Modernising toxicity monitoring in chronic myeloid leukaemia

Yolba Smit

RADBOUD UNIVERSITY PRESS

Radboud Dissertation Series

Modernising toxicity monitoring in chronic myeloid leukaemia

This research was funded in part by grant 516022524 from the Netherlands Organization for Health Research and Development (ZonMw), which included unrestricted educational grants from AbbVie, AstraZeneca, and Janssen Pharmaceutical Companies. Funding bodies had no role in the design of the research, collection and analysis of data, or the decision to publish.

For reasons of consistency some terms have been standardised throughout the text. As a consequence, the text may differ from the articles that have been published.

Modernising toxicity monitoring in chronic myeloid leukaemia
Yolba Smit

Radboud Dissertation Series
ISSN: 2950-2772 (Online); 2950-2780 (Print)

Published by RADBOUD UNIVERSITY PRESS
Postbus 9100, 6500 HA Nijmegen, The Netherlands
www.radbouduniversitypress.nl

Design: Proefschrift AIO | Annelies Lips
Cover: Yolba Smit & Guntra Laivacuma
Printing: DPN Rikken/Pumbo

ISBN: 9789465151106
DOI: 10.54195/9789465151106
Free download at: https://doi.org/10.54195/9789465151106

© 2026 Yolba Smit

RADBOUD
UNIVERSITY
PRESS

This is an Open Access book published under the terms of Creative Commons Attribution-Noncommercial-NoDerivatives International license (CC BY-NC-ND 4.0). This license allows reusers to copy and distribute the material in any medium or format in unadapted form only, for noncommercial purposes only, and only so long as attribution is given to the creator, see http://creativecommons.org/licenses/by-nc-nd/4.0/.

Modernising toxicity monitoring in chronic myeloid leukaemia

Proefschrift ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.M. Sanders,
volgens besluit van het college voor promoties
in het openbaar te verdedigen op

vrijdag 13 maart 2026
om 12.30 uur precies

door

Yolba Smit
geboren op 25 juli 1970
te Utrecht

Promotoren:
Prof. dr. N.M.A. Blijlevens
Prof. dr. R.P.M.G Hermens
Prof. dr. ir. A.L.A.J. Dekker (Maastricht University)

Manuscriptcommissie:
Prof. dr. A. Timen
Prof. dr. M.A.W. Merkx
Dr. S.K. Klein (Universitair Medisch Centrum Groningen)

Table of contents

Chapter 1
General introduction

At present, clinical practice guidelines in haemato-oncology primarily focus on efficacy, as demonstrated by randomised controlled trials. However, daily practice is much less standardised and more diverse, with both patients and professionals often not adhering to guideline standards, as these do not always address the real-world dilemmas they face. In this thesis, we explore how to make real-world patient-reported toxicity symptoms available for chronic myeloid leukaemia (CML) care and the CML guideline.

In this General introduction, we first describe the process of clinical guideline development, and how it is embedded within the broader medical evidence ecosystem. We then address some of the key weaknesses of this ecosystem and how they result in the exclusion of patients' experiences from guidelines. Finally, we introduce CML and its evidence ecosystem as a specific example on which we focus our research.

Clinical practice guidelines

Clinical practice guidelines are defined as *"... statements that include recommendations intended to optimise patient care that are informed by a systematic review of evidence and an assessment of the benefits and harms of alternative care options"*[1], *"supplemented with expertise and experiences from health care professionals and health care users" [2].* Their goal is to support and improve clinician and patient health care decisions [1, 3], ultimately enhancing the quality of health care and patient outcomes [1].

Guidelines are initiated and produced by networks of various local, national and international organisations, governed by different laws and regulations in different parts of the world. In Western countries, guidelines are often developed by national bodies such as NICE (National Institute for Health and Care Excellence) in the United Kingdom, IKNL (Netherlands Comprehensive Cancer Organisation) in the Netherlands, or national medical professional societies (or their networks). International collaboration and development occur within international bodies such as the World Health Organisation [4], and through initiatives such as the Guidelines International Network [5].

The evidence ecosystem

Guidelines form an integral part of a learning health care system: the evidence ecosystem. This is a feedback loop in which evidence is synthesised, guidance is created, disseminated, implemented and evaluated, ultimately informing the production of new evidence [6, 7]. In its essence, it is a plan-do-check-act cycle where evidence or knowledge is developed through research, interpreted in guidelines, and implemented and evaluated in clinical practice (Figure 1, and Figure 2, Chapter 5) [3]. A cyclic and integrated loop can increase the value of health care outcomes and reduce unnecessary health care procedures, within a true learning health care environment [6].

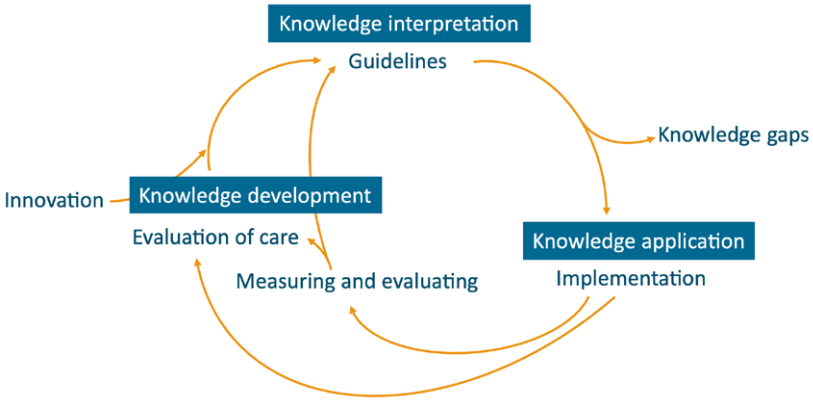


Figure 1 Plan-do-check-act cycle of the evidence ecosystem of learning health care environments, adapted from [3]

Guideline development heavily relies on evidence-based medicine [1], which is also reflected in guideline development manuals and the appraisal of guidelines themselves, both of which are subjected to rigorous scientific methodology [2, 4, 8-12]. Evidence-based medicine is commendable and remains the best system we have for identifying, evaluating and weighing all evidence, and to translate that evidence into guidance. However, it has inherent weaknesses that can lead to a lack of applicability and insufficient stakeholder involvement [13]; and to guidelines biased towards 'hard' medicotechnical information and guidance, rather than focusing on the 'soft' aspects of high quality care as patient experience.

Firstly, evidence-based medicine, and consequently guidelines, are dominated by frequency-based reasoning [14] which favours randomised controlled

trials and meta-analyses of randomised controlled trials as the gold standard [15-17]. By their very nature, randomised controlled trials are conducted in controlled environments and often involve highly selected patients, excluding elderly or comorbid patients, or restricting inclusion to those expected to respond the best to treatment [1, 16, 18, 19]. As a consequence, much of the evidence-based guidance does not apply to substantial groups of patients [1], although extrapolation or generalisation of randomised controlled evidence is often applied to patients that would not have met the trial inclusion criteria [20]. Also, randomised controlled trials provide an average treatment estimate for the population included [20], but not for each individual in that trial population, let alone those outside it [21]. Moreover, the populations studied tend to be relatively small, and the study periods are typically short, meaning little is known about rare or long-term efficacy or harm [21].

Secondly, evidence-based medicine tends to focus on 'hard' medicotechnical information (e.g., efficacy), and often relies on the most readily available medical information (e.g., laboratory results, survival) [18]. This rarely includes more 'soft' information, such as the type and severity of symptoms patients experience, comorbidities, compliance, patients' treatment goals, and perspectives. This 'hard' clinical information does not adequately capture patient preferences [22]. Physicians, and of course patients themselves, know much more about patients' circumstances than is reflected in the average treatment effect of the trial [20], and patients' experiential knowledge is unique [22] and cannot be fully captured by other stakeholders. Although these issues are not new [18], the current era of value-based health care, with its emphasis on value for patients, offers a new opportunity to address them. In particular, real-world data and patient-reported outcome measures can contribute to solving the inherent issues described. To move forward, we focus on CML and its evidence ecosystem.

