The proportion of different BCR-ABL1 transcript types in chronic myeloid leukemia. An international overview

THE PROPORTION OF DIFFERENT BCR-ABL1 TRANSCRIPT TYPES IN CHRONIC MYELOID LEUKEMIA. AN INTERNATIONAL OVERVIEW.

Running title: BCR-ABL1 transcript type in CML

Michele BACCARANI[1], Fausto CASTAGNETTI[1], Gabriele GUGLIOTTA[1], Gianantonio ROSTI[1], Simona SOVERINI[1], Ali ALBEER[2] and Markus PFIRRMANN[2] for the INTERNATIONAL BCR-ABL1 STUDY GROUP[3]

[1]Institute of Hematology "L. and A. Seràgnoli", Department of Experimental, Diagnostic and Specialty Medicine, S.Orsola-Malpighi University Hospital, University of Bologna, Bologna, Italy; [2]Institut fuer Medizinische Informationsverarbeitung, Biometrie und Epidemiologie (IBE), Ludwig-Maximilians-Universitaet München, Munich, Germany

[3]The members of the International BCR-ABL1 Study Group are listed in alphabetical order, by continent, by country, by town, and by name:

AFRICA: **Algeria:** Mohamed-Amine Bekadja, Badra Entasoltan, Mourad Nachi, EHU 1er Novembre, Oran; **Egypt:** Ashraf Elghandour, Manal El Sorady, Alexandria University, Alexandria; Raafat Abdelfattah, Yasser El Nahass, Mohamed Samra , NCI, Cairo; Mohammed Azzazi, Ezat Elsobki, Mohamed Moussa, Ain Shams University, Cairo; Omar Fahmy, Mervat Mattar, Cairo University, Cairo; Samir Eid Shehata, Asuit University, Cairo; Emad Azmy, Mansoura University and Cancer Center, Mansoura; **Nigeria**: Rahman A. Bolarinwa, Obafemi Awolowo University, Ile-Ife; **Tunisia**: Samir Eid, Abderrhaim Khelif, Farhat Hached, Teaching Hospital, Sousse; Samia Menif, Pasteur Institute, Tunis.

ASIA: **Bangladesh**: Hafizur Rahman, International Center for Diarrhoeal Disease, Dhaka; **China**: Xiaojun Huang, Qian Jiang, Peking University People's Hospital and Peking University Inst. of Hematology, Bejin; Yuanxin Ye, Huanling Zhu, West China Hospital, Sichuan University, Sichuan; Suning Chen, The First Affiliated Hospital of Soochow University, Suzhou; **India:** Neelam Varma, PGIMER, Chandigarh; Prasanth

Ganesan, Cancer Inst. (WIA), Chennai; Sadashivudu Gundeti, Nizam's Inst. of Medical Sciences, Hyderabad; Hemant Malhotra, RK Birla Cancer Center, Jaipur; Vivek S. Radhakrishnan, Tata Medical Center, Kolkata; Lalit Kumar, Surender Kumar Sharawat, All India Inst. of Medical Science, New Delhi; Tulika Seth, AIIMS, New Delhi; B.V. Ausekar, Shiddheshwar Cancer Hospital and Research Center, Solapur; Poonkuzhali Balasubramanian, Christian Medical College, Vellore; **Iran**: Behzad Poopak**,** Islamic Azada University and Payvand Clinical and Specialty Laboratory, Tehran**; Japan**: Koiti Inokuchi, Nippon Medical School, Tokyo; **Korea S**.: Dong-Wook Kim, Seoul St. Mary's Hospital, Leukemia Research Unit, The Catholic University of Korea, Seoul; **Oman**: Salam Al Kindi, Sultan Qaboos University, Muscat; **Philippines:** Angelina Mirasol, The Philippines General Hospital, Manila; **Saudi Arabia**: Mohammed Qari, King Abdul Aziz University Hospital, Jeddah; **Singapore**: Yeow Tee Goh, Singapore General Hospital, Singapore; **Taiwan**: Lee-Yung Shih, Chang Gung University and Chang Gung Memorial Hospital, Linkou.

AUSTRALIA: Susan Branford, University of Adelaide and University of South Australia, Adelaide.

EUROPE: **Austria**: Thomas Lion, Children's Cancer Research Institute, Vienna; Peter Valent, Medical University of Vienna, Vienna; Sonja Burgstaller, Joseph Thaler, Klinikum Wels-Grieskirchen, Wels; **Croatia:** Boris Labar, Renata Zadro, University Hospital Center Zagreb, Zagreb; **Czech Republic**: Jiri Mayer, Daniela Zackova, University Hospital Brno and Masaryk University, Brno; Edgar Faber, Palacky University, Olomouc; **Denmark:** Niels Pallisgaard, Sjaellands Universitethospital, Roskilde; **France:** François Xavier-Mahon, University of Bordeaux, Bordeaux; Eric Lippert, CHU, Brest; Jean Michel Cayuela, Delphine Réa, Hopital Saint Louis, Paris; Frederic Millot; CHU de Poitiers,Poitiers; **Germany**: Meinolf Suttorp, Universitaetsklinikum Carl Gustav Carus, Dresden; Andreas Hochhaus, Universitaetsklinikum Jena, Jena; Dietger Niederwieser, University of Leipzig, Leipzig; Susanne Saussele, University Clinic Mannheim, Mannheim; Torsten Haferlach, Sabine Jeromine, MLL Muenchner Leukaemielabor, Muenchen; **Greece:** Panayiotis Panayiotidis, University of Athens and Laikon General Hospital, Athens; **Ireland:** Eibhlin Conneally, Steve Langabeer, St. James Hospital, Dublin; **Israel:** Arnon Nagler, Chaim Sheba Medical Center and Tel Aviv University, Tel Aviv; **Italy**: Serena Rupoli,

University of Ancona, Ancona; Nicola Santoro, Azienda Ospedaliero-Universitaria "Policlinico Consorziale", Bari; Francesco Albano, University of Bari, Bari; Fausto Castagnetti, Emanuela Ottaviani, University of Bologna and S.Orsola-Malpighi University Hospital, Bologna; Alessandro Rambaldi, University of Milano and AAST Papa Giovanni XXIII, Bergamo; Fabio Stagno, University of Catania, Catania; Stefano Molica, A. Ospedale Pugliese Ciaccio, Catanzaro; Caterina Biagiotti, Barbara Scappini, University of Florence, Florence; Roberto Lemoli, University of Genova, Genova; Alessandra Iurlo, Fondazione IRCCS Ca' Granda and University of Milano, Milano; Ester Pungolino, Ospedale Niguarda, Milano; Giuseppe Menna, Fara Petruzziello, Ospedale Pausillipon, Napoli; Fabrizio Pane, University Federico II, Napoli; Enrico Gottardi, Giovanna Rege-Cambrin, S. Luigi Gonzaga Hospital and University of Torino, Orbassano; Gianni Binotto, University of Padova, Padova; Maria Caterina Putti, Clinica di Emato-oncologia Pediatrica, Azienda Ospedaliera, Padova; Franca Falzetti, University of Perugia, Perugia; Giuseppe Visani, Ospedale S.Salvatore, Pesaro; Sara Galimberti, University of Pisa, Pisa; Pellegrino Musto, IRCCS-CROB, Rionero in Vulture; Elisabetta Abruzzese, S. Eugenio Hospital, Roma; Massimo Breccia, Fiorina Giona, Sapienza University, Roma; Patrizia Chiusolo, Simona Sica, Policlinico Gemelli, Roma; Carmen Fava, Dario Ferrero, University of Torino, Torino; Mario Tiribelli, University of Udine, Udine; Massimiliano Bonifacio, University of Verona, Verona; **Lithuania:** Laimonas Griskevicius, Vilnius University, Vilnius; **Moldova**: Vasile Musteata, State University of Medicine and Pharmacy “N. Testemitanu”, Chisinau; **Netherlands**: Jeroen Janssen, VU University Medical Center, Amsterdam; **Poland**: Witold Prejzner, Medical University of Gdansk, Gdansk; Tomasz Sacha, Joanna Waclaw, Jagiellonian University Hospital, Krakow; **Portugal:** Antonio Medina Almeida, Inst. Portugues de Oncologia de Lisboa Francisco Gentil, Lisboa; **Russia**: Sergei Kulikov, Anna Turkina, National Research Center for Hematology of Ministry for Healthcare of Russian Federation, Moscow; **Serbia**: Andrjia Bogdanovic, Clinical Center of Serbia, University of Belgrade, Belgrade; **Slovenia**: Irena Zupan, University Clinical Center, Ljubljana; **Spain**: Silvia Marcé, Hospital Germans Trias i Pujol-ICO Badalona, Badalona; Francisco Cervantes, Hospital Clinic Barcelona, IDIBAPS, Barcelona; Juan Luis Steegmann, Hospital Universitario de la Princesa, Madrid; **Ukraine:** Konstyantyn Kotlyarchuk, SI Inst. of Blood Pathology and Transfusion Medicine, Lviv; **United Kingdom**: Benedict J. Milner, Aberdeen

Royal Infirmary, Aberdeen; Susan Rose, Birmingham Women's NHS Foundation Trust, Birmingham; Tim Clench, Bristol Royal Infirmary, Bristol; Paula Waits, Bristol Genetics Laboratory, Bristol; Steve Austin, Cardiff and Vale University Health Board, Cardiff; Caroline Wickham, Royal Devon and Exeter NHS Foundation Trust, Exeter; Richard Clark, Royal Liverpool University Hospital, Liverpool; Jane Apperley, Simone Claudiani, Letizia Foroni, Richard Szydlo, Imperial College London, Hammersmith Hospital, London; Emma Burt, The Royal London Hospital, London; Ruth Bescoby, Leanne Cork, Stephen O'Brien, University of Newcastle, Newcastle; Bethaney Green, Sarah Hawtree, Mark Watson, Sheffield Diagnostics Genetics Services, Sheffield.

