Detection of GD_2 negative cells in bone marrow from patients with high-risk neuroblastoma using machine learning

Leah Kelly, Zoe Bell, Marina Danilenko, David Jamieson, Deborah Tweddle

Introduction:

Neuroblastoma (NB) is the most common extracranial solid tumour in children, accounting for approximately 15% of all paediatric cancer-related deaths(1). Patients with high-risk neuroblastoma (HR-NB) (defined in Europe as metastatic disease over 1 year of age or *MYCN* amplified) are given immunotherapy which targets disialoganglioside (GD_2), a tumour-associated antigen expressed by almost all NB patients(2,3). NB often metastasises to the bone marrow (BM) and the presence of GD_2 expression and the absence of CD45 expression can be used to detect BM metastasis. However, loss of GD_2 expression in NB BM metastases has been reported(2). As anti-GD_2 immunotherapy is routinely used in treatment, it is crucial to be able to detect and characterise GD_2-ve cells. This study involved data collected from a previous study, using the ImageStream X Imaging Flow Cytometer (ISx), and aimed to detect GD_2-ve disseminated tumour cells (DTCs) from BM samples from HR-NB patients using a machine learning (ML) algorithm.

Methods:

Patient samples

This study included 21 patients for analysis of DTCs. ISx data, obtained from BM samples from a previous study, was used(3).

Software used

Amnis IDEAS 6.4 image analysis software was used to analyse patient data files, generated using the ISx from a previous study(3).

Identification of truth and base populations for machine learning

To create the 'base' population for machine learning, round single cells were gated on a scatter plot of area versus aspect ratio intensity, excluding multiple cells and fragments (**Fig. 1A**), and in focus cells were gated on a histogram of brightfield (Bf) gradient root mean square (RMS) (**Fig. 1B**).

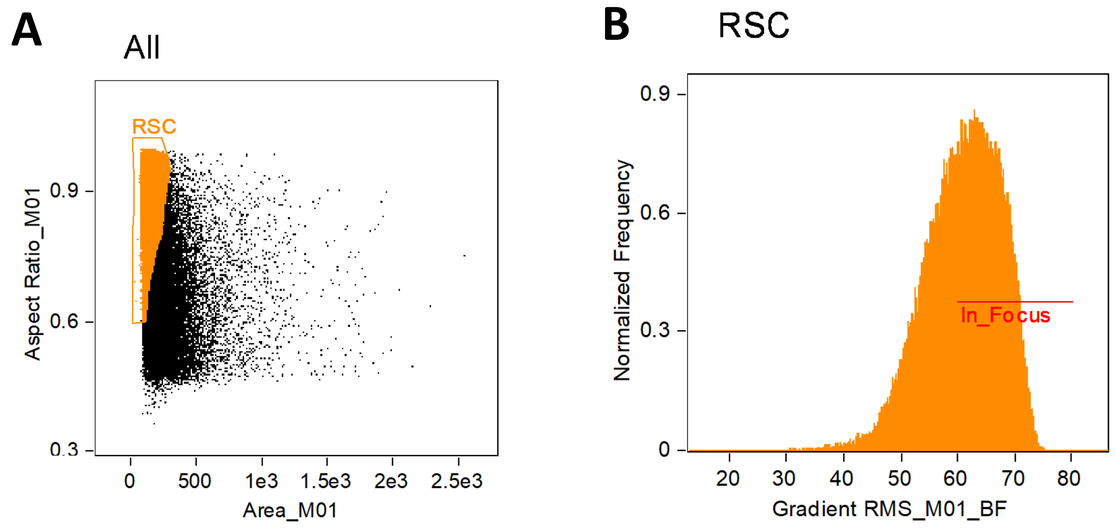

Figure 1. *Creation of 'base' population for ML.* ***A.*** *Scatter plot of area versus aspect ratio intensity, gated to include round single cells. B. Histogram of Bf gradient RMS, gated to include in focus cells.*

25 GD_2+ve/CD45-ve DTCs, with an intact DAPI-stained nucleus, were visually identified from a CD45 versus GD_2 intensity scatter plot (**Fig. 2A**) and tagged as a 'truth' population for machine learning (**Fig. 2B**).

