PERSPECTIVE OPEN

Check for updates

ACUTE LYMPHOBLASTIC LEUKEMIA

Have we been qualifying measurable residual disease correctly?

Yahui Feng[1,2,4], Saibing Qi[1,2,4], Xueou Liu[1,2,4], Li Zhang[1,2,4], Yu Hu[1,2], Qiujin Shen[1,2], Xiaowen Gong[1,2], Wei Zhang[1,2], Junxia Wang[1,2], Wen Yan[1,2], Tiantian Wang[1,2], Huijun Wang[1,2], Zhen Song[1,2], Xiaofan Zhu[1,2✉], Robert Peter Gale[3] and Junren Chen[1,2✉]

Leukemia (2023) 37:2168–2172; https://doi.org/10.1038/s41375-023-02026-4

Someone told me that each equation I included in the book would halve the sales. I therefore resolved not to have any equations at all. In the end, however, I did put in one equation, Einstein's famous equation, E = m c squared. I hope that this will not scare off half of my potential readers.

Stephen Hawking

INTRODUCTION

There is considerable interest in tests quantifying remaining leukaemia cells after therapy, termed measurable residual disease (MRD)-tests, to predict therapy outcomes, leukaemia recurrence and consider potential subsequent interventions [1–10]. Many studies reported a negative MRD-test during or after completing anti-leukaemia therapy independently identifies persons with a low risk of leukaemia relapse compared with those with a positive MRD-test after adjusting for other predictive and prognostic co-variates [5, 11–16]. Other studies recommend specific interventions in someone with a positive MRD-test such as a haematopoietic cell transplant or immune therapy such as chimaeric antigen receptor (CAR)-T-cells. Whether such interventions reduce leukaemia relapse risk in someone with a positive MRD-test can only be proved in a randomized controlled trial [8, 17].

Most MRD-tests focus on detecting a leukaemia-related or -specific immune phenotype, cytogenetic and/or molecular abnormality [1, 2, 18–25]. A perfect MRD-test would precisely quantify only leukaemia cells biologically capable of causing leukaemia relapse and likely to do so within a defined interval after accounting for competing causes of therapy-failure [7, 8]. Routine clinical use of MRD-testing requires refinements and standardization/harmonization of assay platforms and result reporting [1, 2, 21–23].

There is consensus a flow cytometry-based MRD-test should be reproducible at a limit of detection (LoD) of ≤0.01% leukaemia cells in a blood or bone marrow sample [26]. Based on this reasoning it is proposed a multi-parameter flow cytometry (MPFC)-based MRD-test should only be declared positive if ≥ 5 × 10E+5 cells are analysed and if ≥20 or ≥50 cells are positive [27–30]. However, this definition is often unmet in clinical practice. For example, modern MRD-directed, risk-stratified approach to treating childhood acute lymphoblastic leukaemia (ALL) requires an MPFC-based MRD-test done in bone marrow aspirate 2–3 weeks after starting induction chemotherapy, a time when collecting > 5 × 10E+5 bone marrow mononuclear cells is difficult [31, 32]. The same limitation operates in adults receiving intensive induction chemotherapy. How should a physician use results of MRD-testing in these settings?

TYRANNY OF SAMPLING ERROR

Assume in an MPFC-based MRD-test N cells are analysed out of which n cells are identified as *leukaemia cells*. By *leukaemia cells* we mean cells with immune phenotype of the leukaemia, not necessarily cells able to cause relapse within a defined interval. The conventional way to estimate MRD is $\mathrm{MRD}_{\mathrm{conventional}} = \frac{n}{N}$ [33, 34].

When the true proportion of leukaemia cells ("true MRD") is $< \frac{1}{N}$, the standard error of $\mathrm{MRD}_{\mathrm{conventional}}$ has a magnitude even larger than true MRD because of sampling error (Supplementary Methods). Simply put, the $\mathrm{MRD}_{\mathrm{conventional}}$ test can be very imprecise.

To better appreciate the tyranny of sampling error consider the hypothetical example of a haematologist reviewing the following MRD-test result: $N = 50000$ and $n = 0$. Analysing these few cells is not uncommon in practice for reasons we discussed above. Using the conventional approach to quantifying MRD the haematologist interprets this MRD-test as $\mathrm{MRD}_{\mathrm{conventional}} = \frac{0}{50000} = 0\%$. In doing so the haematologist fails to appreciate the result of this MRD-test is compatible with a broad range of true MRD values. In reality, the haematologist can only conclude MRD-test result is $\leq 0.006\%$ with a 5-percent probability true MRD is actually $>0.006\%$.