Chronic myeloid leukaemia

CML is a model disease in many ways. Its course is chronic and relatively straightforward in the vast majority of patients, who are in the chronic phase CML [23]. In this phase, CML remains under control with daily, lifelong oral medication known as tyrosine kinase inhibitors (TKIs). The remaining 10% of patients have acute or blast phase CML, which has a much more aggressive course and falls outside the scope of this thesis. Disease control in chronic

phase CML requires strict, time-bound monitoring of a near-perfect biomarker BCR::ABL1, in order to adjust or switch medication when needed [23]. At present, there are six main types of TKIs available on the market in most countries: imatinib, dasatinib, nilotinib, bosutinib, ponatinib and asciminib [24]. Since the introduction of TKIs in early 2000s, the life expectancy of CML patients has approached that of the general population [25-29]. Around one-third of patients is eligible to attempt a treatment-free interval [30]. Despite good survival, CML patients experience a loss in quality of life, which is approximately one-third lower than that of the general population [31].

Toxicity from oral CML treatment

In the transition from a lethal to a chronic cancer, CML research has understandably focused on the efficacy of the revolutionary TKIs. Now, with survival safeguarded in the vast majority of patients, the clinical focus is shifting towards preventing and managing the symptoms and complications of therapy [32, 33]. TKI therapy invokes considerable and enduring toxic effects [34]. Cardiovascular events and pulmonary hypertension, in particular, have garnered much attention, as these are serious side effects that can and need to be prevented and/or treated [34-36].

But symptoms reported by patients have often been considered less important outcome data, illustrated by the fact that 94% of leukaemia trials did not include any patient-reported outcome measures as a reported outcome [37]. In most TKI trials, side effects were reported by physicians to assess safety, using methods designed for short-course chemotherapy rather than long-term daily oral treatment [37]. Currently, such methods are considered wholly inadequate for evaluating the tolerability of modern therapies [38, 39]. This approach does not adequately reflect a patient's symptom burden, as physicians tend to underestimate the incidence and severity of symptoms compared to patients' self-reports [40-43]. Physicians tend to focus more on 'hard' medico-technical outcomes, such as platelet and neutrophil counts, whereas patients tend to find symptoms – such as fatigue – more distressing than their laboratory results [44, 45]. In addition, haematology trials that have evaluated patient-reported outcome data have sometimes veered towards minimising language in their reporting, by e.g. describing patient-reported symptoms as low-grade [38]. However, evidence suggests that low-grade

toxicity is associated with inferior quality of life, poor treatment tolerability, and early treatment discontinuation [38, 46-49].

CML patient-reported TKI toxicity symptoms

In patients with solid tumours, the routine monitoring of patient-reported symptoms with feedback to professionals has been documented to improve both quality of life and survival [50-53]. This evidence strongly supports the use of patient-reported outcome measures. Monitoring across the care continuum is now recommended by the European Society for Medical Oncology, to detect and manage treatment toxicities early, identify relapses, and tailor care to individual needs [54].

In haemato-oncology, evidence for symptom monitoring is still limited [55]. Even so, there is a clear need to modernise toxicity assessment by incorporating patient-reported outcomes into real-world care in this field as well [37-39, 48, 56, 57]. The *Lancet Haematology Commission on improving adverse event assessment* states that "*Improved use of electronic patient-reported outcome measures in routine care holds promise to generate a rich stream of structured real-world data that could help advance the goal of a true learning health care system*" [56].

In CML, the symptom burden during TKI therapy, as reported by patients, is considerable and strongly influences experienced quality of life, treatment adherence, and consequently treatment response [58-61]. Compared with the general population or noncancer controls, CML patients on TKIs reported worse depression, dyspnoea, fatigue, pain, and composite symptom burden scores (e.g., nausea, diarrhoea, itching, skin changes, and swelling of arms or legs) [62, 63]. This high symptom burden is reflected in lower health-related quality of life compared with the general population or noncancer controls [61-63], with marked impairments in role functioning in younger patients [61] and low-grade adverse events negatively effecting quality of life [47].

It is important to note that a symptom reported by patients does not automatically imply causality to the treatment, in a similar way as symptoms or adverse events reported by professionals do not necessarily imply causality. It is important to note that adverse event attribution is prone to inaccuracies and relies heavily on investigators' interpretation. This is illustrated by the

fact that nearly 50% of adverse events were attributed to the study drug in the placebo arm of two randomised clinical trials [64]. Even so, because CML is well-controlled in the vast majority of patients, symptoms cannot be directly attributed to their CML. The higher symptom burden at group-level, compared to non-CML patients, increases the probability that the CML treatment is involved. Though the benefits of real-world evidence in CML care have been acknowledged [65], real-world CML patient-reported toxicity has not been evaluated systematically. Gathering this data could help understand treatment satisfaction and adherence [66], as well as the impact of toxicities on quality of life [58, 65]. It could also address the limited stakeholder involvement and applicability of current guidelines [13].

CML guidelines

Increasing the value of health care outcomes, and reducing unnecessary health care procedures within a true learning health care environment are of especial importance in fields where the focus of care has shifted, like it did in CML care. In the CML context, the emphasis has moved from survival towards attaining treatment-free remission, towards preventing and managing chronic toxicity, such as finding the minimum effective dose [32, 67]. If the CML guideline continues to focus primarily on survival and disease control alone, it risks losing relevance for patients. For most patients, disease control is adequate, making the management of treatment-related toxicity the primary concern.

Besides enhancing the quality of care, patient-reported toxicity symptom assessment can also be systematically aggregated and evaluated as a quality metric [54], feed the scientific debate, and inform the development of evidence-based guidelines. However, patient-reported outcomes, such as symptoms and quality of life derived from the real-world, are rarely considered in a systematic, aggregated way in the process of haemato-oncology guideline development. CML guidelines [23, 68, 69] are primarily based on TKI registration studies – clinical trials in which early disease control is the primary outcomes, and care and populations are highly standardised, uniform and biased towards younger and healthier patients. Real-world diversity, such as information on actual medication usage, dosage adaptations, experienced (symptoms of) side effects, comorbidities, personal treatment goals, medication switches because of side effects, and the actual monitoring of BCR::ABL1, is hardly integrated into CML guidelines at present.

However, in CML real-world TKI treatment data are increasingly published and highlight findings that were undetected in the original trials, such as the high risk of cardiovascular disease associated with ponatinib [34], and higher non-haematological adverse drug reactions than originally reported for drugs such as dasatinib and imatinib [70, 71]. The European LeukemiaNet 2020 recommendations for treating chronic myeloid leukaemia states: "*we encourage additional research in this area [quality of life] and the use of patient-reported outcome questionnaires which may provide guidance about quantifying and addressing chronic issues faced by CML patients*"[23].

Unwanted practice variations in CML care are common, as its real-world is not as controlled as clinical trials and is much more complex. This is exemplified by the non-adherence to timely monitoring of BCR::ABL1, with only 21-27% of patients in the Netherlands meeting required monitoring milestones [72]. In China, only 43% of patients adhere to the recommended 3-monthly testing schedule [44], and around one-third of Chinese physicians report not following the CML guideline according to the time-related targets [73]. Similarly, non-adherence to treatment and dose adjustments is also common. Due to intolerability, one in five to one in three patients switches to another type of TKI [30, 74, 75].

This highlights that patient-reported outcomes, including toxicity symptoms, are a key factor in shared decision-making in the consultation room and, therefore, play a significant role in treatment outcomes. Yet, at present, we lack evidence-based recommendations on which TKI to switch to when a specific intolerable symptom arises, or on how to predict intolerability, which is necessary for more personalised care. Treatment switches are currently guided by expert knowledge and by available evidence on tolerability, which is mainly physician-reported tolerability. As a result, current CML guidelines fall short in addressing real-world treatment scenarios, as they often overlook the patient-centred context—limiting their practical relevance and personalised application.

