клетка се картографира към съответстваща точка в n -измерено пространство, базирано на измереното множество интензитети на флуоресценция на клетката. Чрез определяне на центроидна линия и радиус съответните точки оформят клъстери в n -измерното пространство, при което всеки клъстер съответства на ниво на зреење в рамките на клетъчна линия. Клъстерите се представят графично в двумерен плот. Това картографиране трябва да се извърши поотделно за "нормални проби", съдържащи нормални клетки от здрав индивид и за "патологични проби", съдържащи туморни клетки от болен индивид. Идентифицирането на туморните клетки става визуално чрез сравнение на картата на "нормалната проба" с картата на "патологичната проба". Описано е изпълнение, при което картата на "нормалната проба" се получава от 25 "нормални проби". Тя може да се извади електронно от картата на патологичната проба и по този начин могат да се визуализират редки патологични клетки, като клетки на МРБ.

В сравнение с нашия полезен модел това изобретение има следните недостатъци: 1) настройването на местоположението на центроида и радиусите е ръчно и съпроводено от субективни решения; 2) за определяне на патологичните клетки, включително клетки на МРБ, трябва да се картографират множество "нормални проби", което е свързано с етични проблеми при вземането на кръв и костен мозък от здрави индивиди, особено деца. Поради различията на флоуцитометрите и техните настройки в отделните лаборатории, картата от "нормални проби" ще е валидна само за лабораторията, която я е създала и няма да е валидна за останалите лаборатории. В допълнение на това, картата от "нормални проби" ще е валидна само за определен панел от антитела срещу клетъчните маркери. При всеки нов панел ще трябва да се изработва нова карта от "нормални проби"; 3) Изпълнението на системата е извършено, чрез събиране само на 200,000 събития (клетки), които лесно могат да се представят като клъстери на двумерен плот. За измерване на МРБ при дВ-ОЛЛ с чувствителност 10^{-5} е необходимо да се съберат поне 4×10^6 събития (клетки). При толкова много събития не е ясно доколко клъстерите ще бъдат различни на двумерния плот и 4) Изпълнението на системата е извършено с данни от 6-параметърна флоуцитометрия. Не е известно, как тя ще работи с данни от 10-16-параметърна флоуцитометрия.

Нашият полезен модел предлага решение, при което: 1) получаването на клъстерите не изисква контрол на оператора; 2) не се изискват "нормални" проби; 3) работи с 4×10^6 клетки и могат да се анализират данни от 16-параметрова флоуцитометрия, което повишава чувствителността на измерването на МРБ при дВ-ОЛЛ.

В научна публикация Jafari, K. и съавт. [9] през 2018 г. описват визуализация чрез Radar Plot на клетъчния състав и съзряване в костния мозък с помощта на 10-цветна поточна цитометрия, през 2021 г. **в научна публикация Kárai B. и съавт.** [10] описват модифициран метод на Radar Plot за измерване на МРБ при дВ-ОЛЛ чрез многопараметърна флоуцитометрия, а през 2022 г. **в научна публикация Shopsowitz K. и съавт.** [11] [12] описват оптимизиране на Radar Plot чрез машинно обучение за измерване на МРБ при В-ОЛЛ чрез МПФЦ.

Radar Plot е графична система за показване на многоизмерни данни в двуизмерна (2D) диаграма. Това е разширение на типична двумерна графика към по-високи измерения. Всички или няколко параметъра се представят на графиката под формата на оси, излизащи от един център. Местоположението на центъра, разположението на осите и тяхната големина могат да се настройват ръчно, с оглед да се визуализират най-голям брой клъстери и те да са максимално отдалечени един от друг. Клъстерите могат да се селектират чрез гейтове за статистически анализ.

В научна публикация Kárai B. и съавт. [10] през 2021 г. описват оптимизиране на разположението на центъра и осите на радар плота, както и тяхната посока и големина, с оглед най-добро разграничаване на клетъчните популации. Те изготвят 6 шаблона за най-добра визуализация на патологичните области при данни от 6-параметърна флоуцитометрия. При МРБ в областите на разположение на областите се търсят клетки на МРБ. Трябва да се подчертае, че шаблоните са специфични само за даден панел от антитела срещу клетъчните маркери.

В научна публикация Shopsowitz K. и съавт. [12] през 2022 г. описват използването на машинно обучение за оптимизиране на параметрите на Radar Plot. Те изготвят библиотека от плотове от проби с В-

ОЛЛ. Всяка радарна графика в библиотеката е използвана за трансформиране на балансирания набор от данни за обучение, след което класификаторите на поддържаща векторна машина (SVM) с радиална базисна функция са научени да класифицират нормални (т.е. зрели В-клетки или хематогони) срещу аномални събития за всяка радарна графика. Чрез ROC curve анализ е изчислена площта под кривата (AUC) на всеки модел на SVM и е използвана като мерило за класиране на кандидат радарните графики в библиотеката. Радарната диаграма с най-добри резултати е избрана за следващия кръг на оптимизация и цикълът на класиране и селекция е повторен, докато ефективността на радарната диаграма достигне плато. По този начин са оптимизирани параметрите на Radar Plot. Следва да се отбележи, че както и при описанието на Bettina Kárai В. и съавт. [10] и тук се използват данни от 6-параметрова цитометрия и че шаблоните са специфични само за даден панел от антитета срещу клетъчните маркери.

В сравнение с нашия полезен модел описаните по-горе решения имат следните ограничения: 1) Оптимизирането на Radar Plot изисква или ръчно настройване на осите или ръчен подбор на шаблон с най-голяма площ под кривата (AUC), което зависи до голяма степен от квалификацията на оператора; 2) Шаблоните са специфични за един панел от антитета срещу левкоцитните маркери и до голяма степен от вида на използвания флоуцитометър. Това би било сериозно ограничение за приложението на подхода в различни лаборатории и 3) в по-горните описания Radar Plot е приложен при данни от 6-параметърна флоуцитометрия. В нашето изпълнение ние постигнахме добри резултати с Radar Plot при 8-параметърни измервания на МРБ. При данни от 14-параметърна флоуцитометрия Radar Plot показва незадоволителни резултати. Това се потвърждава и от наблюденията на Shopsowitz К. и съавт. [12], които споделят, че дизайнът на радарните диаграми се основава на интуиция/проба и грешка, а с големия брой канали, налични при съвременните флоуцитометри, пространството за проектиране е изключително голямо, което прави невъзможно да се знае дали дадена радарна диаграма е оптимална за поставената задача.

В нашия полезен модел клъстерифицирането на многомерните данни е напълно автоматично, системата е приложима за измерване на МРД при дВ-ОЛЛ с 14-16-параметърна флоуцитометрия, което повишава чувствителността на измерването. Освен това системата при нашия полезен модел не зависи от състава на антителята в панела и вида от флоуцитометъра, което ги прави приложими във всяка една лаборатория.

В научна публикация Shopsowitz К. и съавт. [11] през 2024 г. описват система за измерване на МРБ при остра миелоидна левкемия (ОМЛ). Системата, наречена MAGIC-DR се състои от подсистема за контролирано машинно обучение (XGBoost класификатор) и подсистема UMAP (Uniform Manifold Approximation and Projection for Dimension Reduction) за неконтролирано представяне на данните в двумерна графика. XGBoost използва градиентно усилване, при което се създава модел чрез последователно добавяне на решения, докато не могат да бъдат направени допълнителни подобрения в решенията. Като контролирана техника за машинно обучение, класификаторите на XGBoost изискват етикети за действителния състав на клетките по време на началната фаза на обучение. За това XGBoost класификаторът се обучава да класифицира левкемичните клетки, като анализира множество проби, които не съдържат левкемични клетки за ОМЛ (негативни проби) и проби, които съдържат левкемични клетки за ОМЛ (позитивни проби). Процесът на обучението за разпознаване се контролира от оператора. Веднъж

обучен, класификаторът XGBoost, може надеждно да идентифицира имунофенотипно разнообразни ОМЛ бласти/промоноцити с голяма вероятност. XGBoost е свързан с подсистемата UMAP, към която подава информацията за класифицираните ОМЛ клетки. Задача на UMAP е неконтролирано (автоматично) да редуцира многомерните данни до двумерна графика, на която се представят клетъчните популации като отделни клъстери. Левкемичната популация се представя в различен цвят за по-добро визуализиране. Тя може да бъде селектирана чрез гейт за провеждане на статистически анализ.

В сравнение с нашия полезен модел системата MAGIC-DR има следните ограничения: 1) Важно ограничение на подсистемата XGBoost е, че моделите за контролирано машинно обучение са специфични за панелите от антитела срещу клетъчните маркери, използвани за тяхното обучение. Може да се очаква, че дори незначителни модификации на панела или конфигурацията на флоуцитометъра могат да окажат значително въздействие върху работата на системата и ще трябва повторно обучение на XGBoost класификатора с новопридобити данни. Това би било сериозна пречка за внедряване на системата в различни лаборатории; 2) Подсистемата UMAP има също съществено ограничение. Както и други техники за намаляване на размерността, UMAP по принцип позволяват опростена визуализация на данни с големи размери. Тълкуването на резултатите от тези диаграми обаче не винаги е лесно. Например графиките на UMAP често имат много отделни популации, чийто точен състав и позиции варират от проба на проба. Следователно при анализа на МРБ обикновено не е възможно да се знае предварително къде да се търси малка необичайна популация на UMAP диаграмата.

При нашия полезен модел функционирането на системата не зависи от състава на антителата в панела и вида на флоуцитометъра, което ги прави приложими във всяка една лаборатория.

В научна публикация Reiter M. и съавт. [13] през 2019 г. описват система за контролирано машинно обучение, използвайки комбинация от множество модели на смесване на Гаус (Gaussian Mixture Models - GMM), модифицирани като параметричен модел на плътност. Подходът е използван за намиране на теглата на линейна комбинация от множество GMM за представяне на нови "невиждани" проби чрез интерполация на съхранени проби. За контролирано машинно обучение се използват данните от флоуцитометрични проби с ръчно зададени гейтове. Наборът от данни за обучение съдържа 6-параметърни флоуцитометрични данни от 337 проби от костен мозък на педиатрични пациенти с дВ-ОЛЛ, събрани на 15-ия ден от индукционната терапия в три различни лаборатории.

