C-terminal mutation may have slightly poorer EFS and OS than those who harbored an N-terminal mutation (Figures 4c and d), which is not consistent with previous adult AML studies.[12,13] Gene expression profiling suggests that *CEBPA*-single patients harboring a C-terminal mutation are more similar to *CEBPA*-double patients than to *CEBPA*-single patients harboring N-terminal mutations.[10] This latter study was performed in adult AML patients and needs to be validated in pediatric AML patients.

So far, the biological mechanisms underlying a favorable clinical outcome for AML patients harboring *CEBPA* mutations (including relative drug sensitivity) are not clear. Further studies of single and double *CEBPA* mutations and the underlying biology are required to enable better risk assessment and therapeutic approaches in pediatric AML.

CONCLUSION

This is the first nationwide study to examine the clinical significance of *CEBPA* mutations in Japanese pediatric AML patients. The results suggest that *CEBPA* mutations, especially double or triple *CEBPA* mutations, are an independent favorable prognostic factor for pediatric AML patients. *CEBPA*-double patients should be stratified into the favorable risk group, and the prognostic significance of these mutations should be validated prospectively.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

This work was supported by a Grant for Clinical Cancer Research and a Grant-in-Aid for Cancer Research from the Ministry of Health, Labour and Welfare of Japan. We also thank to H Miyachi, N Kiyokawa, T Deguchi, Y Hayashi and all doctors participating in JPLSG AML-05 protocol.

REFERENCES

1 Tenen DG, Hromas R, Licht JD, Zhang DE. Transcription factors, normal myeloid development, and leukemia. *Blood* 1997; **90**: 489–519.
2 Porse BT, Pedersen TA, Xu X, Lindberg B, Wewer UM, Friis-Hansen L *et al.* E2F repression by C/EBPalpha is required for adipogenesis and granulopoiesis *in vivo*. *Cell* 2001; **107**: 247–258.
3 Johansen LM, Iwama A, Lodie TA, Sasaki K, Felsher DW, Golub TR *et al.* c-Myc is a critical target for c/EBPalpha in granulopoiesis. *Mol Cell Biol* 2001; **21**: 3789–3806.
4 Leroy H, Roumier C, Huyghe P, Biggio V, Fenaux P, Preudhomme C. CEBPA point mutations in hematological malignancies. *Leukemia* 2005; **19**: 329–334.
5 Nerlov C. C/EBPalpha mutations in acute myeloid leukaemias. *Nat Rev Cancer* 2004; **4**: 394–400.
6 Pabst T, Mueller BU, Zhang P, Radomska HS, Narravula S, Schnittger S *et al.* Dominant-negative mutations of CEBPA, encoding CCAAT/enhancer binding protein-alpha (C/EBPalpha), in acute myeloid leukemia. *Nat Genet* 2001; **27**: 263–270.
7 Asou H, Gombart AF, Takeuchi S, Tanaka H, Tanioka M, Matsui H *et al.* Establishment of the acute myeloid leukemia cell line Kasumi-6 from a patient with a dominant-negative mutation in the DNA-binding region of the C/EBPalpha gene. *Genes Chromosomes Cancer* 2003; **36**: 167–174.
8 Preudhomme C, Sagot C, Boissel N, Cayuela JM, Tigaud I, de Botton S *et al.* Favorable prognostic significance of CEBPA mutations in patients with *de novo* acute myeloid leukemia: a study from the Acute Leukemia French Association (ALFA). *Blood* 2002; **100**: 2717–2723.
9 Pabst T, Eyholzer M, Fos J, Mueller BU. Heterogeneity within AML with *CEBPA* mutations; only *CEBPA* double mutations, but not single *CEBPA* mutations are associated with favourable prognosis. *Br J Cancer* 2009; **100**: 1343–1346.
10 Wouters BJ, Löwenberg B, Erpelinck-Verschueren CA, van Putten WL, Valk PJ, Delwel R. Double *CEBPA* mutations, but not single *CEBPA* mutations, define a subgroup of acute myeloid leukemia with a distinctive gene expression profile that is uniquely associated with a favorable outcome. *Blood* 2009; **113**: 3088–3091.
11 Dufour A, Schneider F, Metzeler KH, Hoster E, Schneider S, Zellmeier E *et al.* Acute myeloid leukemia with biallelic *CEBPA* gene mutations and normal karyotype represents a distinct genetic entity associated with a favorable clinical outcome. *J Clin Oncol* 2010; **28**: 570–577.
12 Green CL, Koo KK, Hills RK, Burnett AK, Linch DC, Gale RE. Prognostic significance of *CEBPA* mutations in a large cohort of younger adult patients with acute myeloid leukemia: impact of double *CEBPA* mutations and the interaction with *FLT3* and *NPM1* mutations. *J Clin Oncol* 2010; **28**: 2739–2747.
13 Fasan A, Haferlach C, Alpermann T, Jeromin S, Grossmann V, Eder C *et al.* The role of different genetic subtypes of *CEBPA* mutated AML. *Leukemia* 2013; **28**: 794–803.
14 Liang DC, Shih LY, Huang CF, Hung IJ, Yang CP, Liu HC *et al.* CEBPalpha mutations in childhood acute myeloid leukemia. *Leukemia* 2005; **19**: 410–414.
15 Ho PA, Alonzo TA, Gerbing RB, Pollard J, Stirewalt DL, Hurwitz C *et al.* Prevalence and prognostic implications of *CEBPA* mutations in pediatric acute myeloid leukemia (AML): a report from the Children's Oncology Group. *Blood* 2009; **113**: 6558–6566.
16 Hollink IH, van den Heuvel-Eibrink MM, Arentsen-Peters ST, Zimmermann M, Peeters JK, Valk PJ *et al.* Characterization of *CEBPA* mutations and promoter hypermethylation in pediatric acute myeloid leukemia. *Haematologica* 2011; **96**: 384–392.
17 Staffas A, Kanduri M, Hovland R, Rosenquist R, Ommen HB, Abrahamsson J *et al.* Presence of *FLT3*-ITD and high *BAALC* expression are independent prognostic markers in childhood acute myeloid leukemia. *Blood* 2011; **118**: 5905–5913.
18 Mizushima Y, Taki T, Shimada A, Yui Y, Hiraumi Y, Matsubara H *et al.* Prognostic significance of the *BAALC* isoform pattern and *CEBPA* mutations in pediatric acute myeloid leukemia with normal karyotype: a study by the Japanese Childhood AML Cooperative Study Group. *Int J Hematol* 2010; **91**: 831–837.
19 Tomizawa D, Tawa A, Watanabe T, Saito AM, Kudo K, Taga T *et al.* Excess treatment reduction including anthracyclines results in higher incidence of relapse in core binding factor acute myeloid leukemia in children. *Leukemia* 2013; **27**: 2413–2416.
20 Tomizawa D, Tawa A, Watanabe T, Saito AM, Kudo K, Taga T *et al.* Appropriate dose reduction in induction therapy is essential for the treatment of infants with acute myeloid leukemia: a report from the Japanese Pediatric Leukemia/Lymphoma Study Group. *Int J Hematol* 2013; **98**: 578–588.
21 Wouters BJ, Louwers I, Valk PJ, Löwenberg B, Delwel R. A recurrent in-frame insertion in a CEBPA transactivation domain is a polymorphism rather than a mutation that does not affect gene expression profiling-based clustering of AML. *Blood* 2007; **109**: 389–390.
22 Biggio V, Renneville A, Nibourel O, Philippe N, Terriou L, Roumier C *et al.* Recurrent in-frame insertion in C/EBPalpha TAD2 region is a polymorphism without prognostic value in AML. *Leukemia* 2008; **22**: 655–657.
23 Kim S, Kim DH, Jang JH, Jung CW, Jang MA, Ki CS *et al.* Novel mutations in CEBPA in Korean Patients with acute myeloid leukemia with a normal karyotype. *Ann Lab Med* 2012; **32**: 153–157.
24 Bereshchenko O, Mancini E, Moore S, Bilbao D, Månsson R, Luc S *et al.* Hematopoietic stem cell expansion precedes the generation of committed myeloid leukemia-initiating cells in C/EBPalpha mutant AML. *Cancer Cell* 2009; **16**: 390–400.
25 Somervaille TC, Cleary ML. Preview. Mutant CEBPA: priming stem cells for myeloid leukemogenesis. *Cell Stem Cell* 2009; **5**: 453–454.
26 Dufour A, Schneider F, Hoster E, Benthaus T, Ksienzyk B, Schneider S *et al.* Monoallelic CEBPA mutations in normal karyotype acute myeloid leukemia: independent favorable prognostic factor within NPM1 mutated patients. *Ann Hematol* 2012; **91**: 1051–1063.
27 Park SH, Chi HS, Cho YU, Jang S, Park CJ. CEBPA single mutation can be a possible favorable prognostic indicator in NPM1 and FLT3-ITD wild-type acute myeloid leukemia patients with intermediate cytogenetic risk. *Leuk Res* 2013; **37**: 1488–1494.

EVI1 overexpression is a poor prognostic factor in pediatric patients with *mixed lineage leukemia-AF9* rearranged acute myeloid leukemia

The ecotropic viral integration site-1 gene (*EVI1*) encodes a zinc finger protein that functions as a transcriptional regulator of hematopoietic stem cell self-renewal and long-term multilineage repopulating activity.[1,2] The mixed lineage leukemia gene (*MLL*) rearrangements [i.e. t(11q23)] occur at high frequency in pediatric acute myeloid leukemia (AML) patients with *EVI1* overexpression,[3] and *EVI1* is a transcriptional target of MLL oncoproteins.[4] *EVI1* overexpression has been reported in up to 10% of patients with AML and is associated with an adverse prognosis. However, the prognostic value of *EVI1* overexpression has been studied mostly in adult AML.[5-9] Only two studies have examined *EVI1* overexpression in pediatric AML, but a detailed analysis according to the type of leukemia was not performed because of the small sample size.[3,10]

Recent data from an international consortium, including those from our group, suggest that pediatric *MLL*-rearranged AML can be divided into certain risk groups on the basis of different translocation partners.[11] However, clinical outcome data leading to risk stratification of the *MLL*-rearranged subgroups are still scarce and further investigation is necessary to identify new prognostic factors. Here, we retrospectively examined *EVI1* expression levels and clinical outcomes of pediatric *MLL*-rearranged AML patients treated in the Japanese Pediatric Leukemia/Lymphoma Study Group (JPLSG) AML-05 study.

After excluding patients with acute promyelocytic leukemia, Down syndrome, secondary AML, myeloid/natural killer cell leukemia and myeloid sarcoma, 485 AML patients were enrolled in the AML-05 study. Overall, 42 patients were excluded, mainly because of misdiagnosis. Details of the treatment schedules and risk stratification were described in previous publication.[12] This study was conducted in accordance with the principles set down in the Declaration of Helsinki and was approved by the Ethics Committees of all participating institutions. All patients, or the patients' parents/guardians, provided written informed consent.

RNA obtained from diagnostic bone marrow samples was used to analyze the expression of *EVI1* using a previously established *EVI1* quantitative real-time polymerase chain reaction assay that covers the various *EVI1* splice variants.[7] Event-free survival (EFS) was defined as the time from the diagnosis of AML to the last follow up or the first event (failure to achieve remission, relapse, secondary malignancy, or any cause of death). In this study, most of the events were relapses (n=23) and the rest were deaths with sepsis (n=1) and acute respiratory distress syndrome (n=1). Overall survival (OS) was defined as the time from the diagnosis of AML to any cause of death. All tests were two-tailed and $P<0.05$ was considered statistically significant.

Among 443 eligible AML patients, 69 were diagnosed as *MLL*-rearranged AML and diagnostic samples from 50 patients were analyzed for *EVI1* mRNA expression. No significant differences in the characteristics and clinical outcomes were observed between these 50 patients and the 19 patients who did not have *EVI1* data [EFS (P=0.20), OS (P=0.45)]. *EVI1* expression levels were dichotomized based on a cut off of 0.1 relative to SKOV3, an ovarian carcinoma cell line over-expressing *EVI1*: values higher than 0.1 were defined as *EVI1*$^+$ and those lower than 0.1 or undetectable

Table 1. Characteristics of patients categorized according to *EVI1* expression status.