SOUTH AMERICA: **Argentina**: Raquel Maria Bengiò, Irene Larripa, IIHEMA-Academia Nacional de Medicina, Buenos Aires; Carolina Pavlovsky, FUNDALEU, Hospitalization and Clinical Research Center, Buenos Aires; Beatriz Moiraghi, Hospital JM Ramos Mejia, CABA, Buenos Aires; **Brasil:** Cristiane Almeida Requiao de Pinna, Federal University Hospital of Bahia, Bahia; Gustavo Henrique Romani Magalhaes, Minas Gerais Federal University-Hospital das Clinicas, Belo Horizonte; Katia Pagnano, University of Campinas, Campinas; Vaneuza Funke, Federal University of Parana, Curitiba; Renato Sampaio Tavares, Hospital das Clinicas da UFG, Goiania; Adriana Prado, Clinical Hospital of Federal University of Goias, Goias; Alita Andrade Azevedo, Centro de Hematologia e Hematoterapia de Pernambuco, Pernambuco; Laura Fogliatto, Hospital de Clinicas de Porto Alegre, Porto Alegre; Simone Bonecker, CEMO/INCA, Rio de Janeiro; Renato Centrone, Inst. Hemomed of Oncology and Hematology, Rio de Janeiro; Artur Moellman, INCA/MS, Rio de Janeiro; Israel Bendit, University of Sao Paulo, Sao Paulo; Monika Conchon, Hospital Santa Marcelina, Sao Paulo; **Paraguay**: Maria Elida Centurion, Central Hospital of Asuncion, Asuncion; **Uruguay**: Ana-Ines Prado, Hospital Maciel, Montevideo; **Venezuela**: JL Lopez, Banco Municipal de Sangre del DC, Caracas.

Correspondence address to

Michele Baccarani, MD, Institute of Hematology "L. and A. Seràgnoli", S.Orsola-Malpighi University Hospital, Via Massarenti 9, 40138 Bologna, Italy

Tel. +39051390413, Fax +39051398973

e-mail michele.baccarani@unibo.it; c.ferri@unibo.it

Conflict of interest: none

Competing interests: none

Word count: abstract 210, text 3383

Tables: 5 + 5 Supplementary

Figures: 3

ABSTRACT

There are different *BCR-ABL1* fusion genes that are translated into proteins that are different from each other, yet all leukemogenic, causing chronic myeloid leukemia (CML) or acute lymphoblastic leukemia. Their frequency has never been systematically investigated. In a series of 45503 newly diagnosed CML patients reported from 45 countries, it was found that the proportion of e13a2 (also known as b2a2) and of e14a2 (also known as b3a2), including the cases co-expressing e14a2 and e13a2, was 37.9% and 62.1% respectively. The proportion of these two transcripts was correlated with gender, e13a2 being more frequent in males (39.2%) than in females (36.2%), was correlated with age, decreasing from 39.6% in children and adolescents down to 31.6% in patients ≥80 years old, and was not constant worldwide. Other, rare transcripts were reported in 666/34561 patients (1.93%). The proportion of rare transcripts was related with gender (2.27% in females and 1.69% in males) and with age (from 1.79% in children and adolescents up to 3.84% in patients ≥ 80 years old). These data show that the differences in proportion are not by chance. This is important, as the transcript type is a variable that is suspected to be of prognostic importance for response to treatment, outcome of treatment, and rate of treatment-free remission.

INTRODUCTION

A reciprocal translocation between the long arms of chromosome 9 and chromosome 22 [t(9;22)(q34;q1.1)] results in the formation of a fusion hybrid gene (*BCR-ABL1*) that encodes proteins with a constitutively activated tyrosine kinase activity (1,2). The proteins are leukemogenic and cause either chronic myeloid leukemia (CML) or acute lymphoblastic leukemia (ALL, so-called Philadelphia chromosome-positive ALL) (1,2). The site of the breakpoint in the *ABL1* gene (chromosome 9) "*can occur anywhere within a >300-kb segment at the 5′ end of the gene, either upstream of the first alternative intron Ib, between exons Ib and Ia, or downstream of exon Ia"* (1). The site of the breakpoint in the *BCR* gene (chromosome 22) can occur in different regions "*within a 5.8-kb region known as the major breakpoint cluster region (M-bcr), spanning 5 exons historically named b1 to b5, now known to be exon 12 to 16 of the bcr gene"* (1). In more than 90% of patients the transcripts are e13a2 (also known as b2a2) or e14a2 (also known as b3a2), that can be found either alone or together. However, in some patients other types of transcripts can be found, that are defined as "rare" or "atypical" (1,2). The frequency of the different molecular types of CML has never been systematically and thoroughly investigated. The reported proportion of e13a2 ranged between 34.4% and 42.5% in clinical studies, between 33.9% and 49.4% in international, national or regional registries, and between 20.0% and 94.6% in monocentric studies (4-31) **(Table 1S).** The reported proportion of patients co-expressing the two M-bcr transcripts (e14a2 and e13a2) ranged between 2.1% and 17.7% (**Table 1S**). The proportion of atypical, rare transcripts was rarely reported, ranging between 0.9% and 13.0% **(Table 1S).**

The translation of each different RNA transcript results in different protein tyrosine kinases that can potentially affect the biologic characteristics of the disease and the response to treatment. For this reason, we planned a worldwide overview of the prevalence of BCR-ABL1 transcript types.

PATIENTS and METHODS

An invitation to join the study and to contribute data was sent to 201 investigators in 180 centers in Africa, Asia, Australia, Europe, North America, and South America. The investigators were selected on personal basis, or as first or senior authors of clinical papers on CML published in the period 2000 to 2016.

One hundred and thirteen centers (63% of those invited) contributed data, nine from Africa (Algeria, Egypt, Nigeria, and Tunisia), twenty from Asia (Bangladesh, China, India, Iran, Japan, Oman, Philippines, Saudi Arabia, Singapore, South Korea, and Taiwan), one from Australia, sixty-six from Europe (Austria, Croatia, Czech Republic, Denmark, France, Germany, Greece, Ireland, Israel, Italy, Lithuania, Moldova, Netherlands, Poland, Portugal, Russia, Serbia, Slovenia, Spain, Ukraine, and United Kingdom), seventeen from South America (Argentina, Brazil, Paraguay, Uruguay, and Venezuela), and none from North America. Data were provided also by the GIMEMA CML Working Party, the International Childhood CML Registry, and the French, German, and Taiwan CML Study Groups. Three main reasons prevented from participation, particularly of North American centers, namely the inability to distinguish between the two *M-bcr* transcripts, e13a2 and e14a2, for technical (laboratory) reasons, or the absence of an institutional CML database and the lack of a financial support.

Sixty-one of 113 centers were also able to contribute the data of the patients with other, rare transcripts, who were seen during the same period. These centers were from Africa (Algeria and Nigeria), Asia (Bangladesh, China, India, Japan, Singapore, South Korea, and Taiwan), Australia, Europe (Austria, Croatia, Czech Republic, Denmark, France, Germany, Greece, Ireland, Israel, Italy, Lithuania, Moldova, Netherlands, Poland, Portugal, Slovenia, Spain, and United Kingdom), and South America (Brazil and Paraguay).

The study protocol required to report transcript type, gender, and age (by decade) of all patients with BCR-ABL1 positive CML who were newly diagnosed over a defined period of time, i.e. between 1990 and 2016. The

protocol was approved by the Ethic Committee (EC) of the S.Orsola-Malpighi University Hospital, Bologna, and by the EC of the participating institutions, as required by national rules.

The sources of all the data were institutional - clinical or laboratory - databases. The aim of the study was to investigate if the proportion of different transcript types was associated with gender, age, and the region where patients were living. Accordingly, calculations and comparisons were made for gender (females vs males), for different age groups (by 20-year intervals), and for continents. so as to understand if the distribution of the type of the transcript was constant worldwide or could vary on geographic bases.

The data were tabulated and analyzed according to the original transcript identification, as it was reported in the respective databases, as e13a2, e14a2, e14a2/e13a2, or other transcript types, although it is acknowledged that any patients may express more than one transcript, due to alternative mechanisms of splicing (1-3). The transcript that is prevalent is used to define the transcript type of the patient, and, regrettably, there is no agreement on how reporting the other transcripts that can be detected in smaller amounts, depending on the sensitivity of PCR. For all the calculations, we have pooled the patients where a co-expression of both *M-bcr* transcripts (e14a2/e13a2) was reported with the patients who were reported to express only e14a2. In such cases, the breakpoint must be located downstream to exon 14, because an e14a2 can also make an e13a2, but an e13a2 does not contain exon 14. Some patients with the simultaneous expression of an *M-bcr* and an *m-bcr* (e1a2) transcript were also reported by eleven centers. Here again, such co-expression is explained by alternative splicings (1-3,32,33). Therefore, none of the patients with a reported simultaneous expression of an M-bcr and an m-bcr transcript was classified as a case of an atypical or rare transcript type.

Statistics

For descriptive analyses, raw, unadjusted proportions are stated.. To improve comparability of the prevalence of e13a2 between groups, adjusted odds ratios (OR), together with their 95% confidence interval (95% CI), were estimated by multiple logistic regression. The adjustment was done by including gender, age group, and

continent in the regression model. Regarding the analyses of atypical transcripts, continent was not included in multiple model, because information was only based on Asia and Europe. The significance level of the two-sided P-values was 0.05. Since all analyses are descriptive and explorative, p-values were not adjusted for multiple comparisons. Analyses were done with SAS version 9.4.

RESULTS

A summary of the number of patients who were collected, and detailed information on their distribution by continent, by gender, and by age is shown in **Table 2S.** The total number of reported patients was 45503. Females were 48% of the total in Africa, 43% in Australia, Europe, and South America, and only 39% in Asia. Excluding the childhood registry, the male to female ratio (M:F) was 1.33 overall, ranging from 1.10 for Africa to 1.64 for Asia (**Table 3S**). Of note, the M:F ratio decreased with age, down to 1.10 for the very elderly (**Table 3S**). Continent distribution by age **(Figure 1)** was uneven, since the proportion of children and adolescents ranged between 1.3% for Africa and 6.6% for Asia, while the proportion of elderly patients (≥ 60 years old) ranged between 13.1% for Asia and 41.8% for Europe. The majority of the data of young patients (0-19 and 20-39 years old) was contributed by Asia and Europe **(** see also **Table 2S),** while Europe contributed more than 50% of the 40-59-years old patients, and more than 80% of the patients ≥ 80 years old. The simultaneous expression of both M-bcr transcripts (e14a2/e13a2) was reported in 7.6% of the patients overall, ranging from 2.4% for Asia to 18.7% for Australia **(Table 4S).**

M-bcr TRANSCRIPTS

The proportion of patients who were reported to express only the e13a2 transcript was 37.9% overall, but was different by gender, by age, and by continent **(Table 1).** By gender, the proportion was 39.2% for males and 36.2% for females. The age- and continent-adjusted OR of males vs females was 1.149 (95% CI: 1.106-1.194, p < 0.0001). Regarding age, the raw proportions expressing only the e13a2 transcript were similar in the first two age groups (39.6% and 39.9%, respectively), but decreased in the subsequent age groups to 38.7%, 35.6%, and 31.6%, respectively (**Table 2).** With reference to the 0-19-years old, the ORs of the 60-79-years old patients (OR: 0.775, 95% CI: 0.690-0.871, P < 0.0001) and of the patients ≥ 80 years (OR: 0.647, 95% CI: 0.560-0.749, P < 0.0001) were significantly lower. This pattern of decrease was found in males as well as in females (**Figure 2a).** The proportion of patients with e13a2 was the lowest in Asia: 33.2% vs 35.4% in Australia, 37.6% in Europe, and 44.6% in Africa (p = 0.0001, 0.0308, 0.0773, and 0.0001, respectively) (**Table 5S).**

The data from children and adolescents are listed in **Table 3**, distinguishing between children (0-9 years old) and adolescents (10-19 years old), and separately for the data from the International Childhood CML Registry and all other sources. The proportion with e13a2, overall, was as in adults (38.4%), being 37.5% in males and 39.5% in females. No further comparisons were made because of low patient numbers and different sources.

