Flow cytometry of bone marrow aspirates from neuroblastoma patients is a highly sensitive technique for quantification of low-level neuroblastoma [version 1; peer review: 1 approved, 1 approved with reservations]

Neha Jain1*, Shaista Sattar1*, Sarah Inglott1, Susan Burchill2, Jonathan Fisher3, Andreea-Madalina Serban1, Rebecca Thomas1, Chris Connor1, Niharendra Ghara1, Tanzina Chowdhury1, Catriona Duncan1, Giuseppe Barone1, John Anderson3

1Great Ormond Street Hospital for Children, London, WC1N 3JH, UK
2Leeds Institute of Medical Research, St James University Hospital, Leeds, LS9 7TF, UK
3University College London Institute of Child Health, London, WC1N 3JH, UK

* Equal contributors

Abstract

Background: Bone marrow involvement is an important aspect of determining staging of disease and treatment for childhood neuroblastoma. Current standard of care relies on microscopic examination of bone marrow trephine biopsies and aspirates respectively, to define involvement. Flow cytometric analysis of disaggregated tumour cells, when using a panel of neuroblastoma specific markers, allows for potentially less subjective determination of the presence of tumour cells.

Methods: A retrospective review of sequential bone marrow trephine biopsies and aspirates, performed at Great Ormond Street Hospital, London, between the years 2015 and 2018, was performed to assess whether the addition of flow cytometric analysis to these standard of care methods provided concordant or additional information.

Results: There was good concurrence between all three methods for negative results 216/302 (72%). Positive results had a concordance of 52/86 (61%), comparing samples positive by flow cytometry and positive by either or both cytology and histology. Of the remaining samples, 20/86 (23%) were positive by either or both cytology and histology, but negative by flow cytometry. Whereas 14/86 (16%) of samples were positive only by flow cytometry.

Conclusions: Our review highlights the ongoing importance of expert
cytological and histological assessment of bone marrow results. Flow cytometry is an objective, quantitative method to assess the level of bone marrow disease in aspirates. In this study, flow cytometry identified low-level residual disease that was not detected by cytology or histology. The clinical significance of this low-level disease warrants further investigation.

**Keywords**
neuroblastoma, paediatric cancer, bone marrow, flow cytometry, histology, RTqPCR, minimal residual disease

This article is included in the UCL Child Health gateway.

This article is included in the University College London collection.

Corresponding author: John Anderson (j.anderson@ucl.ac.uk)

Author roles: Jain N: Data Curation, Formal Analysis, Investigation, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing; Sattar S: Data Curation, Formal Analysis, Investigation, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing; Inglof S: Methodology, Resources, Supervision, Validation, Writing – Review & Editing; Burchill S: Methodology, Resources, Supervision, Validation, Writing – Review & Editing; Fisher J: Data Curation, Formal Analysis, Investigation, Methodology, Writing – Review & Editing; Serban AM: Data Curation, Investigation, Writing – Review & Editing; Thomas R: Data Curation, Investigation, Writing – Review & Editing; Connor C: Data Curation, Investigation, Writing – Review & Editing; Ghara N: Data Curation, Investigation, Writing – Review & Editing; Chowdhury T: Conceptualization, Methodology, Supervision, Writing – Review & Editing; Duncan C: Conceptualization, Methodology, Supervision, Writing – Review & Editing; Barone G: Conceptualization, Methodology, Supervision, Writing – Review & Editing; Anderson J: Conceptualization, Funding Acquisition, Methodology, Resources, Software, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: The author(s) declared that no grants were involved in supporting this work.

Copyright: © 2021 Jain N et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Jain N, Sattar S, Inglof S et al. Flow cytometry of bone marrow aspirates from neuroblastoma patients is a highly sensitive technique for quantification of low-level neuroblastoma [version 1; peer review: 1 approved, 1 approved with reservations] F1000Research 2021, 10:947 https://doi.org/10.12688/f1000research.53133.1

First published: 21 Sep 2021, 10:947 https://doi.org/10.12688/f1000research.53133.1
Introduction

Neuroblastoma is the most common extracranial solid tumour of childhood (Xie, Onyskio and Morrison, 2018). A combination of stage of disease, patient age, tumour histology and tumour biology are used to risk stratify patients for treatment (Monclair et al., 2008). Metastatic disease in patients more than 18 months of age places a patient in the high-risk category. Consequently, accurate staging at the time of diagnosis is critical. These patients receive multimodal treatment with chemotherapy, myeloablative chemotherapy and autologous stem cell rescue, surgery, radiation therapy and immunotherapy. Approximately 50% of those diagnosed with neuroblastoma have high-risk stage M disease, with poor overall survival of <50% (Tas et al., 2020).