При изследване на чувствителността на подхода, авторите установяват праг на МРБ от 0,05% като граница на разделителна способност за точно количествено определяне на МРБ.

Според оценка на авторите, описаният подход превъзхожда обучението на опорни векторни машини (SVM) и дълбоките невронни мрежи (DNN) и единичен GMM подход в контекста на оценката на МРБ, проведени с флоуцитометрични данни на педиатрични пациенти с В-ОЛЛ.

Предимство на описаната система е, че тя е насочена специално за измерване на МРБ при дВ-ОЛЛ, но тя има следните ограничения: 1) Предпоставка за приложение в клиничната рутина е надеждното прогнозиране, дори ако пробите са получени на различни флоуцитометри с неизвестни настройки и различни панели за оцветяване. Поради тази причина данните, използвани в описания модел, са получени на различни флоуцитометри в различни лаборатории. Установено е, че представянето на модела е най-

добро, ако тестовите и тренировъчни проби са от една и съща система. Когато комплектите за обучение и тест са придобити на различни системи (флоуцитометър и панел за оцветяване), се наблюдава намалена производителност. Това ще затрудни въвеждането на модела в различни лаборатории; 2) Друго ограничение на системата е по отношението на откриването на редки фенотипи на МРБ при дВ-ОЛЛ. Първо, някои проби с редки фенотипове не могат да бъдат представени добре от пробите от набора за обучение. Такива например са МРБ в проби с про-В-ОЛЛ или ОЛЛ с промяна на линията. 3) Обучението на модела е извършено с проби от 6-параметрова флоуцитометрия. Не е известно също така, как ще функционира моделът с данни от МПФЦ (14-16 и повече параметъра) и 4) чувствителността на измерването на МРБ е 0.05%, докато съвременното изискване за чувствителност е най-малко 0.01%

В нашия полезен модел клъстериизирането на многопараметърните данни е напълно автоматично и не зависи от машинно обучение. Поради това моделът може да се внедри в различни лаборатории. Също така с него могат да се открият редки фенотипи. Нашият модел е изпълнен с данни от 14-параметърна флоуцитометрия и чувствителността надвишава 0.01%.

Заключение

От направения преглед на настоящото състояние на техниката се вижда, че съществуващите практики и подходи не успяват да осигурят ефективни, стабилни, надеждни и точни системи за измерване на МРБ при дВ-ОЛЛ. От гореизложеното се разбира, че съществува необходимост от система, която да преодолее недостатъците на съществуващите преди това решения, и която да позволи идентифицирането на малка субпопулация от клетки в голяма популация от клетки, в каквато и да е проба, и което не зависи от референтна проба.

Целта на настоящия полезен модел е да се преодолеят горните неудобства, като се осигури система, позволяваща откриването на редки клетки по възпроизводим начин, без необходимостта от машинно обучение, която система да бъде приложена за измерване на МРБ при дВ-ОЛЛ.

Техническа същност на полезния модел

Проблемите, посочени в раздела за предшестващо състояние на техниката по-горе и по-специално, проблемите от предшестващото състояние на техниката, се решават чрез описаната система за анализ на данни от МПФЦ за измерване на МРБ при дВ-ОЛЛ.

Характеристики и предимства на настоящия полезен модел, както и структурата и работата на различни изпълнения на настоящия полезен модел, са описани подробно по-долу с позоваване на придружаващите чертежи.

Система за измерване и анализ на МРБ при дВ-ОЛЛ чрез многопараметърна флоуцитометрия

На Фиг. 1 е представена концептуалната схема на система за измерване на МРБ при дВ-ОЛЛ чрез МПФЦ 100, съгласно приложението на полезния модел. Системата 100 може да бъде реализирана по различни начини, включително като отделни подсистеми. В едно изпълнение системата 100 се състои от три подсистеми 101, 103 и 106, хранилище на файлове 105 и външен сървър 107. Двете подсистеми 103 и 106, и хранилището на файлове 105 могат да бъдат свързани в мрежа 112, като например кабелни и

безжични комуникационни мрежи, локални мрежи, подмрежа, интранет и др. Подсистема-2 106 се свързва с външния сървър 107 чрез интернет 113.

Подсистема-1 101 служи за измерване на пробата на пациента 102 чрез МПФЦ. Дигиталните данни от измерването се предават чрез интерфейс 112 на Подсистема-2 103 за придобиване на данни. Тя служи за приемане на данните от флоуцитометричното измерване и трансформирането им в стандартен файл със сурови данни 104 (FCS файл). FCS файлът 104 може да се съхрани в подсистема-2 103 или чрез комуникационна мрежа 112 да се изпрати за съхранение в хранилище на файлове 105 или до Подсистема-3 106. Аналитичната Подсистема-3 106 извършва анализа на флоуцитометричните данни. Тя е съставена от 4 модула: модул-РД 108 за редуциране на данните; модул-ИД 109 за изчистване на данните; модул-КД 110 за клъстерифициране на данните и модул-АК 111 за анализ на клъстерите. Подробен примерен изглед на приложението на Система 100 е илюстриран на Фиг. 2-5. и е разгледан по-долу.

Компонентите на Подсистема-1 101 са представени на Фиг. 2 съгласно изпълнението на полезния модел. Тя се състои от многопараметров флоуцитометър 201, преобразувател на събитията в сурови дигитални данни 202 и мрежов интерфейс 103. Функциите на Подсистема-1 101 да измери пробата на пациента по всички параметри, да дигитализира данните и да ги предаде на Подсистема-2 103.

Компонентите на Подсистема-2 103 са представени на Фиг. 3, съгласно изпълнението на полезния модел. Функциите на Подсистема-2 103 са да управлява флоуцитометъра 101, да придобива флоуцитометрични данни за клетките и да генерира FCS файл 104. Тези функции могат да се изпълняват от компютърна система, която се състои от централен процесор 301, графичен процесор 302, потребителски интерфейс и монитор 303, входно устройство 304, основна оперативна памет 305 и вторична памет 306.

Централният процесор 301 изпълнява инструкциите на софтуера за изпълнение на задачи, описани тук. Той обработва постъпващите данни от флоуцитометъра 101 и чрез графичния процесор 302 ги представя на потребителския интерфейс на монитора 303. На потребителския интерфейс 303 се представят данни за работата на флоуцитометъра 101 и данни от измерванията на клетките. Чрез потребителския интерфейс 303 операторът може да управлява работата на флоуцитометъра 101, като например да конфигурира параметрите, да настройва напрежението на детекторите, да регулира скоростта на потока на клетките, да определя праг на измерванията и други. Съгласно изпълненията на настоящия полезен модел, входното устройство 304 може да бъде, но не се ограничава до, например, посочващо устройство, тракбол, устройство за сензорен интерфейс, тъчпад, джойстик, система за управление с гласово активиране, сензорен екран или други входящи устройства 304, използвани за осигуряване на взаимодействие между потребител и потребителски интерфейс 303.

След приключване на измерването с флоуцитометъра 101, Подсистема-2 трансформира придобитите данни в FCS файлов формат 104 и те се записват в основната паметта 305, която може да бъде памет с произволен достъп (RAM) или се записват във вторичната памет 306. Колективно паметта може да съхранява едно или повече софтуерни приложения, данните, получени/използвани от тези приложения, и изходните/генерирани данни от тези приложения. Вторичната памет 306 може да включва, например, устройство с твърд диск (HDD) или полупроводников диск (SSD), устройство за сменяемо устройство за

съхранение, флаш памет и/или други подобни енергонезависими памет. Сменяемостта на устройството за съхранение може да включва флопи дисково устройство, устройство с магнитна лента, оптично дисково устройство, флаш памет или други подобни. Софтуерът на Подсистема-2 може да варира в зависимост от производителя на инструмента за измерване.

Чрез мрежов интерфейс 203 FCS файловете се изпращат до хранилището 105 или до Подсистема-3 106. Хранилището 105 служи за съхранение на FCS файловете 104. То може да включва вътрешно или външно запаметяващо устройство, мрежово устройство за съхранение (NAS), файлов сървър, облъчен сървър и други подобни. От хранилището FCS файловете 104 се изпращат до Подсистема-3 106 за анализ.

На Фиг. 4 са представени компонентите на Подсистема-3 106, съгласно изпълнението на полезния модел и са описани накратко техните функции. Функцията на Подсистема-3 е да анализира многопараметърните флоуцитометрични данни чрез поредица от стъпки, изпълнявани последователно от модул-РД 108, модул-ИД 109, модул-КД 110 и модул-АК 111, което завършва с резултат за измерена МРБ при дВ-ОЛЛ. Тази функция може да се изпълнява от компютърна система, която се състои от централен процесор 401, графичен процесор 402, потребителски интерфейс и монитор 303, входно устройство 304, оперативна памет 305 и вторична памет 306. Модул-РД 108 изпълнява инструкции за редуциране на данните чрез селектиране на В-клетките. Той генерира FCSdr файл 403, който изпраща за съхранение в оперативна памет 305 и вторична памет 306. Модул-ИД 109 изпълнява инструкции за автоматично изчистване на флоуцитометричните данни. Той генерира FCSd файл 404, който изпраща за съхранение в оперативна памет 305 и вторична памет 306. Модул-КД 108 изпълнява инструкции за автоматично клъстерифициране на флоуцитометричните данни. Той генерира CG файл 405 CS файл 406, които изпраща за съхранение в оперативна памет 305 и вторична памет 306. Модул-АК 108 изпълнява инструкции за представяне на корелации между клъстери и параметри 901. Той генерира AG файл 407 AS файл 408, които изпраща за съхранение в оперативна памет 305 и вторична памет 306.

Функциите на модулите може да се изпълняват от компютърна система, която се състои от централен процесор 401, графичен процесор 402, потребителски интерфейс и монитор 303, входно устройство 304, оперативна памет 305 и вторична памет 306.