ATYPICAL, RARE *BCR-ABL1* TRANSCRIPTS

Thirteen different atypical transcripts were reported in 603 of 34561 patients (**Table 4**). Sixty-three additional cases of atypical transcripts were reported, where the exact transcript type was not identified, totaling to 666 cases and a proportion of 1.93%. The distribution of the 666 cases of atypical transcripts, overall and by gender, is shown in Figure 3. Overall, all these transcripts were more frequent in females than in males: 2.27% vs 1.69% (ORs 1.352, 95% CI 1.159-1.576, p = 0.0001). More frequent transcripts were e1a2, e19a2, e13a3, and e14a3. e1a2 was twice more frequent in females than in males: 1.33% vs 0.62% of all 34561 patients (ORs 2.172, 95% CI 1.731-2.724, p < 0.0001), and 58.6% vs 36.5% of all 666 patients with atypical transcripts Other less frequent transcripts were e1a3, e6a2 and e8a2. Other transcripts (e1a1, e8a1, e8a3, e15a2, e23a1) were reported very rarely, in only six patients. These transcripts were not clear exon/exon fusions, and some had additional nucleotides inserted in the chimeric mRNA that may or may not have maintained the reading frame. By age, the proportion of atypical transcripts was similar in the first three age groups, ranging from 1.79% in patients 0-19 years old to 1.58% in patients 40-59 years old and increasing in the elderly up to 3.84% for patients ≥ 80 years old (**Table 5**). The increase with higher age was found in both genders (**Figure 2b).**

For children (0-9 years old) and children and adolescents (10-19 years old) the proportion of patients with atypical transcripts was calculated separately for the data collected by the International Childhood CML Registry and the data collected from all other sources (**Table 3**). Overall, for children the proportion of atypical transcripts was 3.25%. For children and adolescents the proportion was 2.12%. No further comparisons were made because of low patient numbers and different sources of the data.

DISCUSSION

The results of this international overview must be interpreted with caution, taking into account *how* the data were collected. The collection was retrospective, from the clinical and laboratory database of the centers that were invited. Care was taken to ensure the quality of the data, but it is acknowledged that the basis of the review was by necessity so large that few errors or inconsistencies may not have been avoided. Thus, the data cannot provide the exact prevalence of all the transcripts. However, it is justifiable to assume that at least the *M-bcr* transcripts were reported independently of their parameter value, may it have been e13a2 or e14a2 or both. Accordingly, we can at least report an estimate of the "prevalence" of e13a2 in relation to our sample.

This study was planned with the aim of collecting as many data as possible from as many countries as possible, because the goal was to provide an overview of the prevalence of different transcripts, to analyze if it is related with gender and age, and also if it is constant worldwide or varies from region to region. The full achievement of these goals was limited because several countries could not contribute data, and, regrettably, no data could be collected from North and Central Americas.

The first aim was to calculate the proportion of the two *M-bcr* transcripts, that was 37.9% for e13a2, and, accordingly, 62.1% for e14a2 + e14a2/e13a2. Even taking into account the limitations due to the lacking information from many regions, this overview provides the best available estimates of the true prevalence of the two M-bcr transcripts, worldwide.

It was interesting and important to find that the M-bcr transcript proportions differed by gender and age, and also by continent. These differences were previously unknown or unrecognized. Differences by gender and age are well known to affect the incidence of many cancers, including CML, that is more frequent in males and in the elderly. Now, we found that age and gender affect also the distribution of the M-bcr transcript types. The differences by continent cannot have technical explanations, because they were found across gender and age, and because "continents" included different countries and different laboratories. Moreover they cannot depend on the

definition of the transcript, because the patients who were reported to co-express both M-bcr transcripts were always pooled together with the patients who were reported to express only e14a2. Since every continent includes regions with geographic and ethnic differences, and since data collection was far from covering all continents, these explicit proportions must be considered with caution and do not provide reliable estimates for a certain combination of continent, age group or gender in epidemiologic sense. However, they suggest that the distribution of BCR-ABL1 transcript types is not the same worldwide.

In summary, it was found that the proportion of e13a2 was higher in males than in females, that it decreased with age for both genders, and that it has geographic variations. Whether the type of the *M-bcr* transcript affects the biological characteristics of the disease and the response to treatment was a matter of a debate that begun many years ago (34), in the era of chemotherapy and interferon-alfa (5,6,10). A difference in platelet count (lower in e13a2 patients) and in leukocyte count (higher in e13a2 patients) had already been reported (1,2,4,12,20,23). After the introduction of tyrosine kinase inhibitors (TKI), it was found that e13a2 was associated with a lower rate of complete cytogenetic response (CCyR) and a longer time to CCyR (17), and with a lower rate of major molecular response (MMR) and a longer time to MMR in imatinib-treated patients (4,5,20,21,35), and also in nilotinib-treated patients (8), as well as a lower rate of deep MR (MR 4.0) (5,35). Three studies of 1494, 481, and 559 imatinib-treated patients reported an inferior overall survival, leukemia-related survival, progression-free survival and transformation–free survival in e13a2 patients, respectively (15,20,35). Few studies of treatment-free remission (TFR), that is nowadays the most important goal of treatment, have reported or analyzed transcript data. In a study of de-escalation of treatment (36), it was found that the molecular relapse rate was twice higher in e13a2 patients than in e14a2 patients. The e13a2 transcript was reported to have a significant adverse impact on the achievement of a sustained deep molecular response and on the maintenance of TFR (37). In a preliminary analysis of 249 patients who had achieved a deep molecular response and had discontinued treatment, only 62 (25%) were found to have the e13a2 transcript, less than expected from a random distribution, thus supporting the hypothesis that e13a2 patients can achieve a deep molecular response less frequently (38). Moreover, female sex and older age, that are both associated with a

lower incidence of e13a2, were reported to predict stable undetectable molecular response and TFR rates, respectively (39,40). Of interest, in a study of *BCR-ABL1*+ acute lymphoblastic leukemia (ALL) (41) it was found that out of a total of fifteen P210 patients, 13 (86%) had the e13a2 transcript, suggesting that the protein encoded by e13a2 may have different biological properties.

Different transcripts result into slightly different proteins, but could also result in different cellular amounts of the leukemogenic proteins, and in different sensitivity to TKIs. Moreover, since the proteins that are encoded by the two transcripts differ in 25 aminoacids at the fusion junction of *Bcr* and *Abl1*, the two proteins may have a different immunogenicity, and may be recognized by different antibodies or different lymphocytes or natural killer cells, thus affecting the immunological control of minimal residual disease (42-45). Few studies have reported on a relationship between immunogenicity and transcript type and it is of interest that they all focused on the e14a2 transcript. Patients with e14a2, but not e13a2, were reported to produce IFNγ in response to stimulation with Ph+ monocyte-derived dendritic cells (42). An immune response and a clinical benefit were reported in 11 of 16 e14a2 patients who were vaccinated with a peptidic vaccine derived from the sequence p210-e14a2 (43). Nineteen e14a2 imatinib-treated patients were vaccinated with *Bcr-Abl1* peptides spanning the e14a2 fusion junction, and 14 of 19 developed T cell responses to those peptides (44). In a pilot study of vaccination with autologous non-irradiated dendritic cells, T-cells recognizing leukemia-associated antigens became detectable in three of ten patients, and all three had the e14a2 transcript (45).

The so-called atypical or rare transcripts (all transcripts but e13a2 and e14a2) were detected and reported since many years (1,2), with focus on a shorter transcript, e1a2 (4,20,25,46-50), and on a longer transcript, e19a2 (1,2,4,20,25). The frequency of e1a2 reported so far ranged between 0.26% and 3.31% (4,25,46-51). In this overview it was found that the proportion of e1a2 was 0.91%, twice higher in females than in males (1.33% vs 0.62%). e1a2 is the transcript coding for P190, that has different, peculiar, biologic characteristics (52-57), and causes *BCR-ABL1*+ ALL, that is more frequent in the elderly. Several case-reports have suggested that in patients with CML e1a2 is associated with an inferior response to TKIs and an inferior outcome (46-49,58).

Opposed to what was observed for e1a2, it was found that the proportion of the "a3" variants (where the transcript does not contain "a2") was higher in males than in females. The biological characteristics of the “a3” variants, lacking the a2 exon, are not clear, because the partial deletion of the SH3 domain that is encoded by the a2 exon may result either in higher levels of activation and leukemogenicity, or, on the opposite, might reduce signaling and leukemogenicity (59,60). e19a2, coding for P230, has been associated with a particular type of chronic neutrophilic leukemia (1,2,4,25,55,56). Precautionary, all patients with a rare, atypical transcript should be considered as having atypical forms of CML with different sensitivity to treatment and different outcomes.

In conclusion, this overview showed that the distribution of the different transcript types is not likely to be ruled by chance. The data should be considered with caution, only as a best approximate to the true transcript frequency, since the overview could not cover homogeneously all continents, and the data of rare transcripts were collected mainly from Asia and Europe and were not available across all centers. Epidemiologic studies are requested to clarify geographic or ethnic differences. The study was designed to evaluate only the distribution of transcript types, and its relationship with gender and age, and cannot allow to speculate on the mechanisms underlying these relationships. However, since a relationship with gender and age has been established, and since different transcripts may affect the biology of the disease and the rates of response, of outcome, and particularly of TFR, it is suggested that the transcript is a variable that may help refining the strategies of treatment, may influence the outcome, and should be taken into consideration in all studies of CML. The differences shown in this study are small, but even small differences may be important in a disease where current treatment is so successful that the survival of CML patients is already aligned with the survival of non-leukemic people (61,62), so that the ambition is to achieve a TFR.