[1]State Key Laboratory of Experimental Hematology, National Clinical Research Center for Blood Diseases, Haihe Laboratory of Cell Ecosystem, Institute of Hematology & Blood Diseases Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Tianjin, China. [2]Tianjin Institutes of Health Science, Tianjin, China. [3]Centre for Haematology, Department of Immunology and Inflammation, Imperial College of Science, Technology and Medicine, London, UK. [4]These authors contributed equally: Yahui Feng, Saibing Qi, Xueou Liu, Li Zhang. ✉email: xfzhu@ihcams.ac.cn; chenjunren@ihcams.ac.cn

Received: 21 July 2023 Revised: 30 August 2023 Accepted: 5 September 2023
Published online: 13 September 2023

Using Bayesian reasoning, the worst-case (probability <0.05)[1] scenario estimate of MRD, which we denote as MRD_{worst_case}, can be computed using a beta distribution (the formula is "BETA.INV (0.95, $1+n$, $1+N-n$)" in Microsoft Excel; Supplementary Methods) [35].

Table 1 displays the extent to which $MRD_{conventional}$ under-estimates true MRD at different values of N in the worst-case scenario (that is, by how much $MRD_{conventional}$ under-estimates MRD_{worst_case}). Note that when $MRD_{conventional}$ is $\leq 0.01\%$ MRD_{worst_case} is considerably larger than $MRD_{conventional}$ across a broad range of N values. Conversely, when $MRD_{conventional}$ is $\geq 0.1\%$ MRD_{worst_case} is usually very close to $MRD_{conventional}$ unless the number of analysed cells N is < 10E+5.

Typically result of an MRD-test is interpreted as positive or negative based on applying a cut-off threshold to $MRD_{conventional}$. Our analysis of the adverse impact of sampling error (Table 1) suggests any cut-off threshold <0.01% used in $MRD_{conventional}$ would yield unreliable results with many false-negatives. Moreover, when estimating the hazard function of $MRD_{conventional}$ for leukaemia relapse risk false-negatives would cause "flattening" of the estimated curve because the contrast between MRD-positives and -negatives is attenuated by contamination of false-negative MRD-test results.

BORROWING LESSONS FROM DECISION SCIENCE

How to solve this problem when an inaccurate false-negative test result could have adverse clinical consequences? We propose the haematologist should instead rely on MRD_{worst_case} rather than $MRD_{conventional}$ to estimate relapse risk.

Our reasoning follows. When interpreting an MRD-test result to predict relapse the haematologist is essentially playing a "chess game against *nature*". It's his/her 1st move to make, declaring the MRD-test result positive or negative. In response the opponent (*nature*) has two possible moves, causing relapse or not. When MRD_{worst_case} is larger the haematologist is more likely to later regret if he/she declares the MRD-test result negative, because more plausibly *nature* would *play tricks* on the haematologist by causing relapse.

Ranking of people's test results based on MRD_{worst_case} from high to low values minimises the sum of regrets in the worst-case scenario because people whose MRD-test results are more likely to cause regret in case of a negative interpretation are already considered to have a higher risk of relapse. In the language of decision science, MRD_{worst_case} is a *minimax regret* approach to quantifying MRD test results according to Leonard Savage's theory of statistical decision or Herbert Simon's theory of rational choice under uncertainty [36, 37].

A CLINICAL EXAMPLE

To illustrate using MRD_{worst_case} to interpret test results we interrogated data from 883 consecutive children with ALL <16 years (Supplementary Fig. 1; Supplementary Table 1; and Supplementary Methods). The subjects were treated on the Chinese Children's Cancer Group study ALL-2015 (CCCG-ALL-2015) protocol [32]. 618 (70%) and 265 (30%) of the children were low- and intermediate-risk at diagnosis according to the CCCG-ALL-2015 criteria. MPFC-based MRD-testing was done on bone marrow samples 19 days after starting therapy. Median number of analysed cells (N) was $4\times10E+5$ (Interquartile Range [IQR], $2.4–5.0\times10E+5$; Range, $3.4\times10E+3$ to $1.0\times10E+6$). 686 (78%) MRD-tests analysed $<5\times10E+5$ cells, a threshold stipulated by guideline for good laboratory practice (GLP) [27, 28, 30].