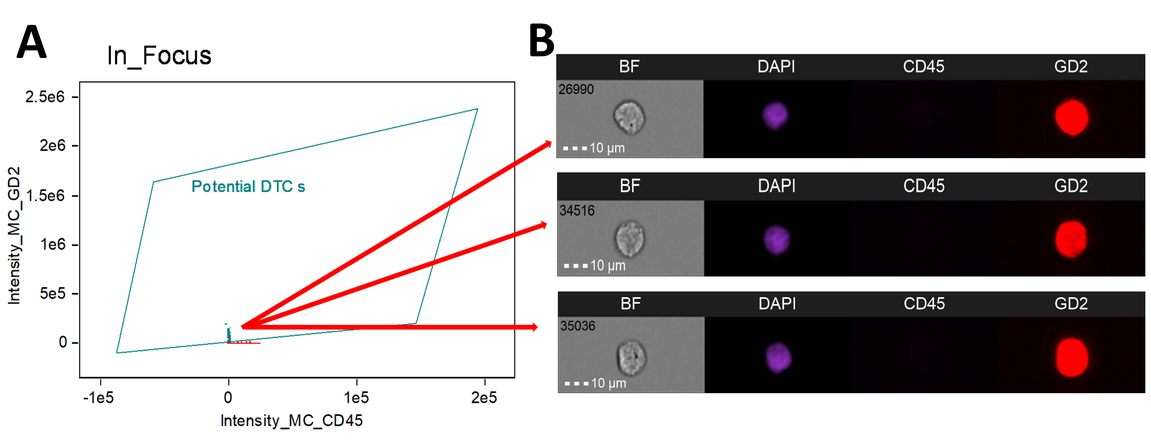

Figure 2. *Identification of DTC 'truth' population.* ***A.*** *Scatter plot of CD45 vs GD_2 intensity, gated to include GD_2+ve/CD45-ve cells.* ***B.*** *Example images of GD_2+ve/CD45-ve cells included in the DTC 'truth' population for ML.*

Initially, 200 cells, visually confirmed as 'non-DTCs', were identified and tagged as another 'truth' population. However, as this separated the 'truth' populations poorly, an additional 3 'truth' populations, corresponding to areas of high CD45, low CD45, and 'false positive' DTCs identified by the initial ML classifier, were included, with 25 cells in each additional 'truth' population.

Machine Learning (ML)

The ML module within IDEAS was used to compute classifier features to separate the manually identified truth populations and combine them into a single score using Linear Discriminant Analysis. 'Truth' and 'base' populations, previously identified, were selected, and all features within the Size, Comparison, Shape, Texture and Location categories for the Bf and DAPI channels were included. The NCL003 ML classifier template maximally separated 'truth' populations and was therefore applied as an inter-sample ML classifier template.

Results:

The NCL003 ML classifier template was applied to all BM files from all 21 patients studied. Of these patients, 3/21 were selected as negative controls, with no DTCs detected in any files; all 3 were low-risk NB cases. In the remaining 18 HR-NBL cases, the NCL003 ML classifier detected GD_2+ve DTCs in 14/18 patients (1-63310 DTCs/mL) and GD_2-ve DTCs in 17/18 patients (1-5448 DTCs/mL) (**Table 1**).

For 16/21 patients studied, GD_2-ve and GD_2+ve circulating tumour cells (CTCs)/mL had previously been detected and counted in matched blood samples (**Table 1**)(4). In the blood, the NCL003 ML classifier detected GD_2+ve CTCs in 9/13 high-risk patients (1-86 CTCs/mL) and GD_2-ve CTCs in 11/13 (1-98 CTCs/mL) (**Table 1**)(4).

Table 1. Summary table comparing the number of GD_2+ve and GD_2-ve DTCs per mL to the number of CTCs detected per mL detected by the NCL003 ML classifier template in relation to risk group, ploidy status, MYCN amplification status and BM involvement.

Patient	Diagnosis/Relapse	Risk	Ploidy status	*MYCN* amplified	BM involved (Yes/No)	GD2+ve DTCs/mL detected by NCL003 ML Template	GD2-ve DTCs detected by NCL003 ML Template	GD2+ve CTCs/mL detected by NCL003 ML Template	GD2-ve CTCs/mL detected by NCL003 ML Template
BRI005	Diagnosis	Low	-	Non-MYCN amplified	No	0	0	0	0
NCL001	Relapse	Low	-	Non-MYCN amplified	No	0	0	0	0
MAN006	Relapse	Low	-	Non-MYCN amplified	No	0	0	0	0
NCL003	Diagnosis	High	Hyperdiploid	MYCN amplified	Yes	16311	5448	86	28
NCL005	Diagnosis	High	Hyperdiploid	Non-MYCN amplified	Yes	242	23	3	1
NCL011	Diagnosis	High	Hyperdiploid	MYCN amplified	No	8245	932	14	4
BRI001	Diagnosis	High	Diploid	Non-MYCN amplified	Yes	41	27	0	1
NCL012	Diagnosis	High	Diploid	Non-MYCN amplified	No	365	34	1	0
BRI004	Diagnosis	High	Diploid	Non-MYCN amplified	Yes	406	2	6	2
MAN003	Diagnosis	High	Hyperdiploid	Non-MYCN amplified	Yes	63310	2816	34	11
MAN004	Diagnosis	High	Diploid	Non-MYCN amplified	Yes	16	129	0	12
MAN005	Diagnosis	High	Diploid	Non-MYCN amplified	Yes	0	765	0	5
GLA001	Diagnosis	High	Diploid	Non-MYCN amplified	Yes	169	11	3	0
GLA002	Diagnosis	High	Diploid	Non-MYCN amplified	Yes	712	156	14	28
NCL004	Relapse	High	Hyperdiploid	Non-MYCN amplified	No	0	1370	0	98
NCL008	Relapse	High	Hyperdiploid	MYCN amplified	No	1	1	6	31
NCL013	Relapse	High	Hyperdiploid	MYCN amplified	No	162	9	-	-
NCL007	Diagnosis	High	Hyperdiploid	MYCN amplified	Yes	0	1	-	-
NCL009	Diagnosis	High	Diploid	Non-MYCN amplified	Yes	12363	239	-	-
MAN002	Diagnosis	High	Diploid	MYCN amplified	Yes	0	0	-	-
GLA004	Relapse	High	Diploid	MYCN amplified	Yes	31	2	-	-
					Total	102374.0	11965.0	167.0	221.0
					Mean	4875.0	569.8	10.4	13.8
					St Dev	14136.6	1308.1	22.0	25.0