	All (n=50)				
	EVI1$^-$ (n=32)		*EVI1*$^+$ (n=18)		*P*
Age (years)					0.03#
median	4.5		6.6		
range	0.1-14.7		0.8-15.1		
Sex, n(%)					0.77*
male	16	(50)	8	(44)	
female	16	(50)	10	(56)	
WBC(x10^9/L)					0.01#
median	48.4		88.7		
range	0.8-459		4.1-322		
Types of *MLL* rearrangement, n(%)					0.96*
MLL-AF6	2	(6)	1	(6)	
MLL-AF9	18	(56)	11	(61)	
MLL-AF10	5	(16)	2	(11)	
MLL-ELL	3	(9)	3	(17)	
MLL-ENL	3	(9)	1	(6)	
MLL-AF17	1	(3)	0	(0)	
FAB, n(%)					<0.0001*
M1	1	(3)	3	(17)	
M2	0	(0)	1	(6)	
M4	2	(6)	6	(33)	
M5	27	(84)	4	(22)	
RAEB-T	0	(0)	3	(17)	
Unclassified	2	(6)	1	(6)	
FLT3-ITD, n(%)	0	(0)	3	(17)	0.04*

*WBC: white blood cell count; FAB: French-American-British. *Fisher's exact test. #Mann-Whitney U test.*

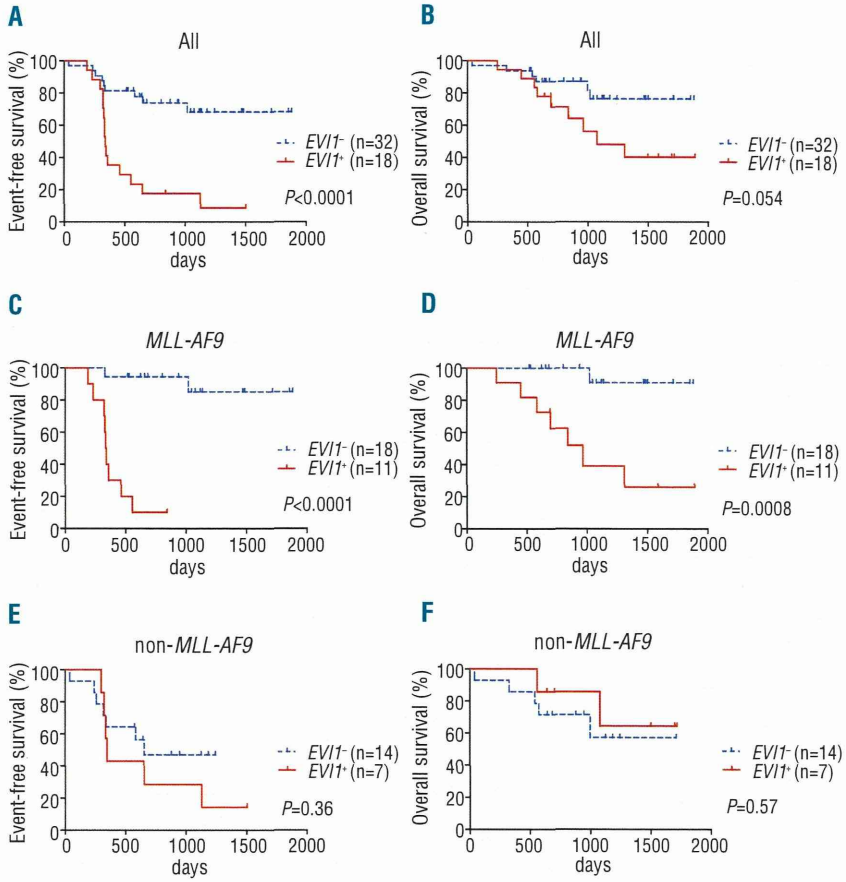

Figure 1. Kaplan-Meier survival curves of event-free survival (EFS) and overall survival (OS) from the time of diagnosis according to *EVI1* expression status. (A) Kaplan-Meier estimates of EFS in the cohort of *MLL*-rearranged AML in *EVI1*$^+$ and *EVI1*$^-$ patients. (B) Kaplan-Meier estimates of OS in the cohort of *MLL*-rearranged AML in *EVI1*$^+$ and *EVI1*$^-$ patients. (C) Kaplan-Meier estimates of EFS in the cohort of *MLL-AF9* in *EVI1*$^+$ and *EVI1*$^-$ patients. (D) Kaplan-Meier estimates of OS in the cohort of *MLL-AF9* in *EVI1*$^+$ and *EVI1*$^-$ patients. (E) Kaplan-Meier estimates of EFS in the cohort of *MLL*-rearranged AML without *MLL-AF9* in *EVI1*$^+$ and *EVI1*$^-$ patients. (F) Kaplan-Meier estimates of OS in the cohort of *MLL*-rearranged AML without *MLL-AF9* in *EVI1*$^+$ and *EVI1*$^-$ patients. *P* values determined using the log rank test.

were defined as *EVI1*$^-$, as described in a previous study.[7] *EVI1*$^+$ was present in 18 patients (36%). *EVI1* expression levels in different *MLL* translocation partners relative to that in SKOV3 cells are shown in *Online Supplementary Figure S1*. The clinical features of *EVI1*$^+$ and *EVI1*$^-$ patients are summarized in Table 1. *EVI1*$^+$ patients were significantly older (P=0.03) and had a higher WBC count (P=0.01) at the time of diagnosis than *EVI1*$^-$ patients. Most of the *MLL*-rearranged AML cases were classified as FAB-M5 or FAB-M4. Specifically, most *EVI1*$^-$ patients (84%) presented with FAB-M5 morphology, which was less frequent in *EVI1*$^+$ patients (22%), consistent with the findings of a previous study.[8] *EVI1*$^+$ was not correlated with sex or *MLL* translocation partners. The frequency of *FLT3*-ITD was significantly higher in *EVI1*$^+$ patients (P=0.04). We also analyzed *CEBPA* and *NPM1* mutations, which are established favorable prognostic factors; however, none of the patients harbored these mutations, except for one *EVI1*$^-$ patient harboring double *CEBPA* mutations.

Next, clinical outcomes were compared between *EVI1*$^+$ patients and *EVI1*$^-$ patients (Figure 1). In the *MLL*-rearranged AML cohort (n=50), *EVI1*$^+$ patients had a significantly worse EFS than *EVI1*$^-$ patients (P<0.0001) (Figure 1A). However, OS did not differ significantly between the two groups (P=0.054) (Figure 1B). Among several types of *MLL*-rearrangements, *MLL-AF9* was the most common translocation (n=29, 58%) (Table 1). Therefore, clinical outcomes in the cohort of *MLL-AF9* positive patients were compared between *EVI1*$^+$ patients (n=11) and *EVI1*$^-$ patients (n=18). The results showed significant differences in EFS (P<0.0001) and OS (P=0.0008) (Figure 1C and D). By contrast, no differences in EFS (P=0.36) or OS (P=0.57) were observed among patients with *MLL*-rearranged AML after excluding *MLL-AF9* positive patients (Figure 1E and F). The clinical outcomes associated with each type of *MLL*-rearrangement could not be analyzed because of the small sample size. Multivariate Cox regression analysis, including *FLT3*-ITD, WBC count, and age identified *EVI1*$^+$ as the only prognostic factor predicting poor EFS in the total cohort of *MLL*-rearranged AML (hazard ratio (HR), 4.94; P<0.01) and in the *MLL-AF9* positive cohort (HR, 33.81; P<0.01), but not OS (*Online Supplementary Table S1*).

These results suggest that *EVI1* overexpression is an independent adverse prognostic factor because of its association with reduced remission duration in pediatric patients with *MLL*-rearranged AML, especially in patients harboring *MLL-AF9*. A recent large study identified several novel prognostic *MLL*-rearranged subgroups, including a favorable-risk *MLL-AF1q* positive subgroup and a poor-risk *MLL-AF6* positive subgroup.[11] However, *MLL-AF9* positive patients are categorized as an intermediate risk group, and this subgroup may be dichotomized as a favorable and poor-risk subgroup based on *EVI1* expression levels. Pretreatment screening for *EVI1* expression should be considered in patients with *MLL*-rearranged AML to enable better risk assessment and alternative consolidation therapies to be considered. Our results need to be confirmed in larger studies because of the limited case numbers.

From a biological viewpoint, the 'evil'-like adverse effects of *EVI1* in patients with *MLL-AF9*-positive AML were partially elucidated in a recent study in which *EVI1* positive cells harboring *MLL-AF9* showed distinct morphological,

molecular, and mechanistic differences from *EVI1* negative cells.[13] Moreover, *EVI1* overexpression has been linked to CD52 overexpression, which could be a therapeutic target for monoclonal antibody treatment.[14] Further investigation is required to identify novel prognostic factors in the various subgroups of *MLL*-rearranged AML and to develop therapeutic strategies effective for patients with *EVI1* overexpression.

Hidemasa Matsuo,[1] Mio Kajihara,[1] Daisuke Tomizawa,[2] Tomoyuki Watanabe,[3] Akiko Moriya Saito,[4] Junichiro Fujimoto,[5] Keizo Horibe,[4] Kumi Kodama,[1] Mayu Tokumasu,[6] Hiroshi Itoh,[1] Hideki Nakayama,[7] Akitoshi Kinoshita,[8] Takashi Taga,[9] Akio Tawa,[10] Tomohiko Taki,[11] Norio Shiba,[12] Kentaro Ohki,[13] Yasuhide Hayashi,[13] Yuka Yamashita,[14] Akira Shimada,[15] Shiro Tanaka,[16] and Souichi Adachi[1]

[1]Department of Human Health Sciences, Kyoto University; [2]Department of Pediatrics, Tokyo Medical and Dental University (TMDU); [3]Department of Nutritional Science, Aichi Gakuin University, Nisshin; [4]Clinical Research Center, National Hospital Organization Nagoya Medical Center; [5]Epidemiology and Clinical Research Center for Children's Cancer, National Center for Child Health and Development, Tokyo; [6]Department of Pediatrics, Kyoto University; [7]Department of Pediatrics, National Hospital Organization Fukuoka-Higashi Medical Center, Koga; [8]Department of Pediatrics, St. Marianna University School of Medicine, Kawasaki; [9]Department of Pediatrics, Shiga University of Medical Science, Otsu; [10]Department of Pediatrics, National Hospital Organization Osaka National Hospital; [11]Department of Molecular Diagnostics and Therapeutics, Kyoto Prefectural University of Medicine; [12]Department of Pediatrics, Gunma University Graduate School of Medicine, Maebashi; [13]Department of Hematology and Oncology, Gunma Children's Medical Center, Shibukawa; [14]Department of Pediatrics, National Hospital Organization Nagoya Medical Center; [15]Department of Pediatrics, Okayama University Hospital; and [16]Department of Pharmacoepidemiology, Kyoto University, Japan

Trial registration: This trial is registered with the UMIN Clinical Trials Registry (UMIN-CTR, URL: http://www.umin.ac.jp/ctr/index.htm; number UMIN000000511).

Correspondence: adachiso@kuhp.kyoto-u.ac.jp
doi:10.3324/haematol.2014.107128

Acknowledgments: We thank all the clinicians participating in the JPLSG AML-05 study.

Funding: This work was supported by a Grant for Clinical Cancer Research and a Grant-in-Aid for Cancer Research from the Ministry of Health, Labour and Welfare of Japan.

Key words: acute myeloid leukemia, pediatrics, MLL, EVI1, prognostic factor.

Information on authorship, contributions, and financial & other disclosures was provided by the authors and is available with the online version of this article at www.haematologica.org.

