Blood proteomics of paediatric bronchiolitis obliterans syndrome after haematopoietic cell transplant

Kasiani C. Myers[1,2], Emily J. Skala[2], Sudhir Ghandikota[3], Anil G. Jegga[1,3,4], Nadeem Mousa[5], Matthew E. Siefert[2], Adam Lane[1,2], Valentina Reffatto[2], Christopher Towe[1,6], Nathan Luebbering[1,2], Sheyar Abdullah[1,2], Azada Ibrahminova[1,2], Alexandra Bridgeland[2], Samuel Goldfarb[7,8], Stella M. Davies[1,2] and Assem G. Ziady[1,2]

[1]Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH, USA. [2]Division of Bone Marrow Transplantation and Immune Deficiency, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA. [3]Division of Biomedical Informatics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA. [4]Department of Computer Science, University of Cincinnati College of Engineering, Cincinnati, OH, USA. [5]Department of Pediatrics, Valley Children's Hospital, Madera, CA, USA. [6]Division of Pulmonary Medicine, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA. [7]Department of Pediatrics, School of Medicine at the University of Minnesota, Minneapolis, MN, USA. [8]Division of Pulmonary Medicine, Masonic Children's Hospital, Minneapolis, MN, USA.

Corresponding author: Kasiani C. Myers (kasiani.myers@cchmc.org)

Shareable abstract (@ERSpublications)
Diagnosing bronchiolitis obliterans syndrome (BOS) after HSCT in children is challenging. A discovery proteomics approach identified and validated plasma biomarkers predictive of paediatric BOS and linked identified proteomic pathways. https://bit.ly/4i9yXPm

Cite this article as: Myers KC, Skala EJ, Ghandikota S, *et al*. Blood proteomics of paediatric bronchiolitis obliterans syndrome after haematopoietic cell transplant. *ERJ Open Res* 2026; 12: 00745-2025 [DOI: 10.1183/23120541.00745-2025].

Received: 26 Jan 2025
Accepted: 15 Nov 2025

Abstract

Background The molecular determinants for lung graft-*versus*-host disease-associated bronchiolitis obliterans syndrome (BOS) are poorly understood, and biomarkers in children following haematopoietic stem cell transplant do not exist. To address this gap, we analysed plasma samples from 21 paediatric stem cell transplant recipients prior to and at diagnosis of BOS.

Methods Participants included three cohorts: seven with BOS; seven sex-, age- and timepoint-matched with severe graft-*versus*-host disease alone; and seven sex-, age- and timepoint-matched transplant recipients without BOS or other graft-*versus*-host disease.

Results Our proteomic approach evaluated the expression of 190–12 588 protein isoforms, depending on statistical stringency, and distinguished the three cohorts of paediatric patients prior to and at the time of BOS diagnosis. Differences included proteins that regulate chromatin modification, acute phase signalling, complement, fibrosis, hypoxia, serine protease inhibition, vitamin transport, glucocorticoid receptor transactivation and blood coagulation pathways. A subset of newly discovered proteins was cross-platform validated by ELISA in a larger cohort of paediatric patients 14 (n=107), 30 (n=108), 60 (n=108) and 100 (n=134) days post-transplant. Pathways analysis highlighted potential therapeutics including azithromycin, statins, pazopanib and cediranib.

Conclusion Our strategy offers a potential for early diagnosis and the identification of interventions for paediatric stem cell transplant-associated graft-*versus*-host disease and BOS.

Introduction

Bronchiolitis obliterans syndrome (BOS) occurs in approximately 10% of children after haematopoietic stem cell transplant (HSCT) and is commonly, although not always, associated with other graft-*versus*-host disease (GvHD) [1]. GvHD is a frequent complication after HSCT, and most children with GvHD will not develop BOS. Currently, the factors that determine whether BOS does or does not occur in those with other GvHD are unknown. Mortality rates from BOS in paediatric HSCT patients are reported as high as 22%, along with significant morbidity, in large part due to a delay in diagnosis [2]. In April 2018, the National Heart Lung and Blood Institute, the National Institute of Child Health and Human Development and the National Cancer Institute convened a workshop of multidisciplinary experts to address the

pulmonary complications of HSCT in children [3]. The workshop concluded there were insufficient prospective and evidence-based studies to define the actual incidence, risk factors and biomarkers for BOS in children. Importantly, the report established that research involving paediatric HSCT recipients is limited by the inability to perform pulmonary function testing with accurate and consistent results.

Currently, spirometry is the strategy of choice for diagnosing post-HSCT pulmonary complications. Several studies in adult HSCT recipients reported that spirometric changes, including forced expiratory flow at 25–75% of forced vital capacity (FVC) ($FEF_{25–75\%}$) and forced expiratory volume in 1 s (FEV_1) in the first 3–6 months after HSCT, predict late onset non-infectious pulmonary complications (LONIPCs) [4–7]. The GvHD scoring system for detecting BOS also incorporates alterations in FEV_1, FEV_1/FVC and $FEF_{25–75\%}$ after HSCT [8]. Spirometry presents major challenges in children due to the inconsistent effort and coordination performing a forced expiratory manoeuvre. Although spirometry may be accurate in children older than 6 years after considerable training, results can often be unsatisfactory. Paediatric HSCT recipients are often acutely ill and even older children are unable to perform reliable spirometry. Although spirometry is used successfully in other paediatric populations, such as cystic fibrosis, those particular patients are introduced to pulmonary function testing earlier in childhood and perform testing regularly, whereas paediatric HSCT recipients have inferior success rates due to sporadic use, even in teenagers [9]. The lack of reliable spirometry in most children after HSCT often delays the identification of pulmonary complications, especially BOS, resulting in irreversible lung injury before therapeutic intervention can be implemented.

Plasma proteomics have been used in a few select studies to examine response to therapy in early acute GvHD [10, 11] leading to the development of valuable predictive algorithms [12–15]. However, there are few studies using proteomics in the study of BOS [16]. Given the limitations of existing diagnostic modalities of BOS and GvHD in children after HSCT, novel approaches are necessary to facilitate early,

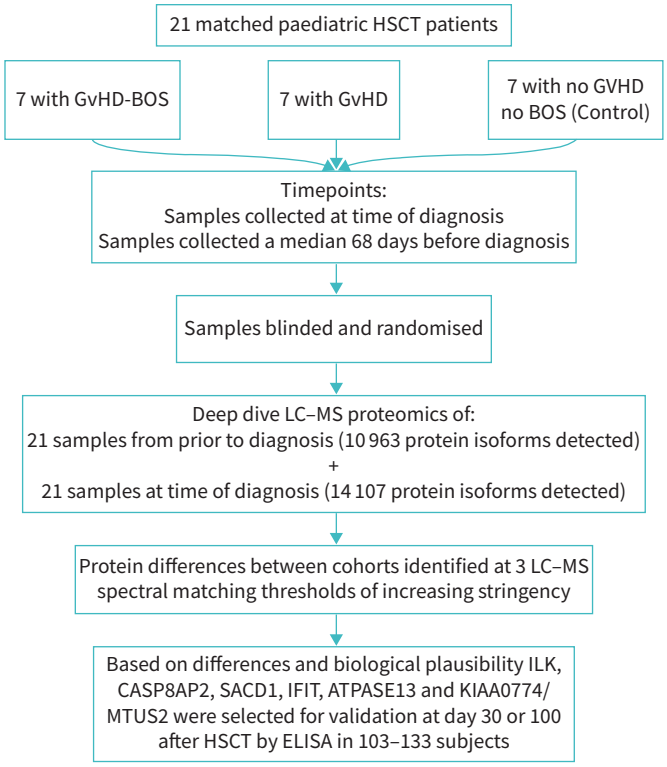


FIGURE 1 Workflow diagram. Summary of the workflow of our studies is shown. The diagram outlines the two phases (discovery and validation) of the presented work. HSCT: haematopoietic stem cell transplant; GvHD: graft-*versus*-host disease; BOS: bronchiolitis obliterans syndrome; LC–MS: liquid chromatography–mass spectrometry; KIAA0774/MTUS2: microtubule-associated tumour suppressor 2; ILK: integrin-linked protein kinase; CASP8AP2: caspase 8-associated protein 2; SAC3D1: SAC3 domain-containing protein 1; IFIT1: interferon-induced protein with tetratricopeptide repeats 1; ATPase13: endoplasmic reticulum transmembrane helix translocase 13A1

accurate diagnosis of BOS and timely treatment. To address this need we used non-biased mass spectrometry (MS) proteomics to identify biomarkers and aberrant pathway signalling in BOS more than 60 days before clinical diagnosis based on lung function and imaging. We aimed to interrogate the expression of thousands of protein isoforms and examine whether patterns emerged in paediatric HSCT patients that go to develop BOS *versus* GvHD without BOS *versus* those that do not develop BOS or GvHD. Furthermore, to begin to establish our approach we aimed to test whether some of the changes would be validated by cross-platform analysis using ELISA in larger cohorts as early as 30 days after HSCT.