Consistent with recommendations from the International Neuroblastoma Risk Group (INRG), evaluation of both bone marrow cores (trephines) and aspirates is reported to most accurately detect bone marrow disease (Aronica et al., 1998); a combination of bilateral cores and aspirates is associated with 94.7% sensitivity (Parsons et al., 2017). Further, addition of immunohistochemistry to histological assessments can lead to increased inter-observer agreement (Parsons et al., 2016). The INRG Staging System (INRGSs) defines bone marrow infiltration as any involvement of bone marrow aspirate or trephines detected by the either or combination of cytology, histology, and/or immunohistochemical techniques, with >10% bone marrow involvement being one of the criteria used to distinguish between stage M and Stage MS disease (Monclair et al., 2008; Beiske et al., 2009; Burchill et al., 2017). A revision of the International Neuroblastoma Response Criteria (INRC) outlined that follow-up bone marrow samples with ≤5% involvement would represent minimal disease (Park et al., 2017).

Alternative methods, including flow cytometry, immunocytochemistry and quantitative reverse transcriptase polymerase chain reaction (RTqPCR), are currently being evaluated as more sensitive and specific methods for the detection of low-level disease than cytological or histological assessment (Tsang et al., 2003; Corrias et al., 2004; Beiske et al., 2005; Ferreira-Facio et al., 2013; Uemura et al., 2019). Flow cytometry is a well-validated method of detecting bone marrow infiltration in haematological malignancies but its role in solid paediatric cancers is not established. Studies have demonstrated that flow cytometry can detect disease at lower levels than histopathology (Komada et al., 1998; Tsang et al., 2003; Szantho et al., 2018). The first triple colour flow cytometry assay to detect neuroblastoma was developed in 1998 (Komada et al., 1998). This has subsequently been optimised and today CD45/CD56+CD81+/GD2+ cells by flow cytometry have been accepted to represent neuroblastoma cells (Swerts et al., 2004; Ferreira-Facio et al., 2013). Disialoganglioside (GD2) is detected in the vast majority of neuroblastoma cells, but also expressed by melanomas, gliomas and focally in rhabdomyosarcomas and osteosarcomas (Beiske et al., 2005; Ferreira-Facio et al., 2013). Importantly, GD2 is not expressed by normal bone marrow cells (Swerts et al., 2004). CD56 antibody is present on a subset of CD4+, and CD8+ T-cells and NK cells in peripheral blood, as well as neural derived cells and tumours (Beiske et al., 2005). CD45 is present on all human leukocytes but absent on neuroblastoma cells. Using flow cytometry, Komada et al. (1998) were able to detect a single neuroblastoma cell in up to 1 × 10⁷/10⁸ mononuclear cells. Szantho et al. (2018) analysed 36 samples from 16 patients and concluded that flow cytometry was highly specific and more sensitive than immunohistochemistry, as more cells can be evaluated. However, other studies have suggested that flow cytometry is 10-fold less sensitive than immunocytochemistry or quantitative reverse transcriptase polymerase chain reaction (RTqPCR) (Swerts et al., 2004; Uemura et al., 2019).