References

1. Morishita K, Parker DS, Mucenski ML, Jenkins NA, Copeland NG, Ihle JN. Retroviral activation of a novel gene encoding a zinc finger protein in IL-3-dependent myeloid leukemia cell lines. Cell. 1988;54(6):831-40.
2. Kataoka K, Sato T, Yoshimi A, Goyama S, Tsuruta T, Kobayashi H et al. Evi1 is essential for hematopoietic stem cell self-renewal, and its expression marks hematopoietic cells with long-term multilineage repopulating activity. J Exp Med. 2011;208(12):2403-16.
3. Ho PA, Alonzo TA, Gerbing RB, Pollard JA, Hirsch B, Raimondi SC et al. High EVI1 expression is associated with MLL rearrangements and predicts decreased survival in paediatric acute myeloid leukaemia: a report from the children's oncology group. Br J Haematol. 2013;162(5):670-7.
4. Arai S, Yoshimi A, Shimabe M, Ichikawa M, Nakagawa M, Imai Y et al. Evi-1 is a transcriptional target of mixed-lineage leukemia oncoproteins in hematopoietic stem cells. Blood. 2011;117(23):6304-14.
5. Barjesteh van Waalwijk van Doorn-Khosrovani S, Erpelinck C, van Putten WL, Valk PJ, van der Poel-van de Luytgaarde S, Hack R et al. High EVI1 expression predicts poor survival in acute myeloid leukemia: a study of 319 de novo AML patients. Blood. 2003; 101(3):837-45.
6. Lugthart S, van Drunen E, van Norden Y, van Hoven A, Erpelinck CA, Valk PJ et al. High EVI1 levels predict adverse outcome in acute myeloid leukemia: prevalence of EVI1 overexpression and chromosome 3q26 abnormalities underestimated. Blood. 2008;111(8):4329-37.
7. Gröschel S, Lugthart S, Schlenk RF, Valk PJ, Eiwen K, Goudswaard C et al. High EVI1 expression predicts outcome in younger adult patients with acute myeloid leukemia and is associated with distinct cytogenetic abnormalities. J Clin Oncol. 2010;28(12):2101-7.
8. Vázquez I, Maicas M, Cervera J, Agirre X, Marin-Béjar O, Marcotegui N, et al. Down-regulation of EVI1 is associated with epigenetic alterations and good prognosis in patients with acute myeloid leukemia. Haematologica. 2011;96(10):1448-56.
9. Gröschel S, Schlenk RF, Engelmann J, Rockova V, Teleanu V, Kühn MW et al. Deregulated expression of EVI1 defines a poor prognostic subset of MLL-rearranged acute myeloid leukemias: a study of the German-Austrian Acute Myeloid Leukemia Study Group and the Dutch-Belgian-Swiss HOVON/SAKK Cooperative Group. J Clin Oncol. 2013;31(1):95-103.
10. Balgobind BV, Lugthart S, Hollink IH, Arentsen-Peters ST, van Wering ER, de Graaf SS et al. EVI1 overexpression in distinct subtypes of pediatric acute myeloid leukemia. Leukemia. 2010; 24(5):942-9.
11. Balgobind BV, Raimondi SC, Harbott J, Zimmermann M, Alonzo TA, Auvrignon A et al. Novel prognostic subgroups in childhood 11q23/MLL-rearranged acute myeloid leukemia: results of an international retrospective study. Blood. 2009;114(12):2489-96.
12. Tomizawa D, Tawa A, Watanabe T, Saito AM, Kudo K, Taga T et al. Appropriate dose reduction in induction therapy is essential for the treatment of infants with acute myeloid leukemia: a report from the Japanese Pediatric Leukemia/Lymphoma Study Group. Int J Hematol. 2013;98(5):578-88.
13. Bindels EM, Havermans M, Lugthart S, Erpelinck C, Wocjtowicz E, Krivtsov AV et al. EVI1 is critical for the pathogenesis of a subset of MLL-AF9-rearranged AMLs. Blood. 2012;119(24):5838-49.
14. Saito Y, Nakahata S, Yamakawa N, Kaneda K, Ichihara E, Suekane A et al. CD52 as a molecular target for immunotherapy to treat acute myeloid leukemia with high EVI1 expression. Leukemia. 2011; 25(6):921-31.

Yang et al. BMC Cancer 2014, 14:654
http://www.biomedcentral.com/1471-2407/14/654

BMC Cancer

RESEARCH ARTICLE Open Access

Prognostic model for predicting overall survival in children and adolescents with rhabdomyosarcoma

Limin Yang[1,2*], Tetsuya Takimoto[1] and Junichiro Fujimoto[1]

Abstract

Background: The purpose of this study was to develop a prognostic model for the survival of pediatric patients with rhabdomyosarcoma (RMS) using parameters that are measured during routine clinical management.

Methods: Demographic and clinical variables were evaluated in 1679 pediatric patients with RMS registered in the Surveillance, Epidemiology, and End Results (SEER) program from 1990 to 2010. A multivariate Cox proportional hazards model was developed to predict median, 5-year and 10-year overall survival (OS). The Akaike information criterion technique was used for model selection. A nomogram was constructed using the reduced model after model selection, and was internally validated.

Results: Of the total 1679 patients, 543 died. The 5-year OS rate was 64.5% (95% confidence interval (CI), 62.1-67.1%) and the 10-year OS was 61.8% (95%CI, 59.2-64.5%) for the entire cohort. Multivariate analysis identified age at diagnosis, tumor size, histological type, tumor stage, surgery and radiotherapy as significantly associated with survival ($p < 0.05$). The bootstrap-corrected c-index for the model was 0.74. The calibration curve suggested that the model was well calibrated for all predictions.

Conclusions: This study provided an objective analysis of all currently available data for pediatric RMS from the SEER cancer registry. A nomogram based on parameters that are measured on a routine basis was developed. The nomogram can be used to predict 5- and 10-year OS with reasonable accuracy. This information will be useful for estimating prognosis and in guiding treatment selection.

Keywords: Rhabdomyosarcoma, Cancer, Nomogram, Overall survival

Background

Rhabdomyosarcoma (RMS) is the most common soft-tissue sarcoma in children and adolescents and accounts for 3% of all pediatric tumors [1]. Approximately 350 children are diagnosed with RMS in the United States every year [2]. Incidence peaks at a very young age. Because RMS is derived from immature striated skeletal muscle, this disease can occur at any site in the body. Prognosis of RMS has improved significantly, with multidisciplinary management accounting for most of the increase in survival rate. Since 1972, the Intergroup Rhabdomyosarcoma Study Group (IRSG) has conducted a series of clinical trials and published a series of treatment guidelines for different primary sites. As a result, the long-term survival rate of these patients has nearly tripled from approximately 25% in 1970 to more than 70% in the 1990s [3,4].

The rarity of this disease means that most information regarding survival is derived from these clinical trials. However, overall survival (OS) results differ between clinical trials and population-based cancer registries because of important differences between patients treated in routine practice and those treated in clinical trials. For example, IRSG reports showed a 5-year OS of 70% in the 1990s [3,4], while, even in the 2000s, the 5-year OS was only approximately 50% in children with RMS according to population-based data [5]. Clinical trials may select participants based on strict inclusion criteria, which consider the extent of disease, previous history of treatment, comorbidities, psychosocial conditions and

* Correspondence: yo-r@ncchd.go.jp
[1]Epidemiology and Clinical Research Center for Children's Cancer, National Center for Child Health and Development, 2-10-1 Okura, Setagaya-ku, Tokyo 157-8535, Japan
[2]Division of Allergy, Department of Medical Subspecialties, Medical Support Center for Japan Environment and Children's Study (JECS), National Center for Child Health and Development, 2-10-1 Okura, Setagaya-ku, Tokyo 157-8535, Japan

other factors [6]; patients in a poor condition may thus be excluded from the protocol. OS in trials may therefore not reflect the prognosis of patients who receive treatment in a community setting.

Individualized estimation of the prognosis could be useful for counseling cancer patients on treatment selection and for optimizing therapeutic approaches [7]. However, to the best of our knowledge, there is currently no such estimation tool for RMS based on patients from the general population. In this study, we analyzed the OS in children and adolescents with RMS using population-based data collected by the Surveillance, Epidemiology, and End Results (SEER) program of the National Cancer Institute (NCI) [5], and constructed a nomogram based on variables collected from the routine cancer registry, with the aim of providing clinicians and patients with a practical clinical tool to predict survival.

Methods

Study population

The data were derived from the SEER program, which collects demographic, diagnostic and treatment information on all newly diagnosed cancer patients residing within specific US geographic regions. Registry data are submitted without personal identifiers to the NCI, and these data are publicly available for research purpose. Because all information in public-use SEER database remains de-identified, approval by an ethics committee was not necessary to perform the analysis [8]. All authors have signed the data-use agreement and got permission from SEER program to use this data.

Using the SEER registry public database, we identified patients with RMS diagnosed from 1990 to 2010 [5]. Children diagnosed with malignant, first primary RMS and aged 0–19 years were eligible for this analysis. In this study, eligible RMS cases had International Classification of Childhood Cancer (ICCC) code IXa, corresponding to ICO-O-3 morphology codes: 1) RMS not otherwise specified 8900/3; 2) pleomorphic RMS adult-type 8901/3; 3) mixed-type RMS 8902/3; 4) embryonal RMS 8910/3; 5) spindle cell RMS 8912/3; 6) alveolar RMS 8920/3; or 7) embryonal sarcoma 8991/3. Patients were excluded from the analysis if the diagnosis was made at autopsy or by death certificate only. Patients with no confirmation of diagnosis by microscopy were also excluded. After selection, there were 1679 cases left in the cohort. The flow chart for data selection is shown in Figure 1.

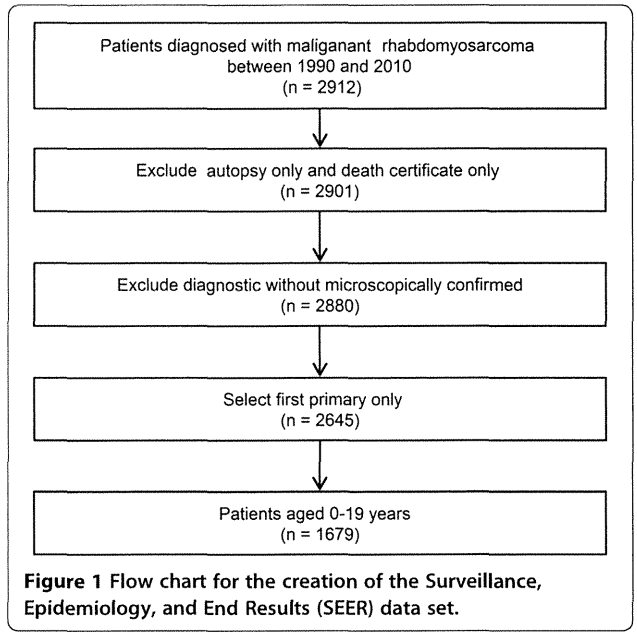

Figure 1 Flow chart for the creation of the Surveillance, Epidemiology, and End Results (SEER) data set.

Data analysis

In the description of variables and calculation of OS, age at diagnosis was classified as 0–4, 5–9, 10–14 or 15–19 years. Age at diagnosis was treated as a continuous variable in multivariate analysis. Other clinical factors included primary tumor site, histologic tumor subtype, tumor stage, tumor size, surgery and radiotherapy (RT). Primary tumor sites were classified as favorable or unfavorable based on the criteria for staging of pediatric tumors [9]. The head and neck (nonparameningeal), genitourinary (non-bladder/prostate), and bile duct regions were defined as favorable sites, all other sites were defined as unfavorable, and an unknown site was regarded as a missing value. Histology was classified as embryonal, alveolar or other histological subtype. Histological subtypes with RMS not otherwise specified were treated as missing values. Tumor stage was classified according to the SEER historic staging system. Cases with insufficient information to define the stage were regarded as having a missing value. Tumor size was truncated at 20 cm and was grouped into three levels for both character description and calculation of OS: 1) 0–4 cm; 2) 5–9 cm; and 3) ≥10 cm. Size was treated as a continuous variable in the multivariate model.