Methods

Banked sample selection

The Cincinnati Children's Hospital Medical Centre's Institutional Review Board approved the study. Patients with a clinical diagnosis of BOS and available plasma samples were identified in our institutional HSCT repository. Our studies include two phases: a proteomics discovery phase and a validation of biologically plausible significant findings by ELISA phase. The workflow for our studies is summarised in figure 1 and a quality control schematic is shown in supplementary figure S1.

We conducted our proteomic discovery studies in 21 patients representing three cohorts: seven patients with GvHD-BOS (GvHD-BOS cohort, designated cases); seven sex-, age ±2 years- and timepoint-matched patients with severe grade 3–4 GvHD without BOS (GvHD cohort); and seven sex-, age ±2 years- and timepoint-matched transplant recipient controls with no GvHD or BOS (control cohort). GvHD and GvHD-BOS diagnoses were made by the treating physician according to consensus criteria [17–19], and the diagnoses were supported with tissue biopsies whenever feasible. The entire proteomics analyses included 42 samples, with 21 participants providing a sample at two timepoints: one timepoint was designated "prior to diagnosis" and was a median of 68 (38–188) days before clinical diagnosis of chronic GvHD or BOS in respective cohorts. Samples analysed from controls for the "prior to diagnosis" timepoint were the available sample closest to the timepoint after transplant of the matched GvHD or BOS "prior to diagnosis" samples. The second timepoint, termed "time of diagnosis" (median 203 days after transplant; range 92–361) was a sample collected closest to the time of diagnosis of GvHD or BOS, and the closest sample available at a similar post-transplant timepoint for the matched control cohort. Demographics for the subjects in our proteomics discovery study are shown in supplementary table S1.

For our validation studies, we used an ELISA to examine plasma samples from a larger cohort of >100 consecutive paediatric transplant recipients, including the cases tested by proteomics. ELISA was performed on samples collected 14, 30, 60 and 100 days after transplant for six proteins selected for their biological plausibility and significant changes prior to diagnosis of GvHD and/or GvHD-BOS. Demographics for the subjects in our ELISA validation study are shown in supplementary table S2.

The remaining methods are described in the supplementary material.

Results

Proteomic discovery and data summary

Our discovery MS analyses generated data on a maximum of 9870 (prior to diagnosis) and 12 588 (at time of diagnosis) plasma protein isoforms that were present in at least 57% of the control, GvHD or GvHD-BOS cohorts across the entire study. Total isoforms considered varied from the maximum number of proteins depending on comparison. To identify differences, we analysed these data at three statistical thresholds with increasing stringency, relaxed (XCorr range 1.00–6.81), intermediate (XCorr range 1.50–6.81) or strict (XCorr range 1.50–6.83+percolator). Importantly, several isoforms were present or absent in only one cohort. table 1 summarises the numbers of isoforms that differed at $p<0.05$ or were present/absent. Of these overall differences, the largest number of changes $\geqslant$1.5 fold existed when controls were compared with GvHD-BOS or GvHD (figure 2a–f, volcano plots; $p<0.05$, $\geqslant$twofold change) at the intermediate threshold, which supports expectations based on clinical presentation. Comparisons at the relaxed threshold showed differences in an expanded set of protein isoforms (supplementary figure S2a–f, volcano plots; $p<0.05$, $\geqslant$twofold change), whereas those at the strict threshold reduced the set (supplementary figure S2a–f, volcano plots; $p<0.05$, $\geqslant$twofold change). Importantly, significant changes were observed at the "prior to" and "at the time of diagnosis" timepoints at all statistical thresholds (table 1, figure 2 and supplementary figures S2 and S3), supporting the potential of plasma proteins as predictive biomarkers of GvHD and/or GvHD-BOS.

Supervised heatmap analysis of plasma protein isoform expression levels across subjects in all three cohorts revealed strong association of expression with phenotype at all statistical thresholds. Prior to diagnosis correlation of protein expression forms clearly contrasting patterns observed for GvHD-BOS,

TABLE 1 Summary for cohort comparisons of plasma protein expression differences or presence/absence identified by mass spectrometry at three statistical thresholds[#]

		Comparison					
		GvHD-BOS *versus* control		GvHD-BOS *versus* GvHD		GvHD *versus* control	
Timepoint	MS threshold (XCorr)	Different (total) (p<0.05)	Present/ absent	Differences (total) (p<0.05)	Present/ absent	Differences (total) (p<0.05)	Present/ absent
Prior	1.00–6.81	1439 (9433)	139/292	267 (8247)	152/128	1356 (9345)	129/369
	1.50–6.81	819 (7255)	115/317	233 (5750)	120/94	876 (7154)	98/448
	1.50–6.83 w/ percolator	37 (306)	1/14	24 (271)	0/0	109 (314)	8/13
At	1.00–7.52	1137 (12 180)	156/178	325 (11 776)	121/131	1692 (12 216)	161/211
	1.50–7.52	808 (10 075)	142/161	292 (9703)	129/115	1198 (10 069)	147/183
	1.50–7.51 w/ percolator	16 (233)	2/8	14 (188)	0/2	56 (229)	8/45

GvHD: graft-*versus*-host disease; BOS: bronchiolitis obliterans syndrome; MS: mass spectrometry; w/: with [#]: Summary of results for comparisons of plasma protein expression levels in GvHD-BOS *versus* control or GvHD or GvHD *versus* control. Comparisons were conducted at three MS spectral matching thresholds with increasing stringency. Number of proteins shown in "different (total)" columns are those differences that achieved significance (p<0.05), whereas numbers in the "present/absent" columns are those that were either only detected (present) or not detected (absent) in the primary comparand *versus* the secondary comparand in the comparison conducted.

GvHD or control cohorts with the exception of one GvHD-BOS subject (figure 2g; p<0.05). At the time of diagnosis, patterning continued to be observed and the one case of GvHD-BOS that tracked more with control at the "prior to diagnosis" timepoint now tracked with GvHD-BOS (figure 2h; p<0.05). Unsupervised analysis of differences by 1−Spearman's correlation with complete-linkage agglomerative hierarchical clustering successfully identified control and partially identified GvHD-BOS and GvHD individuals at the relaxed and intermediate thresholds at both timepoints (supplementary figure S4). Importantly, data missingness reduced the success of unsupervised analysis for the stringent threshold (supplementary figure S4), supporting our approach of considering all three thresholds for further studies. Notably, clear patterns for the presence or absence of protein isoform expression (unsupervised analysis) were observed, including the presence or absence of proteins in individual cohorts prior to (figure 2i) and at the time of diagnosis (figure 2j). Similar results were obtained for the relaxed (supplementary figure S2) and stringent threshold analyses (supplementary figure S3).

Taken together, volcano plot and heatmap analysis (supervised and unsupervised) revealed that a significant portion of plasma protein isoforms are significantly altered prior to and at the time of clinical diagnosis of GvHD or GvHD-BOS. Furthermore, these changes form clear patterns that can distinguish patients that develop GvHD-BOS from those that develop GvHD or no disease. Our analysis of the data also justified the examination of multiple thresholds for statistical stringency and revealed that highly stringent analyses may paint an incomplete picture of the patient phenotype.

Pathway analysis: prior to diagnosis

To better understand the molecular determinants of the development of GvHD-BOS compared with control, we analysed biological pathways enriched for by differences between these two cohorts at the stringent threshold. Prior to diagnosis we observed increased GvHD-BOS associations with thrombophilia ($p=5.42\times10^{-74}$, false discovery rate (FDR) of 4.38×10^{-71}), complement factor I deficiency ($p=4.86\times10^{-50}$, FDR of 7.87×10^{-48}), respiratory viral infections ($p=2.24\times10^{-48}$, FDR of 2.26×10^{-46}), myocardial infarction ($p=1.30\times10^{-45}$, FDR of 6.56×10^{-44}), and vasculopathy ($p=2.61\times10^{-44}$, FDR of 1.11×10^{-42}) (supplementary table S3). Networks that contributed to these disease associations included alterations in complement mediated inflammation ($p=8.10\times10^{-37}$, FDR of 2.51×10^{-35}), kallikrein–kinin system regulation ($p=9.07\times10^{-16}$, FDR of 1.41×10^{-14}), blood coagulation ($p=1.30\times10^{-7}$, FDR of 1.35×10^{-6}), proteolysis of extracellular matrix ($p=2.29\times10^{-3}$, FDR of 1.18×10^{-2}) and interleukin (IL)-6 inflammatory signalling ($p=7.63\times10^{-3}$, FDR of 2.63×10^{-2}) (supplementary table S4). Furthermore, angiotensin system maturation ($p=2.89\times10^{-11}$, FDR of 4.80×10^{-10}; supplementary table S5) was dysregulated in patients later diagnosed with GvHD-BOS compared with controls. Gene Ontology (GO) process analysis confirmed highly significant changes in proteolysis, complement activation and coagulation (supplementary table S6).