ACKNOWLEDGEMENTS

The cooperation, the advices and the comments of the members of the International Childhood CML Registry, the GIMEMA CML Working Party, the French, German, and Taiwan CML Study Groups, the Arab LeukemiaNet, and of Mohammed Azzazi, Susan Branford, Richard Clark, Nicholas Cross, Letizia Foroni, Torsten Haferlach, Quian Jiang, Eric Lippert, Junia Melo, and Dietger Niederwieser, are kindly acknowledged. The skilled technical support of Michela Apolinari and Chiara Ferri is kindly acknowledged. This work was supported in part by the ELN Foundation.

Conflict of interest: none

Competing interests: none

Supplementary information is available at Leukemia's website at the end of the article and before the references.

REFERENCES

1. Melo JV. The diversity of BCR-ABL fusion proteins and their relationship to leukemia phenotype. Blood 1996;88(7):2375-2384.

2. Chereda B and Melo JV. The biology and pathogenesis of chronic myeloid leukemia. In, Hehlmann R ed, Chronic myeloid leukemia, Springer, 2016, pp.17-40.

3. Verschraegen CF, Kantarjian HM, Hirsch-Ginsberg C, Lee MS, O'Brien S, Rios MB, et al. The breakpoint cluster region site in patients with Philadelphia chromosome-positive chronic myelogenous leukemia. Clinical, laboratory and prognostic correlations. Cancer 1995;76:992-997.

4. Hanfstein B, Lauseker M, Hehlmann R, Saussele S, Erben P, Dietz C, et al. Distinct characteristics of e13a2 versus e14a2 BCR-ABL1 driven chronic myeloid leukemia under first-line therapy with imatinib. Haematologica 2014;99(9):1441-1447.

5. The ICSG on CML. Chronic myeloid leukemia, BCR/ABL transcript, response to α-interferon, and survival. Leukemia 1995;9:1648-1651.

6. Baccarani M, Rosti G, de Vivo A, Bonifazi F, Russo D, Martinelli G, et al. A randomized study of interferon-α versus interferon-α and low dose arabinosyl cytosine in chronic myeloid leukemia. Blood 2002;99 (5);1527-1535.

7. Castagnetti F, Gugliotta G, Breccia M, Stagno F, Iurlo A, Albano F, et al. Long-term outcome of chronic myeloid leukemia patients treated frontline with imatinib. Leukemia 2015;29:1823-1831.

8. Castagnetti F, Gugliotta G, Palandri F, Breccia M, Stagno F, Levato L, et al. The BCR-ABL1 transcript type does not influence the response and the outcome of chronic myeloid leukemia patients treated frontline with nilotinib. 54th Annual Meeting of the American Society of Hematology, Blood 2012, abstract 1680.

9. Cervantes F, Lopez-Garrido P, Montero M-I, Jonte F, Martinez J, Hernandez-Boluda I-C, et al. Early intervention during imatinib therapy in patients with newly diagnosed chronic-phase chronic myeloid leukemia: a study of the Spanish PETHEMA Group. Haematologica 2010;95(8):1317-1324.

10. Shepherd P, Suffolk R, Halsey J and Allan N. Analysis of molecular breakpoint and m-RNA transcripts in a prospective randomized trial of interferon in chronic myeloid leukaemia: no correlation with clinical features, cytogenetic response, duration of chronic phase, or survival. Br J Haematol 1995;89:546-554.

11. Bonifacio M, Scaffidi L, Binotto G, De Marchi F, Maino E, Calistri E, et al. Predictive factors of stable deep molecular response in chronic myeloid leukemia patients treated with standard dose imatinib: a study from the "Gruppo Triveneto LMC". 57th Congress of the American Society of Hematology, Blood 2015, abstract 597.

12.Castagnetti F, Di Raimondo F, De Vivo A, Spitaleri A, Gugliotta G, Fabbiano F, et al. A population-based study of chronic myeloid leukemia patients treated with imatinib in first line. Am J Hematol 2017;92:82-87.

13. Osorio S, Casado LF, Giraldo P, Maestro N, Andrade M, Redondo S, et al. Chronic myeloid leukaemia in Spain: its presentation characteristics have changed. Spanish Section of the EUTOS population-based registry. Rev Clin Esp 2016;216:293-300.

14. Hoffmann VS, Baccarani M, Hasford J, Lindoerfer D, Burgstaller S, Sertic D, et al. The EUTOS population-based registry: incidence and clinical characteristics of 2904 CML patients in 20 European countries. Leukemia 2015;29:1336-1343.

15. Pfirrmann M, Evtimova D, Saussele S, Castagnetti F, Cervantes F, Janssen J, et al. No influence of BCR-ABL1 transcript types e13a2 and e14a2 on long-term survival: results in 1494 patients with chronic myeloid leukemia treated with imatinib. J Cancer Res Clin Oncol 2017;143(5):843-850.

16. Kim D, Goh HG, Kim S-H, Choi S-Y, Park S-H, Jang E-J, and Kim D-W. Comprehensive therapeutic outcomes of frontline imatinib mesylate in newly diagnosed chronic phase chronic myeloid leukemia patients in Korea: feasibility assessment of current ELN recommendations. Int J Hematol 2012;96:47-57.

17. Lucas CM, Harris RJ, Giannoudis A, Davies A, Knight K, Watmough SJ, et al. Chronic myeloid leukaemia patients with the e13a2 BCR-ABL fusion transcript have inferior responses to imatinib compared to patients with the e14a2 transcript. Haematologica 2009;94(10):1362-1367.

18. de Lavallade H, Apperley JF, Khorashad JS, Milojkovic D, Reid AG, Bua M, et al. Imatinib for newly diagnosed patients with chronic myeloid leukemia: incidence of sustained responses in an intention-to-treat analysis. J Clin Oncol 2008;26(20):3359-3363.

19. Marin D, Gabriel IH, Ahmad S, Fioroni L, de Lavallade H, Clark R, et al. KIR2DS1 genotype predicts for complete cytogenetic response and survival in newly diagnosed chronic myeloid leukemia patients treated with imatinib. Leukemia 2012;26:296-302.

20. Jain P, Kantarjian H, Patel KP, Nogueras Gonzales G, Luthra R, Shamanna RK, et al. Impact of BCR-ABL transcript type on outcome in patients with chronic-phase CML treated with tyrosine kinase inhibitors. Blood 2016;127(10):1269-1275.

21. Pagnano KBB, Miranda EC, Delamain MT, Duarte GO, de Paula EV, Lorand-Metze I, and de Souza CA. Influence of BCR-ABL transcript type on outcome in patients with chronic phase chronic myeloid leukemia treated with imatinib. Clinical Lymphoma, Myeloma & Leukemia 2017 17(11):728-733.

22. Paz-Y-Mino C, Burgos R, Morillo SA, Santos JC, Fiallo BF, and Leone PE. BCR-ABL rearrangement frequencies in chronic myeloid leukemia and acute lymphoblastic leukemia in Ecuador. Cancer Gen Cytogen 2002;132:65-67.

23. Mondal BC, Bandyopadhyay A, Majumdar S, Mukhopadhyay A, Chandra S, Chaudury U, et al. Molecular profiling of chronic myeloid leukemia in Eastern India. Am J Hematol 2006;81(11):845-849.

24. Sharma P, Kumar L, Mohanty S, and Kochupillai V. Response to imatinib mesylate in chronic myeloid leukemia patients with variant BCR-ABL fusion transcripts. Ann Hematol 2010;89:241-247.

25. Arun AK, Senthamizhselvi A, Mani S, Vinodhini K, Janet NB, Lakshmi KM, et al. Frequency of rare BCR-ABL1 fusion transcripts in chronic myeloid leukemia patients. Int Jnl Lab Hem 2017;39:235-242.

26. Yaghmaie M, Ghaffari SH, Ghavamzadeh A, Alimoghaddam K, Jahani M, Mousavi SA, et al. Frequency of BCR-ABL fusion transcripts in Iranian patients with chronic myeloid leukemia. Arch Iran Med 2008;11(3):247-251.

27. Osman E-AI, Hamad K, Elmula IMF, and Ibrahim ME. Frequencies of BCR-ABL1 fusion transcripts among Sudanese chronic myeloid leukaemia patients. Genet Mol Biol 2010;33(2):229-231.

28. Adler R, Viehman S, Kulisch E, Martiniak Y, Rottgers S, Harbott J, and Suttorp M. Correlation of BCR/ABL transcript variants with patients' characteristics in childhood chronic myeloid leukemia. Europ J Haematol 2009;82(2):112-119.

29. Giona F, Putti MC, Micalizzi C, Menna G, Moletti ML, Santoro N, et al. Long-term results of high dose imatinib in children and adolescents with chronic myeloid leukemia in chronic phase: the Italian experience. Br J Haematol 2015;170:398-407.

30. Hasan SK, Sazawal S, Kumar B, Chaubey R, Mishra P, Chaudry VP, et al. Childhood CML in India: b2a2 transcript is more common than b3a2. Cancer Genet Cytog 2006;169:76-77.

31. Millot F, Suttorp M, Guilhot J, Sedlacek P, De Bont E, Li CK, et al. The International Registry for chronic myeloid leukemia in children and adolescents (I-CML-Ped-Study): objectives and preliminary results. 54th Annual Meeting of the American Society of Hematology, Blood 2012; 120(21), abstract 3741.

32. van Rhee F, Hochhaus A, Lin F, Melo JV, Goldman JM, and Cross NC. p190 BCR-ABL mRNA is expressed at low levels in p210-positive chronic myeloid and acute lymphoblastic leukemias. Blood 1996;87:5213-5217.

33. Lichty BD, Keating A, Callum J, Yee K, Croxford R, Corpus G, et al. Expression of p210 and p190 BCR-ABL due to alternative splicing in chronic myelogenous leukaemia. Br J Haematol 1998;103:711-715.

34. Mills KL, Mackenzie ED, and Birnie GD. The site of the breakpoint within the bcr is a prognostic factor in Philadelphia-positive CML patients. Blood 1988;72:1237-1241.