Statistical methods

All missing values were imputed with the 'transcan' function of the rms package [10]. OS was calculated by the Kaplan-Meier product-limited method. Survival curves were compared using the log-rank test. Cox proportional hazard regressions were performed to assess the effects of covariates on OS. For continuous variables, we fitted restricted cubic splines with three knots at 10%, 50% and 90% empirical quantiles. We also considered the interaction effect between surgery and RT. The proportional hazard assumption was justified by examining residual plots. The Akaike information criterion was

utilized for model selection. We constructed a nomogram with the beta coefficients of variables in the reduced model.

The model was internally validated. We generated 200 bootstrap samples to determine the calibration and discrimination of the model. Calibration refers to the ability of a model to make unbiased estimates of outcome. Calibration was assessed using a calibration curve generated by plotting the model-predicted 5-year and 10-year survival probabilities against the observed probability, as calculated by the Kaplan-Meier method. The prognostic accuracy of the model was quantified by computing the concordance index (c-index) described by Harrell et al. [11]. The c-index is a discrimination measure that estimates the probability that, of two randomly chosen patients, the patient with the higher predicted survival will outlive the patient with the lower predicted survival. The c-index ranges from 0.5 (no discrimination) to 1.0 (perfect discrimination).

All statistical analyses were conducted using R version 3.1.0 software (Institute for Statistics and Mathematics, Vienna, Austria; www.r-project.org) [12]. The model and nomogram were constructed using the R package rms [10]. All statistical tests were two-sided, and values of $p < 0.05$ were considered significant.

Results

Patient demographics are listed in Table 1. A total of 1679 pediatric patients with RMS were included in the study. Approximately 38.1% of the subjects were aged 0–4 years, 23.2% were 5–9 years, 20.6% were 10–14 years and 18.1% were 15–19 years. There were 974 (58.0%) boys, and 705 (42.0%) girls. The majority of patients were white (75.9%). Approximately 61.1% of RMS occurred at unfavorable sites. Around 59.0% of patients were diagnosed with embryonal RMS, 33.2% with alveolar RMS and 7.7% with other RMS. Based on SEER staging, 33.4% of patients had localized tumors, 34.9% had regional RMS and 31.7% had metastasis. More than half (62.8%) of the patients had received RT, and 59.1% received surgery.

The 5-year OS rate for the entire cohort was 64.5% (95% confidence interval (CI), 62.1-67.1%) and the 10-year OS rate was 61.8% (95%CI, 59.2-64.5%). Five- and 10-year OS rates by characteristic are listed in Table 1. Sex and race had no influence on OS. Prognosis worsened with increasing age; young children (0–4 years) had a better prognosis than adolescents (15–19 years), with 5-year OS of 71.3% and 47.9%, respectively. Children with embryonal RMS had a longer survival than those with alveolar RMS, with estimated 5-year OS of 73.5% and 46.3%, respectively. Patients with localized tumors had a better prognosis (5-year OS of 84.0%) than those with regional disease (72.4%) or distant metastasis (35.7%). RMS at favorable sites had a better prognosis than that at unfavorable sites ($p < 0.001$). Patients with surgery had improved survival compared with those without surgery ($p < 0.001$). RT showed a weak but significant association with prognosis; 5-year OS was 65.6% in patients with RT compared with 62.7% in those without RT ($p = 0.045$).

Multivariate analysis was performed using a Cox proportional hazards regression model. We pre-specified nonlinearity for age at diagnosis and tumor size variables, and considered the effect on prognosis of the interaction between surgery and RT. Residual plots indicated that the proportional hazards assumption held. After model selection, we obtained a reduced model. Beta coefficients and hazard ratios of variables are listed in Table 2.

The nomogram included age at diagnosis, size, tumor site, stage, histological type, surgery and RT (Figure 2). To use the nomogram, we drew a vertical line to the point row to assign point values for each variable, summed the point values for each variable to obtain total points, and then dropped a vertical line from the total points row to get the 5- and 10-year OS rates.

The model was internally validated. Discrimination suggested good accuracy with a bootstrap-corrected c-index of 0.74, which denotes 74% probability that, of two randomly selected patients, the patient who survives longer will have a higher survival probability than the patient with shorter survival. The calibration plots for 5- and 10-year OS are shown in Figure 3. Points in the calibration plot were close to the 45° line, which suggested that the model was well-calibrated for all predictions.

Discussion

The current study evaluated OS among pediatric patients with newly-diagnosed RMS in a population-based dataset, and constructed a nomogram to predict 5- and 10-year OS. This prognostic tool will be useful for estimating prognosis and guiding treatment selection.

The rarity of this disease means that most published studies are retrospective analyses of clinical studies, or small, single-institution, observational studies. Results from a single institution often fail to identify a true relationship between outcome and risk factors because of the small sample size and short follow-up period. Our analyses were based on the SEER database, which is considered to be the largest cancer registry. Reports from a population-based cohort have the advantage of including many more patients, thus increasing the power to estimate the true effects of risk factors on survival. Moreover, unlike most results from clinical studies, analysis of a population-based database includes not only those treated using formal protocols, but also those excluded from protocols because of comorbidity, tumor stage, or

Table 1 Patient demographics and overall survival

	All patients		5 Years OS (%)		10 Years OS (%)		p
Characteristics	No.	Events	Rate	95%CI	Rate	95%CI	
Entire cohort	1679	543	64.5	62.1-67.1	61.8	59.2-64.5	
Age (years)							<0.001
0-4	639	173	71.3	67.5-75.2	69.1	65.2-73.2	
5-9	390	97	73.2	68.5-78.2	68.8	63.6-74.5	
10-14	346	134	56.4	50.9-62.6	52.4	46.6-59.0	
15-19	304	139	47.9	42.1-54.6	47.3	41.5-54.0	
Tumor size (cm)							<0.001
0-4	618	116	79.5	76.0-83.1	77.1	73.4-81.1	
5-9	675	237	61.6	57.7-65.7	57.6	53.5-62.1	
≥10	386	190	45.9	40.7-51.7	44.4	39.1-50.3	
Sex							0.311
Male	974	306	65.7	62.5-69.1	62.5	59.1-66.1	
Female	705	237	63.0	59.2-67.0	60.8	56.8-65.0	
Race							0.359
White	1274	407	65.2	62.4-68.1	62.2	59.2-65.3	
Black	277	88	64.0	58.0-70.7	62.2	55.9-69.2	
Others	128	48	58.7	50.0-69.0	56.6	47.5-67.5	
Site							<0.001
Unfavorable	1026	406	56.6	53.2-59.9	53.5	50.1-57.1	
Favorable	653	137	77.1	73.6-80.7	74.6	70.8-78.5	
Stage							<0.001
Localized	561	83	84.0	80.7-87.5	81.1	77.3-85.1	
Regional	586	152	72.4	68.6-76.5	68.5	64.3-73.1	
Distant	532	308	35.7	31.5-40.5	34.4	30.1-39.2	
Histology							<0.001
Embryonal	991	249	73.5	70.6-76.5	70.8	67.6-74.1	
Alveolar	558	263	46.3	41.9-51.2	43.2	38.7-48.3	
Others	130	31	73.1	64.9-82.3	71.4	62.9-81.0	
Surgery							<0.001
None	686	294	52.0	48.0-56.2	50.4	46.4-54.7	
Surgery	993	249	73.2	70.2-76.3	69.9	66.3-73.0	
Radiotherapy							0.045
None	625	213	62.7	58.7-67.0	60.8	56.6-65.3	
Radiation	1054	330	65.6	62.5-68.9	62.3	59.1-65.8	

OS, overall survival; CI, confidence interval.

other factors. The reported results thus represent the full spectrum of the disease. Furthermore, SEER data are high quality and are collected in a uniform manner with uniform data standards. Quality control ensures that the SEER program has a relatively low rate of errors in the cancer registry [6].

Our study cohort included 1679 RMS patients and 543 deaths, which sample size was adequate to establish a reasonable model. The outcome measure of OS is one of the most useful pieces of information for counseling and is commonly used to develop staging schemes. Although the nomogram is not perfectly accurate, the error bars in the calibration plot suggest that predictions from the nomogram are within approximately 5% of the actual probability on average, and the bootstrap-corrected c-index of 0.74 suggests that the nomogram

Table 2 Cox proportional hazards multivariate regression model parameters

Covariate	Beta coefficient	Hazard ratio	95% CI	p
Age	−0.037*	-**	-	0.154
Age′	0.089*	-**	-	0.013
Size	0.006†	-††	-	0.095
Size′	−0.004†	-††	-	0.359
Favorable site	−0.204	0.82	0.65-1.02	0.076
Stage				
Regional	0.404	1.50	1.13-1.98	0.004
Distant	1.259	3.52	2.64-4.70	<0.001
Histology				
Alveolar	0.497	1.64	1.35-2.00	<0.001
Other	−0.135	0.87	0.59-1.29	0.499
Received surgery	−0.612	0.54	0.40-0.72	<0.001
Received RT	−0.632	0.53	0.42-0.68	<0.001
Interaction terms				
Surgery × RT	0.564	1.75	1.24-2.50	0.002

CI, confidence interval; RT, radiotherapy.
*Age was modeled using a restricted cubic spline function with three knots, which yields two independent beta coefficients, annotated as Age and Age′.
**The hazard ratio varies continuously with age.
†Size was modeled using a restricted cubic spline function with three knots, which yields two independent beta coefficients, annotated as Size and Size′.
††The hazard ratio varies continuously with size.

has good ability to discriminate among patients. This accuracy is comparable with most published nomograms for cancer prognosis.

The prognostic nomogram is a model-based tool to predict patient outcome. It directly quantifies the prognosis of individual patients based on proven prognostic factors. Different from a staging or scoring system, a nomogram considers multiple commonly available prognostic variables simultaneously, including continuous variables. Individual predictions are expressed on a probability scale, making it more easily understood by patients and clinicians than relative rates or hazard ratios [13]. A nomogram has the potential to stratify patients for clinical studies, meaning that treatment regimens can be tested in more homogenous populations. Selecting high-risk patients based on predictions from a nomogram can also help to improve trial efficiency; for example, trials evaluating a treatment strategy could target patients with poor prognoses. Identifying high-risk patients for trial recruitment using a nomogram will increase the power to detect differences among treatment effects, thus reducing the required sample size. This method has been used in prospective randomized trials [14].

There is increasing interest in personalized medicine. A number of cancer nomograms to predict prognosis have been published in recent decades, such as for prostate, breast, soft-tissue sarcoma, and other cancers, including our previous nomogram for thyroid cancer [7,15-23]. To the best of our knowledge however, the nomogram constructed in this study represents the first OS nomogram for pediatric RMS that is generalizable to the population.

The results of the Cox model identified age at diagnosis, stage of tumor, tumor size, histological subtype and treatment as important predictors of RMS survival in pediatric patients. The findings are expressed consistently in the nomogram. For example, adolescence, distant disease, large tumor size, alveolar RMS and no treatment, which were associated with a reduced survival based on the model, were given larger points in the nomogram. Meanwhile, a larger total point indicates a lower OS.

Simplicity is a strength of our model. Unlike models that aim to identify associations between prognosis and risk factors, predictive models should focus more on accuracy and parsimony [24]. Complex models including a number of variables may be abandoned in clinical practice. In contrast, the nomogram developed in this study relies on limited variables that are routinely available from the tumor registry, making it easy for clinicians to use to calculate survival for individual patients.