Table 1. To what extent $MRD_{conventional}$ under-estimates true MRD at different numbers of analysed cells N in the worst-case scenario.

		$MRD_{conventional}$				
		10%	**1%**	**0.1%**	**0.01%**	**0.002%**
N	50000	−2%	−7%	−21%	−52%	−79%
	100000	−2%	−5%	−15%	−41%	−68%
	200000	−1%	−4%	−11%	−31%	−56%
	300000	−1%	−3%	−9%	−26%	−49%
	400000	−1%	−3%	−8%	−23%	−45%
	500000	−1%	−2%	−7%	−21%	−41%
	600000	−1%	−2%	−7%	−19%	−38%
	700000	−1%	−2%	−6%	−18%	−36%
	800000	−1%	−2%	−6%	−17%	−34%
	900000	−1%	−2%	−5%	−16%	−33%
	1000000	0%	−2%	−5%	−15%	−31%

294 (33%) children had $MRD_{conventional}<0.01\%$ on day 19, 274 (93%) of whom had zero values (*i.e.* no leukaemia cell was detected [$n=0$]). The remainder (20 [7%]) had 8–24 leukaemia cells detected. Because most children with $MRD_{conventional}<0.01\%$ had no leukaemia cells detected in the sample, $MRD_{conventional}$ could not identify relative relapse risk in these children. The C-statistic (the probability of pairwise agreement with relapse time [38]) of MRD_{worst_case} (0.57) was significantly higher (P <0.001; 2-sided Wilcoxon test on 500 bootstrap samples [39]) compared with C-statistic of $MRD_{conventional}$ (0.50). In short, MRD_{worst_case} was a better predictor of relapse than $MRD_{conventional}$ when $MRD_{conventional}$ was close to zero (Fig. 1A). In contrast, for the 589 (67%) children who had $MRD_{conventional}\geq 0.01\%$ on day 19, C-statistics of MRD_{worst_case} (0.58) and $MRD_{conventional}$ (0.58) were similar ($P=0.61$).

We estimated non-linear hazard functions of $MRD_{conventional}$ and MRD_{worst_case} for relapse by fitting restricted cubic spline curves using Markov chain Monte Carlo [40–43]. Since MRD_{worst_case} is always larger than $MRD_{conventional}$, all else being equal, switching from $MRD_{conventional}$ to MRD_{worst_case} should induce a *right-shift* of the hazard function curve. Instead, we observed the hazard function of MRD_{worst_case} rose more steeply than the hazard function of $MRD_{conventional}$ (Fig. 1B). Inaccuracies in MRD-estimation using the conventional approach distorted the critical range of MRD for discriminating low- from high-risks of cumulative incidence of relapse (CIR).

Combining MRD_{worst_case} on day 19 with estimated relapse risk at diagnosis further improved risk-stratification of the children whose $MRD_{conventional}$ on day 19 was $<0.01\%$ with a C-statistic of 0.73. This was significantly better than using MRD_{worst_case} alone (0.73 vs. 0.57 [$P<0.001$; 2-sided Wilcoxon test on 500 bootstrap samples]) or using relapse risk at diagnosis alone (0.73 vs. 0.68 [$P<0.001$]; Fig. 1C). 214 children (73%) with $MRD_{conventional}<0.01\%$ on day 19 were low-risk at diagnosis and all subsequently received low-intensity therapy. The remainder (80 [27%]) were intermediate-risk at diagnosis and all received high-intensity therapy. Consequently, therapy-intensity did not confound results within each therapy cohort.

Interestingly, point-estimates for relapse at 1.5 years for high- and low-MRD_{worst_case} cohorts were similar and their relapse curves only diverged after 1.5 years (Fig. 1A, C). Because MRD_{worst_case} corrected for (probable) under-sampling of leukaemia cells at

[1]Strictly speaking, the worst possible value of true MRD is always ≈1 even when $MRD_{conventional}=0$. (The chance of true MRD ≈1 might be practically zero but the probability of this unlikely event is never zero.) Defining the worst-case scenario estimate as "not likely (probability ≤ 0.05) to exceed this value" is more useful for comparing MRD-test results.

