

## Discriminator Level を用いた異常サンプルの検出方法についての検討

根室事業所

### 1. 目的

FossomaticFC による体細胞数検査において、測定機が正常であっても、サンプルの性状によって正しい測定ができない場合があることがわかった。そのようなサンプルを測定すると、PHA の波形に異常が見られ、繰り返し測定したときに体細胞数のバラツキが大きく、Discriminator Level (以下、DL) の値が通常よりも高いといった特徴が見られた(図1)。また、こうしたサンプルは正常なサンプルに比べ、次のサンプルへのキャリーオーバーの影響が大きいことが窺えた。

DL とは PHA 上において、バックグラウンドの明るさと体細胞数のパルスを区別するための閾値であり、ソフトウェア上でサンプル毎に自動計算されるものである。DL を下回るパルスは全てノイズと見なされ、体細胞としてカウントされない。上記のような異常サンプルでは DL が正しく計算されずに高い値となるので、パルスのほとんどがノイズとして処理され、結果として体細胞数は低い値で算出される。

つまり、こうしたサンプルの測定値は信頼性が低いと考えられ、またキャリーオーバーの恐れもあることから、検査中に何らかの形で検出し対応する必要がある。

こうしたサンプルは初乳のような粘性の高いものに多く見られるものの、正常に見受けられるサンプルであっても同様の現象が発生するため、事前に目視で検出することは困難である。また、現状の精度管理基準では、このような異常サンプルを検出することが出来ない。

そこで、DL が通常よりも高いという特徴を利用し、ソフトウェアでリミットチェックを設けることで異常サンプルを自動的に検出できないか検討を行った。

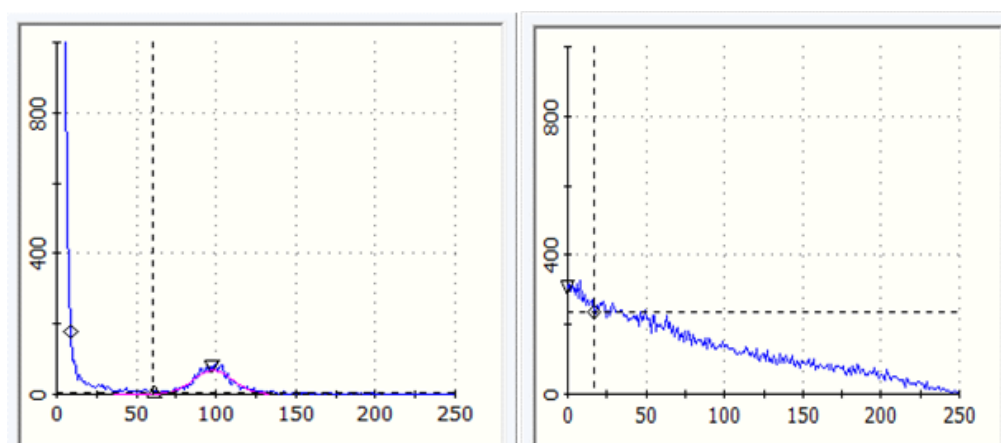


図1 正常な PHA(左)と異常な PHA(右)

## 2.方法

### 試料

平成 27 年 8 月～10 月の 2 ヶ月間において、根室事業所で検査された体細胞数依頼検査サンプル、40,467 検体のデータを用いて分析を行った。

### 分析方法

#### . 仮閾値の設定

全てのサンプルについて詳細な調査を行うことは不可能であるため、DL に仮の閾値を設定し、それを超えるサンプルについてデータを収集して分析することとした。本試験に入る前の事前調査として、5,904 検体の体細胞数依頼検査サンプルの DL の分布を調査したところ、その 99%は 60 以下であったため（表 1）、仮の閾値を「60」として設定し、DL 上位 1%のサンプルについてデータを収集した。

区間	0～10	11～20	21～30	31～40	41～50	51～60
頻度	18	23	818	2125	2344	517
累積%	0.54%	0.69%	14.55%	50.54%	90.24%	99.00%

表 1 DL の度数分布 (5,904 検体)

区間	61～70	71～80	81～90	91～100	101～
頻度	44	6	1	1	7
累積%	99.75%	99.85%	99.86%	99.88%	100.00%

#### . 対象サンプルのデータ収集

仮閾値を越えたサンプルについて、2 回以上繰り返し測定を行い、そのときの体細胞数、DL、PHA 波形についてデータを収集した。

#### . データの分類

各サンプルについて、測定機による測定が正しく行われているかを判別するために、3 つの条件を用いてサンプルを振り分け、最終的にサンプルを 4 種類に分類した。分類の条件は以下の通りである（図 2）。