Adult RMS was not included in this model. Pediatric and adult RMS have different clinical characteristics and prognoses. For example, pleomorphic RMS is common among adult patients, but is seldom seen in pediatric patients. Additionally, adult patients have a poorer response to chemotherapy. Research has suggested that increased levels of a resistance-related protein in adult embryonal and pleomorphic RMS compared with pediatric RMS may explain the reduced response to chemotherapy [25]. Information regarding chemotherapy and variables in the protein level was not available in the current study and it was therefore not possible to adjust for these potential effects on prognosis in the SEER cohort. Moreover, adult RMS may have lower pathologic accuracy compared with pediatric RMS [9]. We therefore excluded adult RMS from the current analysis to avoid these confounders and bias and to increase the accuracy of the model.

Although our nomogram showed reasonable accuracy for predicting OS, care should be taken when using a nomogram for counseling. Because it is impossible to include all risk factors, the prognostic predictive value of a nomogram should not be used as the sole basis for selecting a treatment regimen; treatment should be selected based on not only the expected value from the nomogram, but also taking into account other prognostic factors and quality of life.

There were some limitations to our study. First, the SEER public dataset does not include information on

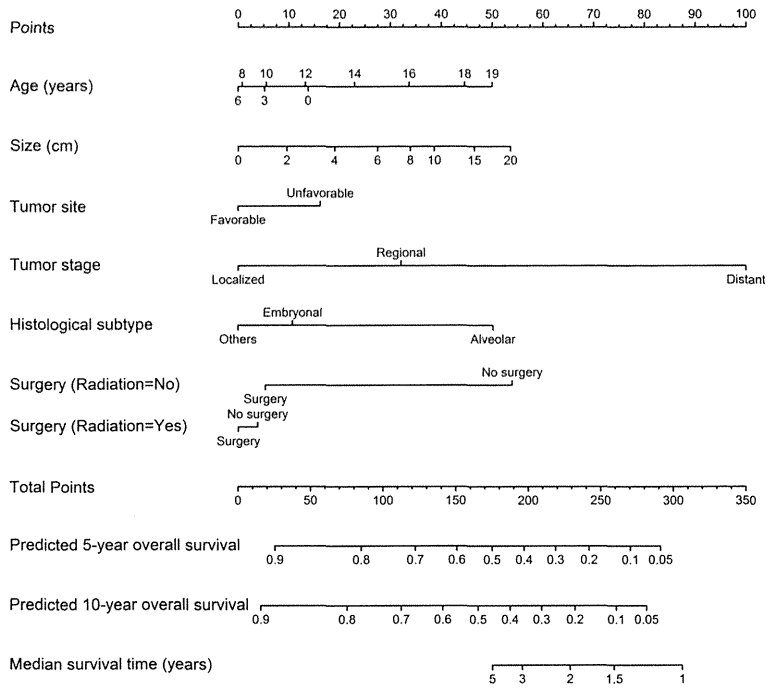

Figure 2 Nomogram for predicting 5- and 10-year overall survival and median survival time. Instructions: Locate the patient's characteristic on the variable row, draw a vertical line straight upward to the points row to obtain a points value for the variable. Move to the next row of variables, and repeat this process to get points for each variable. Sum the total points and drop a vertical line from the total points row to assign the values for overall survival rates.

chemotherapy, comorbidity and surgical margins, which are viewed as important prognostic variables. This information would be useful for refining the predictive model. Second, although we restricted our cohort to patients diagnosed after 1990, the study period still spanned approximately two decades, during which time there have been improvements in surgery, chemotherapy and RT. Our nomogram thus tends to underestimate current OS. Third, unlike IRSG clinical trials, the SEER program does not utilize a central pathology review to minimize misclassification [6]. Finally, we used internal validation to evaluate the accuracy of the model, and external validation based on independent data would be useful to validate the model further.

Conclusions

In conclusion, we used a population-based dataset to establish and internally validate a model to estimate the

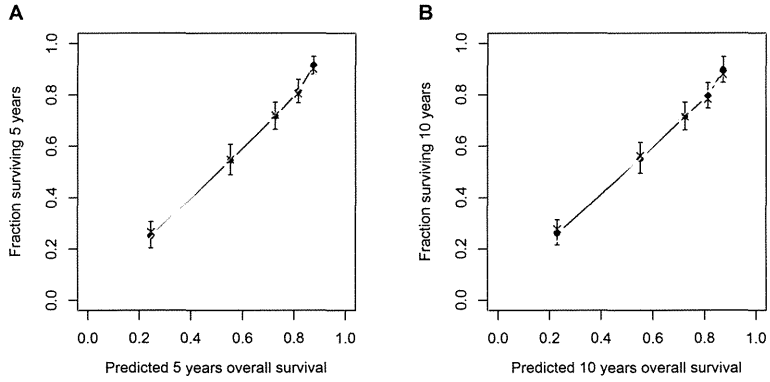

Figure 3 Calibration plot. (A) Five-year overall survival; **(B)** 10-year overall survival. The grey line is the "ideal" line if there is a perfect match between predicted and observed survivals. Vertical arrows represent 95% confidence intervals of observed survival. Dots correspond to apparent predictive accuracy. X marks the bootstrap-corrected estimates.

probability that a pediatric patient will be alive 5 and 10 years after being diagnosed with RMS. This study represents an objective analysis of all currently available data from the SEER cancer registry. The model shows good ability to discriminate among patients, with a c-index of 0.74. This predictive tool may be useful for patient counseling and to enable more individualized treatment planning.

Abbreviations

RMS: Rhabdomyosarcoma; SEER: Surveillance, Epidemiology, and End Results; OS: Overall survival; NCI: National Cancer Institute; IRSG: Intergroup Rhabdomyosarcoma Study Group; ICCC: International Classification of Childhood Cancer; c-index: Concordance index.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

LY designed and performed analyses and drafted the paper; TT prepared the data and created the figure; JF edited the paper and commented on the interpretation of the results. All authors read and approved the final manuscript.

Acknowledgement

This research was not funded. The authors would like to thank SEER for open access to their database.

Received: 6 March 2014 Accepted: 2 September 2014
Published: 5 September 2014

References

1. Miller RW, Young JL Jr, Novakovic B: **Childhood cancer.** *Cancer* 1995, **75**(1 Suppl):395–405.
2. Ries L, Smith M, Gurney J, Linet M, Tamra T, Young J, Bunin G: *Cancer Incidence and Survival among Children and Adolescents: United States SEER Program 1975–1995.* Bethesda, MD: National Cancer Institute, SEER Program; 1999. NIH Pub. No. 99–4649.
3. Crist W, Gehan EA, Ragab AH, Dickman PS, Donaldson SS, Fryer C, Hammond D, Hays DM, Herrmann J, Heyn R, Jones PM, Lawrence W, Newton W, Ortega J, Raney RB, Ruymann FB, Tefft M, Webber B, Wiener E, Wharam M, Vietti TJ, Maurer HM: **The Third Intergroup Rhabdomyosarcoma Study.** *J Clin Oncol* 1995, **13**(3):610–630.
4. Pappo AS, Shapiro DN, Crist WM, Maurer HM: **Biology and therapy of pediatric rhabdomyosarcoma.** *J Clin Oncol* 1995, **13**(8):2123–2139.
5. *Surveillance, Epidemiology, and End Results (SEER) Program SEER*Stat Database: Incidence - SEER 9 Regs Research Data, Nov 2011 Sub (1973–2010) <Katrina/Rita Population Adjustment> – Linked To County Attributes - Total U. S., 1969–2010 Counties, National Cancer Institute, DCCPS, Surveillance Research Program, Surveillance Systems Branch,released April 2013, based on the November 2012 submission.* [www.seer.cancer.gov].
6. Punyko JA, Mertens AC, Baker KS, Ness KK, Robison LL, Gurney JG: **Long-term survival probabilities for childhood rhabdomyosarcoma. A population-based evaluation.** *Cancer* 2005, **103**(7):1475–1483.
7. Barnholtz-Sloan JS, Yu C, Sloan AE, Vengoechea J, Wang M, Dignam JJ, Vogelbaum MA, Sperduto PW, Mehta MP, Machtay M, Kattan MW: **A nomogram for individualized estimation of survival among patients with brain metastasis.** *Neurooncol* 2012, **14**(7):910–918.
8. Lee CM, Szabo A, Shrieve DC, Macdonald OK, Gaffney DK: **Frequency and effect of adjuvant radiation therapy among women with stage I endometrial adenocarcinoma.** *JAMA* 2006, **295**(4):389–397.
9. Sultan I, Qaddoumi I, Yaser S, Rodriguez-Galindo C, Ferrari A: **Comparing adult and pediatric rhabdomyosarcoma in the surveillance, epidemiology and end results program, 1973 to 2005: an analysis of 2,600 patients.** *J Clin Oncol* 2009, **27**(20):3391–3397.
10. *rms: Regression Modeling Strategies. R package version 4.2-0.* [http://CRAN.R-project.org/package=rms]
11. Harrell FE Jr, Lee KL, Mark DB: **Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors.** *Stat Med* 1996, **15**(4):361–387.
12. R Core Team: *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing; [http://www.R-project.org/.]
13. Porter CR, Suardi N, Capitanio U, Hutterer GC, Kodama K, Gibbons RP, Correa R Jr, Perrotte P, Montorsi F, Karakiewicz PI: **A nomogram predicting prostate cancer-specific mortality after radical prostatectomy.** *Urol Int* 2010, **84**(2):132–140.
14. Eastham JA, Kelly WK, Grossfeld GD, Small EJ: **Cancer, Leukemia Group B; Cancer and Leukemia Group B (CALGB) 90203: a randomized phase 3 study of radical prostatectomy alone versus estramustine and docetaxel before radical prostatectomy for patients with high-risk localized disease.** *Urology* 2003, **62**(Suppl 1):55–62.
15. Stephenson AJ, Scardino PT, Kattan MW, Pisansky TM, Slawin KM, Klein EA, Anscher MS, Michalski JM, Sandler HM, Lin DW, Forman JD, Zelefsky MJ, Kestin LL, Roehrborn CG, Catton CN, DeWeese TL, Liauw SL, Valicenti RK, Kuban DA, Pollack A: **Predicting the outcome of salvage radiation therapy for recurrent prostate cancer after radical prostatectomy.** *J Clin Oncol* 2007, **25**(15):2035–2041.
16. Rouzier R, Pusztai L, Delaloge S, Gonzalez-Angulo AM, Andre F, Hess KR, Buzdar AU, Garbay JR, Spielmann M, Mathieu MC, Symmans WF, Wagner P, Atallah D, Valero V, Berry DA, Hortobagyi GN: **Nomograms to predict pathologic complete response and metastasis-free survival after pre-operative chemotherapy for breast cancer.** *J Clin Oncol* 2005, **23**(33):8331–8339.
17. Brennan MF, Kattan MW, Klimstra D, Conlon K: **Prognostic nomogram for patients undergoing resection for adenocarcinoma of the pancreas.** *Ann Surg* 2004, **240**(2):293–298.
18. Kattan MW, Karpeh MS, Mazumdar M, Brennan MF: **Postoperative nomogram for disease-specific survival after an R0 resection for gastric carcinoma.** *J Clin Oncol* 2003, **21**(19):3647–3650.
19. Yang L, Shen W, Sakamoto N: **Population-based study evaluating and predicting the probability of death resulting from thyroid cancer and other causes among patients with thyroid cancer.** *J Clin Oncol* 2013, **31**(4):468–474.
20. Kutikov A, Egleston BL, Wong YN, Uzzo RG: **Evaluating overall survival and competing risks of death in patients with localized renal cell carcinoma using a comprehensive nomogram.** *J Clin Oncol* 2010, **28**(2):311–317.
21. Kattan MW, Leung DH, Brennan MF: **Postoperative nomogram for 12-year sarcoma-specific death.** *J Clin Oncol* 2002, **20**(3):791–796.
22. Gronchi A, Miceli R, Shurell E, Eilber FC, Eilber FR, Anaya DA, Kattan MW, Honore C, Lev DC, Colombo C, Bonvalot S, Mariani L, Pollock RE: **Outcome prediction in primary resected retroperitoneal soft tissue sarcoma: histology-specific overall survival and disease-free survival nomograms built on major sarcoma center data sets.** *J Clin Oncol* 2013, **31**(13):1649–1655.
23. Chisholm JC, Marandet J, Rey A, Scopinaro M, de Toledo JS, Merks JH, O'Meara A, Stevens MC, Oberlin O: **Prognostic factors after relapse in nonmetastatic rhabdomyosarcoma: a nomogram to better define patients who can be salvaged with further therapy.** *J Clin Oncol* 2011, **29**(10):1319–1325.
24. Zini L, Cloutier V, Isbarn H, Perrotte P, Capitanio U, Jeldres C, Shariat SF, Saad F, Arjane P, Duclos A, Lattout JB, Montorsi F, Karakiewicz PI: **A simple and accurate model for prediction of cancer-specific mortality in patients treated with surgery for primary penile squamous cell carcinoma.** *Clin Cancer Res* 2009, **15**(3):1013–1018.
25. Komdeur R, Klunder J, van der Graaf WT, van den Berg E, de Bont ES, Hoekstra HJ, Molenaar WM: **Multidrug resistance proteins in rhabdomyosarcomas: comparison between children and adults.** *Cancer* 2003, **97**(8):1999–2005.

doi:10.1186/1471-2407-14-654
Cite this article as: Yang *et al.*: Prognostic model for predicting overall survival in children and adolescents with rhabdomyosarcoma. *BMC Cancer* 2014 **14**:654.