Context of this thesis

Our research is embedded within the Dutch CML evidence ecosystem and is shaped by three key drivers: (1) the patient organisation Hematon, (2) the HOVON CML-MPN (Dutch-Belgian Cooperative Trial Group for Haematology-

Oncology) working group, and (3) the digital care platform CMyLife. We briefly discuss these three drivers here.

Firstly, the Dutch CML patient organisation, Hematon, has prioritised the development of an overview of each patient's symptoms to benefit patients. This is due to the significant impact that patients' symptoms have on their daily lives, and the relative underrecognition of these symptoms in the medical field. This clear priority has strengthened the focus of our research aim.

Secondly, within the Dutch CML evidence ecosystem, a vital role is played by the HOVON CML-MPN working group of the Dutch Society for Haematology. Not only does it develop the Dutch CML guideline [68], but it also initiates and coordinates clinical trials. Its members, all haematologists or patient representatives from patient organisation Hematon, are deeply involved in CML care. Through these combined activities, the CML-MPN working group links clinical care to knowledge gaps, research, and guideline development within the Dutch CML evidence ecosystem. As a result, The HOVON CML-MPN working group is a key driver and end-user of our research and its findings, under the supervision of the Dutch Society of Haematology.

Finally, the nationwide digital care platform CMyLife was established to provide patients with the tools to manage and monitor their own CML, interpret BCR::ABL1 results, and take appropriate action. CMyLife includes a patient-portal (www.cmycml.nl) and a guideline-application [76, 77]. Patients using the platform have access to their own electronic personal health environment, which connects individual hospitals' electronic patients records CMyLife's nationwide guideline-application. Dutch CML patients are treated in eight academic hospitals and 68 general hospitals [78], making CMyLife a unique platform for research, as patients can now be reached directly.

Aim of this thesis

The main aim of this thesis is to explore how real-world patient-provided toxicity symptoms can be integrated into CML care and the CML guideline, in a way that modernises and enhances the useability of the CML guideline. This would be an initial step towards creating a personalisation of the CML guideline, with a central focus on the patient's perspective. We therefore addressed:

- What is already known about patient-reported toxicity symptoms, both per symptom and per TKI (**Chapter 2**)?
- Which instruments are available to measure patient-reported toxicity symptoms in CML patients, and what is their content validity (**Chapter 3**)?
- Given the lack of instruments with sufficient content validity, can we develop and validate a new instrument to capture CML patient-reported toxicity symptoms? (**Chapter 4**)
- Can we explore pilot data to evaluate and compare the toxicity burden? (**Chapter 4**)
- What are the benefits and limitations of real-world patient-reported toxicity monitoring for guidelines and care, as perceived by patients, clinical professionals, and guideline developers (**Chapter 5**)?
- Which knowledge gaps in CML care can be addressed by real-world patient-reported toxicity data, and how should these be prioritised (**Chapter 5**)?

Thesis outline

In **Chapter 2**, we systematically search for and meta-analyse the prevalence of patient-reported toxicity, focusing on individual symptoms and TKIs. Limited evidence and non-standardised instruments prompt us to investigate existing instruments to assess patient-reported toxicity. We therefore systematically search for and critically assess the content validity of existing instruments (**Chapter 3**). As there is room for improvement we set out to develop and validate a new instrument. Additionally, we collect pilot data to explore differences in toxicity burden (**Chapter 4**).

To prepare for the practical implementation of our instrument in care and guideline development, we assess the benefits and limitations as perceived by end-users regarding its useability in both contexts (**Chapter 5**). We also identify and prioritise knowledge gaps that could be addressed through real-world patient-reported toxicity data. These gaps translate our research findings into actionable points for CML guideline development Finally, in **Chapter 6,** we conclude with a general discussion, reflecting on our findings and considering their implications for future practice and research.

Terminology

Throughout this thesis we refer to *guidelines* as clinical guidelines, as defined by the Institute of Medicine in the United States [1]. In addition, we refer to *toxicity symptoms* as symptoms that may indicate toxic effects or an adverse event resulting from treatment. It is important to note that a toxicity symptom does not necessarily imply a causal relationship with the treatment received.

References

1. *Clinical practice guidelines we can trust*. 2011: Institute of Medicine of the National Academies.
2. *AQUA-Guideline [In Dutch]*. 2021, Advisory and expert group Quality standards (AQUA).
3. *Medical specialist guidelines 3.0 [In Dutch]*. 2023, Federatie Medisch Specialisten.
4. *WHO handbook for guideline development 2nd edition*. 2014, World Health Organization.
5. Ollenschlager, G., et al., *Improving the quality of health care: using international collaboration to inform guideline programmes by founding the Guidelines International Network (G-I-N).* Qual Saf Health Care, 2004. **13**(6): p. 455-60.
6. Vandvik, P.O. and L. Brandt, *Future of evidence ecosystem series: evidence ecosystems and learning health systems: why bother?* J Clin Epidemiol, 2020. **123**: p. 166-170.
7. Cartabellotta, A. and J.K. Tilson, *The ecosystem of evidence cannot thrive without efficiency of knowledge generation, synthesis, and translation.* J Clin Epidemiol, 2019. **110**: p. 90-95.
8. *Developing NICE guidelines: the manual*. 2024, National Institute for Health and Care Excellence (NICE).
9. *Guideline for guidelines. 20 Criteria for developing and implementing a clinical guideline [In Dutch]*. 2012, Regieraad Kwaliteit van Zorg: Den Haag.
10. *Appraisal of guidelines for research and evaluation (AGREE II)*. 2017, The AGREE Research Trust.
11. Guyatt, G.H., et al., *GRADE: an emerging consensus on rating quality of evidence and strength of recommendations.* BMJ, 2008. **336**(7650): p. 924-6.
12. *SIGN 50. A guideline developer's handbook*. 2019, Scottish Intercollegiate Guidelines Network.
13. Moleman, M., et al., *Evidence-basing for quality improvement; bringing clinical practice guidelines closer to their promise of improving care practices.* J Eval Clin Pract, 2022.
14. Wieringa, S., et al., *Different knowledge, different styles of reasoning: a challenge for guideline development.* BMJ Evid Based Med, 2018. **23**(3): p. 87-91.
15. Sackett, D.L., et al., *Evidence based medicine: what it is and what it isn't.* BMJ, 1996. **312**(7023): p. 71-2.
16. Deaton, A. and N. Cartwright, *Reflections on randomized control trials.* Soc Sci Med, 2018. **210**: p. 86-90.
17. Zuiderent-Jerak, T., F. Forland, and F. Macbeth, *Guidelines should reflect all knowledge, not just clinical trials.* BMJ, 2012. **345**: p. e6702.
18. Feinstein, A.R. and R.I. Horwitz, *Problems in the "evidence" of "evidence-based medicine".* Am J Med, 1997. **103**(6): p. 529-35.
19. Jin, S., R. Pazdur, and R. Sridhara, *Re-evaluating eligibility criteria for oncology clinical trials: analysis of investigational new drug applications in 2015.* J Clin Oncol, 2017. **35**(33): p. 3745-3752.
20. Deaton, A. and N. Cartwright, *Understanding and misunderstanding randomized controlled trials.* Soc Sci Med, 2018. **210**: p. 2-21.
21. Frieden, T.R., *Evidence for Health Decision Making - Beyond Randomized, Controlled Trials.* N Engl J Med, 2017. **377**(5): p. 465-475.