GvHD-BOS compared with GvHD showed increased disease association with amyloidosis ($p=1.85\times10^{-15}$, FDR of 7.09×10^{-13}), complement component 5 deficiency ($p=2.28\times10^{-10}$, FDR 1.75×10^{-8}), wounds and injuries ($p=1.58\times10^{-8}$, FDR of 6.04×10^{-7}), hyperlipidaemia/cholesterol metabolism ($p=4.32\times10^{-8}$, FDR

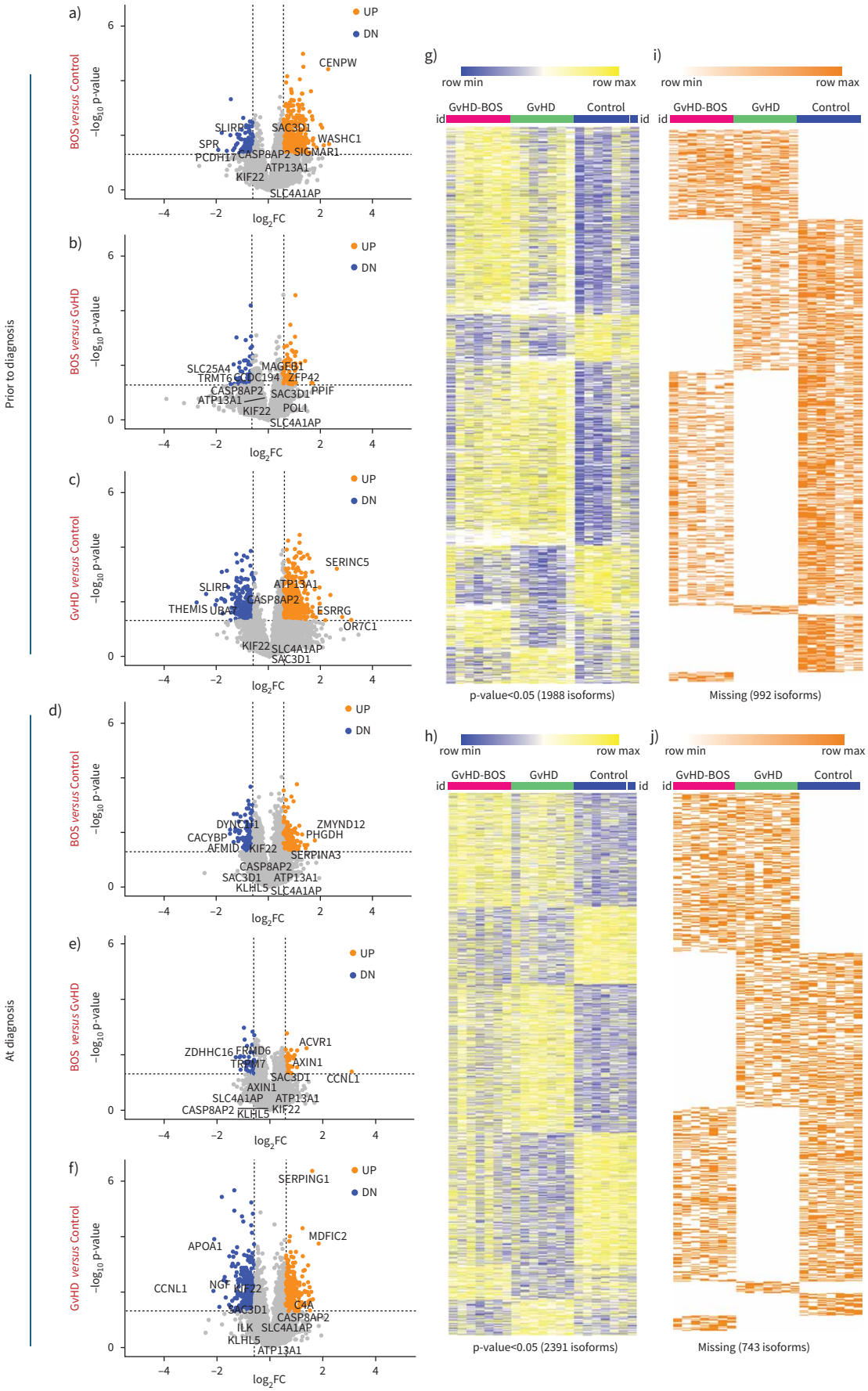

Prior to diagnosis
At diagnosis
a)
BOS versus Control
−log10 p-value
log2FC
UP
DN
CENPW
SLIRP
SAC3D1
WASHC1
SPR
PCDH17
CASP8AP2
SIGMAR1
ATP13A1
KIF22
SLC4A1AP
b)
BOS versus GvHD
SLC25A4
MAGEB1
TRMT6
CCDC194
ZFP42
CASP8AP2
SAC3D1
PPIF
ATP13A1
KIF22
POLI
SLC4A1AP
c)
GvHD versus Control
SERINC5
SLIRP
ATP13A1
CASP8AP2
THEMIS
ESRRG
OR7C1
KIF22
SLC4A1AP
SAC3D1
d)
BOS versus Control
DYNC1I1
ZMYND12
CACYBP
PHGDH
AFMID
KIF22
SERPINA3
CASP8AP2
SAC3D1
ATP13A1
KLHL5
SLC4A1AP
e)
BOS versus GvHD
ZDHHC16
FAM6
ACVR1
TRPM7
AXIN1
SAC3D1
CCNL1
AXIN1
SLC4A1AP
ATP13A1
CASP8AP2
KLHL5
KIF22
f)
GvHD versus Control
SERPING1
MDFIC2
APOA1
CCNL1
NGF
KIF22
C4A
SAC3D1
CASP8AP2
ILK
SLC4A1AP
KLHL5
ATP13A1
g)
row min
row max
GvHD-BOS
GvHD
Control
id
p-value<0.05 (1988 isoforms)
i)
Missing (992 isoforms)
h)
p-value<0.05 (2391 isoforms)
j)
Missing (743 isoforms)

FIGURE 2 Summary of discovery liquid chromatography–mass spectrometry (LC–MS) plasma proteomics. Non-biased blinded and randomised LC–MS label-free quantitation was used to quantitate plasma protein isoform expression prior to (average 68 days) or at the time of diagnosis in banked samples from paediatric haematopoietic stem cell transplant (HSCT) patients that went on to develop graft-*versus*-host disease (GvHD)-bronchiolitis obliterans syndrome (BOS), GvHD or no GvHD or BOS (control). Intermediate statistical analysis is shown here (relaxed and stringent threshold analyses are shown in supplementary figures S2 and S3, respectively). a–f) Volcano plots of all data (dashed lines indicate vertical at −1.5 or +1.5-fold change; horizontal at p=0.05). Prior to diagnosis: GvHD-BOS *versus* control (a); GvHD-BOS *versus* GvHD (b); and GvHD *versus* control (c). At diagnosis: GvHD-BOS *versus* control (d); GvHD-BOS *versus* GvHD (e); and GvHD *versus* control (f). g–i) Supervised heatmaps of the expression of protein isoforms that segregated GvHD-BOS, GvHD and control paediatric HSCT patients. Prior to diagnosis: differentially expressed proteins (p<0.05) (g) and present/absent protein expression (i). At diagnosis: differentially expressed proteins (h) and present/absent protein expression (j). Unsupervised heatmaps at all three thresholds are shown in supplementary figure S4. UP: increased protein expression; DN: decreased protein expression; FC: fold change.

of 1.26×10^{-6}), DNA viral infections ($p=1.07\times10^{-8}$, FDR of 4.55×10^{-7}) and insulin resistance ($p=4.45\times10^{-8}$, FDR of 1.26×10^{-6}) (supplementary table S7). GO process analysis confirmed dysregulation of lipid transport and metabolism (supplementary table S8) in GvHD-BOS.

Nine complement factors are significantly elevated prior to diagnosis in both GvHD-BOS and GvHD *versus* control (supplementary table S9). These proteins contribute to important changes in complement pathway signalling (figure 3). Complement activation alone did not discriminate GvHD from GvHD-BOS prior to diagnosis; however, complement C5, C5a, C5b and C6 and α-2 macroglobulin were significantly different (supplementary table S7), and C2 isoform 5 and C4-A proteins were only present in patients who would go on to develop GvHD-BOS. Furthermore, inhibitors of complement signalling, such as C1 inhibitor, factor I and vitronectin were significantly elevated, which may suggest feedback mechanisms are activated to control stimulation of complement signalling in both GvHD and GvHD-BOS (figure 3).