Nishimura et al. Clinical and Translational Medicine 2014, 3:38
http://www.clintransmed.com/content/3/1/38

Clinical and Translational Medicine
a SpringerOpen Journal

PERSPECTIVE Open Access

Clinical initiatives linking Japanese and Swedish healthcare resources on cancer studies utilizing Biobank Repositories

Toshihide Nishimura[1], Takeshi Kawamura[2], Yutaka Sugihara[3], Yasuhiko Bando[4], Shigeru Sakamoto[5], Masaharu Nomura[1], Norihiko Ikeda[1], Tatsuo Ohira[1], Junichiro Fujimoto[6], Hiromasa Tojo[7], Takao Hamakubo[2], Tatsuhiko Kodama[2], Roland Andersson[8], Thomas E Fehniger[9], Harubumi Kato[1] and György Marko-Varga[1,3,9*]

Abstract

The Tokyo Medical University Hospital in Japan and the Lund University hospital in Sweden have recently initiated a research program with the objective to impact on patient treatment by clinical disease stage characterization (phenotyping), utilizing proteomics sequencing platforms. By sharing clinical experiences, patient treatment principles, and biobank strategies, our respective clinical teams in Japan and Sweden will aid in the development of predictive and drug related protein biomarkers.

Data from joint lung cancer studies are presented where protein expression from Neuro- Endocrine lung cancer (LCNEC) phenotype patients can be separated from Small cell- (SCLC) and Large Cell lung cancer (LCC) patients by deep sequencing and spectral counting analysis. LCNEC, a subtype of large cell carcinoma (LCC), is characterized by neuroendocrine differentiation that small cell lung carcinoma (SCLC) shares. Pre-therapeutic histological distinction between LCNEC and SCLC has so far been problematic, leading to adverse clinical outcome. An establishment of protein targets characteristic of LCNEC is quite helpful for decision of optimal therapeutic strategy by diagnosing individual patients. Proteoform annotation and clinical biobanking is part of the HUPO initiative (http://www.hupo.org) within chromosome 10 and chromosome 19 consortia.

Keywords: Cancer diseases; Protein quantification; Proteomics; Mass spectrometry; MRM; Biobanking; HUPO

TMU/LUH - A Joint clinical center effort

The Tokyo Medical University Hospital, a pioneer in lung cancer treatment and surgery forms a center effort jointly with Lund University Hospital to build an expert capability resource. This joint establishment will intensify the utility of cancer expertise and experiences in both Japan and Sweden to benefit cancer patients. Clinical samples will be sampled from the hospitals with dedicated quality protocols and standard operating procedures (SOPs) by automated processing. These samples will be archived in Biobank storages, and will be built as a resource for R&D studies [1].

Today we have a lack of protein biomarker-, and imaging diagnostics within most cancer disorders. New clinical tools are expected to be used as early indicators of disease, or, as personalized indicator assays for targeted and stratified disease phenotype drug treatments in the near future. There is also a poor understanding of the mode of drug action mechanisms, by commonly used therapies, which is also true for new drugs introduced to the market. The actual targeted cells-, and proteins within disease, and the actual drug interactions are by no means understood for most medicines used in today's therapies. These drug characteristics are needed for both efficacy-, and safety improvements, and also requested by regulatory authorities like; FDA/EMA/MHLW.

New technology developments that can be used to target the specificity and efficacy of new drug treatments within the health care system is progressing with

* Correspondence: gyorgy.marko-varga@bme.lth.se
[1]First Department of Surgery, Tokyo Medical University, 6-7-1 Nishishinjuku Shinjuku-ku, Tokyo 160-0023, Japan
[3]Clinical Protein Science & Imaging, Biomedical Center, Dept. of Biomedical Engineering, Lund University, BMC D13, 221 84 Lund, Sweden
Full list of author information is available at the end of the article

increasing successes, in both Japan, North America and Europe. The new generation of drugs with mono-specific affinity will require more in depth knowledge in order to provide a tailor made cure to all patients taking the drug.

The aim and purpose of the Swedish-Japanese research program is to develop new insights into disease presentation and the disease progression. These objectives are intimately linked to the classification of patients that presents similar pathologies in disease. The stratification of these patients is an added value to both the healthcare sector as well as the drug industry, as it will point to the ability to optimize the treatment of cancer patients. A cornerstone of the Japanese-Swedish hospital initiatives will be to develop new Multiple Reaction Monitoring (MRM) Multiplex analysis methods for cancer, that can be used in combination with CT-imaging and pathophysiology diagnosis. At present, there is a lack of disease specific radiographic markers or protein/peptide biomarkers. These new biomarkers should be introduced into routine clinical practice to support clinical decision making for the management of common diseases that are consequent to life styles, but also to genetic heritage. The developed methods and technologies outlined here will bring new opportunities for describing the indices of pathogenesis that are associated with the processes of early disease development. Our research teams have previously been successful in presenting lung cancer biomarker candidates [2-4]. We propose to continue developing new Protein Biomarker Diagnosis Panels that can measure novel patterns of structural and functional marker expression that precede and predict certain disease development. The biology change within disease will be followed by alterations due to drug intervention. Drug compound localization in pulmonary tissue will be characterized in lung cancer and COPD tissue by imaging mass spectrometry. The aim is to further probe the drug and metabolite(s) content and distribution in tissues. Following administration and drug exposure, drug parent ion (m/z) and fragmented daughter ions will be analyzed by spatially localization, after scanning the histology section in the MALDI instrument at single cell level [5,6]. The joint teams will also focus on excellence in biological and medical mass spectrometry, that provides health care institutions with expertise in clinical proteomics, drug imaging and mass spectrometry aiming at developing new diagnostic tools that can be integrated in routine clinical medicine. We have a major challenge in modern healthcare in both Japan and Sweden, to provide the right medicine to the right patient at the right time. Targeted treatment defined as "Personalized medicine" is becoming the next generation of drugs where drug efficacy and patient safety, is expected to predict the stratified treatment. In these developments, it is expected that patients will benefit by effective curing, and society will benefit by financial resources with increased efficiencies at a lower cost.

Global healthcare and drug developments

There is a highly unmet need within the healthcare sector that calls for an increasing world-wide demand for new medicines and curative therapies to aid in the treatment of cancer patients.

We need to utilize and apply cutting edge research with state-of-art facilities more efficiently and drive opportunities that can be used in order to impact on providing quality care that extend prognosis.

In this respect, the pharmaceutical industry and its drug product provider responsibility is under major restructuring. The main objective for pharmaceutical industry is to provide more specific and more efficacious drugs by targeted treatment, i.e. by Personalized Medicine (PM), and to provide the diagnostic test that will direct the right patient group to the right drug. We outlined the directives of these principles recently in a white paper [7], predicting a future link in-between PM and diagnostic guidance. Today, this is an accepted concept and is being applied in Japan as the first country; The Clinical Practice Guideline, (from The Japan Lung Cancer Society (https://www.haigan.gr.jp/modules/guideline/index.php?content_id=3)).

In practice this means and results in that the Biomarker diagnostic research is increasing. Companion diagnostics is an area of impressive growth, where the pharmaceutical industry, the biotech sector, and academia are investing considerable resources.

The national healthcare expenses are also increasing for most countries in the western world, where the US is in the lead, spending >17% national healthcare expenses/GDP (Shown in Figure 1). The growth is almost 5% over the last two decades. Japan reaches almost 10% which is very similar levels of healthcare budgets as Sweden. The Japanese healthcare expense was 38.6 trillion yen in 2011 and 40 trillion yen in 2013. This cost continues to increase by 1 trillion yen/year, looking at the recent three years. Interestingly, China and Russia are on the lower end of healthcare spending with 5% and 6%, respectively.

The expenses that each country needs to provide for their respective healthcare is tightly associated with the development of new medicines.

Currently, the global pharmaceutical industry is worth $300 Bill, and is expected to be rising to $400 Bill by 2015. In addition, the market size for pharmaceuticals now exceeds $1 trillion and is continuously growing. This industry is very profitable, and currently employing about 1.5 million people within the US and EU-countries.

In this regards, the R&D investments in within this industry have consistently managed to grow with

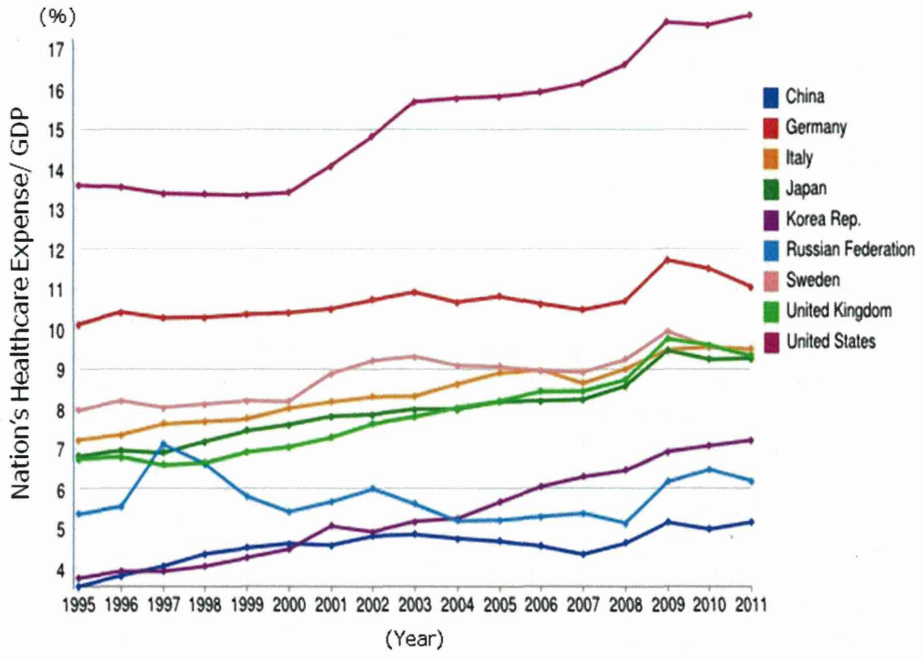

Figure 1 Statistical presentation of national healthcare spending/GDP in countries in-between 1995-2012.

investments made within the biomedical sciences, having a total spend on R&D with more than $100 Bill, per annum.

As the Research and Development phases of drug development is a long-term investment for the pharma industry, taking into account that most drugs fail long before getting to market, there is an incentive strive to find new strategies and operating principles.

Cancer research

Cancer diseases are some of the most costly for our societies. Processes essential to growth and replication of tumor cells and the maintenance of their supportive microenvironment such as the angiogenesis are today target strategies for cancer therapy.

Established chemotherapeutic agents and many novel agents in developments are directed towards various aspects of these essential cellular processes.