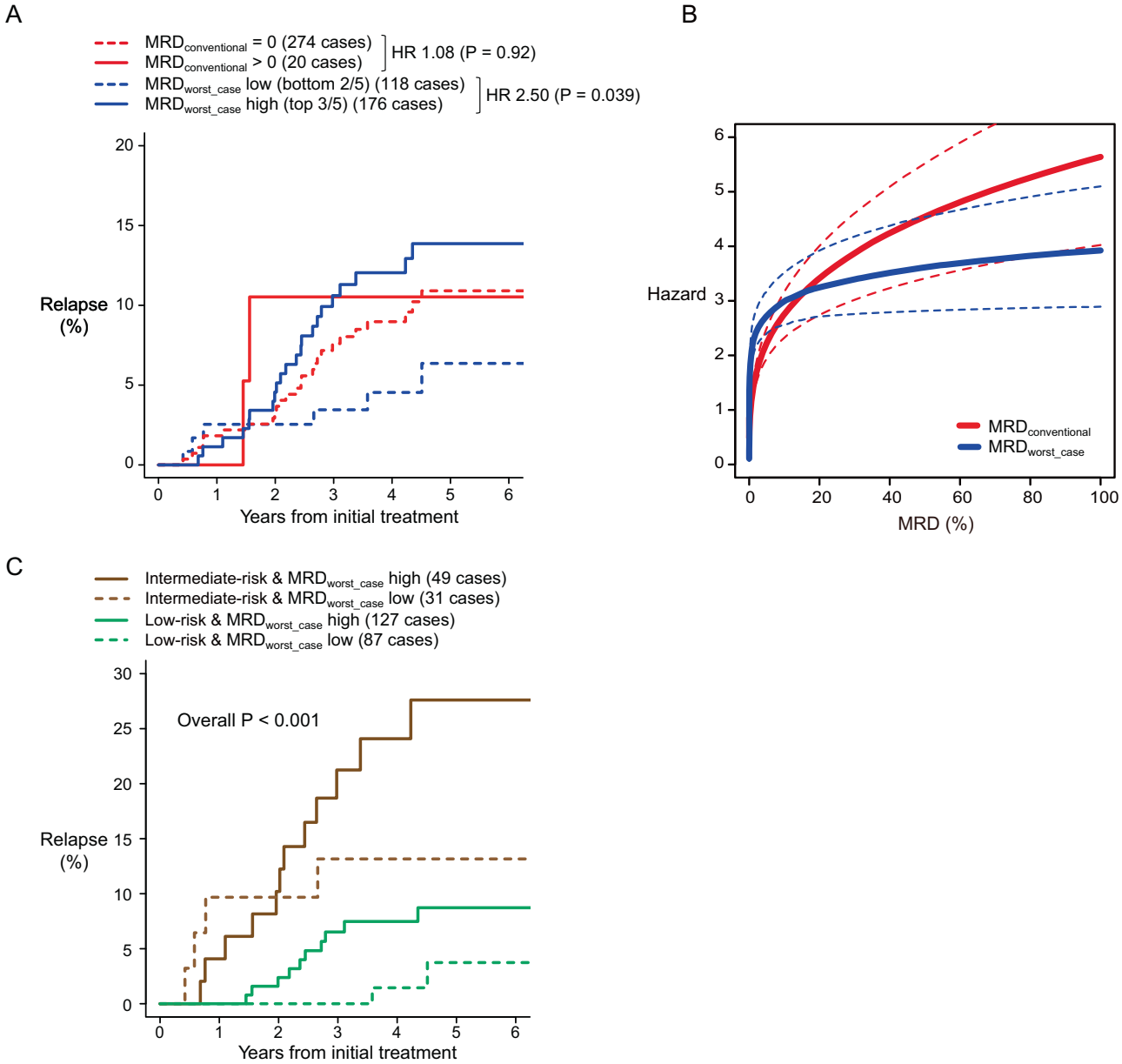

Fig. 1 Using MRD_{worse_case} in a cohort of children with ALL. A Risk-stratifications based on $MRD_{conventional}$ vs. MRD_{worst_case} on day 19 when $MRD_{conventional}$<0.01%. Cut-off threshold for distinguishing "MRD_{worst_case} high" and "MRD_{worst_case} low" is $7.3 \times 10E{-}6$ or 0.00073%. **B** Hazard functions of $MRD_{conventional}$ and MRD_{worst_case} on day 19 for relapse risk. Curve estimation is based on data from the entire cohort of 883 children. Dotted lines indicate 95-percent confidence intervals. **C** Risk-stratification based on joint consideration of estimated relapse risk at diagnosis and MRD_{worst_case} on day 19 when $MRD_{conventional}$<0.01%. Hazard ratios (HRs) and *P*-values are based on the Fine-Gray and Gray methods [48, 49].

therapy start this divergence likely resulted from expansion of pre-existing sub-clones during and/or after the end of low-intensity maintenance therapy (54 to 125 weeks) [32].