35. Castagnetti F, Gugliotta G, Breccia M, Iurlo A, Levato L, Albano F, et al. The BCR-ABL1 transcript type influences response and outcome in Philadelphia chromosome-positive chronic myeloid leukemia patients treated front-line with imatinib. Am J Hematol 2017;92(8):797-805.

36. Clark RE, Polydoros F, Apperley JF, Milojkovic D, Pacock C, Smith G, et al. De-escalation of tyrosine kinase inhibitor dose in patients with chronic myeloid leukaemia with stable major molecular response (DESTINY): an interim analysis of a non-randomized, phase 2 trial. Lancet Haematol 2017;4(7):e310-e316.

37. D'Adda M, Farina M, Schieppati F, Cerqui E, Ruggeri G, Ferrari S, et al. An e13a2 type of BCR-ABL transcript has a significant adverse impact on the achievement of a sustained deep molecular response and on the maintenance of a treatment free remission after stopping tyrosine kinase inhibitors. 59th Annual Meeting of the American Society of Hematology, Blood 2017, abstract 1589.

38. Fava C, personal communication

39. Branford S, Yeung OT, Ross DM, Prime JA, Field CR, Altamura K. et al. Early molecular response and female sex strongly predict stable undetectable BCR-ABL1, the criteria for imatinib discontinuation in patients with chronic myeloid leukemia. Blood 2013;121(19):3818-3824.

40. Mori S, Vagge E, le Coutre P, Abruzzese E, Martino B, Pungolino E, et al. Age and dPCR can predict relapse in CML patients who discontinued imatinib: the ISAV study. Am J Hematol 2015;90:810-814.

41. Ravandi F, O'Brien S, Cortes JE, Thomas DM, Garris R, Faderl S, et al. Long-term follow-up of a phase 2 study of chemotherapy plus dasatinib for the initial treatment of patients with Philadelphia chromosome positive-acute lymphoblastic leukemia. Cancer 2015;121:4158-4164.

42. Yasukawa M, Ohminami H, Kojima K, Hato T, Hasegawa A, Takahashi T, et al. HLA class II-restricted antigen presentation of endogenous bcr-abl fusion protein by chronic myelogenous leukemia-derived dendritic cells to CD4+ T lymphocytes. Blood 2001;98(5):1498-1505.

43. Bocchia M, Gentili S, Abruzzese E, Fanelli A, Iuliano F, Tabilio A, et al. Effect of a p210 multipeptide vaccine associated with imatinib or interferon in patients with chronic myeloid leukaemia and persistent residual disease: a multicentric observational trial. Lancet 2005;365:657-662.

44. Rojas JM, Knight K, Wang L, and Clark RE. Clinical evaluation of BCR-ABL peptide immunisation in chronic myeloid leukaemia: results of the EPIC study. Leukemia 2007;21:2287-2295.

45. Westermann J, Kopp J, van Lessen A, Hecker A-C, Baskaynak G, le Coutre P, et al. Vaccination with autologous non-irradiated dendritic cells in patients with bcr/abl+ chronic myeloid leukaemia. Br J Haematol 2007;137:297-306.

46. Ravandi F, Cortes J, Albitar M, Arlinghaus R, Guo JK, Talpaz M, and Kantarjian HM. Chronic myelogenous leukaemia with p185 BCR/ABL expression: characteristics and clinical significance. Br J Haematol 1999;107:581-586.

47. Verma D, Kantarjian HM, Jones D, Luthra R, Borthakur G, Verstovsek S, et al. Chronic myeloid leukemia with P190 BCR-ABL: analysis of characteristics, outcomes, and prognostic significance. Blood 2009;114(11):2232-2235.

48. Pardanani A, Tefferi A, Litzow MR, Zent C, Hogan WJ, McClure R, Viswanatha D. Chronic myeloid leukemia with p190 BCR-ABL: prevalence, morphology, tyrosine kinase inhibitor response, and kinase domain mutation analysis. Blood 2009;114:3502-3503.

49. Jain P, Romo CG, Khoury HJ, Kantarjian H, Cortes J. Clinical activity of ponatinib in patients with chronic myeloid leukemia in chronic phase with e1a2 transcript. Haematologica 2013;8(11):e141-e142.

50. Lipton JH, Chuah C, Guerci-Bresler A, Rosti G, Simpson D, Assouline S, et al. Ponatinib versus imatinib for newly diagnosed chronic myeloid leukaemia: an international, randomized, open-label, phase 3 trial. Lancet Oncol 2016;17(5):612-621.

51. Qin Y-Z, Jiang Q, Jiang H, Lai Y-Y, Shi H-X, Chen W-M, et al. Prevalence and outcome of uncommon BCR-ABL1 transcripts in patients with CML: data from a single centre. Br J Haematol 2018;doi.org/10.1111/bjh.15453.

52. Li S, Ilaria RL, Million RP, Daley GQ, and Van Etten RA. The P190, P210 and P230 forms of the BCR/ABL oncogene induce a similar chronic myeloid leukemia-like syndrome in mice but have different leukemogenic activity. J Exp Med 1999;189:1399-1412.

53. Hai H, Kizilbash NA, Zaidi SHH, Alruwaili J, and Shahzad K. Differences in structural elements of Bcr-Abl oncoprotein isoforms in chronic myelogenous leukemia. Bioinformation 2014;10(3):108-114.

54. Reckel S, Hamelin R, Georgeon S, Armand F, Jolliet Q, Chiappe D, et al. Differential signaling networks of Bcr-Abl p210 and p190 kinases in leukemia cells defined by functional proteomics. Leukemia 2017;31:1502-1515.

55. Cutler J, Tahir R, Sreenivasamurthy SK, Mitchell C, Renuse S, and Nirujogi RS. Differential signaling through p190 and p210 BCR-ABL fusion proteins revealed by interactome and phosphoproteome analysis. Leukemia 2017;31:1513-1524.

56. Arana-Trejo RM, Sanchez ER, Ignacio-Barra G, De La Fuerte EB, Garces O, Morales EG, et al. BCR/ABL p210, p190 and p230 fusion genes in 250 mexican patients with chronic myeloid leukaemia. Int J Lab Hematol 2002;24(3):145-150.

57. Quackenbush RC, Reuther GW, Miller JP, Courtney KD, Pear WS, and Pendergast AM. Analysis of the biologic properties of p230 Bcr-Abl reveals unique and overlapping properties with the oncogenic p185 and p210 Bcr-Abl tyrosine kinases. Blood 2000;95(9):2913-2921.

58. Molica M, Zacheo I, Diverto D, Alimena G, and Breccia M. Long.term outcome of chronic myeloid leukaemia patients with p210 and p190 co-expression at baseline. Br J Haematol 2015;169:138-153.

59. Barilà D and Superti-Furga G. An intramolecular SH3-domain interaction regulates c-Abl activity. Nature Genetics 1998;18(3):280-282.

60. Smith KM, Yacobi R, and Van Etten RA. Autoinhibition of Bcr-Abl through its SH3 domain. Molecular Cell 2003;12:27.37

61. Sasaki K, Strom SS, O'Brien S, Jabbour E, Ravandi F, Konopleva M, et al. Relative survival in patients with chronic-phase chronic myeloid leukemia in the tyrosine-kinase inhibitor era: analysis of patients data from six prospective trials. Lancet Hematol 2015;2:e186-e193.

62. Bower H, Bjorkholm M, Dickman PW, Hoglund M, Lambert PC, and Andersson TM-L. Life expectancy of patients with chronic myeloid leukemia approaches the life expectancy of the general population. J Clin Oncol 2016;34(24):2851-2857.

FIGURE LEGEND

Figure 1:

Patients distribution by continent and by age. In Africa and in Asia, the age was shift to the left, while in Europe it was shift to the right. To account for this uneven distribution, all the calculations of the odd ratios have been adjusted for age.

Figure 2:

a) Proportion of e13a2 by age, in males and females, all continents. The differences among age groups, are found in both genders. Bars indicate 95% confidence intervals of the proportion.

b) Proportion of atypical, rare transcripts by age, in males and females, all continents. The differences among age groups**,** are found in both genders. Bars indicate 95% confidence intervals of the proportion.

Figure 3:

Distribution of the transcripts in the 666 patients with atypical, rare transcripts, overall, and by gender. The proportion of e1a2 transcript, coding for P190, was 47.3% of all transcripts, 58.6% in females and 36.5% in males. The proportion of other transcripts was as follows: e6a2 3.7% overall, 0.9% in females and 6.4% in males; e8a2 2.7% overall, 2.5% in females and 2.9% in males; e13a3 9.0% overall, 4.95% in females and 12.9% males; e14a3 9.5% overall, 6.5% in females and 12.3% in males; e19a2, coding for P230, 16.4% overall, 17.0 % in females and 15.8% in males. Overall, atypical, rare transcripts were more frequent in females than in males (ORs 1.352, 95% CI 1.159-1.576, $p< 0.0001$). NS = (transcript) Not Specified.

Table 1. Number and proportion of patients with e13a2, by continent, by gender, and by age. By gender, the proportion was 32.9% for males and 36.2% for females, the age- and continent-adjusted ORs of males vs females being 1.149 (95% CI 1.106-1.194, p < 0.0001)