22. Janssens, R., et al., *Opportunities and challenges for the inclusion of patient preferences in the medical product life cycle: a systematic review.* BMC Med Inform Decis Mak, 2019. **19**(1): p. 189.
23. Hochhaus, A., et al., *European LeukemiaNet 2020 recommendations for treating chronic myeloid leukemia.* Leukemia, 2020. **34**(4): p. 966-984.
24. Yeung, D.T., N. Shanmuganathan, and T.P. Hughes, *Asciminib: a new therapeutic option in chronic-phase CML with treatment failure.* Blood, 2022. **139**(24): p. 3474-3479.
25. Sasaki, K., et al., *Relative survival in patients with chronic-phase chronic myeloid leukaemia in the tyrosine-kinase inhibitor era: analysis of patient data from six prospective clinical trials.* Lancet Haematol, 2015. **2**(5): p. e186-93.
26. Bower, H., et al., *Life expectancy of patients with chronic myeloid leukemia approaches the life expectancy of the general population.* J Clin Oncol, 2016. **34**(24): p. 2851-7.
27. Ector, G., et al., *Conditional relative survival among adult patients with chronic myeloid leukemia in the tyrosine kinase inhibitor era: a population-based study in the Netherlands.* Leukemia, 2021. **35**(11): p. 3291-3294.
28. Ector, G., et al., *Primary therapy and relative survival among elderly patients with chronic myeloid leukemia: a population-based study in the Netherlands, 1989-2017.* Leukemia, 2020. **34**(12): p. 3408-3412.
29. Maas, C., et al., *The evolution of the loss of life expectancy in patients with chronic myeloid leukaemia: a population-based study in the Netherlands, 1989-2018.* Br J Haematol, 2022. **196**(5): p. 1219-1224.
30. Geelen, I.G.P., et al., *Treatment outcome in a population-based, 'real-world' cohort of patients with chronic myeloid leukemia.* Haematologica, 2017. **102**(11): p. 1842-1849.
31. Chen, E.Y., et al., *Loss in Overall and Quality-Adjusted Life Expectancy for Patients With Chronic-Phase Chronic Myeloid Leukemia.* Eur J Haematol, 2025. **114**(2): p. 334-342.
32. Steegmann, J.L., et al., *European LeukemiaNet recommendations for the management and avoidance of adverse events of treatment in chronic myeloid leukaemia.* Leukemia, 2016. **30**(8): p. 1648-71.
33. Lipton, J.H. and D. Rea, *Living, not just surviving, with chronic myeloid leukemia - quality of life to the front of the line.* Leukemia, 2025.
34. Dahlen, T., et al., *Adverse outcomes in chronic myeloid leukemia patients treated with tyrosine kinase inhibitors: Follow-up of patients diagnosed 2002-2017 in a complete coverage and nationwide agnostic register study.* Am J Hematol, 2022. **97**(4): p. 421-430.
35. Douxfils, J., et al., *Association between BCR-ABL tyrosine kinase inhibitors for chronic myeloid leukemia and cardiovascular events, major molecular response, and overall survival: a systematic review and meta-analysis.* JAMA Oncol, 2016. **2**(5): p. 625-632.
36. Montani, D., et al., *Pulmonary arterial hypertension in patients treated by dasatinib.* Circulation, 2012. **125**(17): p. 2128-37.
37. Thanarajasingam, G., et al., *Beyond maximum grade: modernising the assessment and reporting of adverse events in haematological malignancies.* Lancet Haematol, 2018. **5**(11): p. e563-e598.
38. Major, A., A.C. Dueck, and G. Thanarajasingam, *Beyond maximum grade: advancing the measurement and analysis of adverse events in malignant haematology trials in the modern era.* Lancet Haematol, 2025. **12**(6): p. e451-e462.

39. Brockelmann, P.J., et al., *Beyond maximum grade: tolerability of immunotherapies, cellular therapies, and targeted agents in haematological malignancies.* Lancet Haematol, 2025. **12**(6): p. e470-e481.
40. Basch, E., et al., *Patient versus clinician symptom reporting using the National Cancer Institute Common Terminology Criteria for Adverse Events: results of a questionnaire-based study.* Lancet Oncol, 2006. **7**(11): p. 903-9.
41. Fromme, E.K., et al., *How accurate is clinician reporting of chemotherapy adverse effects? A comparison with patient-reported symptoms from the Quality-of-Life Questionnaire C30.* J Clin Oncol, 2004. **22**(17): p. 3485-90.
42. Pakhomov, S.V., et al., *Agreement between patient-reported symptoms and their documentation in the medical record.* Am J Manag Care, 2008. **14**(8): p. 530-9.
43. Efficace, F., et al., *Patient- versus physician-reporting of symptoms and health status in chronic myeloid leukemia.* Haematologica, 2014. **99**(4): p. 788-93.
44. Chen, H., et al., *Patient versus physician perspective in the management of chronic myeloid leukemia during treatment with tyrosine kinase inhibitors.* Oncol Ther, 2024. **12**(1): p. 131-145.
45. Henon, C., et al., *Patient-reported tolerability of adverse events in phase 1 trials.* ESMO Open, 2017. **2**(2): p. e000148.
46. O'Connell, N.S., et al., *Importance of Low- and Moderate-Grade Adverse Events in Patients' Treatment Experience and Treatment Discontinuation: An Analysis of the E1912 Trial.* J Clin Oncol, 2024. **42**(3): p. 266-272.
47. Guerin, A., et al., *Impact of low-grade adverse events on health-related quality of life in adult patients receiving imatinib or nilotinib for newly diagnosed Philadelphia chromosome positive chronic myelogenous leukemia in chronic phase.* Curr Med Res Opin, 2014. **30**(11): p. 2317-28.
48. Bhatnagar, V., et al., *Beyond maximum grade: using patient-generated data to inform tolerability of treatments for haematological malignancies.* Lancet Haematol, 2025. **12**(6): p. e463-e469.
49. Efficace, F. and L. Cannella, *The value of quality of life assessment in chronic myeloid leukemia patients receiving tyrosine kinase inhibitors.* Hematology Am Soc Hematol Educ Program, 2016. **2016**(1): p. 170-179.
50. Gibbons, C., et al., *Routine provision of feedback from patient-reported outcome measurements to healthcare providers and patients in clinical practice.* Cochrane Database Syst Rev, 2021. **10**: p. CD011589.
51. Basch, E., et al., *Overall survival results of a trial assessing patient-reported outcomes for symptom monitoring during routine cancer treatment.* JAMA, 2017. **318**(2): p. 197-198.
52. Basch, E., et al., *Effect of electronic symptom monitoring on patient-reported outcomes among patients with metastatic cancer: a randomized clinical trial.* JAMA, 2022. **327**(24): p. 2413-2422.
53. Lai-Kwon, J., et al., *The feasibility, acceptability, and effectiveness of electronic patient-reported outcome symptom monitoring for immune checkpoint inhibitor toxicities: a systematic review.* JCO Clin Cancer Inform, 2023. **7**: p. e2200185.
54. Di Maio, M., et al., *The role of patient-reported outcome measures in the continuum of cancer clinical care: ESMO Clinical Practice Guideline.* Ann Oncol, 2022. **33**(9): p. 878-892.
55. Ruland, C.M., et al., *Effects of a computer-supported interactive tailored patient assessment tool on patient care, symptom distress, and patients' need for symptom management support: a randomized clinical trial.* J Am Med Inform Assoc, 2010. **17**(4): p. 403-10.