Централният процесор 401 изпълнява инструкциите на алгоритмите на модулите на модул-РД 108, модул-ИД 109, модул-КД 110 и модул-АК 111. Процесорът 401 може да бъде процесор със специално предназначение или процесор с общо предназначение. Процесорът 401 може да бъде мултипроцесор, включващ множество процесори (не са показани), които могат да съпоставят една или повече нишки към всеки процесор. Процесорът 401 трябва да бъде с характеристика, позволяваща сложни изчисления на голям обем данни.

Основна задача на графичния процесор 402 е да извършва сложни математически операции по разпределянето на многопараметърните данни в клъстери в многомерното пространство и да представи клъстерите в двумерно и/или тримерно пространство. За целта графичния процесор 402 трябва да има необходимата производителност и размер на видеопамет.

Функциите и характеристиките на оперативната памет 305, вторичната памет 306, потребителския интерфейс 303 и входното устройство 304 са описани по-горе.

За обмяна на данни Подсистема-3 106 може да бъде свързана чрез мрежа 112 с Подсистема-2 103 и хранилището 105. Подсистема-3 106 може да бъде свързана чрез интернет 113 с външен сървър, от който да се изтеглят алгоритмите за анализ.

Функционирането на описаната по-горе система за измерване на МРБ при ДВ-ОЛЛ чрез МПФЦ е представено схематично на Фиг. 5 под формата на блок-диаграма.

В първата стъпка пробата 102 от пациента се измерва от флоуцитометъра 201 на подсистема-1 101, като клетките се пропускат една по една през един или няколко лазерни лъча с различна дължина на вълната. Под въздействието на лазерните лъчи флуорохромите излъчват светлина, която се разлага на цветове чрез система от филтри и всеки цвят се улавя от детектор, наричан "параметър". Съвременните комерсиални флоуцитометри имат между 12 и 50 параметъра. Силата на сигнала на всеки параметър се преобразува в дигитален формат чрез дигитален преобразувател 202. Така за всяка клетка се присвояват данни за силата на флуоресценцията за всеки параметър. Чрез мрежов интерфейс 203 дигиталните данни за всяка клетка се изпращат към Подсистема-2 103.

Във втората стъпка получените данни за клетките от флоуцитометъра 101 чрез мрежовия интерфейс 203 постъпват в Подсистема-2 103. Постъпилите данни се обработват от централния процесор 301 и графичния процесор 302 се представят на потребителския интерфейс 303, графично под форма на хистограми и би-плотове, контур-плот, плътностен плот или като статистически резултат. Чрез входното устройство 304 на хистограмата и би-плота операторът може да очертава гейтове 604, чрез които може да селектира определени клетъчни популации. Данните за клетъчната популация, под формата на FCS файлов формат 104, се записват в на оперативната памет 305 и вторичната памет 306. FCS файлът може да бъде изпратен по мрежата 112 до хранилището за файлове 105 и до Подсистема-3 106 за анализ.

В третата стъпка процесът за анализ на данните започва с получаването от Подсистема-3 на FCS файл 104, който през мрежов интерфейс 203 се изтегля от хранилището 105 или от Подсистема-2 103. Анализ на данните от Подсистема-3 включва 4 подстъпки, изпълнявани от свързани помежду си четири модула: модул-РД 108, модул-ИД 109, модул-КД 110 и модул-АК 111.

В първата подстъпка, Модул-РД 108 обработва FCS файл 104 и представя данни от него на потребителския интерфейс 303 графично под формата на бивариатни плотове, които са илюстрирани на Фиг. 6. Алгоритъмът на Модул-РД 108 може да бъде всеки софтуер за класически анализ на флоуцитометрични данни. Съгласно едно или повече изпълнения, описаните тук процеси позволяват на потребителя да анализира итеративно и показва данните на потребителския интерфейс 303 чрез избиране и/или модифициране на типовете показани графики и параметри, осите и/или гейтове от интерес. Чрез тях операторът може да селектира В-клетъчната популация 614 и да отдели нейните данни от останалите събития 606, като по този начин в значителна степен се редуцират данните за последващ анализ. Редуцираните данни могат да се запазят в оперативната памет 305 и/или да се запишат във вторичната памет 306 под като FCSdr файл 403 за последващо използване.

Втората подстъпка цели изчистването на данните от аномалии 701, появяващи се по време на измерването на пробата на пациента 102 от флоуцитометър 101, като например преминаване на мехурче, агрегат, промяна на скоростта или завихряне на потока. За целта Модул-ИД 190 изтегля файл FCSdr 403

от оперативната памет 305 и/или вторичната памет 306 на Подсистема-3 106 и прилага алгоритми за автоматично изчистване на флоуцитометрични данни, като например flowClean, flowAI, PeasoQC и други, които са свободно достъпни в Проекта Bioconductor на външния сървър 107. Изчистените данни се записват като файл FCScd 404 в оперативната памет 305 и вторичната памет 306 на Подсистема-3 106.

С третата подстъпка се цели многопараметърните флоуцитометрични данни да се групират в клъстери, които да съответстват на отделни клетъчни субпопулации. За целта Модул-КД 110 изтегля файл FCScd 404 оперативната памет 305 и/или вторичната памет 306 на Подсистема-3 106 и прилага алгоритми за автоматично клъстерифициране на многопараметърни флоуцитометрични данни без супервизия, като например FlowSOM, Phenograph, X-shift, flowMeans, flowClust и други, които са свободно достъпни в Проекта Bioconductor на външния сървър 107. Модулът-КД 110 анализира данните в FCScd 404 и автоматично, без контрол на оператора, групира данните в множество клъстери, като всеки клъстер отговаря на клетъчна популация с определени характеристики. Например алгоритъмът на FlowSOM може да групира данните в над 200 клъстера. След това Модул-КД 110 може да редуцира многомерното пространство на клъстерите в двумерно и да представи последното графично на потребителския интерфейс 304 под формата на клъстерно дърво 801, клъстерна карта 802 или топлинна карта 803 (Фиг. 8). За целта Модул-КД 110 може да използва алгоритми за редукция на размера като t-SNE, UMAP, FlowSOM, PHATE и други, които са свободно достъпни в Проекта Bioconductor на външния сървър 107. Двумерните графики могат да се запишат като CG-файл 406 във вторичната памет 306 на Подсистема-3 106. Модул-КД 110 генерира статистическата информация за клъстерите и се записва като CS-файл 406 в оперативната памет 305 и/или вторичната памет 306 на Подсистема-3 106.

Задача на четвъртата подстъпка е да анализира клъстерите, получени в подстъпка три, и да идентифицира клъстера, който съответства на клетките на МРБ при дВ-ОЛЛ, като същевременно измери броя на клетките и техния процент. За целта Модул-КА 111 изтегля файл CS-файл 406 оперативната памет 305 и/или вторичната памет 306 на Подсистема-3 106 и прилага алгоритми за анализ на клъстерите като ClusterExplorer и други, които са свободно достъпни в Проекта Bioconductor на външния сървър 107. Модул-КА 111 използва статистическата информация за клъстерите от CS-файл 406, за да представи клъстерите на потребителския интерфейс под формата на графика 900, представена на Фиг. 9. Клъстерите могат да се селектират, като за всеки маркер 904 може да се постави гейт 906 върху един или няколко клъстера. Гейтовете 905 на два и повече маркера 904 могат да се комбинират чрез операторите AND, OR и NOT. По този начин се селектира клъстерът, който отговаря на клетките на МРБ при дВ-ОЛЛ. Селектираният клъстер се представя на клъстерното дърво 801 и на графика с колонии 906, показващи процента на всеки клъстер. Статистическите данни за клъстера са представени на електронна таблица, генерирана от CS-файл 406 и чрез тях се определя броя на левкемичните клетки на МРБ при дВ-ОЛЛ и техния процент. С това завършва процесът на измерване на МРБ при дВ-ОЛЛ.

Предимствата, които осигурява полезния модел, са:

За разлика от известните технически решения системата, описана в този полезен модел, има следните предимства: а) Не се изисква използването на референтни проби и машинно обучение; б) може

да работи с данни от произволен панел антитета за маркиране на левкоцитните маркери и с произволен флоуцитометър за измерване; в) предимствата посочени в точки а) и б), позволяват настоящия полезен модел да бъде приложен във всяка една лаборатория; г) клъстерифицирането на популациите клетки е автоматично без намеса на оператора; д) позволява да се анализират данни от 16-параметрова цитометрия, което надхвърля настоящите изисквания за клинични флоуцитометри и е) с описаната система се постига чувствителност от 10^{-5} левкемични клетки на МРБ при дВ-ОЛЛ, което е по-високо от известните решения.

Кратко описание на приложените диаграми и фигури

Приложният модел ще бъде описан по-нататък само примерно, с позоваване на приложените диаграми и фигури, където:

фигура 1 е схематично представяне на системата за анализ на данни от многопараметърна флоуцитометрия за измерване МРБ при дВ-ОЛЛ 100;

фигура 2 е опростена блок диаграма на Подсистема-1 101;

фигура 3 е опростена блок диаграма на Подсистема-2 103;

фигура 4 е опростена блок диаграма на Подсистема-3 10;

фигура 5 е представена подробна блок-схема за процеса на анализ надолу по веригата, съгласно изпълнението на полезния модел;

фигура 6 илюстрира графиките, генерирани от Модул-РД 108;

фигура 7 илюстрира графиките, генерирани от Модул-ИД 109;

фигура 8 илюстрира графиките, генерирани от Модул-КД 110;

фигура 9 илюстрира графиките, генерирани от Модул-АК 111;

Фигура 10 илюстрира резултатите от анализа на чувствителността и специфичността на системата и метода 100.