Over the years, we have been involved in developments associated with target proteins, possessing key regulatory functions in cancer [3,8,9]. They are as follows;

- Growth factor signaling such as EGFR
- VEGFR1, VEGFR2 and VEGFR3
- RAS-RAF pathway
- ERK-pathway
- FGFR
- PDGFR
- BRAF
- TGFB

As these growth factors stimulate cell growth by interacting with internal domains of the trans-membrane receptors, we will sequence the variety of targeted proteins in order to get a better understanding of the underlying tumor mechanisms. Previous data has been guiding in this respect [10,11].

Other oncological processes that will be directed within the joint cancer studies incorporate;

a) metastasis
b) angiogenesis that includes VEGF
c) DNA replication, including transcription and repair
d) histone acetylation and de-acetylation that can be mapped by mass spectrometry

Upon growth factor receptor binding to targets such as EGF, FGF, and TGFB, the activation of tumor biology is initiated [3,5,6,8,12]. This ligand-receptor complex formation also imposes conformational changes of the receptor which will result in an activation of the transfer of phosphate groups. The autophosphorylation loop of the kinase is the target area of mass spectrometry deep-mining of these regions within the protein sequence [13-15].

Establishing the identity of the amino acids that are phosphorylated is of high interest.

By dedicated clinical study materials from various patient groupings our objective is to map the receptor-ligand interactions. The glycolytic post-translational-modifications (PTM) of any given ligand activator will

be directly linked to the binding affinity. The PTM status of the receptor itself has also proven to be of major importance to the signaling mechanisms [16]. We recently outlined a strategic overview of key protein targets, playing a key role in pancreatic cancer [16]. The TGF-β receptor and TGF- β ligand has a direct role in the development of lung cancer as well as other cancer forms [15].

On a molecular level, the intracellular targets, with specific assigned function in cancer are enzymes which activities catalysis the signaling cascade through the cellular regions reaching the nucleus. We have previously been working with ERK and MAPKAP signaling, as well as ZAP70. ZAP70 is a key target in a number of cancer types and is responsible for the autophosphorylation, of specific threonine positions within the protein that can be determined [13,14]. The stoichiometric distribution of phosphor-groups being post-translationally modified has been pioneered by our research team. Sequencing of the target peptide of this tyrosine kinase can be accomplished by both ESI- as well as MALDI- mass spectrometry.

We recently reported on the development of biomarkers on the newly identified neuroendocrine phenotype within lung cancer [4], also denamed "The diffuse neuroendocrine system".

We have followed up with new protein sequencing experiments where the expression specificity of the LCNEC phenotype can be identified, as shown in Figure 2. A clear differential expression pattern can be seen in comparison to LCC and SMC. The difference in expression constitutes proteins with both medium-, as well as low-abundant expression (unpublished data).

Interestingly, these LCNEC Carcinoid tumors originate from cells that belong to the diffuse neuroendocrine system. These cells resembles nerve cells in certain ways, but they are also alike hormone-making endocrine cells in other ways. This phenotype of tumor cells are localized throughout the body, found in organs like the lungs, intestines, as well as the stomach [2].

Protein deep sequencing platform and biobanking

Mining the data output from protein deep sequencing studies is an intense research activity where the chromosome initiative has presented impressive novel annotation deliveries [17-21]. The HUPO database; NextProt provides evidence detail on status output of chromosome 10 and 19 (http://www.nextprot.org).

Currently the remaining unknown proteins, coded by the human genome are 27% [17].

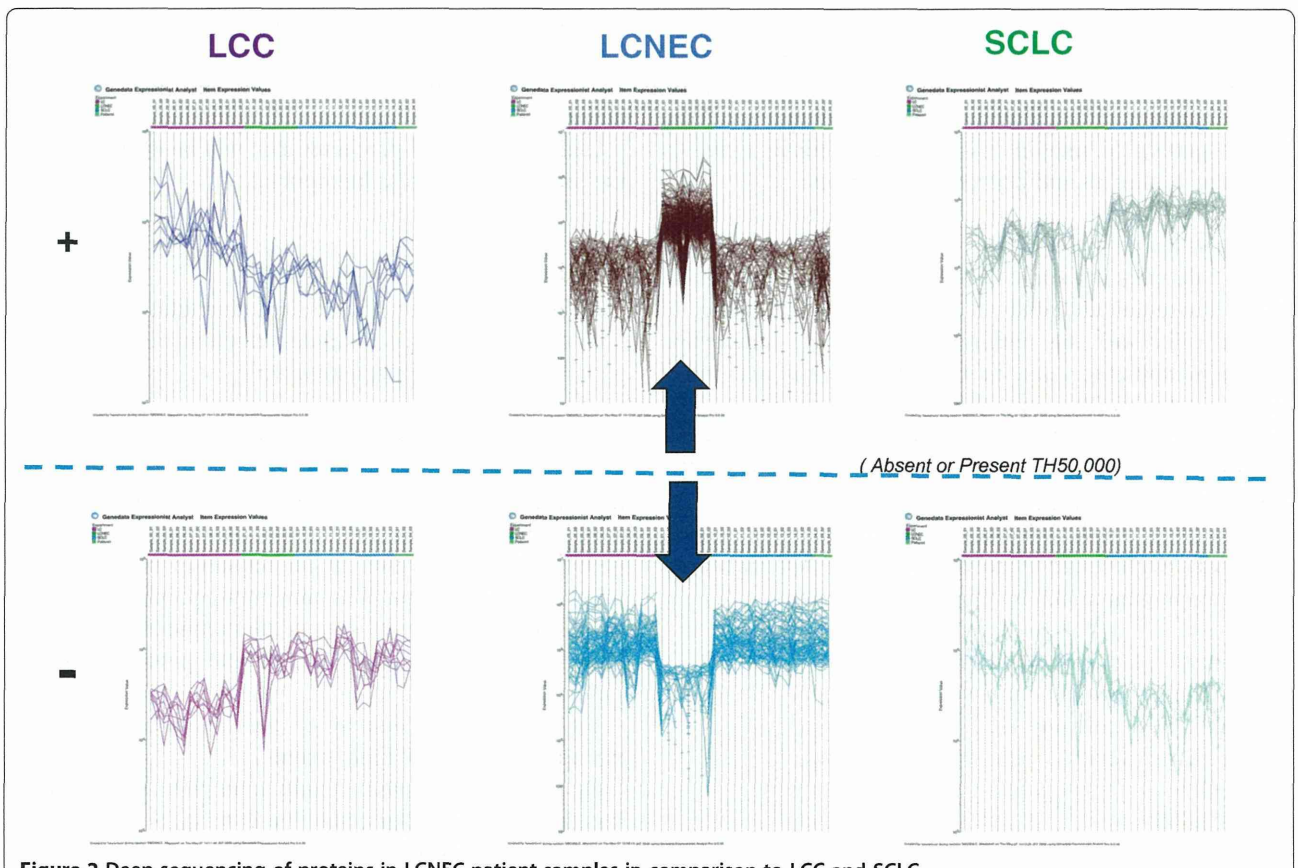

Figure 2 Deep sequencing of proteins in LCNEC patient samples in comparison to LCC and SCLC.

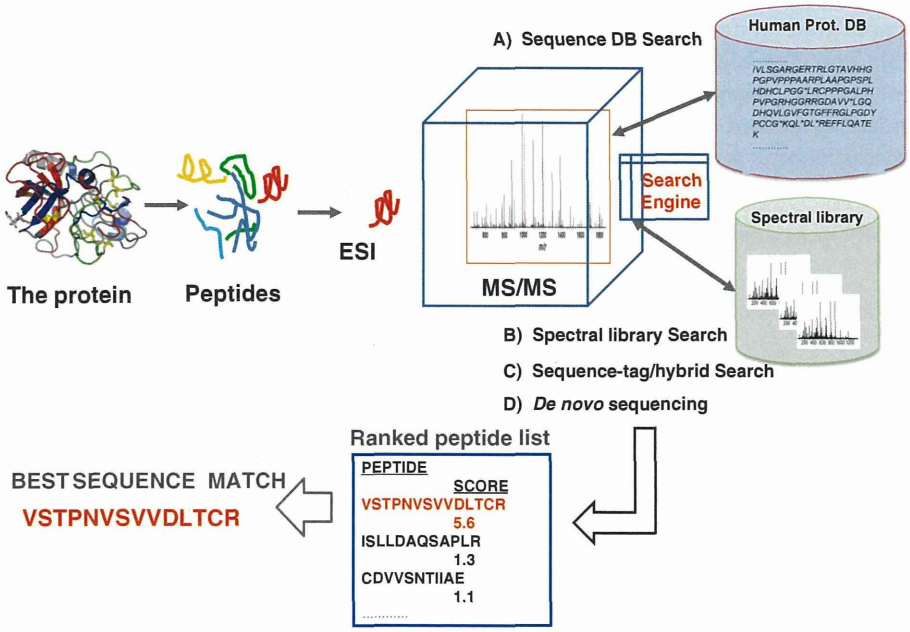

Figure 3 Workflow with LC-MS platform integrated to search engines for the annotation of proteins.

These are recent data generated from the Proteome Exchange database and include post-translational modifications as well as splice variants and protein isomers [22]. Details of the experimental outputs were published in a special issue in 2014, presenting the major developments from the C-HPP chromosome teams [18,19,21,23]. 22 global chromosome team published papers on the novel annotations, including database reports with neXtProt, PeptideAtlas, and CAPER.

A recent development was made jointly by the TMU-LUH research teams, assigning a specific protein deep sequencing platform and work-flow. We are currently applying the process in lung cancer studies where we target novel proteoforms of missing gene coded proteins. The workflow process as we currently use in laboratories in both Tokyo and Lund is depicted in Figure 3.

Large scale Biobanking is presently an intense clinical activity in hospitals around the world. Recently, Biobanking was recognised by the TIME Magazine as One of; "10 Ideas Changing the World Right Now". We will continue to establish biobank capabilitiesis, a first step towards realizing the vision of a comprehensive access to TMU-LUH biobank samples. We will combine these samples with health register data, thereby identifying correlations in-between clinical variables and disease presentations in patients. Currently we are using a resource driven sample collection, as outlined in Figure 4. We are coordinating sample collections in-between TMU and LUH, utilizing robotic -80 °C sample handling systems as previously described [1,24-27]. The ultra-low temperature storage is essential as it ensures the stability of proteins within the clinical samples. Blood samples

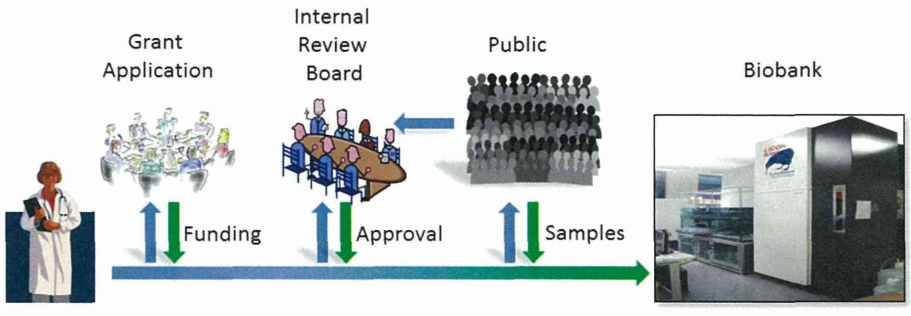

Figure 4 Exploring new governance paradigms for Biobanks.

will be processed by robotic liquid handling and aliquoted in 384-sample tube systems, developed by our research team [25]. The 384 high-density rack system will provide a cost-benefit advantage, and expected to become a new standard in modern biobanking. This allows HUPO to easily process Biobank samples globally.

Competing interests
The authors declare that they have no competing interests.

Author's contributions
All authors equally assisted in drafting the manuscript. All authors read and approved the final version.