IS MRD_{WORST_CASE} AN *INDEX* OR A *METRIC* FOR MRD?

Index is defined as *a number (such as a ratio) derived from a series of observations and used as an indicator or measure*. *Metric* is defined as *a standard of measurement*. Some may argue MRD_{worst_case} is an *index* for MRD whilst $MRD_{conventional} = \frac{n}{N}$ is a *metric*. The distinction between *index* and *metric* is in some measure semantic. Even $MRD_{conventional}$ is a statistical construct for estimating likelihood of relapse. $MRD_{conventional}$ is what statisticians call a maximum-likelihood estimate, which is *not* the same as an estimate for the median *(i.e.* 50th-percentile) value among all the possible values of true MRD conditional on test result (Supplementary Methods). When $MRD_{conventional}$ is zero $MRD_{conventional}$ is actually the 0th-percentile (*i.e.* the lowest possible) value among all the possible values of true MRD conditional on test result! MRD_{worst_case}, on the other hand, is the 95th-percentile value among all the possible values of true MRD conditional on test result.

DISCUSSION

In this Perspective we argue the consensus GLP of MRD-testing is sub-optimal in many instances. Under these circumstances $MRD_{conventional}$ test results are sometimes mis-leading. Our

analyses of data from a large cohort of childhood ALL indicates the *minimax regret* approach (MRD_{worst_case}) improves relapse risk prediction over the current method ($MRD_{conventional}$). MRD_{worst_case} corrects for variation in strength of evidence in MRD-tests when predicting leukaemia relapse. Moreover, non-linear modeling of MRD_{worst_case} hazard function uncovers the critical range of MRD wherein the risk of leukaemia relapse accelerates. Because the true hazard function curve is steeper and operates at a lower range of MRD than previously realised based on $MRD_{conventional}$ it is important to continue developing and using increasingly sensitive (and specific) assays for detecting residual leukaemia cells.

We acknowledge several limitations. Our analyses of the clinical data were retrospective and subject to bias. We focused on MPFC, which enumerates mostly live cells one-by-one and is distinct from other types of assays such as quantitative real time polymerase chain reaction (RT-qPCR) or next generation sequencing (NGS). We also did not analyse false-positive errors in MRD-tests, which are more likely a *biological* than statistical issue as many or perhaps most false-positives are caused by not knowing which leukaemia cells have the biological ability to cause relapse within an observation interval [44–46]. In MPFC some aberrant leukaemia phenotypes may be more confidently identified as *positive* compared with others. Consequently, further refinement of results of MRD-testing is possible. Also, molecular tests such as NGS may increase accuracy of identifying residual leukaemia cells [8, 47]. However, sampling error remains an inherent limitation for any MRD-test as does the current inability to identify leukaemia cells biologically able to cause relapse regardless of detection technology.

We suggest our proposed metric MRD_{worst_case} will help haematologists more accurately predict leukaemia relapse. It is possible to further improve accuracy of predicting leukaemia relapse by considering additional data beyond MRD-tests provided confounding *predictive* and *prognostic* co-variates are adjusted for and the therapy regimen is considered.

DATA AVAILABILITY

Clinical data are available upon reasonable request to the corresponding authors.