Означение на номерата във фигурите

101	Подсистема-1 за измерване	510	Клъстерифициране на данни
102	Проба на пациента	511	Графика на клъстерите -CG-файл
103	Подсистема-2 за придобиване на данни	512	Топлинна карта - НМ-файл
104	FCS файл	513	Анализ на клъстерите
105	Хранилище за файлове	514	Графика на клъстерите

106	Подсистема-3 да анализ на данни	515	Поставяне на гейтове
107	Външен сървър	516	Селекция на клъстер
108	Модул-РД	517	Статистика на клъстера
109	Модул-ИД	518	Резултат за МРБ на дВ-ОЛЛ
110	Модул-КД	601	Би-плот FSC срещу SSC
111	Модул-АК	602	FCS - разсейване напред
112	Комуникационна мрежа (LAN)	603	SSC - разсейване настрани
113	Интернет	605	Левкоцити
201	Флоуцитометър-многопараметърен	606	Отломки на клетки
202	Дигитален преобразувател на данни	607	Би-плот FSC-H срещу FSC-A
203	Мрежов интерфейс	608	FSC-H - интензитет
301	Централен процесор	609	FSC-A - площ
302	Графичен процесор	610	Единични клетки
303	Потребителски интерфейс	611	Дублети
304	Входно устройство	612	Би-плот FSC срещу CD19
305	Оперативна памет	613	CD19
306	Вторична памет	614	В-клетки
401	Централен процесор	701	Аномалия на данните
402	Графичен интерфейс	702	Параметър време
403	FCSdr -файл	801	Клъстерно дърво
404	FCScd -файл	802	Клъстерна карта
405	CG -файл	803	Топлинна карта
406	CS -файл	804	Клъстер на МРБ на дВ-ОЛЛ
407	AG -файл	805	Метаклъстер
408	AS -файл	901	Клъстерна графика
501	Гейт-1 на левкоцити	902	Параметри на флоуцитометъра
502	Гейт-1 на единични клетки	903	Интензитет на флуоресценцията
503	Гейт-3 на В-клетки	904	Клъстерни линии
505	Изчистване на данни	905	Гейт на клъстери
507	Графика на изходни събития	906	Гейт на клъстери
508	Графика на изчистени събития	1001	Чувствителност и специфичност по класически метод с DIVA
509	Статистически резултат	1002	Чувствителност и специфичност чрез полезния модел

Примери за изпълнение на полезния модел

Примерно изпълнение на системата за измерване на МРБ при дВ-ОЛЛ на минимална чрез МПФЦ е илюстрирано на Фиг. 1-10. Полезният модел се илюстрира със следните примери, без да ограничава неговия обхват.

Пример 1. Изпълнение за измерване и анализ на МРБ при дВ-ОЛЛ на минимална чрез 16-параметърна флоуцитометрия.

Изпълнението на пример 1 започва с измерване на пробата на пациента 102 чрез Подсистема-1 101, включваща например 16-параметров флоуцитометър 201. Настройването на конфигурацията на

флуоцитометъра и условията за измерване на пробата се управляват от Подсистема-2 103 чрез специализиран софтуер на производителя на флуоцитометъра. От проба 102 се събират данни за поне 4×10^6 клетки по 14 флуоресцентни параметъра плюс два физически параметъра - FSC и SSC. В това изпълнение са събрани 4,573,827 събития/клетки. Данните за всеки параметър се превръщат в дигитален формат чрез дигитален преобразувател на данни 202 и чрез мрежов интерфейс 103 се изпращат към Подсистема-2 103.

Получените данни от подсистема-1 се обработват от централния процесор 301 на Подсистема-2 103 и чрез графичния процесор се представят 302 на потребителския интерфейс 303 под формата на хистограми и би-плотове. Чрез входното устройство 304 операторът може да настройва параметрите на флуоцитометъра 201, да определя праг на измерванията и да селектира данни на хистограмите и би-плотовете. Селектираните данни се записват в FCS файлов формат 104 в основната памет 305 и/или във вторичната памет 306 на Подсистема-2 103. Данните също така се изпращат по локална мрежа 112 до хранилището за файлове 105.

Анализът на данните се извършва чрез Подсистема-3 106. По локалната мрежа 112 тя извлича FCS файла 104 от Подсистема-2 103 или хранилището на файлове 105 и го изпраща на Модул-РД 108 за обработване. Чрез алгоритъма на Модул-РД 108 данните могат да се представят графично на потребителския интерфейс 303 под формата на би-плот 601 с параметри на флуоцитометъра "разсейване напред" (FSC) 602 и "разсейване настрана" (SSC) 603. На този плот операторът поставя гейт-1 501, чрез който се отделя популацията на левкоцитите 605 от тромбоцитите и отломките 606. В конкретния случай популацията на левкоцитите съдържа 3,963,892 клетки, като този брой служи по-нататък за изчисляване на процента на клетките на МРБ. Селектираните чрез гейт-1 501 популация на левкоцити 605 се представят в нов би-плот 607 с параметри FSC-H 608 и FSC-A 609. На този би-плот са представени популациите на единичните клетки 610 и на дублетите 611. Чрез поставянето на гейт-2 502 на би-плот 607 се селектират единичните клетки 610. След това те се представят на би-плот 612 с параметри FSC 602 и маркера CD19 613, специфичен за В-клетките. На би-плот 612 чрез поставяне гейт-3 503 се селектират CD19 позитивните клетки, които съдържат левкемичните клетки на МРБ при дВ-ОЛЛ. В случая са селектирани 537,374 клетки, които представляват 13.6% от левкоцитите или с други думи е постигната 86.4% редукция на клетките. Тази редукция на данните улеснява значително последващото клъстерифициране на данните от Модул-КД 110. Селектираните данни чрез гейт-3 503 се записват като файл FCSdr 403 в оперативната памет 305 и вторичната памет 306 на Подсистема-3 106.

В следващия етап от анализа чрез Модул-ИД 109 на Подсистема-3 106 става изчистване на данните от аномалии 701, появяващи се по време на измерването на пробата на пациента 102. За целта Модул-ИД 109 извлича данните от файл FCSdr 403 и използва алгоритми за автоматично изчистване на флуоцитометрични данни, в това изпълнение алгоритъмът flowClean, премахва събитията, които са свързани с аномалии 701 при измерване на пробата 102, като например преминаване на мехурче, агрегат, промяна на скоростта или завихряне на потока. След това Модул-ИД 109 представя на потребителския интерфейс 303 на графики с параметъра време 702 по оста X и параметъра SSC 603 по оста Y. Изходните данни са представени на би-плот 507, а изчистените данни на би-плот 508. Заедно с графиките се представя

и броя на събитията 509 в би-плот 507 и би-плот 508. Те са съответно 537,374 и 501,433 събития/клетки, което означава, че са изчистени 35,941 събития/клетки или 6.7%. При необходимост алгоритмът за изчистване на данни може да се изпълни няколко пъти до получаване на чисти данни. Данните с изчистените събития се записват като файл FCScd 404 в оперативната памет 305 и вторичната памет 306 на Подсистема-3 106.

В следващия етап от анализа чрез Модул-КД 110 16-параметърните данни за всяка една от 501,433 клетки се групират в клъстери, които представляват отделни клетъчни субпопулации. За целта Модул-КД 109 изтегля файла FCScd 404 от оперативната памет 305 или вторичната памет 306 на Подсистема-3 106. Модул-КД 109 използва автоматични алгоритми за клъстерифициране на данните във файл FCScd 404, като в това изпълнение е използван алгоритъма FlowSOM, но също така могат да бъдат използвани и други алгоритми за автоматично клъстерифициране. Модул-КД 110 представя на потребителски интерфейс 303 клъстерите от в многомерното пространство под формата на двумерни графики. В това примерно изпълнение са получени 100 единични клъстера и 8 метаклъстера, илюстрирани като клъстерно дърво на Фиг. 8. с графиката 801. На клъстерното дърво 801 е представена връзката между отделните клъстери и тяхната степен на сходство. Генерираните клъстери от многомерното пространство могат също така да бъдат представени графично под формата на клъстерна карта 802 и топлинна карта 803. На топлината карта 803 редовете представляват клъстерите, а колоните представляват параметрите. Стойностите в клетките, в цвetoва скала, показват нивото на експресия на всеки клетъчен маркер при всяко условие. Графиките на клъстерно дърво 801, клъстерна карта 802 и топлинна карта 803 се записват в CG-файл 405 във вторичната памет 306 на Подсистема-3 106. Генерираната статистическа информация за клъстерите се записва като CS-файл 406 в оперативната памет 305 и/или вторичната памет 306 на Подсистема-3 106.

В следващия етап от чрез Модул-КА 111 се извършва анализ на клъстерите с цел да се открие клъстер, с характеристика за левкемични клетки на МРБ при дВ-ОЛЛ. За целта Модул-КА 111 изтегля CS-файл 406 оперативната памет 305 или вторичната памет 306 на Подсистема-3 106 и използва статистическата информация, записана в него. В това изпълнение статистическите данни от CS-файл 406 се обработват от Модул-КА 111 чрез използването на алгоритъма ClusterExplorer, който представя клъстерите на потребителския интерфейс под формата на графика 901. На оста X 902 на графиката са представени всички параметри на флоуцитометъра, заедно със съответните маркери на левкоцитите. На оста Y 903 е представена силата на флуоресценцията в логаритмична скала. Клъстерите са представени като линии 904, които свързват степенята на експресията на всеки маркер по оста X 902, определена по силата на флуоресценцията от оста Y 903. Чрез гейтове 906 е селектиран клъстер № 7 904, който отговаря на характеристиката на левкемичните клетки на МРБ при дВ-ОЛЛ. На клъстерното дърво 801 той се представя като малка популация клетки 804, ясно разграничаваща се от останалите клъстери. Според статистическите данни от CS-файл 405 в клъстер № 7 има 503 клетки.

Изчисленията на процента на МРБ става по следната формула:

$$\% \text{ клетки на МРБ} = (\text{брой клетки в МРБ клъстер} / \text{брой клетки в гейт-1 604}) \times 100.$$

В примерното изпълнение изчисленията са следните $\% \text{ клетки на МРБ} = (86/3,963,892) \times 100 = 0.013\%$

Заклучението от анализа е, че пробата от костен мозък на пациента съдържа 0.013% левкемични клетки, което я прави положителна за МРБ при д-ОЛЛ.