Our findings prior to diagnosis, summarised in figure 4 (stringent threshold), highlighted pathways, such as complement, known to be associated with diagnosed GvHD-BOS. Several novel findings in BOS that may suggest hyperlipidaemia, insulin resistance or dysregulation of angiotensin processing were also revealed as potential contributors to/distinguishers of disease states. We plan to investigate these potential differences as mechanisms of disease in future studies. The proteomic data also provided details on several pathways that can be logically linked to GvHD-BOS. MetaDrug subanalysis indicated that statins and azithromycin (supplementary table S10) may be candidate therapies in children at risk of developing GvHD-BOS. Furthermore, the use of cediranib and/or pazopanib may potentially control fibrosis and vascular abnormalities in GvHD-BOS.

Pathway analysis: time of diagnosis

At the time of diagnosis, comparing children with GvHD-BOS with control, GvHD-BOS was associated with myocardial and vascular abnormalities ($p=1.91\times10^{-50}$, FDR of 1.81×10^{-47}), thrombophilia ($p=1.26\times10^{-42}$, FDR of 1.99×10^{-40}), obstructive lung disease ($p=5.26\times10^{-29}$, FDR of 1.78×10^{-27}), and pulmonary fibrosis ($p=1.50\times10^{-28}$, FDR of 4.90×10^{-27}) (supplementary table S11). Increased complement activation ($p=1.66\times10^{-4}$, FDR of 1.83×10^{-3}; supplementary table S12), altered angiotensin processing ($p=5.91\times10^{-13}$, FDR of 6.03×10^{-11}; supplementary table S13) continued. Furthermore, GvHD-BOS continued to exhibit elevated kallikrein–kinin system signalling ($p=2.79\times10^{-11}$, FDR of 9.21×10^{-10}) and further exhibited coagulopathy ($p=3.73\times10^{-8}$, FDR of 6.15×10^{-7}), and inflammatory signalling driven by IL-6 ($p=9.70\times10^{-4}$, FDR of 8.00×10^{-3}), and integrin priming ($p=8.03\times10^{-3}$, FDR of 3.31×10^{-2}) (supplementary tables S12 and S13). Other inflammatory processes increased in GvHD-BOS included platelet degranulation ($p=7.83\times10^{-29}$, FDR of 1.41×10^{-25}), acute inflammatory responses ($p=5.77\times10^{-22}$, FDR of 3.47×10^{-19}), ERK1 and ERK2 signalling ($p=9.05\times10^{-22}$, FDR of 4.08×10^{-19}), regulation of proteolysis ($p=5.97\times10^{-19}$, FDR of 2.15×10^{-16}), regulation of exocytosis ($p=2.45\times10^{-18}$, FDR of 7.37×10^{-16}) and regulation of iron transport ($p=1.26\times10^{-13}$, FDR of 7.58×10^{-12}) (supplementary table S14).

Compared with GvHD, GvHD-BOS exhibited hypertension ($p=1.64\times10^{-30}$, FDR of 8.27×10^{-28}) and thrombosis ($p=2.34\times10^{-24}$, FDR of 1.47×10^{-22}) (supplementary table S15) that was associated with alterations in angiotensin system maturation ($p=7.23\times10^{-19}$, FDR of 2.24×10^{-17}; supplementary table S16). In addition, bacterial killing ($p=2.72\times10^{-6}$, FDR of 1.91×10^{-3}), vitamin ($p=1.08\times10^{-}$4, FDR of 1.95×10^{-2}) and protein ($p=2.23\times10^{-4}$, FDR of 1.95×10^{-2}) transport are dysregulated in GvHD-BOS compared with GvHD (supplementary table S17).

HIF1α activation was enriched in GvHD-BOS compared with GvHD or control (supplementary figure S5, intermediate threshold). Samples banked at the time of diagnosis exhibited elevation in the levels of the

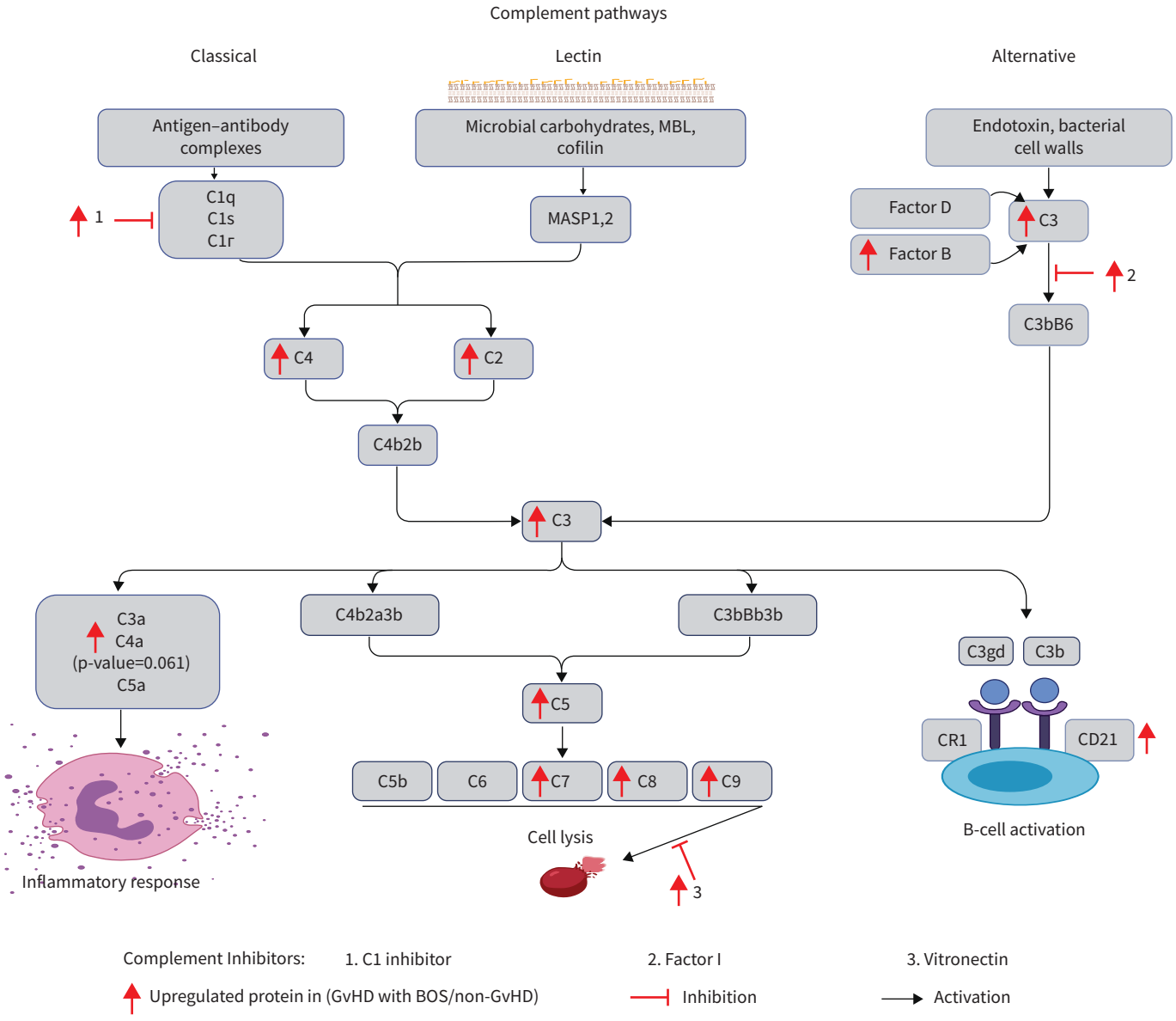


FIGURE 3 Complement activation in graft-*versus*-host disease (GvHD)-bronchiolitis obliterans syndrome (BOS) *versus* control. Samples were collected an average of 68 days before clinical diagnosis of GvHD-BOS compared with controls. Mass spectrometry label-free quantitation was used to quantitate plasma protein isoform expression. Pathway analyses were performed using Clarivate's MetaCore-GeneGo software analysis of proteins that differed ($p<0.05$) between GvHD-BOS and control and adapted from "Complement Pathways", by BioRender.com (https://app.biorender.com/biorender-templates). Schematic shows upregulation of nine components of complement pathways, in addition to upregulation of complement inhibitors: C1 inhibitor (denoted by "1"), factor I (denoted by "2") and vitronectin (denoted by "3") as indicated by the arrows. MBL: mannose binding protein C.

HIF1α activators MAPK1/ERK1 and MAPK2/ERK2 and the HIF1α degradation inhibitor ubiquitin carboxyl-terminal hydrolase isozyme L1. Given HIF1α's functions as a master transcriptional regulator of the protective response to hypoxia, its activation in patients that go on to present with GvHD-BOS is consistent with the airway obstruction presented in supplementary table S11.

