Has the time for first-line treatment with second generation tyrosine kinase inhibitors in patients with chronic myelogenous leukemia already come? Systematic review and meta-analysis

Ronit Gurion,[1,4]* Anat Gafter-Gvili,[1,4]* Liat Vidal,[1,4] Avi Leader,[2] Ron Ram,[1,4] Adi Shacham-Abulafia,[1,4] Mical Paul,[3,4] Isaac Ben-Bassat,[4] Ofer Shpilberg,[1,4] and Pia Raanani[1,4]

[1]Institute of Hematology, Davidoff Cancer Center, Rabin Medical Center; [2]Internal Medicine Department A, Meir Hospital; [3]Infectious Disease Unit, Rabin Medical Center; and [4]Sackler School of Medicine, Tel Aviv University, Israel

 doi:10.3324/haematol.2012.063172

ONLINE SUPPLEMENTARY APPENDIX

Search terms used for the systematic review

The following search terms were used: ("Imatinib"[Mesh] OR "Gleevec"[Mesh] OR "TKI" [Mesh] OR "tyrosine kinase inhibitor" [Mesh]) AND ("Nilotinib"[Mesh] OR "Dasatinib"[Mesh] OR "Sprycel" OR "Tasigna" OR "STI-571" OR "Bosutinib"[Mesh] OR "SKI-606" [Mesh] OR "TKI" [Mesh] OR "tyrosine kinase inhibitor" [Mesh] OR "Ponatinib" [Mesh]) AND ("Leukemia, Myeloid, Chronic"[Mesh] OR "Leukemia, Myelogenous, Chronic"[Mesh] OR CML OR chronic-myeloid-leukemia).

Sokal and Hasford (EURO) prognostic scoring systems for newly diagnosed chronic myeloid leukemia

Various scoring systems have been developed in an attempt to predict disease outcome in CML patients.

(i) The Sokal prognostic score identified four clinical variables: spleen size, percentage of blasts, age, and platelet count over 700x10^9/L. The first three are continuous variables with progressively worse prognosis at higher values. A Sokal risk score calculator is available online at: *http://www.leukemia-net.org/content/leukemias/cml/cml_score/*[1]

(ii) The Hasford or Euro score adds eosinophilia and basophilia to the four clinical variables of the Sokal score. It was developed for CML patients receiving treatment with interferon. A Hasford score calculator is available online at: *http://www.leukemia-net.org/content/leukemias/cml/cml_score/*[2,3]

Statistical terms:[4] fixed effect method

A meta-analysis yields an overall statistic that summarizes the effectiveness of the experimental intervention compared with a control intervention. A pooled intervention effect estimate is calculated as a weighted average of the intervention effects assessed in the individual studies. There are four commonly used methods of meta-analysis for dichotomous outcomes: three fixed-effect methods (Mantel-Haenszel, Peto and inverse variance) and one random-effects method (DerSimonian and Laird). The fixed effect method presumes that each study is estimating the same intervention effect. The Mantel-Haenszel method[5,6] is the most common and is the prefered (the default) fixed-effect method of meta-analysis. Also, this method has been shown to have better statistical properties when there are few events, as in our case. The results are given as relative risks (RRs) and confidence interval. With *positive* outcomes (such as response and survival), the forest plot shows the investigational arm on the right-hand side of the plot. Thus, with positive results, RRs above 1 indicate that 2nd generation TKIs are better. With *negative* outcomes, (for example, mortality and disease progression) the investigational arm is on the left-hand side of the forest plot. Thus, RRs less than 1 indicate that 2nd generation TKIs are better.

χ^2- a statistical measure to assess heterogeneity

It is important to examine heterogeneity in a meta-analysis. A χ^2 test evaluates whether observed differences in results are compatible with chance alone. A low *P* value (or a large χ^2 statistic relative to its degree of freedom) provides evidence of statistical heterogeneity of intervention effects (variation in effect estimates beyond chance). The χ^2 test has low power in meta-analysis when there is a small sample size or few studies, like in the present study. Also, some claim that there will always be heterogeneity in meta-analysis, whether this is detected or not. Methods have been developed to quantify inconsistency across studies. This moves the focus away from testing whether heterogeneity is present to assessing its impact on the meta-analysis. A useful statistic for quantifying inconsistency is I.[2,7,8] This description of the percentage of the variability, in effect, estimates that it is due to heterogeneity rather than sampling error (chance). A rough guide for interpretation would be:[4]

- 0-40%, might not be important;
- 30-60%. may represent moderate heterogeneity;
- 50-90%, may represent substantial heterogeneity;
- 75-100%, considerable heterogeneity.

References

1. Sokal JE, Cox EB, Baccarani M, Tura S, Gomez GA, Robertson JE, et al. Prognostic discrimination in "good-risk" chronic granulocytic leukemia. Blood 1984;63(4):789-99.
2. Hasford J, Pfirrmann M, Hehlmann R, Allan NC, Baccarani M, Kluin-Nelemans JC, et al. A new prognostic score for survival of patients with chronic myeloid leukemia treated with interferon α. J Nat Cancer Inst 1998; 90(11):850-8.
3. Hasford J, Baccarani M, Hoffmann V, Guilhot J, Saussele S, Rosti G, et al. Predicting complete cytogenetic response and subsequent progression-free survival in 2060 patients with CML on imatinib treatment: the EUTOS score. Blood 2011;118(3):686-92.
4. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011] The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.
5. Greenland S, Robins JM. Estimation of a common effect parameter from sparse follow-up data. Biometrics 1985;41(1):55-68.
6. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. J Natl Cancer Inst 1959;22(4):719-48.
7. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002; 21(11):1539-58.
8. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327 (7414):557-60.