The role of minimal residual disease (MRD) in neuroblastoma is increasingly under investigation, although its clinical utility is yet to be defined. In haematological malignancies PCR-based detection of MRD has become part of the routine method for risk stratification and ongoing monitoring of patients during treatment, with an escalation in treatment if there is inadequate MRD response. Corrias et al. (2004) used immunocytochemistry to detect MRD in bone marrow of patients with localised neuroblastoma and found no significant difference in overall survival of patients with MRD compared to those without detectable MRD in bone marrow. In patients with metastatic disease there was no difference in overall survival by bone marrow disease detected by MRD using either immunocytochemistry or PCR techniques. The Children’s Oncology Group (COG) also showed no difference in overall survival for patients with localised disease that had bone marrow involvement detected by immunocytochemistry alone at diagnosis (Seeger et al., 2000). In the same study, COG demonstrated a clear correlation between increasing tumour burden in bone marrow and poor event free survival in those patients with stage M disease, but no difference in survival if bone marrow infiltration was only detected by immunohistochemistry and not by cytology (Seeger et al., 2000). Conversely, others have shown a poorer prognosis in those patients with neuroblastoma detected by flow cytometry but negative by immunophenotyping (Popov et al., 2019) and poor overall survival in those with neuroblastoma detectable by RTqPCR after induction therapy (Druy et al., 2018). These studies have been limited by the small number of analysed samples. Flow cytometry does have an advantage over immunocytochemistry as it helps identify cases that have lost GD2 expression. This is increasingly important as future treatment concentrates on targeting GD2 expression either though GD2-antibodies or experimentally through GD2 targeting CART-cells (Schumacher-Kuckelkorn et al., 2016).
In this study, our aim was to compare flow cytometry with the combination of histological and immunohistological assessment of trephines and cytological review of bone marrow aspirates, to determine if there is a difference in detection of positive results between the various methods and if flow cytometry can provide any additional information.

**Method**

The study was performed as an internal evaluation of bone marrow results by flow cytometry in neuroblastoma, which had been introduced as a standard additional technique at Great Ormond Street Hospital, London in 2015. Samples from consecutive patients diagnosed with neuroblastoma at our institution between June 2015 to March 2018 were evaluated. Samples taken at any time point of treatment/surveillance were included in the review.

Disease stage for each patient was based on the INRG staging system (Monclair et al., 2009) and risk stratification was as per the Children’s Cancer Leukaemia Group (CCLG) Guidelines (Morgenstern et al., 2015). At each time point, samples for cytology of aspirate, flow cytometry of aspirate, and histology/immunocytology of trephine biopsy were taken from the left and/or right side, which were then grouped by side of collection. Bone marrow aspirates and trephines reports issued as part of routine of care were reviewed, which included morphological and flow cytometric assessment of aspirates, and morphology plus immunohistochemical staining of trephine biopsies. Flow cytometry was performed 12–60hrs post collection of bone marrow aspirates. Neuroblastoma cells were identified by using live/dead gating followed by identification of CD45\^Lin neg/CD56\+/GD2\^+ stained populations.

For final analysis any patients with missing data for flow, aspirate or trephine analyses were excluded (Figure 1). Any difference between the results of the trephine histology/immunohistochemistry, aspirate morphology, or flow cytometry were recorded. Significance testing was performed using unpaired t-test with Welch’s correction, with a p-value \( \leq 0.05 \) considered as significant.

Eight patients also had bone marrow aspirates collected for RNA testing, performed by RTqPCR, as part of the European HR-NBL1/SIOPEN trial (ClinicalTrials.gov registration number: NCT01704716) (Viprey et al., 2014). The results from RTqPCR and flow cytometry analyses were compared, in order to establish if there are any correlations between the two assays. RNA was extracted and RTqPCR for the neuroblastoma mRNAs paired like homeobox 2B (PHOX2B) and tyrosine hydroxylase (TH) performed according to standard operating procedures (Viprey et al., 2007, 2014). PHOX2B and TH are established neuroblastoma mRNAs (Stutterheim et al., 2008; Brownhill and Burchill, 2017; Uemura et al., 2019).

For statistical analysis, the Log2 delta Ct values from the RTqPCR were converted to linear values for correlation with flow values by Pearson coefficient and correlation of flow with aspirate morphology or trephine immunohistochemistry was performed using Welch's T test. Statistical analysis we performed using Prism software version 9.

**Results**

A total of 392 bone marrow samples from 72 patients were analysed. Complete bone marrow, trephine and flow cytometry data was available for 302 samples (Figure 1). RTqPCR results were available for 26 samples from 15 patients. A total of 15 samples from eight patients had both flow cytometry and RTqPCR data available (see Underlying data).
Correlation between cytology, histology, and flow cytometry

There was concordance in a negative result across all three modalities for 216/302 samples and a concordance of 38/86 for positive results across all three modalities (Figure 2A), with a further 14/86 (16%) samples positive by flow cytometry.