Author's information
Professor Toshihide Nishimura and Professor György Marko-Varga are principal investigator for the Sweden-Japan research project. Assoc. Professor Takeshi Kawamura, Dr. Yutaka Sugihara, Dr. Yasuhiko Bando, and Dr. Shigeru Sakamoto are experts of mass spectrometry. Dr. Masaharu Nomura, Professor Norihiko Ikeda, Professor Tatsuo Ohira, Professor Junichiro Fujimoto, Professor Hiromasa Tojo, Professor Takao Hamakubo, Professor Tatsuhiko Kodama, Professor Junichiro Fujimoto, Professor Hiromasa Tojo are clinical scientist, Professor Roland Andersson is the vice dean and head of surgery, Professor Thomas E. Fehniger is a pathologist and director of CEBMMS. Harubumi Kato is a head clinician in Japan.

Acknowledgements
We would like to acknowledge The Kamprad Foundation, Thermo, the Swedish Foundation for Strategic Research, "TOTAL-AMI", the Swedish Research Council, Vinnova, Knut and Alice Wallenberg Foundation, Crafoord Foundation.

Author details
[1]First Department of Surgery, Tokyo Medical University, 6-7-1 Nishishinjuku Shinjuku-ku, Tokyo 160-0023, Japan. [2]Laboratory for Systems Biology and Medicine, Research Center for Advanced Science and Technology, The University of Tokyo, 4-6-1, Komaba, Meguro-Ku, Tokyo 153-8904, Japan. [3]Clinical Protein Science & Imaging, Biomedical Center, Dept. of Biomedical Engineering, Lund University, BMC D13, 221 84 Lund, Sweden. [4]Biosys Technology, Daihyaku Seimei Toritsudai Ekimae Bldg 5 F 13-18, Nakane 2, Meguro-ku, Tokyo 152-0031, Japan. [5]ThermoFisher Scientific, 3-9 Moriya-choKanagawa-ku, Yokohama 221-0022, Japan. [6]National Medical Center for Children and Mothers Research Institute, 2-10-1 Okura Setagaya-ku, Tokyo 157-8535, Japan. [7]Dept. of Biophysics and Biochemistry, Osaka University Graduate School of Medicine, 2-2 Yamadaoka, Suita 565-0871, Japan. [8]Department of Surgery, Clinical Sciences Lund, Lund University, and Skåne University Hospital, Lund, Sweden 221 84, Lund, Sweden. [9]Center of Excellence in Biological and Medical mass spectrometry (CEBMMS), 221 84 Lund, Sweden.

Received: 11 September 2014 Accepted: 28 October 2014
Published online: 22 November 2014

References
1. Marko-Varga G, Vegvari A, Welinder C, Lindberg H, Rezeli M, Edula G, Svensson KJ, Belting M, Laurell T, Fehniger TE: **Standardization and utilization of biobank resources in clinical protein science with examples of emerging applications.** *J Proteome Res* 2012, **11**:5124–5134.
2. Nishimura T, Kato H, Ikeda N, Kihara M, Nomura M, Kato Y, Marko-Varga G: **Cancer Phenotype Diagnosis and Drug Efficacy within Japanese Health Care.** *Int J Proteomics* 2012, **2012**:921901.
3. Nyberg F, Ogiwara A, Harbron CG, Kawakami T, Nagasaka K, Takami S, Wada K, Tu HK, Otsuji M, Kyono Y, Dobashi T, Komatsu Y, Kihara M, Akimoto S, Peers IS, South MC, Higenbottam T, Fukuoka M, Nakata K, Ohe Y, Kudoh S, Clausen IG, Nishimura T, Marko-Varga G, Kato H: **Proteomic biomarkers for acute interstitial lung disease in gefitinib-treated Japanese lung cancer patients.** *PLoS ONE* 2011, **6**:e22062.
4. Nomura M, Fukuda T, Fujii K, Kawamura T, Tojo H, Kihara M, Bando Y, Gazdar AF, Tsuboi M, Oshiro H, Nagao T, Ohira T, Ikeda N, Gotoh N, Kato H, Marko-Varga G, Nishimura T: **Preferential expression of potential markers for cancer stem cells in large cell neuroendocrine carcinoma of the lung. An FFPE proteomic study.** *J Clin Bioinformatics* 2011, **1**:23.
5. Fehniger TE, Vegvari A, Rezeli M, Prikk K, Ross P, Dahlback M, Edula G, Sepper R, Marko-Varga G: **Direct demonstration of tissue uptake of an inhaled drug: proof-of-principle study using matrix-assisted laser desorption ionization mass spectrometry imaging.** *Anal Chem* 2011, **83**:8329–8336.
6. Marko-Varga G, Fehniger TE, Rezeli M, Dome B, Laurell T, Vegvari A: **Drug localization in different lung cancer phenotypes by MALDI mass spectrometry imaging.** *J Proteome* 2011, **74**:982–992.
7. Kato H, Nishimura T, Ikeda N, Yamada T, Kondo T, Saijo N, Nishio K, Fujimoto J, Nomura M, Oda Y, Lindmark B, Maniwa J, Hibino H, Unno M, Ito T, Sawa Y, Tojo H, Egawa S, Edula G, Lopez M, Wigmore M, Inase N, Yoshizawa Y, Nomura F, Marko-Varga G: **Developments for a growing Japanese patient population: facilitating new technologies for future health care.** *J Proteome* 2011, **74**:759–764.
8. Marko-Varga G, Ogiwara A, Nishimura T, Kawamura T, Fujii K, Kawakami T, Kyono Y, Tu HK, Anyoji H, Kanazawa M, Akimoto S, Hirano T, Tsuboi M, Nishio K, Hada S, Jiang H, Fukuoka M, Nakata K, Nishiwaki Y, Kunito H, Peers IS, Harbron CG, South MC, Higenbottam T, Nyberg F, Kudoh S, Kato H: **Personalized medicine and proteomics: lessons from non-small cell lung cancer.** *J Proteome Res* 2007, **6**:2925–2935.
9. Westergren-Thorsson G, Larsen K, Nihlberg K, Andersson-Sjoland A, Hallgren O, Marko-Varga G, Bjermer L: **Pathological airway remodelling in inflammation.** *Clin Respir J* 2010, **4**(Suppl 1):1–8.
10. Mok TS: **Personalized medicine in lung cancer: what we need to know.** *Nat Rev Clin Oncol* 2011, **8**:661–668.
11. Kirk R: **Genetics: Personalized medicine and tumour heterogeneity.** *Nat Rev Clin Oncol* 2012, **9**:250.
12. Malmstrom J, Lindberg H, Lindberg C, Bratt C, Wieslander E, Delander EL, Sarnstrand B, Burns JS, Mose-Larsen P, Fey S, Marko-Varga G: **Transforming growth factor-beta 1 specifically induce proteins involved in the myofibroblast contractile apparatus.** *MCP* 2004, **3**:466–477.
13. Miliotis T, Ericsson PO, Marko-Varga G, Svensson R, Nilsson J, Laurell T, Bischoff R: **Analysis of regulatory phosphorylation sites in ZAP-70 by capillary high-performance liquid chromatography coupled to electrospray ionization or matrix-assisted laser desorption ionization time-of-flight mass spectrometry.** *J Chromatogr B Biomed Sci Appl* 2001, **752**:323–334.
14. Miliotis T, Kjellstrom S, Nilsson J, Laurell T, Edholm LE, Marko-Varga G: **Ready-made matrix-assisted laser desorption/ionization target plates coated with thin matrix layer for automated sample deposition in high-density array format.** *RCM* 2002, **16**:117–126.
15. Vegvari A, Magnusson M, Wallman L, Ekstrom S, Bolmsjo G, Nilsson J, Miliotis T, Ostling J, Kjellstrom S, Ottervald J, Franzen B, Hultberg H, Marko-Varga G, Laurell T: **Implementation of a protein profiling platform developed as an academic-pharmaceutical industry collaborative effort.** *Electrophoresis* 2008, **29**:2696–2705.
16. Ansari D, Aronsson L, Sasor A, Welinder C, Rezeli M, Marko-Varga G, Andersson R: **The role of quantitative mass spectrometry in the discovery of pancreatic cancer biomarkers for translational science.** *J Transl Med* 2014, **12**:87.
17. Paik YK, Omenn GS, Thongboonkerd V, Marko-Varga G, Hancock WS: **Genome-wide proteomics, Chromosome-Centric Human Proteome Project (C-HPP), part II.** *J Proteome Res* 2014, **13**:1–4.
18. Marko-Varga G, Omenn GS, Paik YK, Hancock WS: **A first step toward completion of a genome-wide characterization of the human proteome.** *J Proteome Res* 2013, **12**:1–5.
19. Paik YK, Jeong SK, Omenn GS, Uhlen M, Hanash S, Cho SY, Lee HJ, Na K, Choi EY, Yan F, Zhang F, Zhang Y, Snyder M, Cheng Y, Chen R, Marko-Varga G, Deutsch EW, Kim H, Kwon JY, Aebersold R, Bairoch A, Taylor AD, Kim KY, Lee EY, Hochstrasser D, Legrain P, Hancock WS: **The Chromosome-Centric Human Proteome Project for cataloging proteins encoded in the genome.** *Nat Biotechnol* 2012, **30**:221–223.
20. Legrain P, Aebersold R, Archakov A, Bairoch A, Bala K, Beretta L, Bergeron J, Borchers CH, Corthals GL, Costello CE, Deutsch EW, Domon B, Hancock W, He F, Hochstrasser D, Marko-Varga G, Salekdeh GH, Sechi S, Snyder M, Srivastava S, Uhlen M, Wu CH, Yamamoto T, Paik YK, Omenn GS: **The human proteome project: current state and future direction.** *MCP* 2011, **10**:M111.009993.
21. Paik YK, Omenn GS, Uhlen M, Hanash S, Marko-Varga G, Aebersold R, Bairoch A, Yamamoto T, Legrain P, Lee HJ, Na K, Jeong SK, He F, Binz PA, Nishimura T, Keown P, Baker MS, Yoo JS, Garin J, Archakov A, Bergeron J,

Salekdeh GH, Hancock WS: **Standard guidelines for the chromosome-centric human proteome project.** *J Proteome Res* 2012, **11**:2005–2013.
22. Lane L, Bairoch A, Beavis RC, Deutsch EW, Gaudet P, Lundberg E, Omenn GS: **Metrics for the Human Proteome Project 2013-2014 and strategies for finding missing proteins.** *J Proteome Res* 2014, **13**:15–20.
23. Paik YK, Hancock WS: **Uniting ENCODE with genome-wide proteomics.** *Nat Biotechnol* 2012, **30**:1065–1067.
24. Welinder C, Jonsson G, Ingvar C, Lundgren L, Baldetorp B, Olsson H, Breslin T, Rezeli M, Jansson B, Laurell T, Fehniger TE, Wieslander E, Pawlowski K, Marko-Varga G: **Feasibility study on measuring selected proteins in malignant melanoma tissue by SRM quantification.** *J Proteome Res* 2014, **13**:1315–1326.
25. Welinder C, Jonsson G, Ingvar C, Lundgren L, Olsson H, Breslin T, Vegvari A, Laurell T, Rezeli M, Jansson B, Baldetorp B, Marko-Varga G: **Establishing a Southern Swedish Malignant Melanoma OMICS and biobank clinical capability.** *Clinical and translational medicine* 2013, **2**:7.
26. Malm J, Vegvari A, Rezeli M, Upton P, Danmyr P, Nilsson R, Steinfelder E, Marko-Varga G: **Large scale biobanking of blood - the importance of high density sample processing procedures.** *J Proteome* 2012, **76 Spec No**:116–124.
27. Marko-Varga G: **BioBanking as the central tool for translational medicine CTM issue 2013.** *Clin Trans Med* 2013, **2**:4.

doi:10.1186/s40169-014-0038-x
Cite this article as: Nishimura *et al.*: **Clinical initiatives linking Japanese and Swedish healthcare resources on cancer studies utilizing Biobank Repositories.** *Clinical and Translational Medicine* 2014 **3**:38.