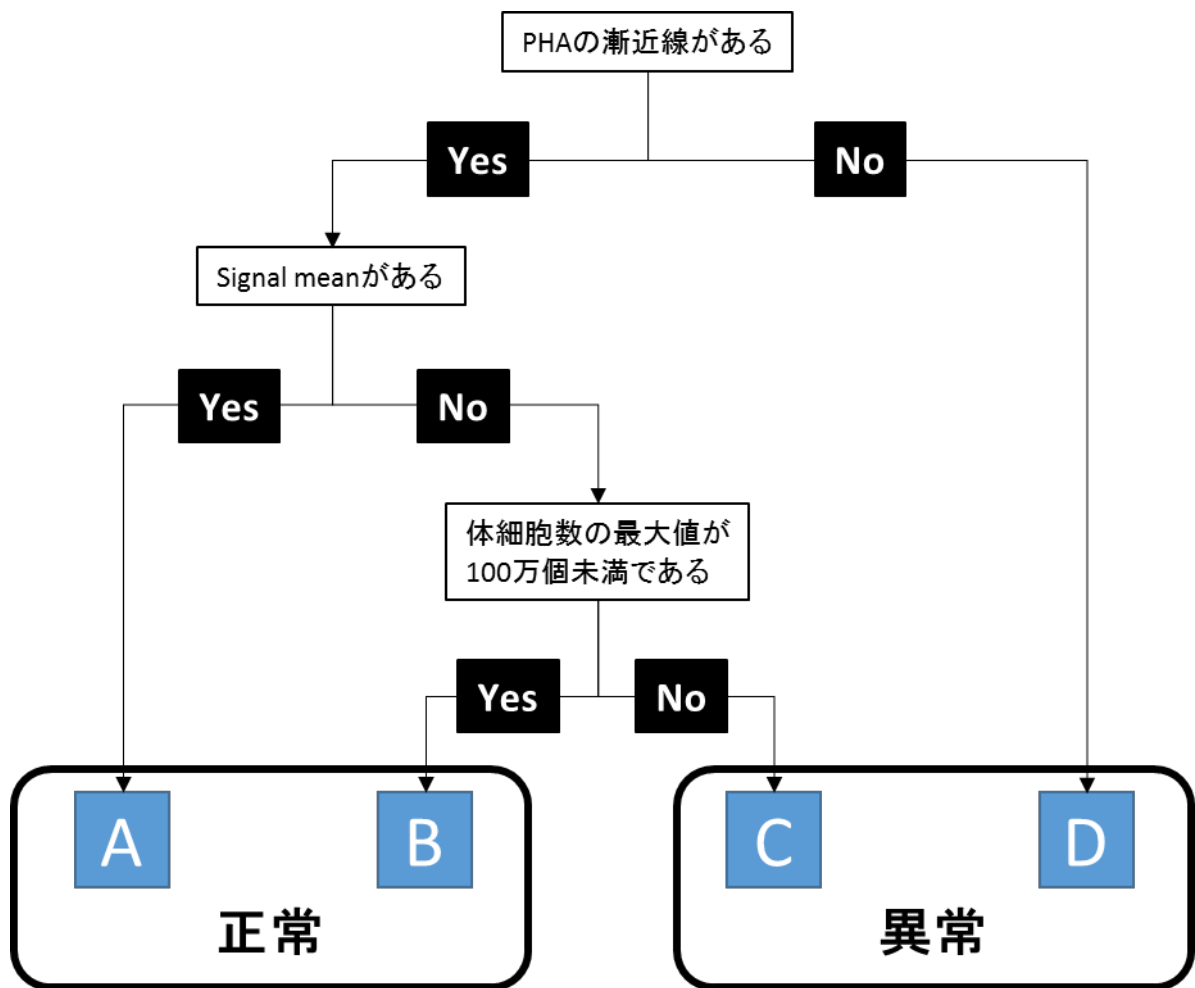


図2 分類わけの模式図

### 第1条件 PHAの漸近線がある

PHAの漸近線とは、バックグラウンドの明るさを示す部分である。この条件を満たさないサンプルはバックグラウンドと体細胞パルスの識別ができないということになるため、異常なサンプル(パターンD)と判断した。

### 第2条件 Signal Meanがある

第1条件を満たすもののうち、Signal Mean(以下SM)が計算されているものは、体細胞パルスを測定機の仕様通りに識別出来ているといえるため、正常(パターンA)であるとみなした。

### 第3条件 繰り返し測定時も含め、体細胞数が100万個/mL未満である

SMが算出されないこと自体は、そもそもの体細胞数が少ないサンプルの場合に起こりうることである。ただし一定以上の体細胞数でありながらSMが算出されないことは、何かしらの不良が疑われる。そこで第1条件を満たし、第2条件を満たさないサンプルについては体細胞数で分類することとした。ここではキャリーオーバーの危険性も

加味し、繰り返し測定したときに1回も体細胞数が100万個/mLを超えないものは正常（パターンB）とみなし、繰り返し測定時に1回でも100万個/mLを超えるものについては異常なサンプル（パターンC）であると判定した。