Continent (gender)	0-19 years n (%)	20-39 years n (%)	40-59 years n (%)	60-79 years n (%)	>=80 years n (%)	Total (%)
Africa (female)	19 (47.5)	575 (47.4)	548 (43.6)	176 (40.7)	18 (38.3)	1336 (44.7)
Africa (male)	17 (42.5)	591 (46.3)	584 (43.6)	222 (43.3)	24 (42.86)	1438 (44.6)
Africa (total)	**36 (45.0)**	**1166 (46.9)**	**1132 (43.6)**	**398 (42.1)**	**42 (40.8)**	**2774 (44.6)**
Asia (female)	62 (31.3)	424 (32.3)	411 (29.7)	128 (28.3)	10 (31.3)	1035 (30.7)
Asia (male)	137 (35.0)	830 (34.8)	738 (35.3)	207 (32.6)	18 (39.1)	1930 (34.8)
Asia (total)	**199 (33.7)**	**1254 (33.9)**	**1149 (33.1)**	**335 (30.8)**	**28 (35.9)**	**2965 (33.2)**
Australia (female)	1 (33.3)	26 (36.6)	43 (36.4)	20 (30.3)	2 (100.0)	92 (35.4)
Australia (male)	6 (66.7)	35 (31.8)	47 (33.1)	36 (41.9)	0 (0.0)	124 (35.4)
Australia (total)	**7 (58.3)**	**61 (33.7)**	**90 (34.6)**	**56 (36.8)**	**2 (40)**	**216 (35.4)**
Europe (female)	89 (42.8)	683 (37.5)	1540 (36.8)	1285 (32.7)	260 (32.8)	3857 (35.3)
Europe (male)	121 (44.5)	1147 (42.2)	2272 (41.4)	1827 (37.3)	253 (28.6)	5620 (39.4)
Europe (total)	**210 (43.8)**	**1830 (40.3)**	**3812 (39.4)**	**3112 (35.2)**	**513 (30.6)**	**9477 (37.6)**
South America (female)	51 (48.6)	232 (42.7)	283 (35.9)	162 (35.0)	23 (31.9)	751 (38.1)
South America (male)	41 (38.0)	350 (42.6)	438 (41.8)	220 (39.8)	20 (39.2)	1069 (41.4)
South America (total)	**92 (43.2)**	**582 (42.6)**	**721 (39.3)**	**382 (37.6)**	**43 (35.0)**	**1820 (40.0)**
Females (total)	**222 (40.1)**	**1940 (39.1)**	**2825 (36.6)**	**1771 (33.2)**	**313 (33.1)**	7071 (36.2)
Males (total)	**322 (39.2)**	**2953 (40.4)**	**4079 (40.3)**	**2512 (37.5)**	**315 (30.3)**	10181 (39.2)
Grand Total	**544 (39.6)**	**4893 (39.9)**	**6904 (38.7)**	**4283 (35.6)**	**628 (31.6)**	**17252 (37.9)**

TABLE 2: Proportion of e13a2 patients, and adjusted odd ratios (ORs) (95% confidence intervals) of pairwise comparisons between age groups. Adjusted ORs were estimated from multiple logistic regression with continent, age group and gender as common explanatory variables. The differences among age groups are all significant, with the exception of the first two age groups (0-19 and 20-39 years), where the proportion of e13a2 was almost identical (39.6% and 39.9% respectively).

AGE GROUP	0-19 years	20-39 years	40-59 years	60-79 years	≥ 80 years
Proportion with e13a2 transcript	**39.6**	**39.9**	**38.7**	**35.6**	**31.6**
ODD RATIOS (95% confidence intervals)					
0-19 years vs all other age groups	///	0.948 (0.845-1.064)	0.893 (0.797-1.000)	0.775 (0.690-0.871)	0.647 (0.560-0.749)
		p = 0.3640	p = 0.0505	p < 0001	p < 0.0001
20-39 years vs 40-59, 60-79, and ≥ 80	///	///	0.942 (0.898-0.989)	0.817 (0.774-0.863)	0.683 (0.616-0.737)
		///	p = 0.0140	p < 0.0001	p < 0.0001
40-59 years vs 60-79 and ≥ 80	///	///	///	0.868 (0.827-0.911)	0.725(0.656-0.801)
		///	///	p < 0.0001	p < 0.0001
60-79 years vs ≥ 80	///	///	///	///	0.835 (0754-0.925)
		///	///	///	p < 0.0005

Table 3. Proportions of e13a2 transcript and of atypical, rare transcripts in children (0-9 years) and in adolescents (10-19 years). Note that the Childhood Registry reported only on patients who had been referred to Children Hospitals, while all the other sources reported mainly on adults, and on children data only occasionally.

Source	Age (years)	Females			Males			Total		
		e13a2	**Total**	**%**	**e13a2**	**Total**	**%**	**e13a2**	**Total**	**%**
Childhood Registry	0-9	13	40	32.5	25	51	49.0	38	91	41.7
Other sources	0-9	30	89	33.7	37	100	37.0	67	189	35.4
Total	**0-9**	**43**	**129**	**33.3**	**62**	**151**	**41.0**	**105**	**280**	**37.5**
Childhood Registry	10-19	34	89	38.2	36	131	27.5	70	220	31.8
Other sources	10-19	192	469	40.9	281	717	39.2	473	1189	39.9
Total	**10-19**	**226**	**558**	**40.5**	**313**	**848**	**37.4**	**543**	**1406**	**38.6**

Source	Age (years)	Females			Males			Total		
		Rare	**Total**	**%**	**Rare**	**Total**	**%**	**Rare**	**Total**	**%**
Childhood Registry	0-9	0	40	NC	2	51	3.92	2	91	2.20
Other sources	0-9	4	84	4.76	2	71	2.82	6	155	3.87
Total	**0-9**	**4**	**124**	**3.22**	**4**	**122**	**3.28**	**8**	**246**	**3.25**
Childhood Registry	10-19	2	89	2.25	6	131	4.58	8	220	3.64
Other sources	10-19	5	366	1.37	9	593	1.52	14	959	1.46
Total	**10-19**	**7**	**455**	**1.54**	**15**	**724**	**2.07**	**22**	**1179**	**1.86**

Table 4. List of atypical, rare transcripts. Numbers and percentages, overall and by gender. Percentages are calculated on all patients. Other transcripts were reported as e1a1, e8a1, e8a3, e15a2 and e23a1, that are not all clear exon/exon fusions. Some of them had additional nucleotides inserted in the chimeric mRNA that may or may not have maintained the reading frame.

Transcript type	Total (n = 34561)		Females (n = 14291)		Males (n = 20270)	
	n	**%**	**n**	**%**	**n**	**%**
e1a2	315	0.91	190	1.33	125	0.62
e1a3	6	0.02	2	0.01	4	0.02
e6a2	25	0.07	3	0.02	22	0.11
e8a2	18	0.05	8	0.06	10	0.05
e13a3	60	0.17	16	0.11	44	0.22
e14a3	64	0.18	21	0.15	43	0.21
e19a2	109	0.31	55	0.38	54	0.27
Other	6	0.02	3	0.02	3	0.01
Non-specified	63	0.18	26	0.18	37	0.18
Total	**666**	**1.93**	**324**	**2.27**	**342**	**1.69**

Table 5. Proportion of patients with atypical, rare transcripts and adjusted odd ratio (ORs) (95% confidence intervals) of pairwise comparisons between age groups. Adjusted ORs were estimated from multiple logistic regression with age group and gender as common explanatory variables. As almost only European and Asian data were available, the variable "continent" was not considered. The proportion was fairly constant in the first three age groups, until the age of 60, but increased significantly in the elderly.

AGE GROUP	0-19 years	20-39 years	40-59 years	60-79 years	≥ 80 years
Proportion with rare transcripts	**1.79**	**1.32**	**1.58**	**2.63**	**3.84**
ODD RATIOS (95% confidence intervals)					
0-19 years vs all other age groups	///	0.740 (0.458-1.196)	0.876 (0.551-1.391)	1.466 (0.926-2.319)	2.136 (1.278-3.569)
		p = 0.2188	p = 0.5734	p < 1027	p < 0.0038
20-39 years vs 40-59, 60-79, and ≥ 80	///	///	1.183 (0.940-1.488)	1.980 (1.585-2.474)	2.886 (2.096-3.975)
		///	p = 0. 1521	p < 0.0001	p < 0.0001
40-59 years vs 60-79 and ≥ 80	///	///	///	1.674 (1.394-2.010)	2.440(1.819-3.273)
		///	///	p < 0.0001	p < 0.0001
60-79 years vs ≥ 80		///	///	///	1.457 (1.093-1.944)
	///	///	///	///	p < 0.0103

Table 1S. Reported frequency of the M-*bcr* transcript types and of other, atypical, transcript types, from clinical studies, from international, national or regional registries, and from single centers. Children data are listed separately.

NR = Not Reported

Ref	Source	Patients n	e13a2 %	e14a2 %	e14a2/ e13a2 %	Other %
4	Study, Germany	1454	40.4	44.8	14.8	NR
5	Study, Italy	146	42.5	57.5	NR	NR
6	Study, Italy	452	35.4	62.8	NR	1.8
7	Study, Italy	559	36.3	51.9	10.7	1.1
8	Study, Italy	201	40.3	59.7	NR	NR
9	Study, Spain	194	37.1	49.0	12.9	1.0
10	Study, UK	119	34.4	60.5	5.0	NR
14	Studies, Europe	1494	38.0	49.0	13.0	NR
11	Registry, Italy	320	40.9	47.5	10.0	1.6
12	Registry, Italy	337	43.0	56.1	NR	0.9
13	Registry, Spain	243	49.4	42.4	NR	8.2
15	Registry, Europe	1533	38.9	56.6	12.8	4.5
16	Registry, Korea S.	363	33.9	63.6	1.1	1.4
17	Single center UK	78	41.0	50.0	3.8	5.1
18	Single center, UK	204	45.6	42.1	11.3	1.0
19	Single center, UK	87	37.9	46.0	16.1	NR
20	Single center, USA	481	41.6	40.7	17.7	NR
21	Single center, Brasil	170	32.9	55.3	11.8	1.1
22	Single center, Ecuador	40	94.6	5.4	NR	NR
23	Single center, India	112	29.5	61.6	5.3	3.6
24	Single center, India	87	39.0	53.0	8.0	NR
25	Single center, India	26	34.3	60.0	NR	1.8
26	Single center, Iran	75	20.0	62.0	5.0	13.0
27	Single center, Sudan	43	53.5	41.9	4.6	NR
28	Children, Germany	146	37.7	36.3	26.9	NR
29	Children, Italy	47	27.6	70.2	2 .1	NR
30	Children, India	47	68.1	31.9	NR	NR
31	Children, Int.registry	100	40.0	51.0	7.0	2.0

Table 2S. Total patient numbers and distribution of all 45503 patients with data on M-bcr transcripts.