56. Thanarajasingam, G., et al., *Reaching beyond maximum grade: progress and future directions for modernising the assessment and reporting of adverse events in haematological malignancies.* Lancet Haematol, 2022. **9**(5): p. e374-e384.
57. Thanarajasingam, G., et al., *Beyond maximum grade: introduction to The Lancet Haematology Adverse Events Reporting Series.* Lancet Haematol, 2025. **12**(6): p. e403-e406.
58. Cella, D., C.J. Nowinski, and O. Frankfurt, *The impact of symptom burden on patient quality of life in chronic myeloid leukemia.* Oncology, 2014. **87**(3): p. 133-47.
59. Eliasson, L., et al., *Exploring chronic myeloid leukemia patients' reasons for not adhering to the oral anticancer drug imatinib as prescribed.* Leuk Res, 2011. **35**(5): p. 626-30.
60. Jiang, Q., et al., *Variables associated with patient-reported outcomes in persons with chronic myeloid leukemia receiving tyrosine kinase-inhibitor therapy.* J Cancer Res Clin Oncol, 2017. **143**(6): p. 1013-1022.
61. Efficace, F., et al., *Health-related quality of life in chronic myeloid leukemia patients receiving long-term therapy with imatinib compared with the general population.* Blood, 2011. **118**(17): p. 4554-60.
62. Foulon, S., et al., *Health state utility and quality of life measures in patients with chronic myeloid leukemia in France.* Qual Life Res, 2021. **30**(7): p. 2021-2032.
63. Phillips, K.M., et al., *Quality of life outcomes in patients with chronic myeloid leukemia treated with tyrosine kinase inhibitors: a controlled comparison.* Support Care Cancer, 2013. **21**(4): p. 1097-103.
64. Hillman, S.L., et al., *Evaluation of the value of attribution in the interpretation of adverse event data: a North Central Cancer Treatment Group and American College of Surgeons Oncology Group investigation.* J Clin Oncol, 2010. **28**(18): p. 3002-7.
65. Webster, J. and B.D. Smith, *The case for real-world evidence in the future of clinical research on chronic myeloid leukemia.* Clin Ther, 2019. **41**(2): p. 336-349.
66. Efficace, F., et al., *Patient-reported symptom monitoring and adherence to therapy in patients with newly diagnosed chronic myeloid leukemia.* Cancer, 2023.
67. Russo, D., et al., *Twenty years of evolution of CML therapy: how the treatment goal is moving from disease to patient.* Ther Adv Hematol, 2023. **14**: p. 20406207231216077.
68. Janssen, J.J.W.M., et al., *Guideline chronic myeloid leukemia [In Dutch]*. 2023, HOVON.
69. Shah, N.P., et al., *Chronic Myeloid Leukemia, Version 2.2024, NCCN Clinical Practice Guidelines in Oncology.* J Natl Compr Canc Netw, 2024. **22**(1): p. 43-69.
70. Adattini, J.A., et al., *Real-world efficacy and safety outcomes of imatinib treatment in patients with chronic myeloid leukemia: An Australian experience.* Pharmacol Res Perspect, 2022. **10**(5): p. e01005.
71. Adattini, J.A., et al., *Insights into dasatinib use and outcomes in real-world patients with chronic myeloid leukemia.* Haematologica, 2023. **108**(8): p. 2224-2228.
72. Ector, G., et al., *Adherence to quality indicators in chronic myeloid leukemia care: results from a population-based study in The Netherlands.* Leuk Lymphoma, 2023. **64**(2): p. 424-432.
73. Chen, H., et al., *Evaluation of physician guideline adherence and areas for improvement in managing patients with chronic myeloid leukemia: a cross-sectional survey.* Leuk Lymphoma, 2023. **64**(8): p. 1458-1464.
74. Busque, L., et al., *Switching TKIs during CML therapy is frequent, mostly driven by intolerance, and does not affect survival: a prospective Quebec registry study.* Blood Cancer J, 2025. **15**(1): p. 35.

75. Flynn, K.E., et al., *HSR25-182: Patient-Physician Communication About Adverse Events and Impact on Quality of Life to Foster Shared Decision-Making in Chronic Myeloid Leukemia.* J Natl Compr Canc Netw, 2025. **23**(3.5).

76. Ector, G.I., et al., *The development of a web-Based, patient-centered intervention for patients with chronic myeloid leukemia (CMyLife): design thinking development approach.* J Med Internet Res, 2020. **22**(5): p. e15895.

77. Verweij, L., et al., *Effectiveness of digital care platform CMyLife for patients with chronic myeloid leukemia: results of a patient-preference trial.* BMC Health Serv Res, 2023. **23**(1): p. 228.

78. Geelen, I.G.P., et al., *Impact of hospital experience on the quality of tyrosine kinase inhibitor response monitoring and consequence for chronic myeloid leukemia patient survival.* Haematologica, 2017. **102**(12): p. e486-e489.

1

Chapter 2

Patient-reported toxicity symptoms during tyrosine kinase inhibitor treatment in chronic myeloid leukaemia: a systematic review and meta-analysis

*Yolba Smit, *Pien Scheuter, Myrthe Lange, Jeroen Janssen, Eduardus Posthuma, Charlotte Bekker, Rosella Hermens, Nicole Blijlevens

*PS and YS contributed equally to this study

Supportive Care in Cancer. 2025, 33(5);446.

Abstract

Purpose

One in five chronic myeloid leukaemia (CML) patients experiences such intolerability that they switch tyrosine kinase inhibitor (TKI) treatment within three years. Information on tolerability is needed to guide shared decision-making. However, an overview of symptoms patients experience per TKI is lacking, and physician-graded toxicity underestimates patients' experiences.

Methods

We systematically searched PubMed and Embase from inception to February 2025 and conducted a meta-analysis on the prevalence of patient-reported symptoms in CML per TKI. This study follows the Preferred Reporting Items for Systematic Reviews (PRISMA) guideline for systematic reviews.

Results

We included 11 studies with 2,987 patients, reporting on 47 different symptoms of any severity. The low-grade patient-reported symptom burden was high. No data were available on asciminib and ponatinib, and minimal data were available for bosutinib. In indirect, unadjusted comparisons, 13 out of 47 symptoms (of any severity) showed significant differences in prevalence between common TKI types.

Conclusion

Our findings provide essential information to guide treatment decisions in cases of intolerability. However, there is a clear need for further research with standardised instruments, especially in second and third generation TKI types, including direct comparisons and comparisons adjusted for covariates.

Introduction

For two decades, patients with chronic myeloid leukaemia (CML) have been effectively treated with tyrosine kinase inhibitors (TKIs). Currently, six types of TKIs are available in most countries (imatinib, dasatinib, nilotinib, bosutinib, ponatinib and asciminib) [1]. As patients reach a near-normal life expectancy when they attain an optimal response [2, 3], quality of life during the often lifelong TKI treatment has become increasingly important. The symptom burden during TKI therapy, as reported by patients, has been shown to strongly influence quality of life, treatment adherence, and consequently, treatment response [4-6]. However, current CML guideline recommendations are primarily based on clinical trials in which survival or disease control is the primary outcome. In these trials, toxicity is primarily evaluated using the Common Terminology Criteria for Adverse Events (CTCAE) [7], graded by physicians. The focus is on clinically relevant toxicities that require medical intervention, such as neutropenia and thrombocytopenia, as well as complications like cardiovascular events [8]. Patient-reported symptoms, such as fatigue and skin problems, are often underreported, as they are deemed less relevant for haemato-oncological management and thus easily overlooked by physicians. Additionally, symptoms experienced by CML patients during therapy are underestimated by physicians, both in severity and prevalence [9]. However, intolerance leads to switching TKI treatment in one in five patients within the first three years treatment initiation [10]. In view of lifelong treatment, needed for most CML patients, efforts to optimise quality of life should be urgently addressed. As part of this, an overview of patient-reported symptoms per TKI is needed, as well as the differences in symptom burden between different TKI types, to adequately support patients, maximizing effectiveness while minimizing symptom burden and informing shared decision-making for treatment choices. Although many reviews, primarily narrative, summarise adverse effects, no quantitative summary of patients' experiences exists. We therefore systematically reviewed and meta-analysed the prevalence of toxicity-symptoms during each type of TKI treatment, as reported by CML patients.

Materials and methods

Data sources and search strategy

We searched PubMed and Embase/Ovid in English (inception - January 2022, updated in July 2023 and February 2025) with MeSH-terms or similar, plus free

text terms - including synonyms and brand names -, for CML and the six current TKI types. The full search strategy is given in the Supplementary material. This study is reported according to the Preferred Reporting Items for Systematic Reviews (PRISMA) format [11].