REFERENCES

1. Lucio P, Parreira A, van den Beemd MW, van Lochem EG, van Wering ER, Baars E, et al. Flow cytometric analysis of normal B cell differentiation: a frame of reference for the detection of minimal residual disease in precursor-B-ALL. Leukemia. 1999;13:419–27.
2. Gabert J, Beillard E, van der Velden VH, Bi W, Grimwade D, Pallisgaard N, et al. Standardization and quality control studies of 'real-time' quantitative reverse transcriptase polymerase chain reaction of fusion gene transcripts for residual disease detection in leukemia - a Europe Against Cancer program. Leukemia. 2003;17:2318–57.
3. Mullighan CG, Goorha S, Radtke I, Miller CB, Coustan-Smith E, Dalton JD, et al. Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. Nature. 2007;446:758–64.
4. Vardiman JW, Thiele J, Arber DA, Brunning RD, Borowitz MJ, Porwit A, et al. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. Blood. 2009;114:937–51.
5. Conter V, Bartram CR, Valsecchi MG, Schrauder A, Panzer-Grumayer R, Moricke A, et al. Molecular response to treatment redefines all prognostic factors in children and adolescents with B-cell precursor acute lymphoblastic leukemia: results in 3184 patients of the AIEOP-BFM ALL 2000 study. Blood. 2010;115:3206–14.
6. Papaemmanuil E, Gerstung M, Bullinger L, Gaidzik VI, Paschka P, Roberts ND, et al. Genomic classification and prognosis in acute myeloid leukemia. N. Engl J Med. 2016;374:2209–21.
7. Estey E, Gale RP. How good are we at predicting the fate of someone with acute myeloid leukaemia? Leukemia. 2017;31:1255–8.
8. Hourigan CS, Gale RP, Gormley NJ, Ossenkoppele GJ, Walter RB. Measurable residual disease testing in acute myeloid leukaemia. Leukemia. 2017;31:1482–90.
9. Ceppi F, Rizzati F, Colombini A, Conter V, Cazzaniga G. Utilizing the prognostic impact of minimal residual disease in treatment decisions for pediatric acute lymphoblastic leukemia. Expert Rev Hematol. 2021;14:795–807.
10. Dohner H, Wei AH, Appelbaum FR, Craddock C, DiNardo CD, Dombret H, et al. Diagnosis and management of AML in adults: 2022 recommendations from an international expert panel on behalf of the ELN. Blood. 2022;140:1345–77.
11. Terwijn M, van Putten WL, Kelder A, van der Velden VH, Brooimans RA, Pabst T, et al. High prognostic impact of flow cytometric minimal residual disease detection in acute myeloid leukemia: data from the HOVON/SAKK AML 42A study. J Clin Oncol. 2013;31:3889–97.
12. Freeman SD, Virgo P, Couzens S, Grimwade D, Russell N, Hills RK, et al. Prognostic relevance of treatment response measured by flow cytometric residual disease detection in older patients with acute myeloid leukemia. J Clin Oncol. 2013;31:4123–31.
13. Chen X, Xie H, Wood BL, Walter RB, Pagel JM, Becker PS, et al. Relation of clinical response and minimal residual disease and their prognostic impact on outcome in acute myeloid leukemia. J Clin Oncol. 2015;33:1258–64.
14. Othus M, Wood BL, Stirewalt DL, Estey EH, Petersdorf SH, Appelbaum FR, et al. Effect of measurable ('minimal') residual disease (MRD) information on prediction of relapse and survival in adult acute myeloid leukemia. Leukemia. 2016;30:2080–3.
15. Berry DA, Zhou S, Higley H, Mukundan L, Fu S, Reaman GH, et al. Association of minimal residual disease with clinical outcome in pediatric and adult acute lymphoblastic leukemia: a meta-analysis. JAMA Oncol. 2017;3:e170580.
16. Dillon LW, Gui G, Page KM, Ravindra N, Wong ZC, Andrew G, et al. DNA sequencing to detect residual disease in adults with acute myeloid leukemia prior to hematopoietic cell transplant. JAMA. 2023;329:745–55.
17. Campbell M, Kiss C, Zimmermann M, Riccheri C, Kowalczyk J, Felice MS, et al. Childhood acute lymphoblastic leukemia: results of the randomized acute lymphoblastic leukemia Intercontinental-Berlin-Frankfurt-Munster 2009 Trial. J Clin Oncol. 2023;41:3499–511.
18. Beillard E, Pallisgaard N, van der Velden VH, Bi W, Dee R, van der Schoot E, et al. Evaluation of candidate control genes for diagnosis and residual disease detection in leukemic patients using 'real-time' quantitative reverse-transcriptase polymerase chain reaction (RQ-PCR) - a Europe against cancer program. Leukemia. 2003;17:2474–86.
19. van der Velden VH, Hochhaus A, Cazzaniga G, Szczepanski T, Gabert J, van Dongen JJ. Detection of minimal residual disease in hematologic malignancies by real-time quantitative PCR: principles, approaches, and laboratory aspects. Leukemia. 2003;17:1013–34.
20. Loken MR, Alonzo TA, Pardo L, Gerbing RB, Raimondi SC, Hirsch BA, et al. Residual disease detected by multidimensional flow cytometry signifies high relapse risk in patients with de novo acute myeloid leukemia: a report from Children's Oncology Group. Blood. 2012;120:1581–8.
21. Kalina T, Flores-Montero J, van der Velden VH, Martin-Ayuso M, Bottcher S, Ritgen M, et al. EuroFlow standardization of flow cytometer instrument settings and immunophenotyping protocols. Leukemia. 2012;26:1986–2010.
22. van Dongen JJ, Lhermitte L, Bottcher S, Almeida J, van der Velden VH, Flores-Montero J, et al. EuroFlow antibody panels for standardized n-dimensional flow cytometric immunophenotyping of normal, reactive and malignant leukocytes. Leukemia. 2012;26:1908–75.
23. Grimwade D, Freeman SD. Defining minimal residual disease in acute myeloid leukemia: which platforms are ready for "prime time"? Blood. 2014;124:3345–55.
24. Ladetto M, Bruggemann M, Monitillo L, Ferrero S, Pepin F, Drandi D, et al. Next-generation sequencing and real-time quantitative PCR for minimal residual disease detection in B-cell disorders. Leukemia. 2014;28:1299–307.
25. Pulsipher MA, Carlson C, Langholz B, Wall DA, Schultz KR, Bunin N, et al. IgH-V(D)J NGS-MRD measurement pre- and early post-allotransplant defines very low- and very high-risk ALL patients. Blood. 2015;125:3501–8.
26. Saygin C, Cannova J, Stock W, Muffly L. Measurable residual disease in acute lymphoblastic leukemia: methods and clinical context in adult patients. Haematologica. 2022;107:2783–93.
27. Roschewski M, Stetler-Stevenson M, Yuan C, Mailankody S, Korde N, Landgren O. Minimal residual disease: what are the minimum requirements? J Clin Oncol. 2014;32:475–6.
28. Schuurhuis GJ, Heuser M, Freeman S, Bene MC, Buccisano F, Cloos J, et al. Minimal/measurable residual disease in AML: a consensus document from the European LeukemiaNet MRD Working Party. Blood. 2018;131:1275–91.
29. Paiva B, Puig N, Cedena MT, Rosinol L, Cordon L, Vidriales MB, et al. Measurable residual disease by next-generation flow cytometry in multiple myeloma. J Clin Oncol. 2020;38:784–92.
30. Buccisano F, Palmieri R, Piciocchi A, Arena V, Maurillo L, Del Principe MI, et al. Clinical relevance of an objective flow cytometry approach based on limit of detection and limit of quantification for measurable residual disease assessment