Continent (gender)	0-19 years n (%)	20-39 years n (%)	40-59 years n (%)	60-79 years n (%)	>=80 years n (%)	Total (%)
Africa (female)	40 (1.3)	1212 (40.6)	1256 (42.0)	433 (14.5)	47 (1.6)	2988 (48.1)
Africa (male)	40 (1.2)	1277 (39.6)	1341 (41.6)	513 (15.9)	56 (1.7)	3227 (51.9)
Africa (total)	**80 (1.3)**	**2489 (40.1)**	**2597 (41.8)**	**946 (15.2)**	**103 (1.7)**	**6215 (13.7)**
Asia (female)	198 (5.9)	1312 (38.9)	1382 (40.9)	452 (13.4)	32 (1.0)	3376 (37.8)
Asia (male)	392 (7.1)	2384 (43.0)	2092 (37.7)	635 (11.4)	46 (0.8)	5549 (62.2)
Asia (total)	**590 (6.6)**	**3696 (41.4)**	**3474 (38.9)**	**1087 (12.2)**	**78 (0.9)**	**8925 (19.6)**
Australia (female)	3 (1.2)	71 (27.3)	118 (45.4)	66 (25.4)	2 (0.8)	260 (42.6)
Australia (male)	9 (2.6)	110 (31.4)	142 (40.6)	86 (24.6)	3 (0.9)	350 (57.4)
Australia (total)	**12 (2.0)**	**181 (29.7)**	**260 (42.6)**	**152 (24.9)**	**5 (0.8)**	**610 (1.3)**
Europe (female)	208 (1.9)	1820 (16.7)	4181 (38.7)	3928 (35.9)	792 (7.3)	10929 (43.4)
Europe (male)	272 (1.9)	2716 (19.0)	5494 (38.5)	4904 (34.4)	884 (6.2)	14270 (56.6)
Europe (total)	**480 (1.9)**	**4536 (18.0)**	**9675 (38.4)**	**8832 (35.1)**	**1676 (6.7)**	**25199 (55.4)**
South America (female)	105 (5.3)	544 (27.6)	788 (40.0)	463 (23.5)	72 (3.7)	1972 (43.3)
South America (male)	108 (4.2)	821 (31.8)	1049 (40.6)	553 (21.4)	51 (2.0)	2582 (56.7)
South America (total)	**213 (4.7)**	**1365 (30.0)**	**1837 (40.3)**	**1016 (22.3)**	**123 (2.7)**	**4554 (10.0)**
Females (total)	**554 (2.8)**	**4959 (25.4)**	**7725 (39.6)**	**5342 (27.4)**	**945 (4.8)**	19525 (42.9)
Males (total)	**821 (3.2)**	**7308 (28.1)**	**10118 (39.0)**	**6691 (25.8)**	**1040 (4.0)**	25978 (57.1)
Grand Total	**1375 (3.0)**	**12267 (27.0)**	**17843 (39.2)**	**12033 (26.4)**	**1985 (4.4)**	**45503 (100)**

Table 3S. Continent distribution, and male to female ratios (M:F), within each age group.

Continents	0-19		20-39		40-59		60-79		≥ 80		Total	
	%	M:F	%	M:F	%	M:F	%	M:F	%	M:F	%	M:F
Africa	5.8	1.00	20.3	1.05	14.5	1.07	7.9	1.18	5.2	1.19	13.7	1.08
Asia.	42.9	1.98	30.1	1.82	19.5	1.51	9.0	1.40	3.9	1.44	19.6	1.64
Australia	0.9	3.00	1.5	1.55	1.5	1.20	1.3	1.30	0.3	1.50	1.3	1.35
Europe	34.9	1.31	37.0	1.49	54.2	1.31	73.4	1.25	84.4	1.10	55.4	1.31
South America	15.5	1.03	11.1	1.51	10.3	1.33	8.4	1.19	6.2	0.71	10.0	1.31
Total	**100.0**	**1.48**	**100.0**	**1.47**	**100.0**	**1.31**	**100.0**	**1.25**	**100.0**	**1.10**	**100.0**	**1.33**

Table 4S. Proportion of the patients who were reported to co-express both M-*bcr* transcript types (e14a2 and e13a2). Notice that in all the analyses, all patients co-expressing both transcripts were pooled and counted together with patients expressing e14a2 only.

Continents	Total (n)	Co-expressing e14a2 and e13a2	
		n	**%**
Africa	6215	405	6.5
Asia	8925	212	2.4
Australia	610	114	18.7
Europe	25199	2336	9.3
South America	4554	407	8.9
Total	**45503**	**3474**	**7.6**
Childhood Registry	311	35	11.2

TABLE 5S: Proportion of e13a2 patients, and adjusted odd ratios (ORs) (95% confidence intervals) of pairwise comparisons between continents. Adjusted ORs were estimated from multiple logistic regression with continent, age group and gender as common explanatory variables. The proportion was lowest for Asia and highest for Africa.

CONTINENT	South Am	Asia	Australia	Europe	Africa
Proportion with e13a2 transcript	**40.0**	**33.2**	**35.4**	**37.6**	**44.6**
ODD RATIOS (95% confidence intervals)					
South Am. vs all other continents	///	0.720 (0.668-0.776)	0.823 (0.690-0.982)	0.943 (0.83-1.006)	1.201 (1.111-1.299)
		p < 0.0001	p = 0.0308	p =0.0773	p < 0.0001
Asia vs Australia, Europe, and Africa	///	///	1.143 (0.963-1.358)	1.310 (1.242-1.381)	1.668 (1.560-1.784)
		///	p = 0. 1271	p < 0.0001	p < 0.0001
Australia vs Europe and Africa	///	///	///	1.145 (0.968-1.355)	1.459(1.227-1.736)
		///	///	p < 0. 1140	p < 0.0001
Europe vs Africa	///	///	///	///	1.274 (1.203-1.349)
		///	///	///	p < 0.0001

Figure 1

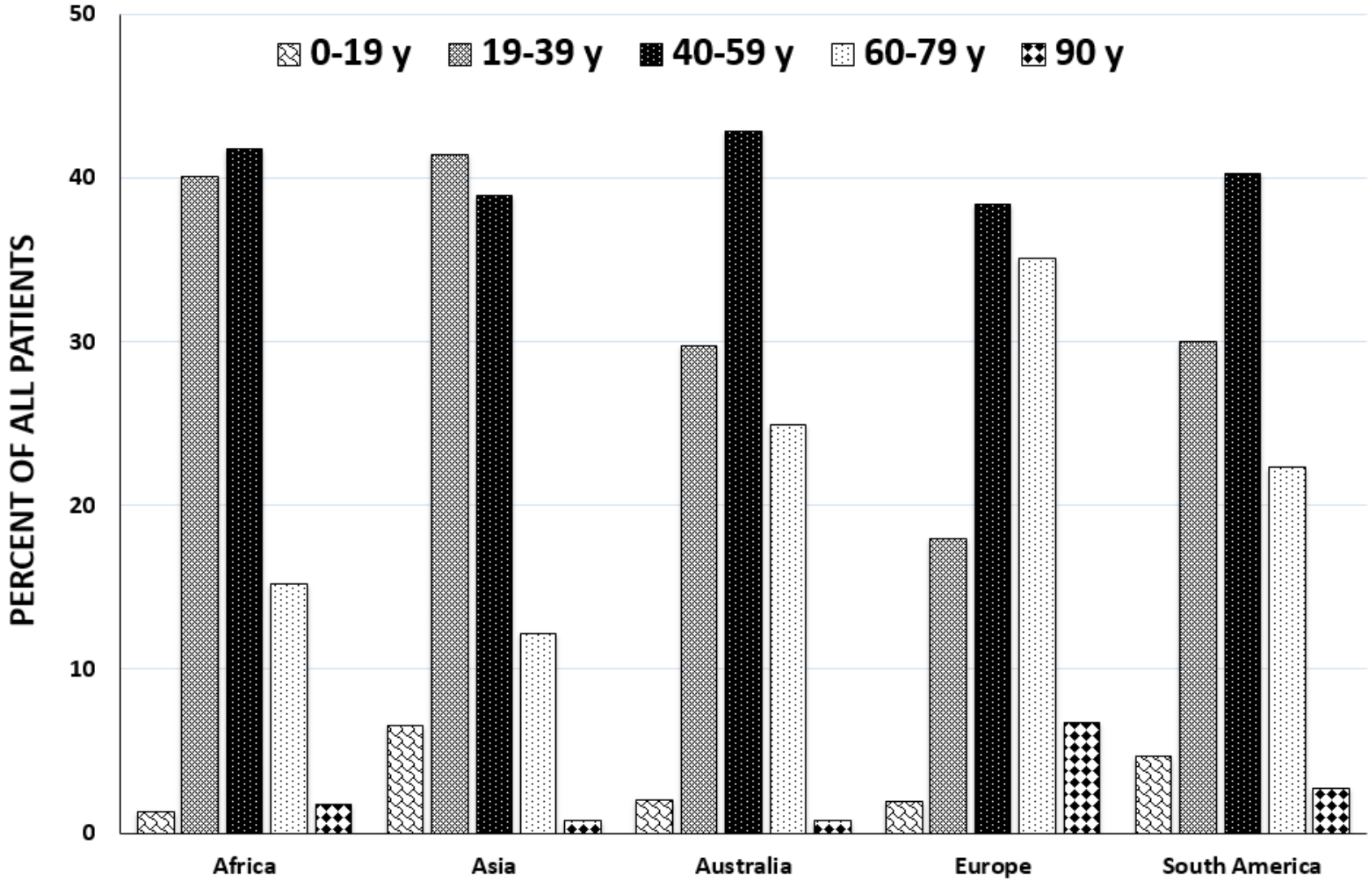

0-19 y
19-39 y
40-59 y
60-79 y
90 y
PERCENT OF ALL PATIENTS
50
40
30
20
10
0
Africa
Asia
Australia
Europe
South America