Eligibility criteria

Articles had to be on chronic phase CML patients ≥18 years, treated with imatinib, nilotinib, dasatinib, bosutinib, ponatinib or asciminib and report the prevalence (proportion) of patients who experienced a symptom separately for each TKI, and for each symptom.

Study selection

Two authors (YS and either PS or ML) independently selected studies, first on title and abstract, and subsequently full text. Prospective studies that mentioned symptoms or adverse effects in their abstract were always screened full text to determine whether these included patient-reported symptoms. Full-text selections were compared between authors and differences were discussed until consensus. Reference lists of included articles and systematic reviews were checked.

Data extraction

Data were extracted or calculated by a single researcher (PS or ML) and included the proportion of patients who reported a symptom of any severity (prevalence). If needed, proportions were calculated or extracted visually from figures. For the analysis of moderate-severe symptoms, the proportion of patients who scored one of the top two response options on a four-point scale was extracted. If >1 proportion was available over time, 12 months was taken, as this was the last measuring point available for all included articles. Symptoms that were described in various ways in different articles were brought together under a general term (see Supplementary material). When available, general symptoms (e.g., pain) were described in more detail (e.g., musculoskeletal pain, abdominal pain). In addition, study type, type of questionnaire(s) used, population size, treatment characteristics, median age, gender, comorbidity and comedication were extracted.

Critical appraisal

Two researchers independently assessed study quality (PS and ML or YS), using the adapted version of the Newcastle Ottawa Scale [12], resolving differences of opinion through discussion. There is no ideal quality assessment

tool designed specifically for prevalence studies [13]; the Newcastle Ottawa Scale has been described as the best option [14] with an adapted version available for prevalence studies [12, 15]. It assesses sampling, sample size, response rate, measurement tool and outcome assessment. We considered the assessment of confounding factors and statistical tests as not applicable to our data, and a score <5 as low-quality.

Statistical analysis

We performed meta-analyses on the prevalence of symptoms per TKI with a random effects model using Stata 17. Zero prevalence was imputed as 0.5 to be included in meta-analysis. The standard error of prevalences was calculated using population size and prevalence. Statistical heterogeneity was assessed using the I-squared (I^2) estimate, and considered high (>50%), moderate (25-50%) or low (<25%) [16]. We evaluated between-group differences in prevalence, also in indirect comparisons between studies. To limit findings due to chance, because of the high number of comparisons, we imposed $p<0.01$ instead of $p<0.05$ as a significancy limit for between-group differences.

We performed sensitivity analysis for the three most common symptoms of each TKI, excluding low-quality studies. For nilotinib sensitivity analysis was also performed on anxiety and depression and pain as these were the only symptoms one of the low-quality studies described. If significant ($p<0.05$), low-quality studies would be excluded from meta-analysis.

We explored heterogeneity in a random effects meta-regression analysis, for symptoms described by ≥5 articles, testing separately for multiple covariates related to population characteristics (median age (if needed mean age), male proportion, median treatment duration, comorbidity prevalence, comedication prevalence). If ≥10 studies with a selected covariate were available, we would perform a multivariate meta-regression [17].

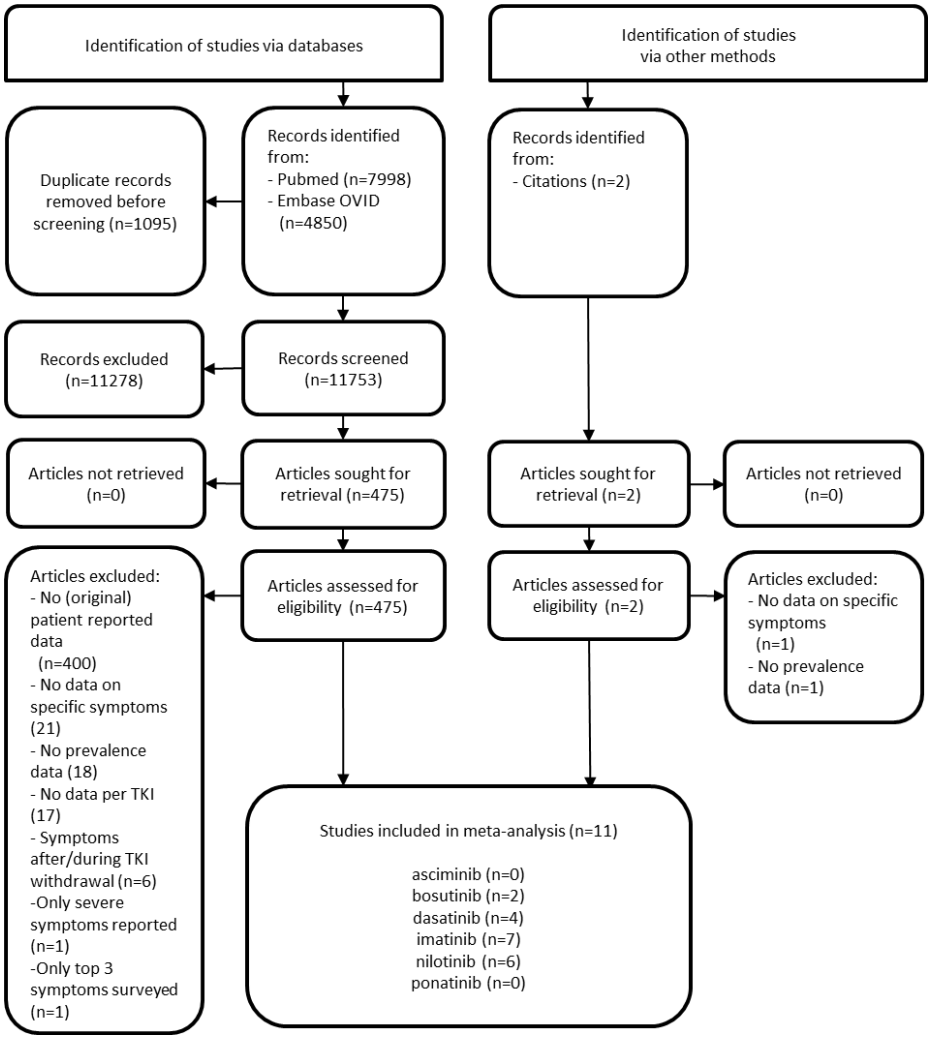
Results

Search results and study selection

We selected 11 studies with 2,987 patients: seven on imatinib [18-24] (1,795 patients), four on dasatinib [18, 21-23] (233 patients), six on nilotinib [18, 22, 23, 25-27] (509 patients), two on bosutinib [19, 28] (450 patients) and zero on asciminib or ponatinib (Figure 1). 66 studies were ineligible

because they did not report (original) patient data; no data on specific symptoms were reported; prevalence data was not provided, or not provided per symptom/per TKI, or only gathered for the top-3 symptoms; symptoms after TKI withdrawal were reported on; or only severe (and not moderate-severe) symptoms were reported on.

Figure 1 PRISMA flow diagram of study selection



Study characteristics

Four out of eleven included articles were published in the past five years (Table -). The questionnaires that were used were the European Organization for Research and Treatment for Cancer (EORTC) Quality of Life Questionnaire C30, CML24 (EORTC QLQ-C30, EORTC QLQ-CML24), the EuroQol 5D (EQ-5D), the functional assessment of cancer therapy leukaemia (FACT-Leu), the MD Anderson Symptom Inventory for chronic myeloid leukaemia (MDASI-CML), the Patient Health Questionnaire-9 (PHQ-9) and generic questionnaires. Population-sizes ranged from nine to 859 per TKI across studies. Treatment duration differed greatly: some studies included patients starting TKI treatment, while maximum median treatment duration was five years. In most studies which reported on the line of therapy, at least half of the population had been treated with a different TKI previously. Median age ranged from 40 to 63 years, and the percentage of male patients ranged from 46 to 63%. Comorbidities ranged from 26 to 60%, reported by five articles. Prevalence of comedication was described by three articles and ranged from 15 to 73%.