Notably, the NCL003 ML classifier was able to detect GD_2-ve, hyperdiploid DTCs and CTCs in both the BM and the blood for NCL004, a high-risk, relapse patient with no GD_2+ve DTCs or CTCs.

Detection of GD_2-ve DTCs using ML

GD_2-ve DTCs were visually confirmed to have similar Bf and nuclear morphology as GD_2+ve DTCs. Ploidy status of both the GD_2-ve and GD_2+ve populations was confirmed to be the same as the primary tumour **(Fig. 3B**).

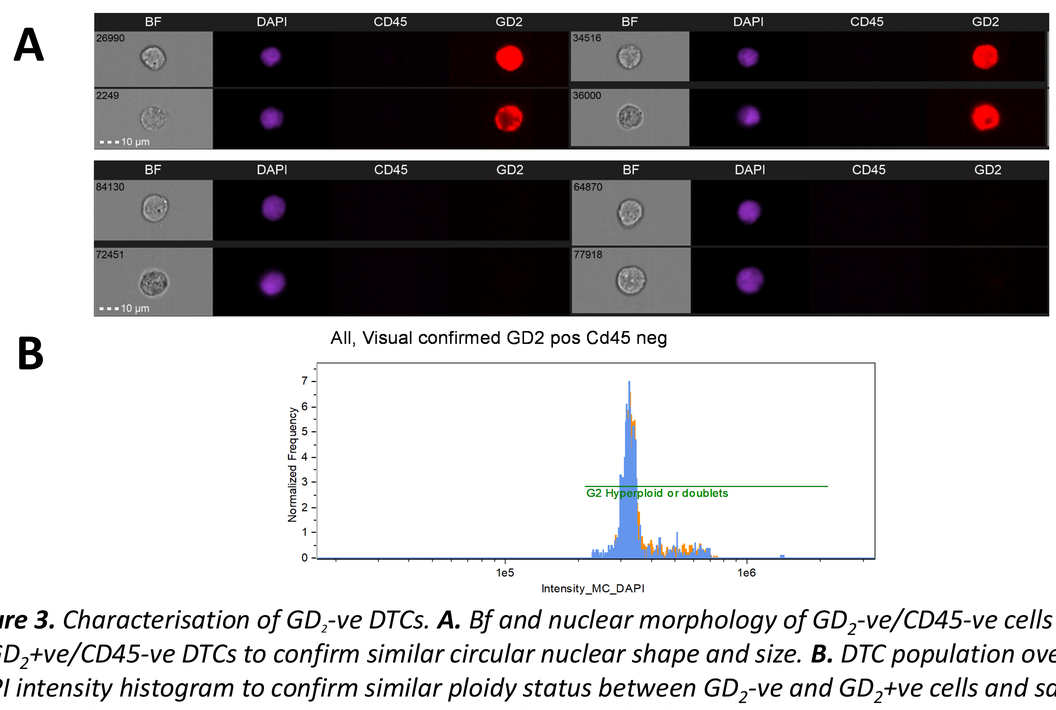

Figure 3. *Characterisation of GD_2-ve DTCs.* ***A.*** *Bf and nuclear morphology of GD_2-ve/CD45-ve cells compared to GD_2+ve/CD45-ve DTCs to confirm similar circular nuclear shape and size.* ***B.*** *DTC population overlaid onto DAPI intensity histogram to confirm similar ploidy status between GD_2-ve and GD_2+ve cells and same ploidy status as primary tumour. In this case the tumour is hyperdiploid.*

Conclusions:

This study demonstrated that GD_2-ve DTCs can be detected in BM samples from HR-NB patients, indicating that this method could be utilised for the detection of GD_2 loss following anti-GD_2 directed immunotherapy. However, the findings should be extended to a larger cohort of patients to confirm clinical relevance. This would allow the prognostic significance of a GD_2-ve cell population to be evaluated.

Acknowledgements:

With thanks to my supervisors Professor Deborah Tweddle, Dr David Jamieson and Dr Marina Danilenko, as well as all patients and families for taking part in this study, and to Newcastle University for a Research Scholarship to support this project.

References:

1. Colon NC, Chung DH. Neuroblastoma. Adv Pediatr. 2011;58(1):297-311.
2. Schumacher-Kuckelkorn R, et al. Pediatr Blood Cancer. 2017;64(1):46-56.
3. Merugu S, et al. Clin Cancer Res. 2020 Jan 1;26(1):122-134.
4. Gordon F. Final Year Undergraduate Project Dissertation. 2022.