Taken together, our studies indicate that at the time of diagnosis features exhibited by GvHD-BOS prior to diagnosis, such as complement activation, coagulopathy, serine protease inhibition, vascular leak and angiotensin processing continue, whereas marked increases in acute inflammation, pulmonary fibrosis and airway obstruction related signalling developed. These findings are consistent with the clinical presentation of GvHD-BOS.

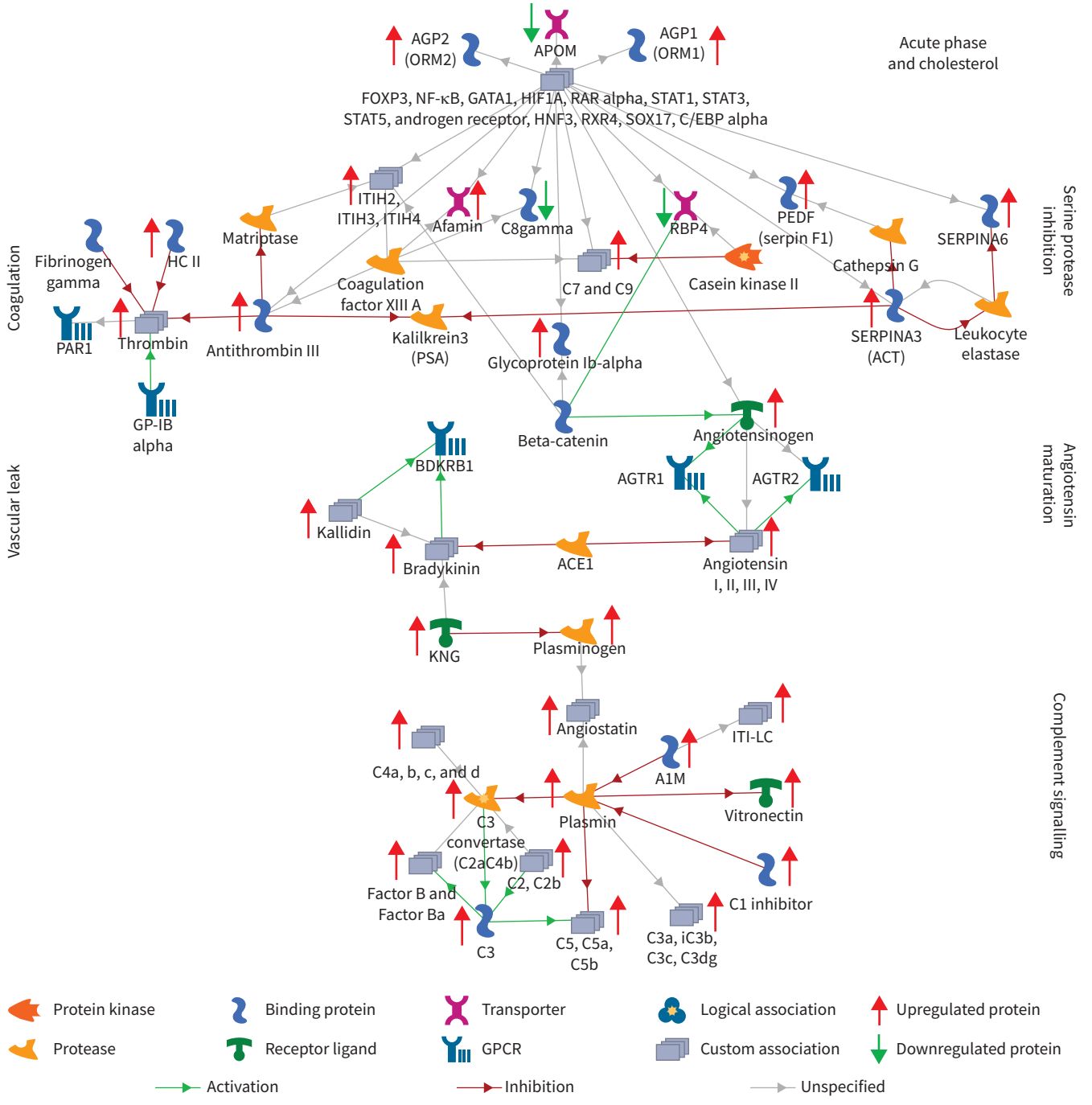


FIGURE 4 Network analysis of differential protein expression in graft-*versus*-host disease (GvHD)-bronchiolitis obliterans syndrome (BOS) and control. Non-biased blinded and randomised proteomic analyses of banked plasma samples collected from paediatric haematopoietic stem cell transplant (HSCT) patients an average of 68 days prior to clinical diagnosis of GvHD-BOS compared with control. Mass spectrometry label-free quantitation was used to quantitate plasma protein isoform expression and the candidates that segregated patients that went on to develop GvHD-BOS *versus* control at the stringent threshold. Proteins that were present in >57% of at least one cohort and either differed ($p<0.1$) or were only present in one cohort were analysed for network interaction and associated processes. Pathway analyses were performed using Clarivate's MetaCore-GeneGo software analysis of protein differences and created with BioRender.com.

Identifying potential biomarkers prior to diagnosis

To reduce the number of marker candidates that may predict the risk of GvHD-BOS and/or GvHD we conducted two sets of protein network analyses: one for proteins only present in each one of the cohorts

and one for the proteins in GvHD-BOS that significantly differed from both GvHD and control. At the relaxed threshold, protein detected only in subjects that went on to develop GvHD-BOS and their transcriptional regulators associated with development, telomere lengthening, responses to hypoxia, responses to oxidative stress, immunity, fibrosis and remodelling, and vesicular trafficking (figure 5a). These features of GvHD-BOS were also observed at the intermediate threshold with additional associations