Пример 2. Определяне на граница на откриване на МРБ чрез системата за измерване на МРБ при дВ-ОЛЛ чрез 16-параметърна флоуцитометрия

За определяне на границата за откриване на МРБ-клетки чрез системата, описана в полезния модел, бяха изготвени пет проби с количество на левкемични клетки 1%, 0.1%, 0.01%, 0.001% и 0.0001%. За целта различни количества левкемични клетки се поставяха в определено количество костен мозък, който не съдържа левкемични клетки. За източник на левкемични клетки беше използван костен мозък на пациент с дВКП-ОЛЛ при диагнозата, съдържащ 80% левкемични клетки. За костен мозък без левкемични клетки бяха използвани проби от педиатрични пациенти със сарком на Юинг. Приложена беше следната процедура:

В 5 епруветки за флоуцитометрия се поставят по 4.5×10^6 клетки от костен мозък без левкемични клетки. След това в тези епруветки се поставя различно количество левкемични клетки, както следва:

- В епруветка № 1 - 45,000 левкемични клетки = 1% (10^{-2})
- В епруветка № 2 - 4,500 левкемични клетки = 0.1% (10^{-3})
- В епруветка № 3 - 450 левкемични клетки = 0.01% (10^{-4})
- В епруветка № 4 - 45 левкемични клетки = 0.001% (10^{-5})
- В епруветка № 5 - 5 левкемични клетки = 0.0001% (10^{-6})

Клетките в пробите бяха оцветени с панел от антитела за 16-параметърна флоуцитометрия.

Таблица 2. Процент на откритите левкемични клетки в експеримента за чувствителност

№ епруветка	% Поставени левкемични клетки	Открити левкемични клетки		
		Общ брой събрани клетки	Брой открити левкемични клетки	% Левкемични клетки
1	1%	560114	6462	1.153620%
2	0.1%	894356	837	0.093640%
3	0.01%	2841754	271	0.009527%
4	0.001%	2927109	36	0.001236%
5	0.0001%	3987292	0	0.000000%

Резултатите са представени на Таблица 2. В епруветки № 1 - № 4 процентът на откритите левкемични клетки е приблизително равен на броя на клетките, поставени в съответната епруветка. В епруветка № 5 не се откриват левкемични клетки.

От тези резултати може да се заключи, че чувствителността на измерването със системата, описана в този полезен модел, е 0.0012% (10^{-5}), която е с един порядък по-висока от известната за други панели чувствителност от 0.010% (10^{-4}).

Пример 3. Определяне на чувствителността и специфичността на системата за измерване и анализ на МРБ при дВ-ОЛЛ чрез 16-параметърна флоуцитометрия.

Целта на този пример е да сравни измерването на МРБ при дВ-ОЛЛ при данни от 14-параметърна флоуцитометрия за измерване и анализ по класически начин и чрез системата от настоящия полезен модел.

За целта проби от костен мозък бяха взети от 77 деца с В-ОЛЛ в период на ремисия. Пробите бяха обработени успоредно с панел от антитела за 16-цветна флоуцитометрия.. Пробите бяха измерени с флоуцитометър FACSAria III. FCS файловете на всички проби бяха анализирани от един оператор успоредно по класическия начин със софтуера BD DIVA на флоуцитометъра FACSAria III и със Система на настоящия полезен модел. За класическия начин на анализ беше използван шаблон, представен на Фиг. 10, който включва минимум 45 би-плота и максимум 60 би-плота. Те се анализират визуално ръчно един след друг чрез поставяне на гейтове и селектиране на популация клетки чрез всеки гейт. Анализът на файловете чрез Система 100 на настоящия полезен модел беше извършен чрез процедурата, описана в Пример 1.

Получените резултати бяха статистически обработени с модула "ROC curve analysis" на софтуера "MedCalc" (<https://www.medcalc.org/manual/roccurves.php>). Резултатите от анализа са представени на Фиг. 10. При използването на класическия метод за анализ с BD DIVA 1001 е отчетена чувствителност 72.2 % и специфичност 66.7 % ($p = 0.001$). При използването на Система 100 1002 от този полезен модел за анализ е отчетена значително по-висока чувствителност 100 % и специфичност 83,3% ($p < 0.001$). Тези резултати показват значително по-високите технически параметри на Система 100 1002 от този полезен модел в сравнение с класическия метод за анализ при измерване на МРБ при дВ-ОЛЛ.

Разкритият досега предмет обаче, може да бъде изпълнен в различни форми и не трябва да се тълкува като ограничен единствено до изпълненията, описани тук. Тези изпълнения са предоставени по-скоро, за да може представянето на полезния модел да бъде изчерпателно и пълно, и да предаде напълно обхвата на изпълненията на специалистите в областта.

Претенции

1. Система за измерване и анализ на минимална резидуална болест при детска В-прекурсорна остра лимфобластна левкемия чрез многопараметърна флоуцитометрия (МПФЦ), характеризираща се с това, че се състои от: подсистема-1 (101), служеща за измерване на пробата на пациента чрез МПФЦ (201) и трансформираща данните в дигитален формат чрез дигитален преобразувател (202); подсистема-2 (103), служеща за придобиване на данни от флоуцитометъра и преобразуването им в FCS файлов формат; подсистема-3 (106) за анализ на многопараметърните флоуцитометрични данни, състояща се от четири последователно взаимосвързани модула, като модул-РД (108) е за отделяне на В-клетките и тяхното редуциране; модул-ИД (109) е с алгоритъм за автоматично изчистване на данните от аномалии; модул-КД (110) е с алгоритъм за автоматично клъстерифициране на данните и модул-АК (111) е с алгоритъм за анализ на клъстерите и за представяне на многомерното пространство на клъстерите в двумерни графики; хранилище за файлове (105); външен сървър (107) с алгоритми за анализ и комуникационна мрежа (112), свързваща подсистема-1 (101), подсистема-2 (103), подсистема-3 (106), хранилището за файлове (105) и външния сървър (107).

2. Система за измерване на минимална резидуална болест при детска В-прекурсорна остра лимфобластна левкемия чрез многопараметърна флоуцитометрия съгласно претенция 1, характеризираща се с това, че подсистема-2 (103) и подсистема-3 (106) представляват компютърни устройства, състоящи се от централен процесор (301), графичен процесор (302), оперативна памет (305), вторична памет (306), мрежов интерфейс (203), потребителски интерфейс (303) и входно устройство (304).