Figure 2. Comparison of positive results by flow cytometry, cytology and immunohistochemistry. A) Venn diagram of all positive cases. B) Negative trephine results in blue, positive trephine results in pink. 216 negative samples by all modalities excluded from analysis. Box and whisker plot showing the Mean and standard deviation of results *p-value 0.0056 by Welch’s t-test. C) Negative cytology results in red, positive cytology results in green. 216 negative samples by all modalities excluded from analysis. Box and whisker plot showing the Mean and standard deviation of results, **p-value 0.0027 by Welch’s t-test.
and either cytology of aspirates or histology of trephine. Of the 86 samples that were positive by at least one test, 14/86 (16%) were positive by flow cytometry alone. Taken together, trephine and aspirate morphology detected 20/86 (23%) positives that were negative by flow cytometry (trephine only n = 11, cytology of aspirates only n = 3, both trephine and cytology of aspirates n = 6) (Figure 2A).

Flow cytometry provides the additional benefit of allowing enumeration of the neuroblastoma cells within the bone marrow sample by calculating the positively gated events and negative gated events. We performed an absolute numerical comparison of flow cytometry results against the binary trephine and aspirate results (Figure 2B and 2C) to determine if numerical flow cytometry results correlate with the aspirate morphology or trephine categorisation. Bone marrow samples that were positive by analysis of trephines were significantly more likely to be positive than negative on flow cytometry (p = 0.0027) and the same was true for samples positive for cytology (p = 0.0056), suggesting a good concordance between these modalities. When comparing trephine and flow cytometry, 18 samples were positive by flow cytometry but not positive on trephine histology. These samples had a percentage detection range of 0.0130% to 5.3% (Figure 2B). Similarly, when comparing flow cytometry and cytology, there were 24 samples positive by flow cytometry, which were negative by cytology (Figure 2C). These samples had a percentage detection range from 0.0041% to 3.75%. Therefore, flow cytometry of bone marrow aspirates detects low-level disease not reported after analysis of trephines or cytology of bone marrow aspirates.

Patient disease course of flow cytometry-only positive samples
A total of 14 samples from nine patients were positive solely on flow cytometry. These patients represent potential cases where flow cytometry may be useful for detecting bone marrow disease below the combined threshold of cytology and trephines. All nine of these patients were diagnosed as high-risk (Table 1). The level of disease detected by flow cytometry was low ranging from 0.008% to 2.37%. Only two patients (patient 5 and 7) had no radiological evidence of metastatic skeletal disease at the time of bone marrow sampling. Patient 5 had radiological localised disease and had bone marrow sample taken at diagnosis. This patient was treated as high-risk due having a MYCN amplified tumour and is now 42 months post diagnosis with no evidence of progression or relapse. Patient 7 had stage M high-risk neuroblastoma; the bone marrow sample was taken after completion of high dose chemotherapy with busulfan and melphalan. This patient is now 38 months post diagnosis with no evidence of relapse. Thus, the clinical follow up of these cases with low-level disease by flow cytometry does not provide any support for altering staging or treatment in such patients.

Correlation between flow cytometry and RTqPCR
To further evaluate the results of flow cytometry, we compared 15 samples from eight patients who had corresponding RTqPCR performed for mRNA using PHOX2B and TH markers. We performed simple linear regression modelling on