#### .DL 閾値の妥当性検討

上記の条件でサンプルを4種類に分類した後、表2の分類方法を基にしてDLリミットを任意の値に変更したときのサンプル数をそれぞれ集計した。また、そのときの感度・特異度・有効度・偽陽性率・偽陰性率を計算し、妥当と考えられるDL閾値を検討した。

表2 DL 閾値妥当性検討のための分類方法と各パラメータの計算方法

	正常 (パターン A or B)	異常 (パターン C or D)
DL リミット未満	真陰性	偽陰性
DL リミット以上	偽陽性	真陽性

$$\text{感度} = \text{真陽性} \div (\text{偽陰性} + \text{真陽性})$$

$$\text{特異度} = \text{真陰性} \div (\text{真陰性} + \text{偽陽性})$$

$$\text{有効度} = (\text{真陰性} + \text{真陽性}) \div \text{サンプル総数}$$

$$\text{偽陽性率} = \text{偽陽性} \div \text{サンプル総数}$$

$$\text{偽陰性率} = \text{偽陰性} \div \text{サンプル総数}$$

ここでの集計において、サンプルのDLには繰り返し測定時の最大のDLをそのサンプルのDLとして採用し、データ分析を行った。

例：あるサンプルを3回測定し、そのときのDLがそれぞれ20、80、140であった場合、このサンプルのDLは140として集計する。

またパターンCおよびDの異常と判定されたサンプルにおいて、繰り返し測定時に体細胞数が1000万個/mLを超えたものは、現状の精度管理基準下で機器エラーとなってサンプルを検出することが可能なため、設定したDLリミットに関わらず全て真陽性としてカウントした。このようなサンプルはパターンDにおいて1検体が該当した。

### 3. 結果

調査期間における対象サンプル40,467検体のうち、仮閾値を越えたものは509検体であった。仮閾値を越えなかった残りの39,958検体については全て正常なサンプルであるとみなした。DLリミットを任意の値に設定し、図2および表2の方法でサンプルを分類したときのサンプルの実本数は表3の通りである。これを元に計算される各パラメータの数値については表4の通りである。

表 3 対象サンプルを分類わけしたときの実本数

DL リミット 設定値	分類	正常		異常	
		パターン A	パターン B	パターン C	パターン D
60	60 未満	39958	0	0	0
	60 以上	314	171	1	23
70	70 未満	40266	108	0	0
	70 以上	6	63	1	23
80	80 未満	40270	121	0	0
	80 以上	2	50	1	23
90	90 未満	40272	127	0	0
	90 以上	0	44	1	23
100	100 未満	40272	132	0	1
	100 以上	0	39	1	22
110	110 未満	40272	140	0	1
	110 以上	0	31	1	22
120	120 未満	40272	149	0	1
	120 以上	0	22	1	22
130	130 未満	40272	163	0	1
	130 以上	0	8	1	22
140	140 未満	40272	170	1	1
	140 以上	0	1	0	22
150	150 未満	40272	170	1	2
	150 以上	0	1	0	21
160	160 未満	40272	171	1	2
	160 以上	0	0	0	21
170	170 未満	40272	171	1	3
	170 以上	0	0	0	20

表 4 任意の DL でリミットを設けたときの各パラメータ

DL リミット値	感度(%)	特異度(%)	有効度(%)	偽陽性率(%)	偽陰性率(%)
60	100.00	98.80	98.80	1.199	0.000
70	100.00	99.83	99.83	0.171	0.000
80	100.00	99.87	99.87	0.128	0.000
90	100.00	99.89	99.89	0.109	0.000
100	95.83	99.90	99.90	0.096	0.002
110	95.83	99.92	99.92	0.077	0.002
120	95.83	99.95	99.94	0.054	0.002
130	95.83	99.98	99.98	0.020	0.002
<b>140</b>	<b>91.67</b>	<b>100.00</b>	<b>99.99</b>	<b>0.002</b>	<b>0.005</b>
150	87.50	100.00	99.99	0.002	0.007
160	87.50	100.00	99.99	0.000	0.007
170	83.33	100.00	99.99	0.000	0.010

分析の結果、有効度が最も高くなるのは DL リミット値を 140 以上に設定したときであった。リミット値を 130 以下にした場合、有効度が下がり、偽陽性のサンプル数が増加する結果となった。リミット値 140 以上は有効度 99.99% で変わらないが、150 以上に設定すると 140 に設定したときと比べ偽陰性のサンプル数が増加する結果となった。170 以上に設定した場合はリミット値 170 に設定したときと同じ結果であった。

以上のことから、DL リミットによるエラー検出を最も有効的に活用できる設定値は 140 であると考えられる。またこのとき、偽陽性のサンプル数が 1 検体、偽陰性のサンプル数が 2 検体となっているが、総数 40,467 検体に対して極々わずかであり問題にならないものと考えられる。

DL にリミットを新たに設けることで、間接的にサンプルの性状を監視することが出来、依頼元に対して信頼性の低い検査結果を返すリスクを回避することが可能である。

(中川大輔、中野まどか)