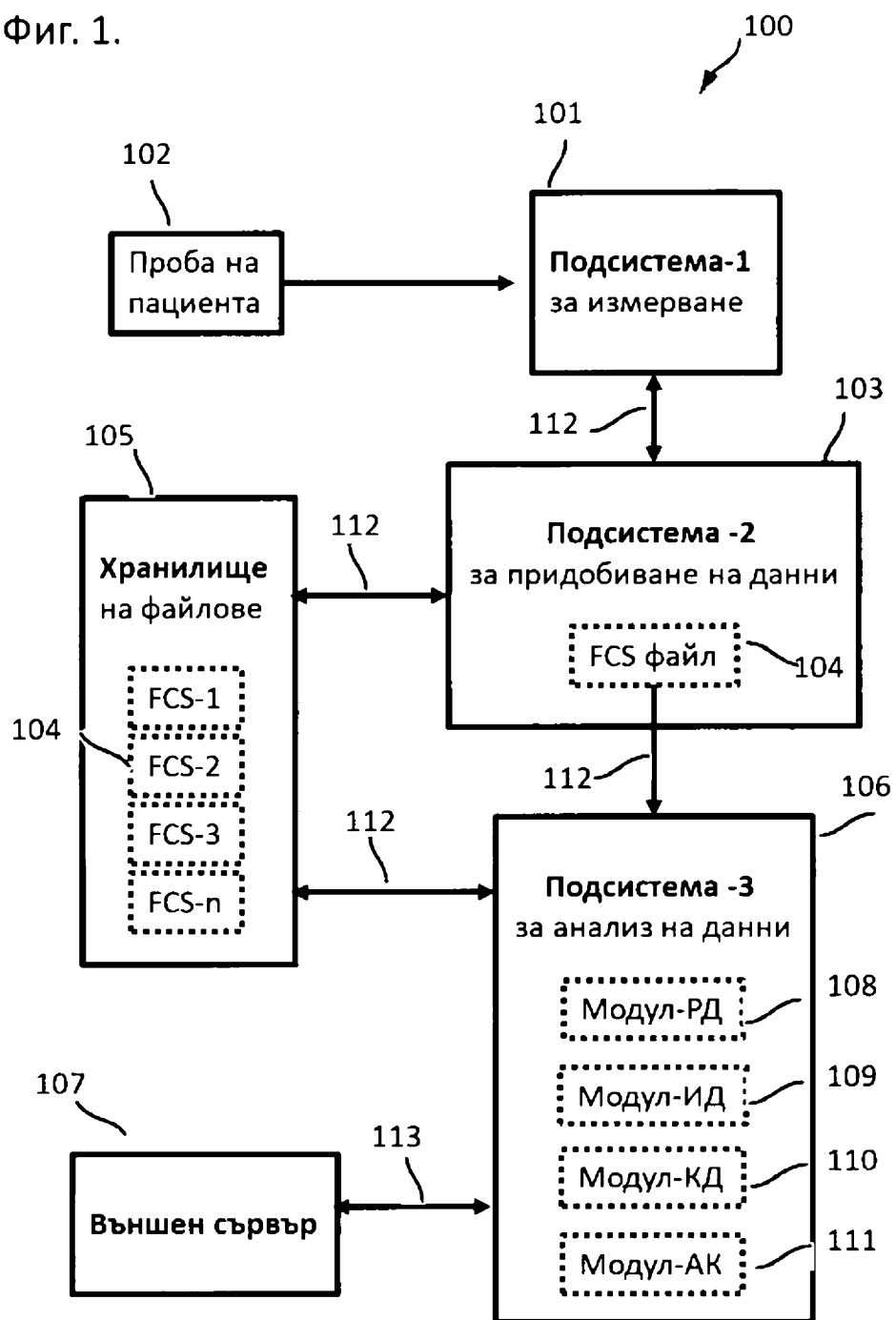
Приложение: 10 фигури

Литература

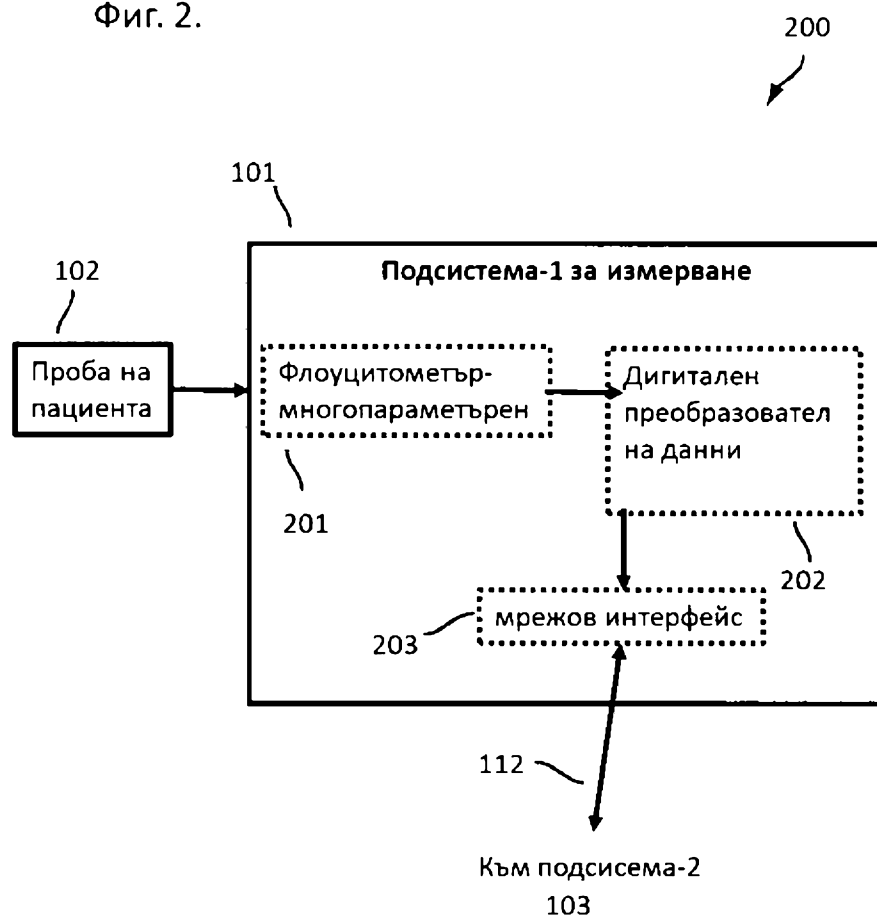
1. Guldberg, S. M., et al., Computational Methods for Single-Cell Proteomics. *Annu Rev Biomed Data Sci*, 2023. 6: p. 47-71.
2. Montante, S. and R. R. Brinkman, Flow cytometry data analysis: Recent tools and algorithms. *Int J Lab Hematol*, 2019. 41 Suppl 1: p. 56-62.
3. EP 2347352 B1, Interactive tree plot for flow cytometry data.
4. US 2003078703 A1 Cytometry analysis system and method using database-driven network of cytometers. 2003.
5. US 5605805 A2 Automatic lineage assignment of acute leukemias by flow cytometry 1997.
6. EP 1785899 B1, Methods for identifying discrete populations (e.g. clusters) of data within a flow cytometer multi-dimensional data set.
7. US 8214157 B2, Method and apparatus for representing multidimensional data.
8. WO 2006089190 A2, System, method, and article for detecting abnormal cells using multi-dimensional analysis.
9. Jafari, K., et al., Visualization of Cell Composition and Maturation in the Bone Marrow Using 10-Color Flow Cytometry and Radar Plots. *Cytometry B Clin Cytom*, 2018. 94(2): p. 219-229.

10. Karai, B., et al., A Novel Method for the Evaluation of Bone Marrow Samples from Patients with Pediatric B-Cell Acute Lymphoblastic Leukemia-Multidimensional Flow Cytometry. *Cancers (Basel)*, 2021. 13(20).
11. Shopsowitz, K., et al., MAGIC-DR: An interpretable machine-learning guided approach for acute myeloid leukemia measurable residual disease analysis. *Cytometry B Clin Cytom*, 2024.
12. Shopsowitz, K. E., et al., Machine learning optimized multiparameter radar plots for B-cell acute lymphoblastic leukemia minimal residual disease analysis. *Cytometry B Clin Cytom*, 2022. 102(5): p. 342-352.
13. Reiter, M., et al., Automated Flow Cytometric MRD Assessment in Childhood Acute B-Lymphoblastic Leukemia Using Supervised Machine Learning. *Cytometry Part A*, 2019. 95(9): p. 966-975.

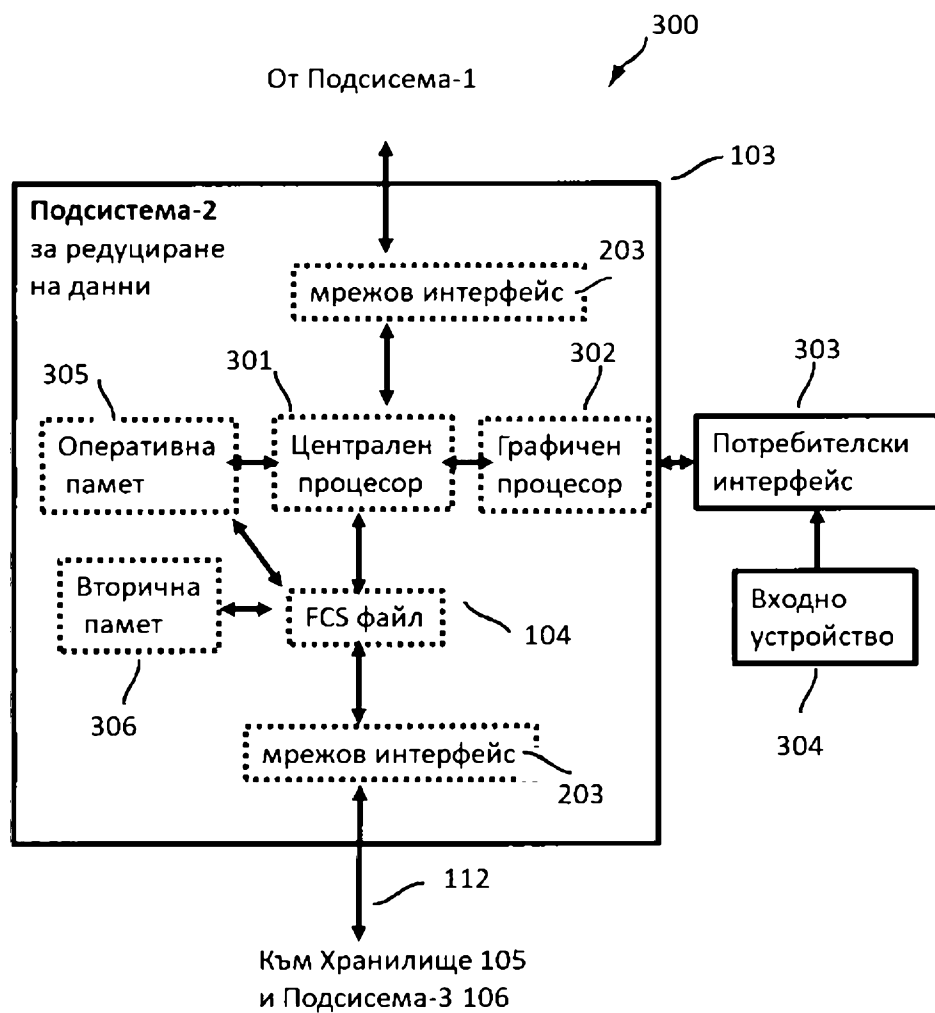
Фиг. 1.



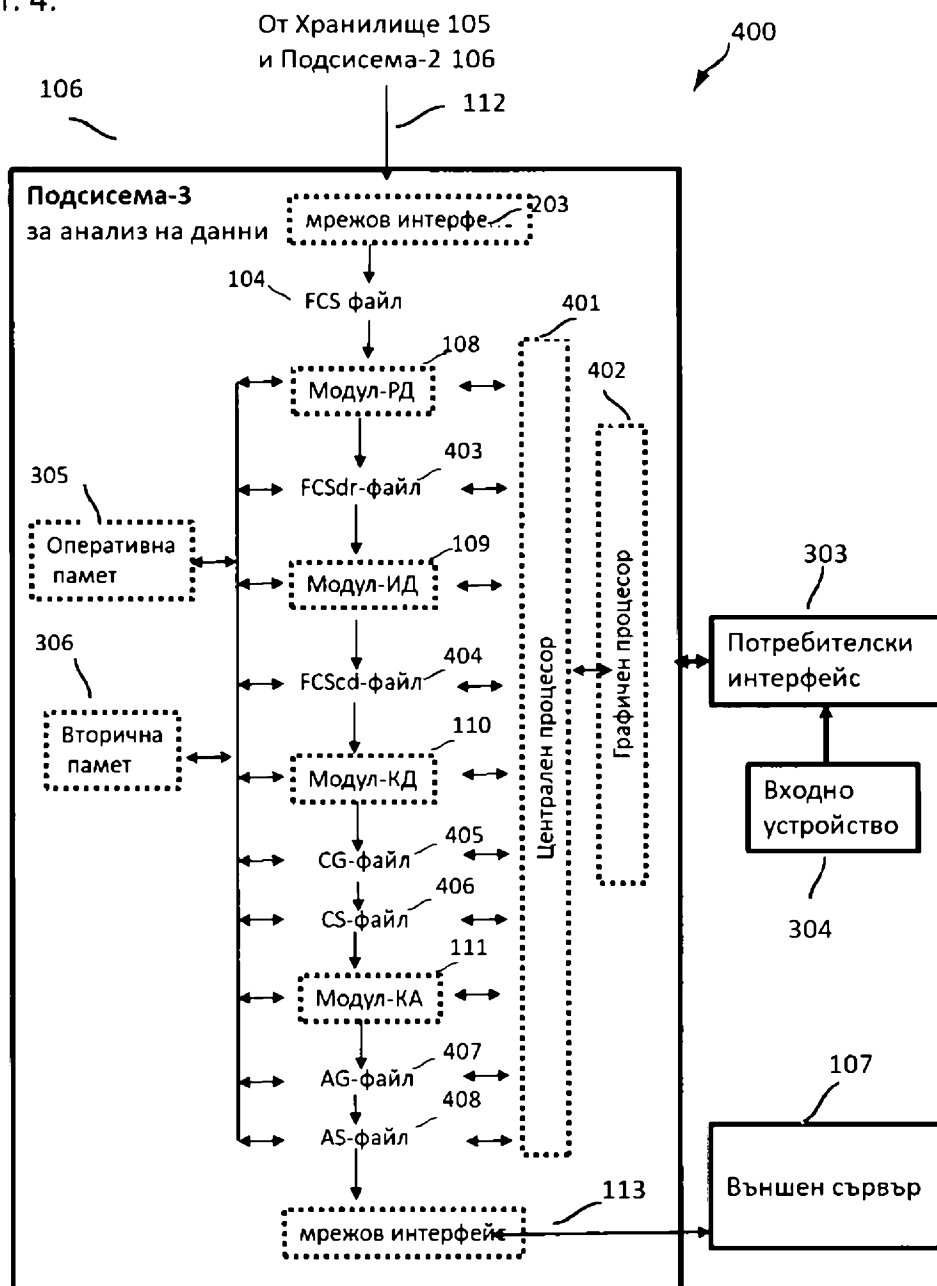
Фиг. 2.