<table>
<thead>
<tr>
<th>Participant</th>
<th>Laterality of sample</th>
<th>Flow cytometry %</th>
<th>Stage</th>
<th>Risk Stratification</th>
<th>MIBG/PET scan evaluation at time of bone marrow sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>R</td>
<td>0.02600</td>
<td>M</td>
<td>HR</td>
<td>Multiple skeletal metastasis</td>
</tr>
<tr>
<td>2</td>
<td>L</td>
<td>0.00820</td>
<td>M</td>
<td>HR</td>
<td>Multiple skeletal metastasis</td>
</tr>
<tr>
<td>2</td>
<td>R</td>
<td>0.07700</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>L</td>
<td>0.04000</td>
<td>M</td>
<td>HR</td>
<td>Low grade uptake in skeletal metastasis</td>
</tr>
<tr>
<td>4</td>
<td>R</td>
<td>0.10000</td>
<td>M</td>
<td>HR</td>
<td>Multiple skeletal metastasis</td>
</tr>
<tr>
<td>4</td>
<td>L</td>
<td>0.06200</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>R</td>
<td>0.31000</td>
<td>L2</td>
<td>HR</td>
<td>No skeletal metastasis</td>
</tr>
<tr>
<td>5</td>
<td>L</td>
<td>0.01300</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>R</td>
<td>0.35000</td>
<td>M</td>
<td>HR</td>
<td>Multiple skeletal metastasis</td>
</tr>
<tr>
<td>6</td>
<td>L</td>
<td>0.68000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>L</td>
<td>0.36700</td>
<td>M</td>
<td>HR</td>
<td>No skeletal metastasis</td>
</tr>
<tr>
<td>8</td>
<td>L</td>
<td>2.37000</td>
<td>M</td>
<td>HR</td>
<td>Multiple skeletal metastasis</td>
</tr>
<tr>
<td>9</td>
<td>R</td>
<td>0.36100</td>
<td>M</td>
<td>HR</td>
<td>Multiple skeletal metastasis</td>
</tr>
<tr>
<td>9</td>
<td>L</td>
<td>0.02800</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 3. Correlation between reverse transcription quantitative polymerase chain reaction (RTqPCR) and flow cytometry results. Simple linear regression between RTqPCR and flow cytometry for fifteen samples A) $R^2$ co-efficient 0.8090 (p-value < 0.0001) and B) $R^2$ co-efficient was 0.8697 (p-value < 0.0001). Where PHOX2B is homeobox 2B, and TH is tyrosine hydroxylase.
RTqPCR and flow cytometry data for matched samples (Figure 3). For PHOX2B the $R^2$ co-efficient was 0.8090 (p-value < 0.0001) and for TH $R^2$ co-efficient was 0.8697 (p-value < 0.0001). This excellent correlation between RTqPCR and flow cytometry further validates the flow cytometry results.

**Discussion and conclusion**

In comparing flow cytometry, histology and cytology of aspirates results, our investigations show a good concordance across all three modalities for negative samples (72%). Taking positivity for either trephine and/or cytology of aspirates samples together, there was also good concordance for positive results 52/86 (61%), though both flow cytometry (23%) or combination of histology/cytology (16%) did miss samples that were positive by the other modality. Furthermore, there was also good correlation between RTqPCR and flow cytometry results where both were available, providing further validation to flow cytometry results. Results positive by flow cytometry alone generally had low-level disease. There is increasing literature to suggest that clinically significant MRD in neuroblastoma can be detected using RTqPCR for neuroblastoma mRNAs (Burchill et al., 2001; Viprey et al., 2014; Druy et al., 2018; van Wezel et al., 2016). These studies commonly include the two markers we have investigated, TH and PHOX2B mRNA. The persistence of bone marrow positivity is associated with poorer prognosis (Horibe et al., 2001; Druy et al., 2018; Popov et al., 2019).

Flow cytometry is a routine test in diagnostic laboratories, which does require the development of expertise for analysis of results. Our results show some discordance between cytology/histology/flow cytometry. This discordance could be related to sampling differences, as different samples may be taken for analysis by various parts of diagnostic laboratories. Further, neuroblastoma cells have a propensity to aggregate. During flow cytometry analysis, clots are removed and samples filtered, which may lead to removal of some neuroblastoma aggregates. Bone marrow aspirates and trephine samples are not disaggregated, which may account for some disparity in results. Further, an element of subjectivity is present in the histological/cytology analysis of bone marrow trephines and aspirates, whereas flow cytometry provides an unequivocal characterisation of individual neuroblastoma cells.

Flow cytometry may be particularly useful for defining disease in patients who do not have adequate trephine biopsies or cells available for review on aspirates. It could serve as an additional quick and cost-effective tool for detection of low-threshold disease in patients with neuroblastoma. However, the presence of 20/86 samples with positivity by either cytology or histology analysis but no detectable neuroblastoma by flow cytometry, whilst may be accountable by sampling differences, highlights the importance of expert haematological and histopathological analysis of samples from these children. The clinical significance of low-level disease, detected using different methods, in neuroblastoma continues to be explored globally and remains to be seen.