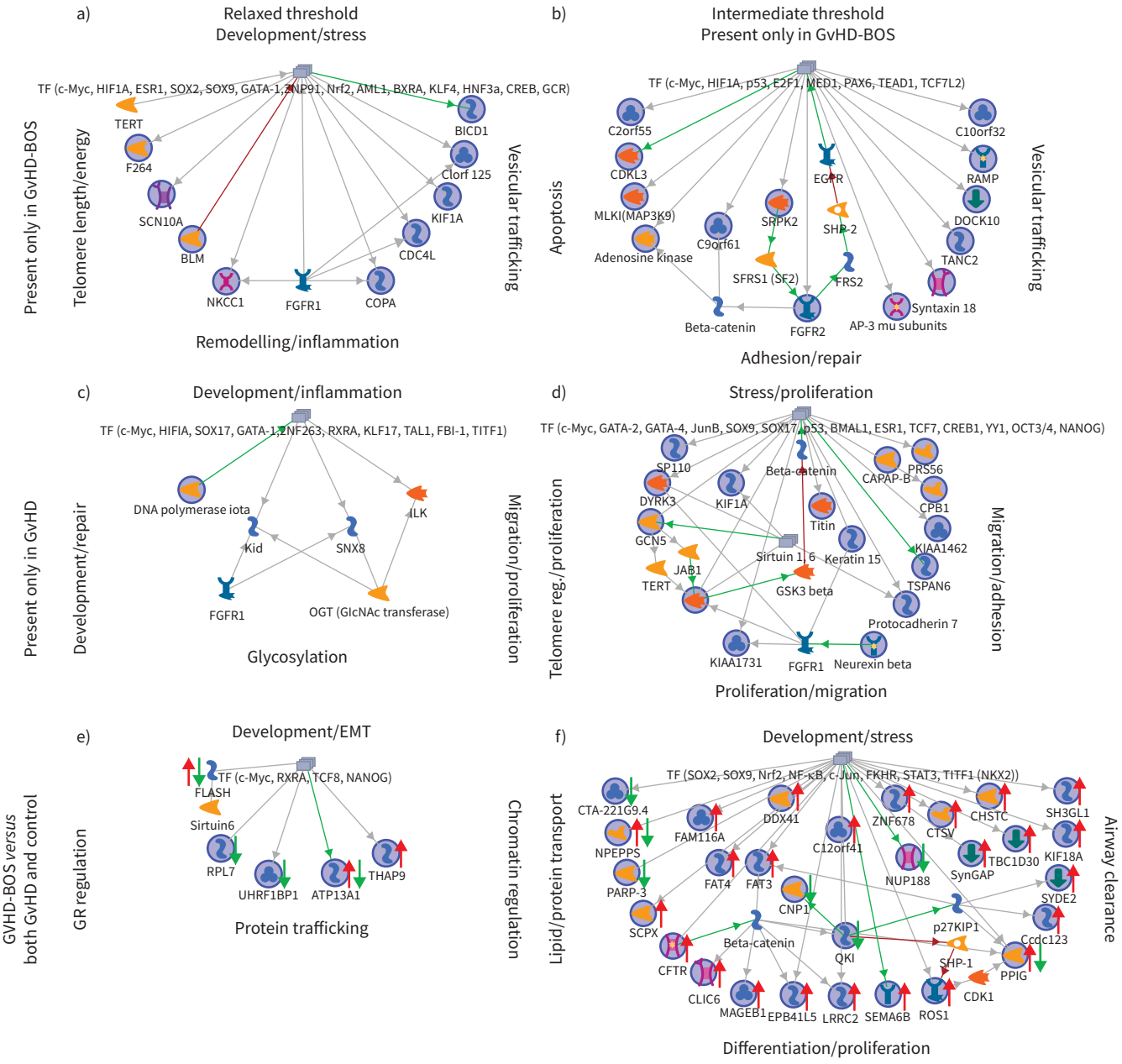


FIGURE 5 Network analysis of markers that segregate graft-*versus*-host disease (GvHD)-bronchiolitis obliterans syndrome (BOS), GvHD and control. Non-biased blinded and randomised proteomic analyses of banked plasma samples collected from haematopoietic stem cell transplant (HSCT) paediatric patients that went on to develop GvHD-BOS, GvHD or no GvHD (control) an average of 68 days prior to clinical diagnosis. Mass spectrometry label-free quantitation was used to quantitate plasma protein isoform expression and the candidates that segregated patients at the relaxed threshold: a) present only in GvHD-BOS; c) present only in GvHD; and e) different in GvHD-BOS *versus* both GvHD and control (0.05). Analyses at the intermediate threshold: b) present only in GvHD-BOS; d) present only in GvHD; and f) different in GvHD-BOS *versus* both GvHD and control (0.05). Processes associated with identified protein isoforms are denoted on panels. Analyses for controls at the relaxed and intermediate thresholds are shown in supplementary figure S6. Network and process analyses was performed by Clarivate's MetaCore-GeneGo software. EMT; endothelial mesenchymal transition; GR: glucocorticoid receptor; TF: transcription factor.

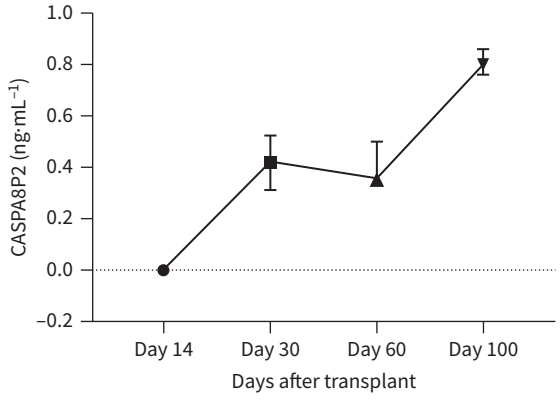
with to apoptosis and purine metabolism (figure 5b). GvHD was associated with increased protein glycosylation, fibrosis, protein trafficking and apoptosis at the relaxed threshold (figure 5c), and with cell proliferation, pancreatic enzyme regulation, cell migration and vesicular trafficking at the intermediate threshold (figure 5d). Controls exhibited increased expression of proteins that negatively regulate T-cells and govern their interaction with epithelial cells, as well as proteins that regulate cell proliferation, inflammation, glycosylation and fibrosis (supplementary figure S6). A notable difference between controls and GvHD-BOS or GvHD is a more pronounced network of β-catenin, sirtuin 1, 5 and 6, FGFR1 and FGFR2. The analysis of the network of proteins that differed in GvHD-BOS from GvHD and controls revealed dysregulation of airway clearance, lipid metabolism, tight junction integrity and cell proliferation in GvHD-BOS (figure 5e, f).

Taken together, the analyses of proteins that were present/absent or differentially expressed in GvHD-BOS *versus* both GvHD and control associated with epithelial cell dysfunction, inflammation, viral infection, proliferation or glucocorticoid receptor regulation, all relevant to GvHD-BOS lung disease. We selected the following six proteins based on their biological relevance to these associations for cross-platform and larger cohort validation: integrin-linked protein kinase (ILK), Kelch-like protein 5 (KLHL5), microtubule-associated tumour suppressor candidate 2 (KIAA0774), SAC3 domain-containing protein 1 (SAC3D1), interferon-induced protein with tetratricopeptide repeats (IFIT), endoplasmic reticulum transmembrane helix translocase (ATP13A1), sorting nexin 8 (SNX8) and caspase 8 associated protein 2 (CASP8AP or FLASH). While the selection of six proteins helped reduce validation work to a manageable number of proteins for the scope of the present study, we recognise the potential for the introduction of confirmation bias. Our selection strategy integrated both statistical rigour and biological plausibility, where we applied predefined statistical thresholds (*e.g.* Xcorr thresholds, p-values, fold changes) to narrow down candidates objectively first, and then to reduce the size of the number of validations, prioritised biomarkers that could be biologically plausible based on biological relevance to known features of the disease, informed by existing literature and mechanistic hypotheses.

ELISA validation of select markers discovered by MS

We first explored the expression of the protein caspase 8 associated protein 2 (a regulator of glucocorticoid receptor transactivation, FLASH in figure 5e) in a cohort of >100 consecutive HSCT patients. Levels of CASP8AP2 were measured by ELISA at days 14 (n=107), 30 (n=108), 60 (n=108) and 100 (n=134) after HSCT (demographics in supplementary table S2). Expression increased from the time of transplant (figure 6a) and analysis of day 100 post-transplant revealed a significant association with GvHD and odds of survival (figure 6b). Day 100 levels were significantly elevated in children that went on to develop GvHD and were similarly elevated in children that went on to develop BOS, but lacked statistical significance (p=0.18), likely due to a smaller number of BOS cases (figure 6b). Our ELISA data also confirmed that ILK, CASP8AP2, SAC3D1, IFIT1, ATPase13 and KIAA0774/MTUS2 plasma levels correlated with GvHD and/or GvHD-BOS, in some cases as early as 14 days post-HSCT (figure 7).

a)



b)

	100 days post transplant	
	Change	p-value
GvHD grades 2–4	↑ in GvHD	0.0001
BO	in BO	0.18
OS 1 year	↑ adverse	0.05

FIGURE 6 ELISA measurement of CASPA8P2 in children (n=134) after haematopoietic stem cell transplant (HSCT). a) Levels of CASPA8P2 (mean±95% CI) were undetectable 14 days after HSCT, but increased steadily through day 100, a conventional post-transplant observational timepoint. b) Elevated CASPA8P2 was associated with increased risk of graft-*versus*-host disease (GvHD) as in the proteomic analysis. BO: GvHD-bronchiolitis obliterans syndrome (BO); OS: overall survival.