Critical appraisal

Seven studies achieved a moderate or higher quality score, with no studies scoring the maximum score of seven because only one point was assigned to a patient-reported assessment of outcome (see the Supplementary material). Four studies [22, 25, 27, 28] scored low on study quality because there were doubts on the representativeness of the sample, the number of non-respondents was high, or the instrument used to register symptoms was not validated.

Table 1 Characteristics of included studies

Article	Study type (setting)	Questionnaires (eligibility)	Patients included	TKI type (n of patients)	TKI treatment duration reported by studies
Boons 2020 [25]	Observational (six Dutch hospitals)	Generic (patients on nilotinib)	68	nilotinib (68)	43% started treatment 57% median 42 m of treatment at study start Symptoms reported at 3, 6 and 12 m
Bostan 2020 [18]	Cross-sectional (university hospital, Turkey)	EORTC QLQ-CML24 MDASI-CML (on TKI treatment)	121	dasatinib (30) imatinib (61) nilotinib (30)	Median 31 m
Cortes 2019 [19]	RCT (multicentre, international)	EQ-5D FACT-Leu (ECOG performance status 0 or 1)	385	bosutinib (194) imatinib (191)	Treatment started for study Symptoms reported at 12 m
Efficace 2011 [20]	Cross-sectional (26 Italian centres)	Generic (imatinib as 1st-line ≥3 years, in complete cytogenetic response)	422	imatinib (422)	Median 5 y
Efficace 2020 [21]	Cross-sectional (38 German and Italian centres)	EORTC QLQ-CML24 (dasatinib or imatinib as 1st-line <3 years, in complete cytogenetic response)	188	dasatinib (94) imatinib (94)	Median 17 m
Huguet 2019 [27]	Observational (multicentre, France)	EQ-5D-3L (recently diagnosed, nilotinib as 1st-line)	98	nilotinib (98)	Treatment started for study Symptoms reported at 24 m
Kantarjian 2018 [28]	RCT (multicentre, international)	EQ-5D FACT-Leu (resistance/ intolerance to imatinib, ECOG 0 or 1)	256	bosutinib (256)	Treatment started for study Symptoms reported at 36, 96, 192 and 360 w
Kapoor 2015 [24]	Cross-sectional (single centre, India)	PHQ-9 (imatinib ≥3 months, <80 years)	100	imatinib (100)	Median 30 m

Line of TKI therapy	Starting dose	Age (median)	Male (%)	Comorbidity (%)	Comedication (%)
51% 1st line 49% 2nd/3rd line	72% 2x300 mg/d 7% 2x150 mg/d 13% 2x400 mg/d 3% 300 mg/d 4% 400 mg/d	Mean: 58	51	56	73
50.4% 1st line 49.6% 2nd line	Not reported Dose reductions in 8% of patients	53	46	91% HCT-CI <3	Not reported
100% 1st line	bosutinib 400 mg/d imatinib 400 mg/d	53	57	Not reported	Not reported
Not reported	91% 400 mg/d 4% <400 mg/d 5% >400mg/d Dose changes in 39% of patients	57	59	36.3% ≥ 1 at diagnosis	Not reported
22.4% previous treatment	dasatinib 85% 100 mg/d imatinib 82% 400 mg/d	63	54	58	59
100% 1st line	98% 600 mg/day	54	53	Not reported	Not reported
45.7% 2nd line 37.7% 3rd line 16.1% 4th line 0.5% 5th line	500 mg/d	Mean: 52	50	Not reported	Not reported
Not reported	81% 400 mg/d 13% 600 mg/d 6% 800 mg/d	40	63	Not reported	15

Table 1 Continued

Article	Study type (setting)	Questionnaires (eligibility)	Patients included	TKI type (n of patients)	TKI treatment duration reported by studies
Kekale 2015 [22]	Cross-sectional (eight secondary and tertiary care hospitals in Finland)	Generic (TKI treatment ≥6 months)	86	dasatinib (9) imatinib (68) nilotinib (9)	Not reported
Nguyen 2022 [29]	Cross-sectional (two Vietnamese centres)	EORTC QLQ-C30 (resistant/ intolerant to imatinib, on nilotinib ≥3 months)	121	nilotinib (121)	Mean 2.06 y
Yu 2019 [23]	Cross-sectional (single centre plus patient advocacy organization, China)	Generic (TKI therapy ≥3 months)	1,142	dasatinib (100) imatinib (859) nilotinib (183)	Median 27 m

Abbreviations: d: day; HCT-CI: haematopoietic cell transplantation comorbidity index; m: month; RCT: randomised controlled trial; w: week; y: year.

	Line of TKI therapy	Starting dose	Age (median)	Male (%)	Comorbidity (%)	Comedication (%)
	54.7% 1st line 29.1% 2nd line 15.1% 3rd line 1.2% 4th line	Not reported	59	52	Median 1 per patient	Median 2 per patient
	100% 2nd line	Not reported	Mean: 47	59	60% ≥1 (21% 1, 24% 2, 15% ≥3)	Not reported
	70% 1st line 30% 2nd/3rd line	Not reported	42	63	26% (15% cardiovascular, 11% other)	Not reported

Meta-analysed prevalence of patient-reported toxicity-symptoms

Eleven studies reported on 47 symptoms of any severity during imatinib, dasatinib, nilotinib or bosutinib use, with a meta-analysed prevalence range of 5.0 to 71.2% across symptoms (overview and detailed forest plots for all meta-analyses in the Supplementary material). The prevalence of the symptom burden of imatinib was reported by 7 different articles on 1,795 patients [18-24]. Fatigue (71.2%, 95% confidence interval [95% CI]: 59.7-82.7%, I^2 = 95.2, five studies), oedema (69.3%, 95% CI: 60.8-77.9%, I^2 = 89.5%, five studies) and muscle soreness (65.2%, 95% CI: 47.2-82.5%, I^2 = 97.8, five studies) were the symptoms with the highest reported prevalence of any severity across studies. Symptoms during dasatinib use were reported by four studies [18, 21-23] with a total population of 233 patients, giving a top three symptom burden of any severity of fatigue (64.1%, 95% CI: 44.5-83.7%, I^2 = 89.7, four studies), frequent urination (53.8%, 95% CI: 45.1-62.6%, I^2 = 0%, two studies) and musculoskeletal pain (52.0%, 95% CI: 26.5-77.5%, I^2 = 92.9%, two studies). Prevalence for nilotinib was described by six articles [18, 22, 23, 25-27] with a total population of 509 patients. Its most frequent symptoms of any severity were fatigue (67.0%, 95% CI: 59.0-75.1], I^2 = 60%, five studies), frequent urination (63.3%, 95% CI: 46.1-80.1%, one study) and itchy skin (53.5%, 95% CI: 44.7-62.3%, I^2 = 42.3%, two studies). Bosutinib was described by 2 articles [19, 28], giving a total population of 450. Only for pain and anxiety/depression of any severity, one or more studies were available, with effect estimates of respectively 37.1% (95% CI: 30.9-43.2%, I^2 = 51.6%, two studies) and 44.1% (95% CI: 24.1-64.1%, I^2 = 96.3%, two studies).

Heterogeneity between studies was high for most symptoms of any severity, resulting in a median heterogeneity and range of 87.3% (0.0 – 97.8) for imatinib, 55.3% (0.0 - 97.6) for dasatinib, 62.0% (0.0 – 90.9) for nilotinib and 73.9% (51.6 - 96.3) concerning bosutinib.