**Data availability**

**Underlying data**

Open Science Framework: Flow cytometry analysis of neuroblastoma bone marrow. <registration DOI>

This project contains the following underlying data:

- De-identified HRNBL PCR and flow data.xlsx

<License statement>

https://osf.io/5c7ts/

**Ethic statement**

The evaluation of results from bone marrow flow cytometry was a routine retrospective evaluation of standard of care procedures and not a formal research study. As such it did not require ethics committee approval. Consent for marrow aspirates and standard of care analysis was obtained from all patients using standard hospital consent procedures.

**Competing interests**

No competing interests were disclosed.

**Grant information**

The author(s) declared that no grants were involved in supporting this work.
References


This is an interesting article that describes an analysis of bone marrow flow cytometry (FC) to identify bone marrow (BM) minimal residual disease (MRD) in consecutive patients diagnosed with neuroblastoma (NB) from a single centre, Great Ormond Street Hospital (GOSH). This is a laboratory-based comparative study focused on the comparison of NB BM MRD with morphology on BM aspirate, BM trephine and in a small number of cases to correlate with qRT-PCR for NB specific transcripts.

MRD has an established role in the management of acute lymphoblastic leukaemia (ALL) and is routinely used in ALL clinical management. MRD in combination with clinical and genetic factors remains one of the most important clinical tools to guide therapy and risk stratification in both newly diagnosed and relapsed ALL. The two major techniques used to measure MRD in ALL are flow cytometry or quantitative PCR. The European Scientific foundation for Laboratory Hemato Oncology (ESLHO) has established consortia to promote the innovation, standardisation, quality control and education of laboratory diagnostics including flow MRD (EuroFlow) and quantitative PCR MRD (EuroMRD). The International Neuroblastoma Risk Group (INRG) Task Force developed standard recommendations for the detection of residual neuroblastoma using immunocytology on BM cytospins or smears and/or qRT-PCR on BM. The INRG Task Force felt that the volume of bone marrow for analysis post-chemotherapy by flow cytometry may limit the sensitivity to detect residual NB to approximately 1 NB per 10⁴ BM cells, however, this is a level of sensitivity similar to what can be achieved using FC MRD in ALL MRD. However, the use of FC in the assessment of residual NB in BM remains clinically underdeveloped and it is a technology that may be easily and widely adapted to assessing treatment response in HR-NB.

Currently, MRD and liquid biopsy are at an earlier stage of development in paediatric solid tumours and NB compared to ALL. The role of MRD in solid tumours continues to be explored.
within clinical research and/or clinical trials but is not currently used to modify risk classification or treatment. There are many evolving technologies that have the capacity to detect and quantify either residual tumour cells (e.g. immunocytology, flow cytometry, qRT-PCR for genes expressed in NB cells\textsuperscript{12}, DNA based assays to detect NB cells\textsuperscript{13,14,15}) or products secreted by tumour cells (e.g. cell-free tumour DNA\textsuperscript{16,17}, exosomal miRNA\textsuperscript{18,19}) which could be used to assess NB treatment response. So a challenge in the field is to clearly identify the methods and techniques of MRD/liquid biopsy which are prognostic and/or predictive in NB and which can be implemented in the clinical setting.

The strengths of the current study are that it describes using FC to measure BM MRD in a consecutive cohort of 72 children with NB treated at Great Ormond Street Hospital. 392 BM samples were collected from these patients at various time points during their treatment. The study team has compared FC with immunocytology (BM aspirate) and histology (BM trephine). There were a limiting number of patients where qRT-PCR analysis was also available. In terms of the patient cohort, there is very limited detail on the clinical characteristics of the cohort, treatment or exact timing of the BM sample in relation to treatment received and clinical outcome. The only clinical information that is presented is for a subset of 9 patients with high-risk neuroblastoma (HR-NB) with FC positive marrows. In terms of FC methodology, the timing of FC varied from 12-60 hours post collection of the BM sample. The major experimental aim was a comparison of results between the different methodologies, so there is a focus on the comparative analysis of different techniques on individual samples. For negative samples (216) there was good concordance across all three techniques used (FC, trephine and aspirate morphology). The discordances arise in 86 samples which were found to be positive by at least one technique, with only 38/86 (44%) samples positive by all three assays. Each assay identified positive samples which were negative by both other techniques: FC 14/86 (16%), trephine histology 11/86 (13%), and 3/86 (3.5%) by aspirate. For the small number of patients who had both FC and qRT-PCR results, there was a good correlation between both techniques.