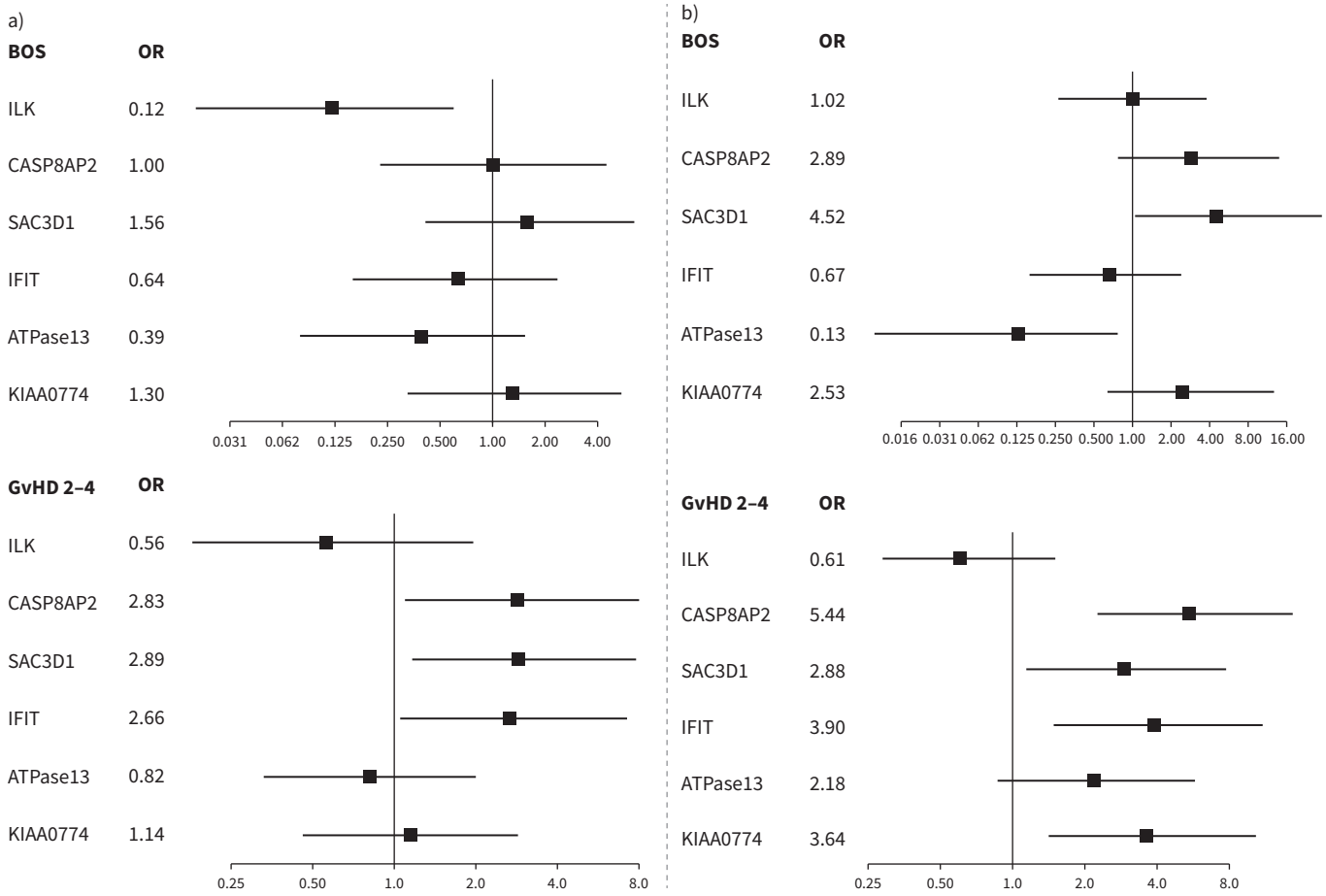


FIGURE 7 ELISA validation of potential plasma protein biomarkers. Banked samples, a) 30 and b) 100 days post-HSCT were analysed. Odds ratios (ORs) and corresponding 95% confidence intervals for graft-*versus*-host disease (GvHD)-bronchiolitis obliterans syndrome (BOS) and GvHD grades 2–4 are shown. KIAA0774/MTUS2: microtubule-associated tumour suppressor 2; ILK: integrin-linked protein kinase; CASP8AP2: caspase 8-associated protein 2; SAC3D1: SAC3 domain-containing protein 1; IFIT1: interferon-induced protein with tetratricopeptide repeats 1; ATPase13: endoplasmic reticulum transmembrane helix translocase 13A1. n=103–133, depending on timepoint.

Discussion

The central finding of our study is that MS proteomics identified blood proteins that can serve as biomarkers for the development of GvHD-BOS and/or GvHD in paediatric HSCT recipients, the first such report in children with HSCT. We identified significant shifts in the levels of ILK, CASP8AP2, SAC3D1, IFIT1, ATPase13 and KIAAA0774/MTUS2 ∼68 days prior to diagnosis, and validated this with ELISA in a larger cohort of patients as early as 30 and/or 100 days after HSCT. We selected this small number of potential markers based on both statistical thresholds first and relevance to the known biology of the disease based on pathways to which our selections belonged. While we recognise the potential for confirmation bias, we note that we did not select markers for validation based solely on any prior expectations but grounded our selection in data-driven criteria first and then biologically meaningful criteria. Markers that differed in GvHD-BOS and GvHD, as early as 30 days after HSCT, indicated changes in pathways that can be connected to disease including complement signalling, vesicular and protein trafficking, angiotensin maturation, protease and antiprotease regulation, fat-soluble vitamin transport and processing, fibrosis, vasculopathy, inflammation and wound healing. These pathways can all be logically linked to clinical features of GvHD-BOS and may provide insight into the underlying mechanisms of disease, many of which far precede the clinical diagnosis at which time treatment is applied under current standards of care.

GvHD and BOS are closely interrelated, with most children with BOS also having GvHD in other organs in addition to the lung. Conversely, only a minority of children that present with non-lung GvHD go on to develop BOS. This is consistent with our findings that identified several processes and pathways that were

abnormal in HSCT recipients with both GvHD-BOS and GvHD, including activation of complement in both cohorts prior to and at diagnosis. Other findings, such as the strong association with angiotensinogen maturation, provided intriguing support for the frequent clinical observation of hypertension in paediatric HSCT recipients diagnosed with GvHD-BOS [20].

In current clinical practice, it is not possible to forecast which HSCT paediatric recipients will go on to develop GvHD or transition from GvHD to GvHD-BOS. Furthermore, the mechanism of BOS initiation is poorly understood and difficult to study in children due to small case numbers and the fact that the disease is not commonly diagnosed until its well established (reflected in lung function decline) and organ injury that may be irreversible (*e.g.* fibrosis) has already occurred. Our inclusion of the GvHD and control cohorts in our wide range proteomic analysis revealed that there are strong plasma protein expression patterns that track with phenotype, even months before clinical diagnosis. This added understanding of the molecular phenotypes of children that do not develop GvHD or go on to develop GvHD or GvHD-BOS is promising to the development of approaches that can forecast risk, define interventions and/or assess responsiveness to therapy. The soundness of our approach and novel discoveries (*e.g.* altered processing of lipids, angiotensin, protein glycosylation or responses to hypoxia) is supported by other findings that were expected (*e.g.* complement activation). For example, our data indicated early alterations in fat-soluble vitamin metabolism and transport in GvHD-BOS, which supports our previous work which reported an association between vitamin A levels and GvHD [21]. Further work is needed to investigate the significance of many of the differences we identified in paediatric HSCT that go on to develop GvHD-BOS.

Important progress has been made by others in the identification and broad validation of biomarkers predicting response to interventions for early acute (a) GvHD, and these biomarkers are now incorporated into clinical practice [12, 22]. A number of secreted markers have been validated for chronic (c) GvHD following BMT, including CXCL10 and MMP3, and demonstrated that the combination of markers improves risk prediction [23]. Other studies have shown the prognostic value of CXCL10 and the number of dendritic cells for predicting cGvHD following HSCT [24]. Metabolites have also been examined as potential biomarkers in cGvHD and late aGvHD, with α-ketoglutaric and fumaric acid levels predicting cGvHD and glutamic and citric acid levels for predicting aGvHD [25]. Our studies can support the addition of new markers including those related to hypoxia (HIF 1α), protein transport (ATPase13A1) and/or TNF-α-mediated modulation of glucocorticoid receptor (CASP8AP2), some of which are supported by cross-platform validated in larger cohorts of >100 consecutive HSCTs at multiple early timepoints after transplant (30 and 100 days after transplant). With the average onset of GvHD-BOS being ∼1 year after transplant, our identification of very early changes provide promise for early risk stratification, which would transform care for these paediatric patients. Additionally, our data provide a resource for identifying drug repurposing and novel therapies.

Our study has strengths and limitations. The depth of our proteomic analysis of ∼10 000–14 000 protein isoforms is more extensive than prior work on GvHD-BOS and GvHD in paediatric HSCT recipients, where existing research is very limited. While this analysis allowed for the identification of important changes in low abundance proteins, MS proteomics is semi-quantitative, and therefore a manageable ELISA validation was added to the experimental design to strengthen the confidence of our findings. Validation of our other protein findings is warranted and ongoing. Our prospective storage of clinical samples was a key resource that allowed all our studies to detect early changes prior to clinical diagnosis when disease is typically established or advanced. Moreover, frequent prospective sample collection, ELISA analysis and association of biomarker data with clinical outcomes supports the approach. Multiple comparisons were made in this analysis, although the ELISA analyses were carefully selected based on the preliminary data of the proteomics. However, as demonstrated by the strong patterning of the protein expression data at relaxed, as well as stringent, statistical thresholds and the validation of some of these, our study suggests that highly stringent data reduction is not beneficial if more relaxed considerations are ignored. Another limitation was a small sample size, which is due to the rarity of paediatric GvHD-BOS sampling, and our single-centre study design can induce bias. Further validation in larger populations from multiple institutions is necessary to establish the potential value and generalisability of the identified potential biomarkers and the additional important proteins identified here that should be further explored. Finally, while we based the study on lung function metrics that are used clinically to diagnose pulmonary disease after HSCT, we did not integrate longitudinal pulmonary function data or radiographic data with our biomarker data. We plan to examine the integration of our MS biomarker data with pulmonary function and radiographical data with large sample sets presently being acquired from the TRANSPIRE multicentre study, which is collecting both plasma and pulmonary function and radiographical data across eight paediatric HSCT centres.