Figure 2a

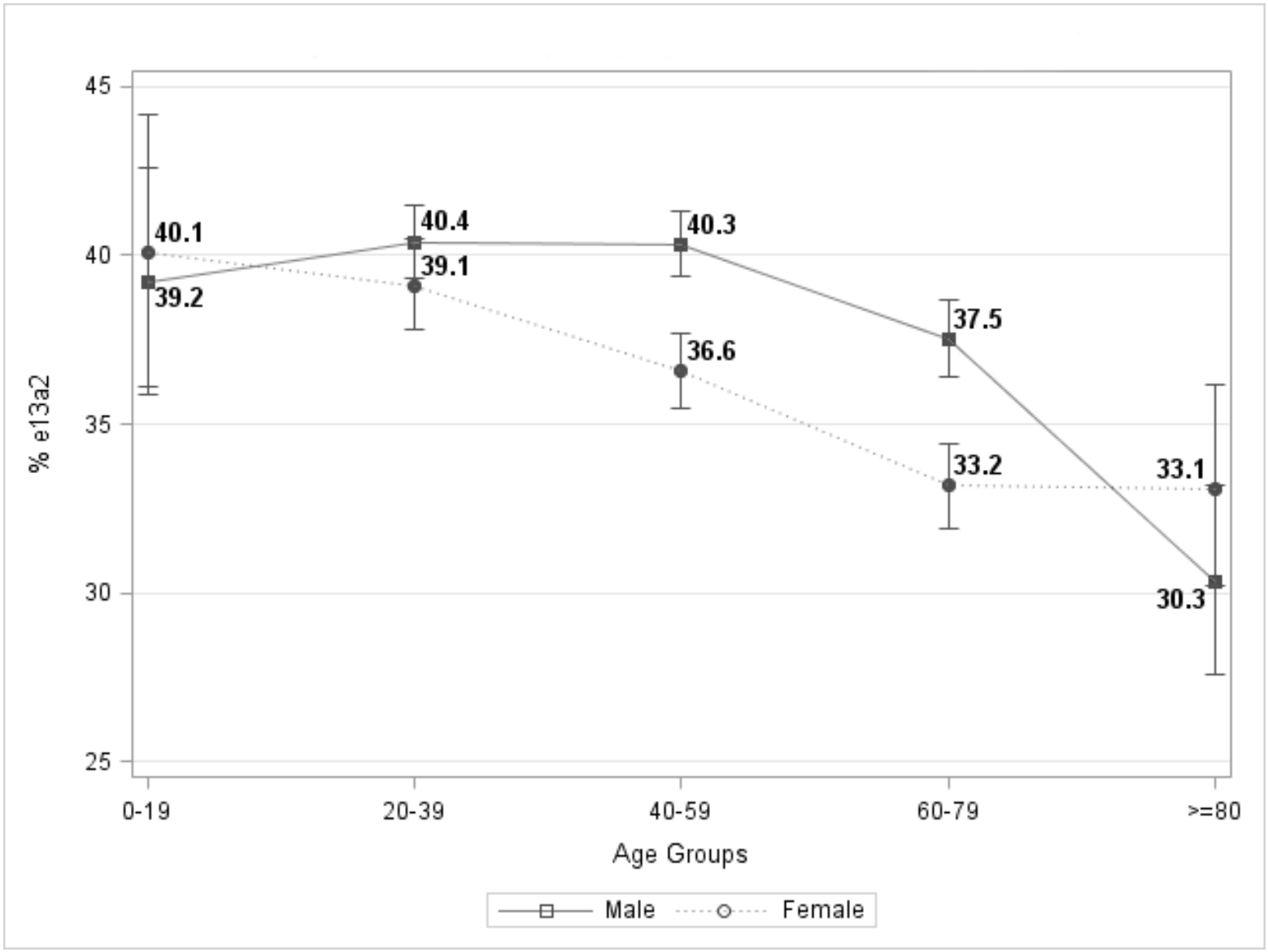

45
40
35
30
25
% e13a2
40.1
39.2
40.4
39.1
40.3
36.6
37.5
33.2
33.1
30.3
0-19
20-39
40-59
60-79
>=80
Age Groups
Male
Female

Figure 2b

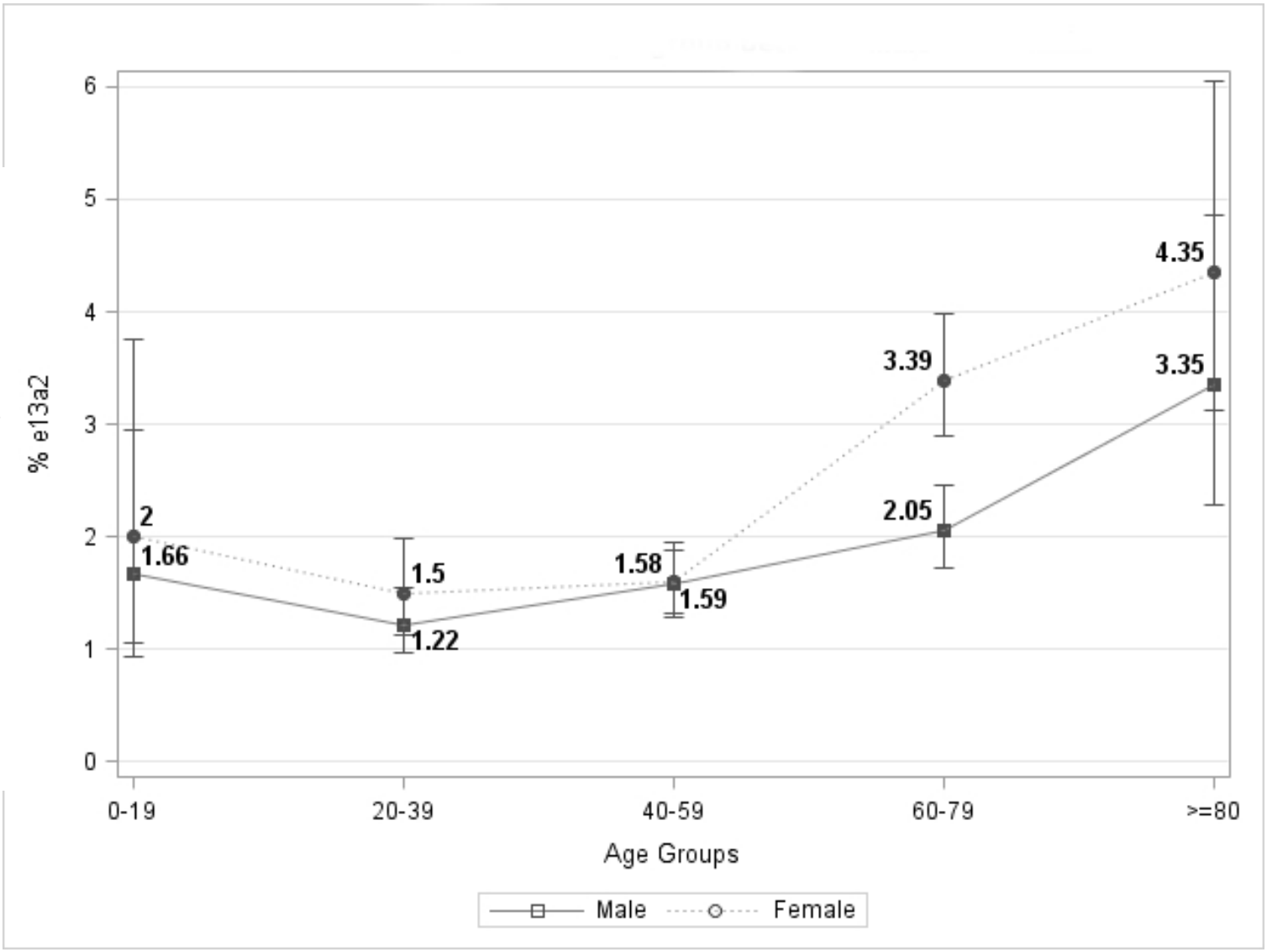
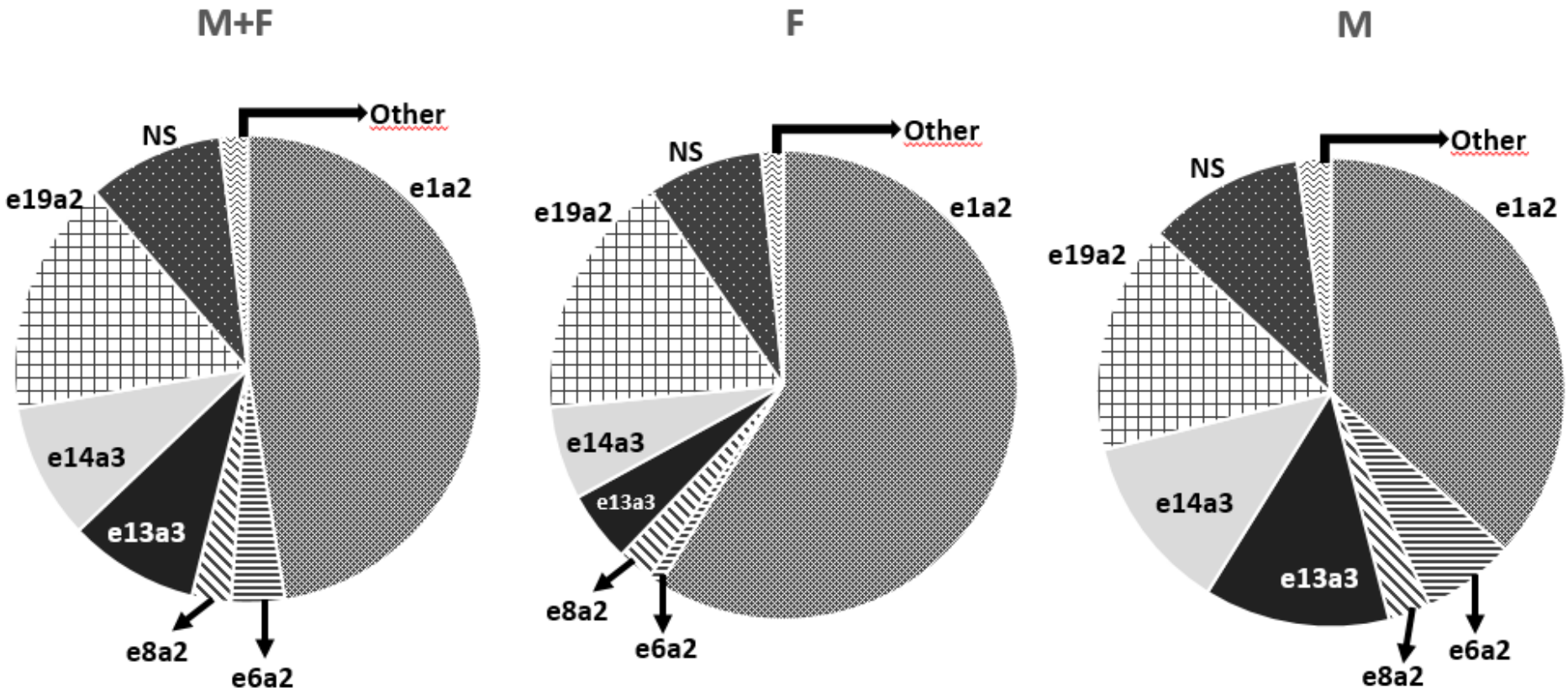

PERCENT TRANSCRIPTS
% e13a2
6
5
4
3
2
1
0
2
1.66
1.5
1.22
1.58
1.59
3.39
2.05
4.35
3.35
0-19
20-39
40-59
60-79
>=80
Age Groups
Male
Female

Figure 3


M+F
Other
NS
e1a2
e19a2
e14a3
e13a3
e8a2
e6a2
F
Other
NS
e1a2
e19a2
e14a3
e13a3
e8a2
e6a2
M
Other
NS
e1a2
e19a2
e14a3
e13a3
e6a2
e8a2