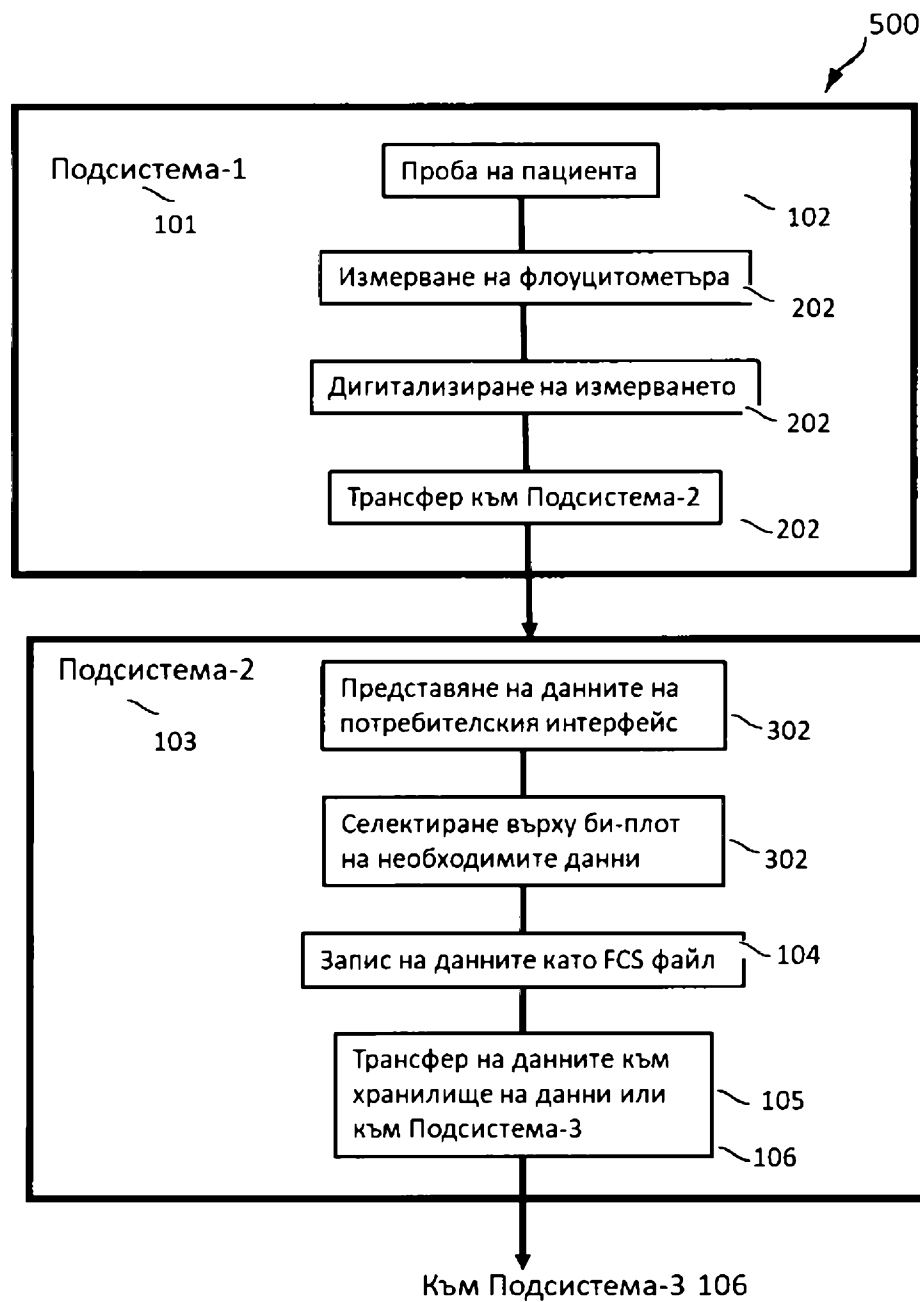
Фиг. 3.



Фиг. 4.