in acute myeloid leukemia. A post-hoc analysis of the GIMEMA AML1310 trial. Haematologica. 2022;107:2823–33.
31. Jeha S, Pei D, Choi J, Cheng C, Sandlund JT, Coustan-Smith E, et al. Improved CNS control of childhood acute lymphoblastic leukemia without cranial irradiation: St Jude Total Therapy Study 16. J Clin Oncol. 2019;37:3377–91.
32. Yang W, Cai J, Shen S, Gao J, Yu J, Hu S, et al. Pulse therapy with vincristine and dexamethasone for childhood acute lymphoblastic leukaemia (CCCG-ALL-2015): an open-label, multicentre, randomised, phase 3, non-inferiority trial. Lancet Oncol. 2021;22:1322–32.
33. Theunissen P, Mejstrikova E, Sedek L, van der Sluijs-Gelling AJ, Gaipa G, Bartels M, et al. Standardized flow cytometry for highly sensitive MRD measurements in B-cell acute lymphoblastic leukemia. Blood. 2017;129:347–57.
34. Modvig S, Hallbook H, Madsen HO, Siitonen S, Rosthoj S, Tierens A, et al. Value of flow cytometry for MRD-based relapse prediction in B-cell precursor ALL in a multicenter setting. Leukemia. 2021;35:1894–906.
35. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. Bayesian Data Analysis. 3rd ed. Boca Raton, FL: Chapman and Hall/CRC; 2013.
36. Simon HA. A behavioral model of rational choice. Q J Econ. 1955;69:99–118.
37. Savage LJ. The theory of statistical decision. J Am Stat Assoc. 1951;46:55–67.
38. Pencina MJ, D'Agostino RB Sr. Evaluating discrimination of risk prediction models: the C statistic. JAMA. 2015;314:1063–4.
39. Efron B. Bootstrap methods: another look at the jackknife. Ann Stat. 1979;7:1–26.
40. Kirkpatrick S, Gelatt CD Jr, Vecchi MP. Optimization by simulated annealing. Science. 1983;220:671–80.
41. Green PJ, Silverman BW. Nonparametric regression and generalized linear models: a roughness penalty approach. London: Chapman & Hall; 1994.
42. Gauthier J, Wu QV, Gooley TA. Cubic splines to model relationships between continuous variables and outcomes: a guide for clinicians. Bone Marrow Transpl. 2020;55:675–80.
43. Chen J, Gale RP, Feng Y, Hu Y, Qi S, Liu X, et al. Are haematopoietic stem cell transplants stem cell transplants, is there a threshold dose of CD34-positive cells and how many are needed for rapid posttransplant granulocyte recovery? Leukemia. 2023. https://doi.org/10.1038/s41375-023-01973-2.
44. Song J, Mercer D, Hu X, Liu H, Li MM. Common leukemia- and lymphoma-associated genetic aberrations in healthy individuals. J Mol Diagn. 2011;13:213–9.
45. Farina M, Rossi G, Bellotti D, Marchina E, Gale RP. Is having clonal cytogenetic abnormalities the same as having leukaemia. Acta Haematol. 2016;135:39–42.
46. Young AL, Challen GA, Birmann BM, Druley TE. Clonal haematopoiesis harbouring AML-associated mutations is ubiquitous in healthy adults. Nat Commun. 2016;7:12484.
47. Zhang Y, Wang S, Zhang J, Liu C, Li X, Guo W, et al. Elucidating minimal residual disease of paediatric B-cell acute lymphoblastic leukaemia by single-cell analysis. Nat Cell Biol. 2022;24:242–52.
48. Gray RJ. A class of K-sample tests for comparing the cumulative incidence of a competing risk. Ann Stat. 1988;16:1141–54.
49. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc. 1999;94:496–509.