Overall, the data shows that there is promise in using FC to assess NB MRD. As noted by the study team, FC is widely available and routinely used to assess haematologic malignancy. Undertaking these studies within the context of routine clinical practice is important as ultimately it can facilitate transfer into clinical practice. The EuroFlow, EuroMRD, and INRG consortia highlight the importance of the standardisation of sample collection, processing, workflow, analysis and reporting of samples\textsuperscript{8,9,10}. In this study, it’s not clear whether the time between collection and running the FC to assess NB MRD, between 12-60 hours, may have had an impact on the results. Specifically, are there differences in samples run at 12 versus 60 hours?

In the GOSH cohort, it’s not clear how the FC assay performs with different input BM cell numbers and whether there were adequate cell numbers for each sample at each time point for the analysis. Assay performance and correlation are likely to be related to the amount of BM sample available for analysis and the levels of residual disease present within the sample. In ALL FC MRD, setting a threshold for the minimum number of BM cells to be analysed by FC is a key feature to permit sensitive detection of rare ALL cells after starting treatment\textsuperscript{10}. The performance of the FC, IC and qRT-PCR assays can be directly compared using spike-in experiments with NB cells and BM MNC, which, although not without limitations, allows an estimation of assay performance under different controlled conditions (e.g. BM cell number and dilution of NB cells).

As discussed, it's known that solid tumours and NB form clumps in the BM and it's possible that a
false negative FC result may occur from NB clumps being filtered out of the sample prior to the FC analysis (identified by the study team as a limitation) or as a consequence of NB being a “patchy” disease within the BM, so there may have been minimal or no NB at the site sampled by BMA, yet there may be clear evidence of BM disease at other sites on functional scanning (e.g. MIBG or FDG-PET CT scan). It would be helpful to understand how each of the marrow assessment methodologies compares with the measured INRC metastatic response in the cohort.

Whether there is additional value arising from FC NB MRD can't really be assessed in this cohort due to the limited amount of clinical and laboratory data. It's likely that the potential impact of NB MRD measured by any technique will be strongly influenced by the underlying genetics and risk classification of the individual patient. It's known in ALL, MRD cutoffs are strongly influenced by the underlying ALL biology and genetics. It's not unreasonable to hypothesise that the same holds true of NB, so the potential utility of NB FC BM MRD will be different in INRG Ms infants (low risk), in INRG M infants (intermediate risk) and in INRG M older children (high risk). An analysis of the data stratified by the patient's risk classification may be informative, depending on the numbers. However, it's possible to hypothesise that the role of FC to assess residual BM NB cells will be of most value in patients with high-risk disease.

Although GOSH had introduced BM FC as a routine standardised assessment from 2015 onwards, the underlying objective(s) of assessing NB MRD by BM FC isn't clearly stated in the submission. In managing ALL, assessing treatment response by MRD forms an integral component of the final ALL risk classification which occurs at the end of induction chemotherapy (i.e. after treatment has commenced). In ALL, rapid clearance of ALL and MRD negativity at the end of induction chemotherapy identifies patients with improved survival, and conversely slow ALL response with persisting high MRD levels at end induction identifies patients with decreased survival. In NB, risk classification is completed at diagnosis, prior to treatment initiation and is not modified by treatment response. Therefore, a potential role for NB MRD (measured by any methodology) to modify INRG risk stratification may not be likely. However, the clinical assessment of NB treatment response is critical and determines post-induction treatment allocation. So analogous to treatment response measurement in ALL, the tempo and completeness of treatment response in HR-NB is important and prognostic. HR-NB patients with inadequately responding disease have a poor chance of survival. One of the major clinical issues facing clinicians managing HR-NB is the early identification of patients with (i) inadequately responding disease or (ii) those who ultimately have a high risk of disease recurrence and poor survival. So standardised response assessment and MRD assessment at earlier time points in HR-NB induction chemotherapy might provide an opportunity to stratify patients to different treatments before the end of induction chemotherapy. Conversely, the role of MRD at later time points in treatment will be different, e.g. the early detection of relapsing disease in patients with a prior good response to treatment. So, in this context, whilst it's important to assess the concordance of different laboratory assays which measure the same variable, it remains difficult to define the clinical significance of NB BM FC in this heterogenous cohort with samples taken at multiple time points and where the linked clinical data isn't available. The value of this dataset will come into its own when it is more comprehensively linked and analysed with matching clinical and response information, specifically with clinically used and/or validated endpoints, such as INRC disease response, progression-free survival and overall survival.