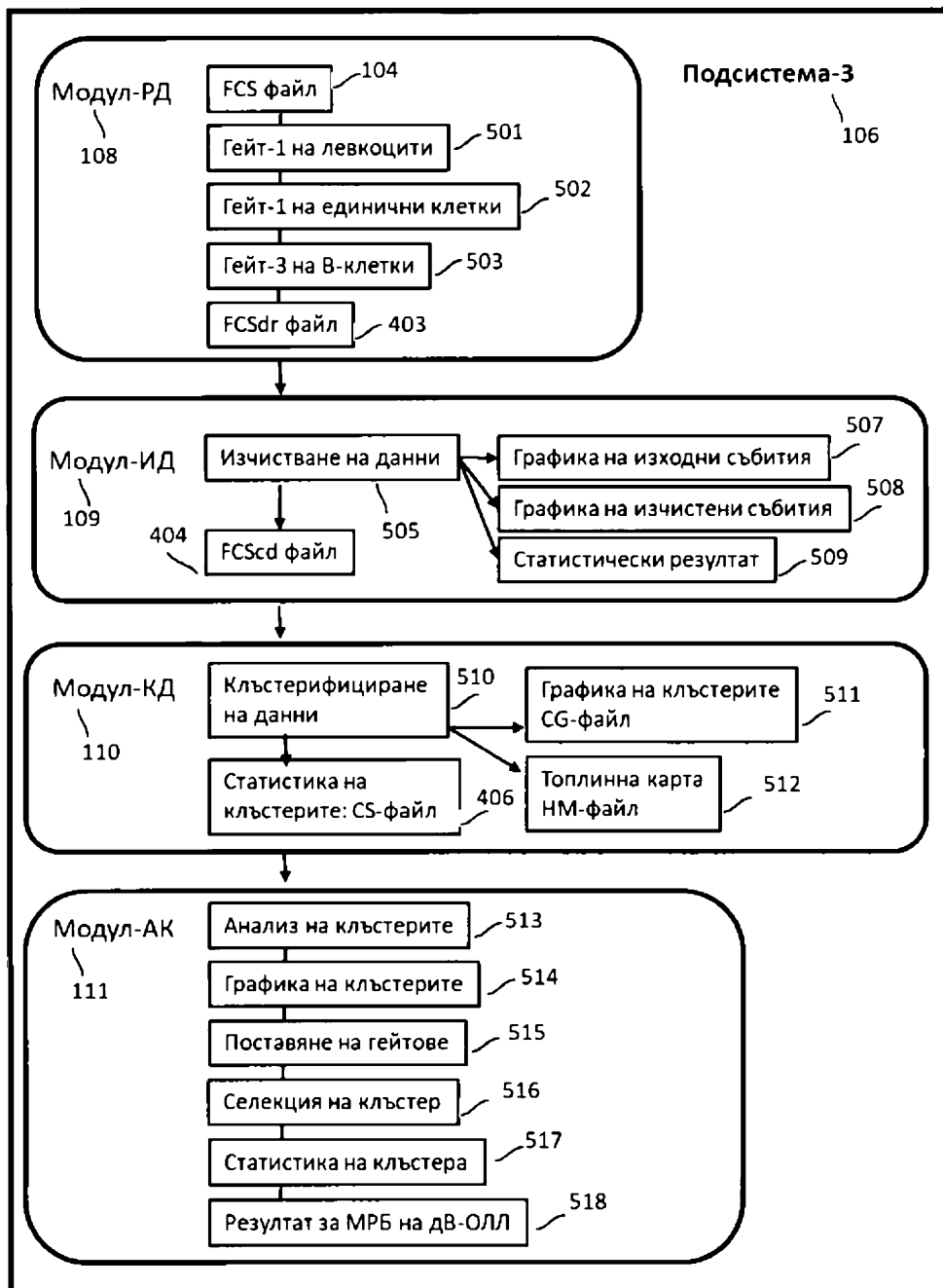


Фиг. 5

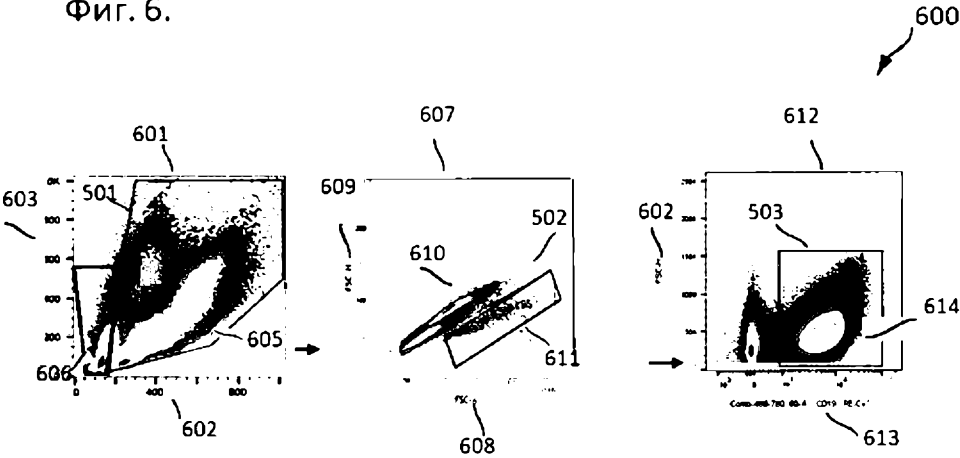


Продължава на следващата страница

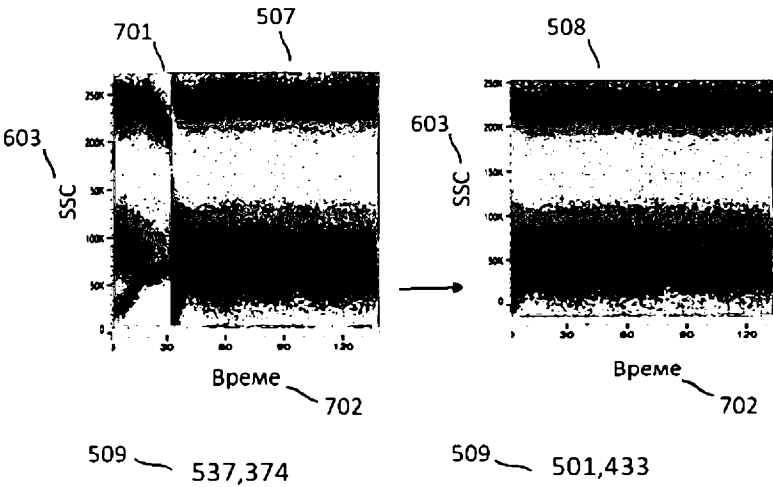
Фиг. 5. продължение



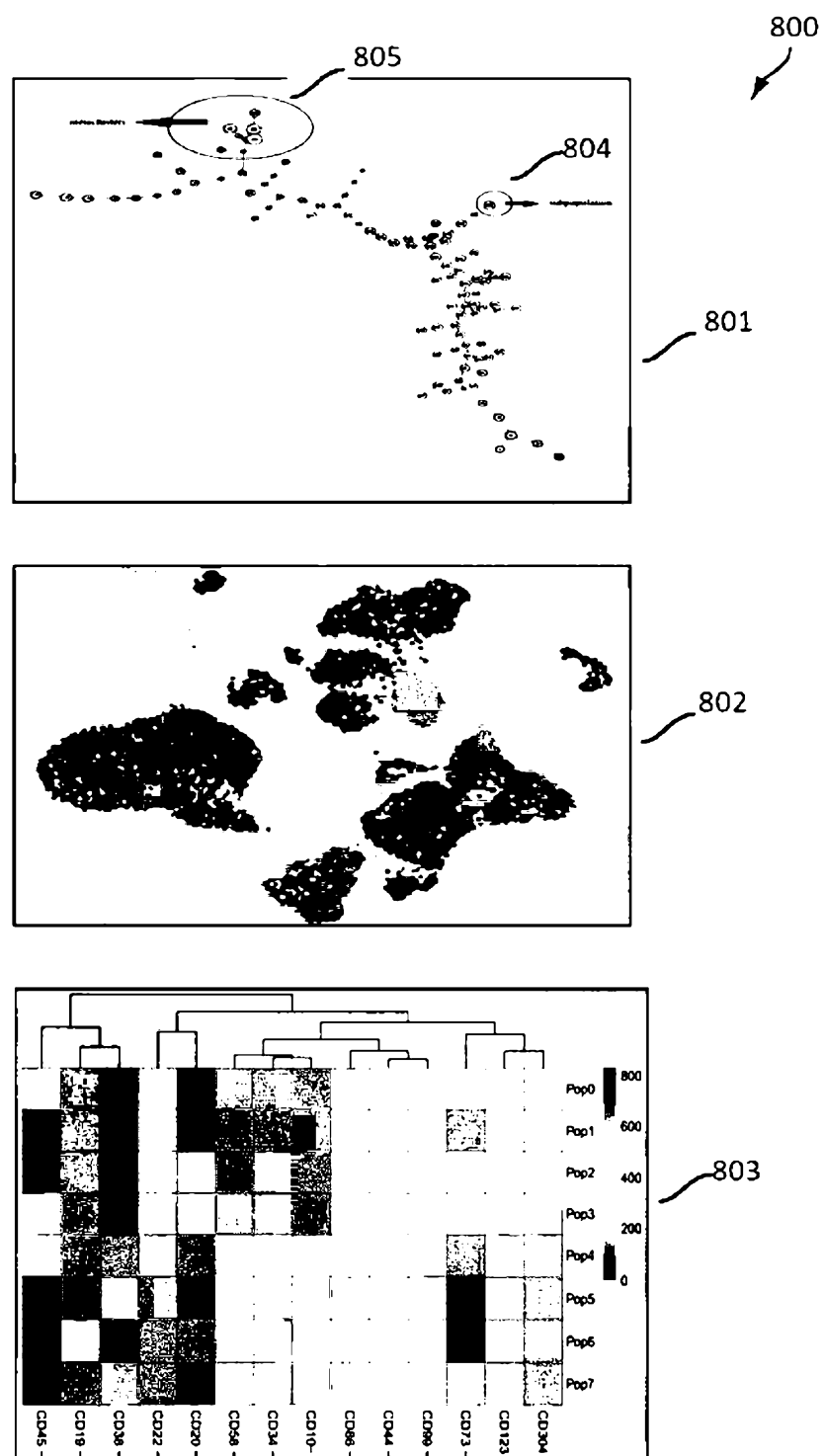
Фиг. 6.



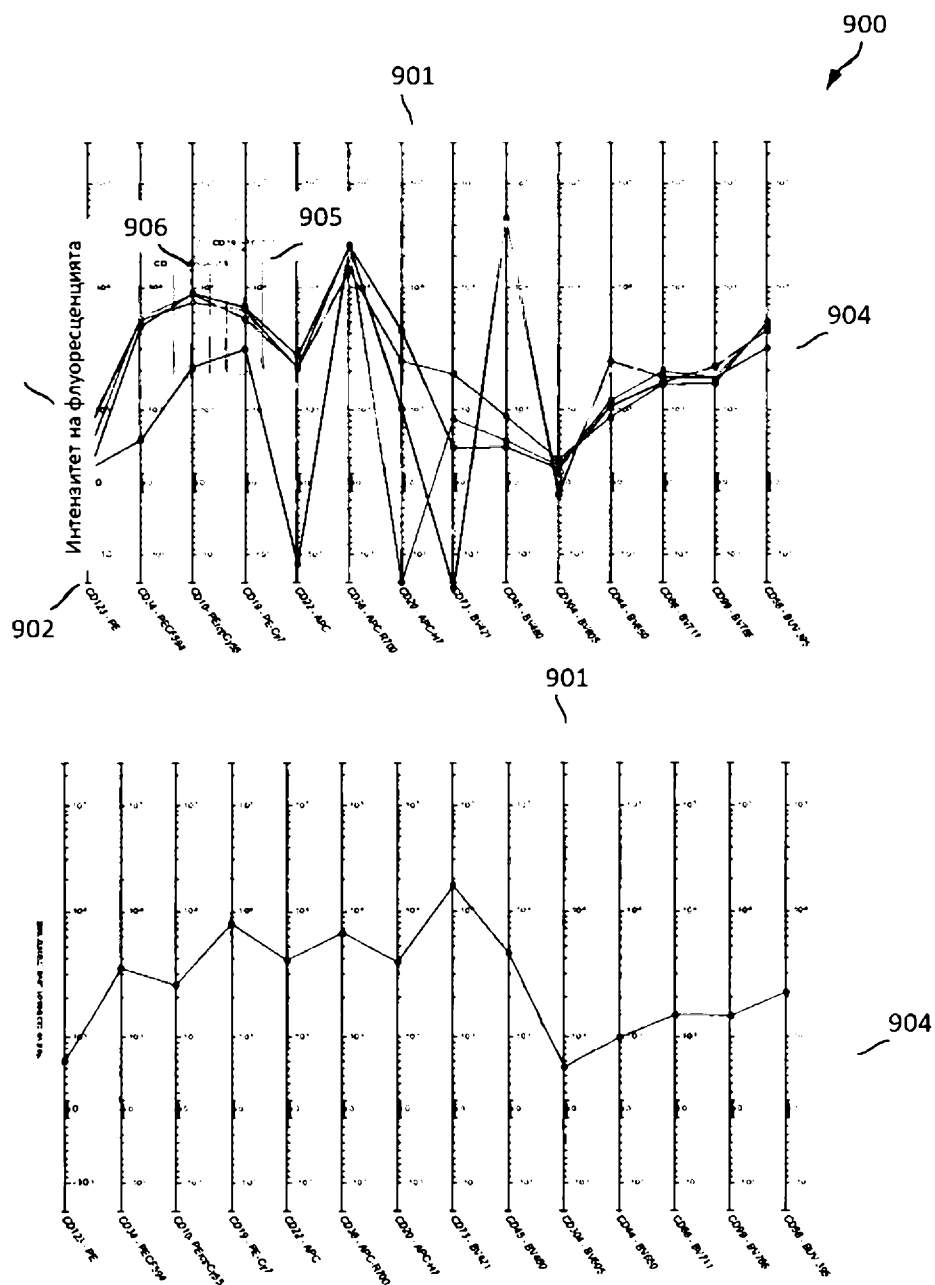
Фиг. 7.



Фиг. 8.



ФИГ. 9



Фиг. 10.