ACKNOWLEDGEMENTS

Profs. Nick Cross (University of Southampton), Christopher Hourigan (National Institutes of Health) and Alec Morley (Flinders University) kindly reviewed the typescript. RPG acknowledges support from the National Institute of Health Research (NIHR). JC acknowledges support from the Institute of Hematology, Chinese Academy of Medical Sciences (IHCAMS).

AUTHOR CONTRIBUTIONS

JC and RPG conceived the study. XZ co-led the CCCG-ALL-2015 study, assisted by LZ and JW. HW led the team that performed MRD-testing. XL and QS compiled and curated the data, assisted by YH, WY, TW and ZS. JC developed the alternative MRD metric. YF, SQ, XL, YH, XG and WZ did the computation and developed the graphs and tables. JC and RPG prepared the typescript. All the authors reviewed the typescript, take responsibility for the content and agreed to submit for publication.

FUNDING

Supported, in part, by grants from the Ministry of Science and Technology of China (2019YFA0110803; XZ), the Chinese Academy of Medical Sciences (CAMS) Innovation Fund for Medical Sciences (2021-I2M-1-001 and 2022-I2M-2-003; JC), the National Institute of Health Research (NIHR) Biomedical Research Centre (RPG) and the Ministry of Science and Technology of China (84000-51200002; RPG).

COMPETING INTERESTS

RPG is a consultant to Antengene Biotech LLC, Ascentage Pharma Group and NexImmune Inc.; Medical Director, FFF Enterprises Inc.; Board of Directors: Russian Foundation for Cancer Research Support; and Scientific Advisory Boards, Nanexa AB and StemRad Ltd.

ETHICS APPROVAL

Approved by the Academic Committee (IIT-NI2020001) and Ethics Review Committee (NI2020001-EC-1) of the Institute of Hematology, Chinese Academy of Medical Sciences (IHCAMS). Subjects gave written informed consent consistent with precepts of the Helsinki Declaration.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41375-023-02026-4.

Correspondence and requests for materials should be addressed to Xiaofan Zhu or Junren Chen.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.