The data presented here represent an incremental step in using NB FC for MRD assessment. However, to take the field of liquid biopsy/MRD assessment forward in neuroblastoma, the ideal
will be to prospectively define clear objectives for MRD assessment at different treatment time points (e.g. early or mid-induction versus later in therapy), collect appropriate samples at standardised time points, and directly compare multiple liquid biopsy/MRD methodologies against each other and also with the use of clinically validated endpoints.

References


**Is the work clearly and accurately presented and does it cite the current literature?**

Yes

**Is the study design appropriate and is the work technically sound?**

Partly

**Are sufficient details of methods and analysis provided to allow replication by others?**

Partly

**If applicable, is the statistical analysis and its interpretation appropriate?**

Partly

**Are all the source data underlying the results available to ensure full reproducibility?**

Partly
Are the conclusions drawn adequately supported by the results?
Partly

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Neuroblastoma, acute lymphoblastic leukaemia, minimal residual disease detection

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 22 October 2021

https://doi.org/10.5256/f1000research.56485.r94970

© 2021 Tytgat G. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Godelieve A. M. Tytgat

Department of Pediatric Oncology, Princess Maxima Center for Pediatric Oncology, Utrecht, 3584 CS, The Netherlands

Flow cytometry of bone marrow aspirates from neuroblastoma patients is a highly sensitive technique for the quantification of low-level neuroblastoma.

In this study, the authors undertake a very nice study to look at the sensitivity and specificity to study NBL cells in the bone marrow by flow cytometry. As there have been studies demonstrating the use of this technique, larger studies that actually compare this technique to other techniques on the same samples are lacking.

In this work, the group studies the detection of NBL cells in a total of 392 samples of 72 patients. The set-up of the paper and the analyses are well done and well described. The article reads well and has a nice flow.

There are some small questions/comments:
- In the methods, the authors describe that 8 patients had BM aspirates collected for RNA testing, and a few sentences later in the results they describe that PCR results were available for 26 samples from 15 patients. Please correct and align the actual numbers here.
- Furthermore, in the results, patient disease course of flow cytometry-only positive samples, the authors state that in their cohort the clinical follow-up of these cases with low-level disease by flow does not provide any support for altering staging or treatment in such patients. The study is not designed nor powered for such an analysis. We know that in patients with no BM metastases, detected by conventional techniques, low levels of mRNA
or DNA and even NBL cells by flow can be detected, which has no impact on the outcome. So I agree with this conclusion and would like to emphasize that this can even be addressed stronger.

- The time points of sampling can be added to the table.
- Looking at Table 1, I wonder what were the results of earlier and later BM samples for patient 7? Did they become negative at a later time point for flow? Please add this information to the text.
- Finally, did the authors look at the paired samples of each patient (left and right side) and investigate the con- and discordances? Also, with respect to the different techniques?
- We test the BM involvement from 2 to 4 sides for each time point. How do the authors deal with the 2 sides? Did they merge the results? Please add this information to the paper.

**Is the work clearly and accurately presented and does it cite the current literature?**
Yes

**Is the study design appropriate and is the work technically sound?**
Yes

**Are sufficient details of methods and analysis provided to allow replication by others?**
Yes

**If applicable, is the statistical analysis and its interpretation appropriate?**
Yes

**Are all the source data underlying the results available to ensure full reproducibility?**
Yes

**Are the conclusions drawn adequately supported by the results?**
Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Pediatric oncologist, translational research PIs

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.
The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com