The results of the current study are the first report of biomarker and pathway examination in children after HSCT and are the beginning of a clinical strategy of screening and diagnosing GvHD-BOS, prior to irreversible lung injury or damage occurring. Earlier detection of this severe pulmonary complication of HSCT in children will allow for stratification of screening and risk assessment-based therapeutic interventions. Significant improvement of post-HSCT care with marker-informed intervention prior to the onset of GvHD-BOS or GvHD, as well as the monitoring of response to therapy can be achieved with this strategy.

Provenance: Submitted article, peer reviewed.

Ethics statement: Samples and clinical data utilised in this work were collected and used with prior approval of the Cincinnati Children's Hospital Medical Center (CCHMC) institutional review board.

Author contributions: K.C. Myers, S.M. Davies and A.G. Ziady conceived the studies and designed experiments. K.C. Myers and S.M. Davies managed the CCHMC HSCT sample repository and selected samples for proteomic analysis. E.J. Skala and A.G. Ziady led the proteomic analyses and data/figure assembly. K.C. Myers, A.G. Ziady, S.M. Davies, N. Mousa and E.J. Skala wrote the manuscript. E.J. Skala, S. Ghandikota, A.G. Jegga, N. Mousa, A. Bridgeland, M.E. Siefert and A.G. Ziady conducted proteomic biomarker and pathway analyses. N. Luebbering, S. Abdullah and A. Ibrahminova performed ELISA validation studies. C. Towe and S. Goldfarb contributed to data interpretation in the context of the clinical presentation of GvHD and GvHD-BOS. E.J. Skala and A. Lane performed statistical analyses. V. Reffatto contributed to the preparation of the manuscript, development of figures and editing text.

Conflict of interest: K.C. Myers reports grants from Elixirgen Therapeutics and Incyte, and patents planned, issued or pending (Compositions and methods for the treatment of bronchiolitis obliterans – 2019-1001a). E.J. Skala reports patents planned, issued or pending (Compositions and methods for the treatment of bronchiolitis obliterans – 2019-1001a). S. Ghandikota, A.G. Jegga, A. Bridgeland, S. Abdullah, S. Goldfarb, V. Reffatto, M.E. Siefert, A. Lane, A. Ibrahminova and N. Luebbering report no conflicts of interest. N. Mousa reports patents planned, issued or pending (Compositions and methods for the treatment of bronchiolitis obliterans – 2019-1001a). C. Towe reports consultancy fees from Sanofi and Boehringer-Ingelheim and patents planned, issued or pending (Compositions and methods for the treatment of bronchiolitis obliterans – 2019-1001a). S.M. Davies reports patents planned, issued or pending (Compositions and methods for the treatment of bronchiolitis obliterans – 2019-1001a). A.G. Ziady reports patents planned, issued or pending (Compositions and methods for the treatment of bronchiolitis obliterans – 2019-1001a).

Support statement: This study was supported by R01HL153108-01A1 (to K.C. Myers), R01HL142210 (to A.G. Ziady), R61HL154105 (to A.G. Ziady), R01HL151588 (to A.G. Ziady), the CCHMC Research Foundation Research Innovation Program (K.C. Myers), U24 CA55727-11, R01 HD093773, R01 DK125418, R01 HL153723, 1R01HL157392-01 and 1U01AI157620-01 (to S.M. Davies). Funding information for this article has been deposited with the Open Funder Registry.

References

1 Uhlving HH, Buchvald F, Heilmann CJ, *et al.* Bronchiolitis obliterans after allo-SCT: clinical criteria and treatment options. *Bone Marrow Transplant* 2012; 47: 1020–1029.
2 Haider S, Durairajan N, Soubani AO. Noninfectious pulmonary complications of haematopoietic stem cell transplantation. *Eur Respir Rev* 2020; 29: 190119.
3 Tamburro RF, Cooke KR, Davies SM, *et al.* Pulmonary complications of pediatric hematopoietic cell transplantation. A National Institutes of Health Workshop Summary. *Ann Am Thorac Soc* 2021; 18: 381–394.
4 Bergeron A, Chevret S, Peffault de Latour R, *et al.* Noninfectious lung complications after allogeneic haematopoietic stem cell transplantation. *Eur Respir J* 2018; 51: 1702617.
5 Cheng GS, Storer B, Chien JW, *et al.* Lung function trajectory in bronchiolitis obliterans syndrome after allogeneic hematopoietic cell transplant. *Ann Am Thorac Soc* 2016; 13: 1932–1939.
6 Estenne M, Maurer JR, Boehler A, *et al.* Bronchiolitis obliterans syndrome 2001: an update of the diagnostic criteria. *J Heart Lung Transplant* 2002; 21: 297–310.
7 Abedin S, Yanik GA, Braun T, *et al.* Predictive value of bronchiolitis obliterans syndrome stage 0p in chronic graft-versus-host disease of the lung. *Biol Blood Marrow Transplant* 2015; 21: 1127–1131.
8 Jagasia MH, Greinix HT, Arora M, *et al.* National Institutes of Health Consensus development project on criteria for clinical trials in chronic graft-versus-host disease: I. The 2014 diagnosis and staging working group report. *Biol Blood Marrow Transplant* 2015; 21: 389–401.e1.

9 Aurora P, Stocks J, Oliver C, *et al.* Quality control for spirometry in preschool children with and without lung disease. *Am J Respir Crit Care Med* 2004; 169: 1152–1159.

10 Wolff D, Greinix H, Lee SJ, *et al.* Biomarkers in chronic graft-versus-host disease: quo vadis? *Bone Marrow Transplant* 2018; 53: 832–837.

11 Yu J, Storer BE, Kushekhar K, *et al.* Biomarker panel for chronic graft-versus-host disease. *J Clin Oncol* 2016; 34: 2583–2590.

12 Aziz MD, Shah J, Kapoor U, *et al.* Disease risk and GVHD biomarkers can stratify patients for risk of relapse and nonrelapse mortality post hematopoietic cell transplant. *Leukemia* 2020; 34: 1898–1906.

13 Srinagesh HK, Özbek U, Kapoor U, *et al.* The MAGIC algorithm probability is a validated response biomarker of treatment of acute graft-versus-host disease. *Blood Adv* 2019; 3: 4034–4042.

14 Major-Monfried H, Renteria AS, Pawarode A, *et al.* MAGIC biomarkers predict long-term outcomes for steroid-resistant acute GVHD. *Blood* 2018; 131: 2846–2855.

15 Hartwell MJ, Özbek U, Holler E, *et al.* An early-biomarker algorithm predicts lethal graft-versus-host disease and survival. *JCI Insight* 2018; 2: e124015.

16 Liu X, Yue Z, Yu J, *et al.* Proteomic characterization reveals that MMP-3 correlates with bronchiolitis obliterans syndrome following allogeneic hematopoietic cell and lung transplantation. *Am J Transplant* 2016; 16: 2342–2351.

17 Glucksberg H, Storb R, Fefer A, *et al.* Clinical manifestations of graft-versus-host disease in human recipients of marrow from HL-A-matched sibling donors. *Transplantation* 1974; 18: 295–304.

18 Paczesny S, Hakim FT, Pidala J, *et al.* National Institutes of Health consensus development project on criteria for clinical trials in chronic graft-versus-host disease: III. The 2014 biomarker working group report. *Biol Blood Marrow Transplant* 2015; 21: 780–792.

19 Przepiorka D, Weisdorf D, Martin P, *et al.* 1994 consensus conference on acute GVHD grading. *Bone Marrow Transplant* 1995; 15: 825–828.

20 Rotz SJ, Ryan TD, Hayek SS. Cardiovascular disease and its management in children and adults undergoing hematopoietic stem cell transplantation. *J Thromb Thrombolysis* 2021; 51: 854–869.

21 Lounder DT, Khandelwal P, Dandoy CE, *et al.* Lower levels of vitamin A are associated with increased gastrointestinal graft-versus-host disease in children. *Blood* 2017; 129: 2801–2807.

22 Vander Lugt MT, Braun TM, Hanash S, *et al.* ST2 as a marker for risk of therapy-resistant graft-versus-host disease and death. *N Engl J Med* 2013; 369: 529–539.

23 Logan BR, Fu D, Howard A, *et al.* Validated graft-specific biomarkers identify patients at risk for chronic graft-versus-host disease and death. *J Clin Invest* 2023; 133: e168575.

24 Chirumbolo G, Dicataldo M, Barone M, *et al.* A Multiparameter prognostic risk score of chronic graft-versus-host disease based on CXCL10 and plasmacytoid dendritic cell levels in the peripheral blood at 3 months after allogeneic hematopoietic stem cell transplantation. *Transplant Cell Ther* 2023; 29: 302.e1–e8.

25 Subburaj D, Ng B, Kariminia A, *et al.* Metabolomic identification of alpha-ketoglutaric acid elevation in pediatric chronic graft-versus-host disease. *Blood* 2022; 139: 287–299.