Five studies reported on the proportion of 1,973 patients with moderate-severe symptoms on a four point scale [20, 21, 23, 24, 29] (overview in the Supplementary material). The prevalence of moderate-severe symptoms during imatinib was reported by four different articles on 1,475 patients [20, 21, 23, 24]. Eye problems (34.0%, 95% CI: 24.4-43.6%, one study), frequent urination (33%, 95% CI: 23.6-42.4%, one study), and oedema (31.0%, 95% CI: 23.5-38.5%, I^2 = 86.1%, three studies) were the three most reported moderate-severe symptoms under imatinib. Two studies reported on moderate-severe symptoms in 194 patients on dasatinib [23, 29]. Fatigue (25.3%, 95% CI: -1.2-

57.7%, I^2 = 95.2%, two studies), weight change (24.0%, 95% CI: 15.6-32.4%, one study), and frequent urination (19.0%, 95% CI: 11.0-27.0%, one study) were the three most reported moderate-severe symptoms under dasatinib. The same two studies reported on symptoms during nilotinib, with the top three moderate-severe symptoms in 304 patients being fatigue (25.6%, 95% CI: 16.8-34.4%, I^2 = 68.8%, two studies), itchy skin (24.0%, 95% CI: 17.7-30.3%, one study), and pain (22.0%, 95% CI: 14.6-29.5%, one study). The median heterogeneity of the studies for moderate-severe symptoms was 45.8% (0-96.4) for imatinib, 22.1% (0-95.2) for dasatinib and 41.8% (0-91.1) for nilotinib.

Sensitivity analysis

Excluding the four low-quality studies [22, 25, 27, 28] did not significantly alter meta-analysed effect estimates of the most prevalent symptoms for each TKI (data not shown). Hence, no articles were left out of the meta-analysis.

Differences in symptom prevalence between TKI types

Statistically significant more patients on imatinib experienced abdominal distension, abdominal pain, breast pain/swelling (females), a decrease in sexual desire, diarrhoea, oedema, dry eyes, hair colour change, hypomenorrhea, muscle cramps/soreness, nausea, pain or skin colour change (any severity and/or moderate-severe), when compared to both dasatinib and nilotinib (Table 2). Musculoskeletal pain, vomiting, and weight gain were experienced more by patients on imatinib compared to nilotinib. Breast pain/swelling (females), dry eyes, hypomenorrhea and an itchy skin were experienced more by patients using nilotinib than dasatinib. Moderate-severe memory problems were experienced more frequently by patients on dasatinib, compared to imatinib and nilotinib. More clinically relevant differences, tentatively defined as a magnitude of ≥10% or higher, were identified between TKI types, but none of these differences reached statistical significance of $p<0.01$ (data not shown).

Table 2 Prevalence (%) [with 95% confidence intervals] of symptoms with statistically significant differences across TKI types, based on indirect comparisons between studies

	dasatinib	imatinib	nilotinib
Abdominal distension			
Any severity	19.0 [11.4; 26.6]*	30.0 [26.9; 33.1]*/**	18.0 [12.5; 23.5]**
Abdominal pain			
Any severity	23.4 [13.0; 33.9]	32.4 [21.1; 43.7]**	13.0 [7.8; 18.3]**
Moderate to severe	2.4 [0.2; 4.6]*	6.1 [4.8; 7.3]*/**	3.0 [0.5; 5.6]**
Breast pain/swelling (female)			
Any severity	8.0 [2.7; 13.3]*/***	25.0 [22.1; 27.9]*	20.0 [14.1; 25.9]***
Moderate to severe	2.0 [0;5.3]*	8.0 [6.2; 9.8]*/**	3.0 [0.7; 5.4]**
Decrease in sexual desire			
Any severity	18.0 [10.6; 25.5]*	31.0 [27.9; 34.1]*/**	20.0 [14.1; 25.9]**
Diarrhoea			
Any severity	18.6 [13.6; 23.6]*	39.1 [28.1; 50.2]*	18.3 [0; 38.8]
Moderate to severe	6.6 [0.8; 12.4]*	13.9 [12.1; 15.8]*/**	7.0 [4.1; 9.9]**
Dry eyes			
Moderate to severe	2.0 [0; 4.7]*/***	13.0 [10.8; 15.2]*/**	7.0 [3.3; 10.7]**/***
Oedema			
Any severity	42.3 [38.5;46.1]*	69.3 [60.8; 77.9]*/**	35.6 [22.7; 48.4]**
Moderate to severe	16.8 [10.9; 22.6]*/***	31.0 [23.5; 38.5]*/**	8.0 [4.1; 11.9]**/***
Eye problems (e.g., burning, watery, irritated or dry)			
Moderate to severe	12.0 [5.5; 18.5]*	34.0 [24.0; 43.6]*	No data available
Hair colour change			
Any severity	21.0 [13.0; 29.0]	28.0 [25.1; 30.9]**	15.0 [9.9; 20.1]**
Moderate to severe	6.0 [1.3; 10.7]*	12.0 [9.8; 14.2]*/**	5.0 [1.9; 8.1]**
Hypomenorrhea (female <50y)			
Any severity	7.0 [1.9; 12.1]*/***	30.0 [26.9; 33.1]*/**	17.0 [11.5; 22.5]**/***
Moderate to severe	0.0 [0; 9.8]*	17.0 [14.5; 19.6]*/**	8.0 [4.1; 11.9]**
Itchy skin			
Any severity	32.0 [22.8; 41.2]***	38.0 [34.7; 41.3]**	53.5 [44.7; 62.3]**/***
Memory problems			
Moderate to severe	17 [9.6; 24.5]***	22 [19.3; 24.7]**	6.0 [0.7; 11.3]**/***
Muscle cramps/soreness			
Any severity	27.0 [20.5; 33.5]*	65.2 [47.9; 82.5]*	42.2 [25.0; 59.4]
Moderate to severe	4.9 [1.8; 8.0]*	26.8 [18.1; 35.6]*/**	7.0 [3.3; 10.7]**
Musculoskeletal pain			
Any severity	52.0 [26.5; 77.5]	62.6 [46.6; 78.7]**	34.0 [27.1; 40.9]**
Nausea			

Table 2 Continued

	dasatinib	imatinib	nilotinib
Any severity	22.8 [15.9; 29.6]*	35.5 [28.9; 42.2]*/**	17.6 [13.5; 21.7]**
Pain (not specified)			
Moderate to severe	6.0 [1.1; 10.9]***	10.0 [4.12; 15.9]**	22.0 [14.6; 29.5]**/***
Skin colour change			
Any severity	28.0 [19.2; 36.8]*	50.0 [46.7; 53.3]*/**	23.0 [16.9; 29.1]**
Moderate to severe	6.0 [1.3; 10.7]*	23.0 [20.3; 25.7]*/**	8.0 [4.1; 11.9]**
Vomiting			
Moderate to severe	7.0 [1.9; 12.1]	7.0 [5.2; 8.8]**	2.0 [0; 4.2]**
Weight gain			
Any severity	37.0 [27.6; 46.4]	45.0 [41.7; 48.3]**	28.0 [21.5; 34.5]**

*Significant difference between imatinib and dasatinib; **Significant difference between imatinib and nilotinib; *** Significant difference between dasatinib and nilotinib.

Significantly better compared to other TKI(s)
Both significantly better and worse, compared to other TKIs
Significantly worse compared to other TKI(s)

Meta-regression

Multivariable meta-regression was not possible due to the limited number of studies that provided data on covariates. Similarly, too few data were available on comorbidity and comedication to perform meta-regression, whereas treatment duration could only be taken up in the analysis of fatigue for imatinib. Explorative meta-regression was performed on fatigue and nausea for imatinib and nilotinib and on muscle cramp, oedema and diarrhoea for imatinib only, as these were the only symptoms with data on (some) covariates in at least five studies. Concerning fatigue of any severity, as reported by nilotinib users, median age seemed to explain a large part of the heterogeneity (51.8%). With a co-efficient of -1.30, the estimated prevalence of fatigue decreased when age increased. No other significant changes were identified (data not shown).

Discussion

This study critically investigated the symptom burden during TKI therapy, as reported by chronic phase CML patients. The prevalence of patients who experienced symptoms of any severity was reported for 47 different symptoms and ranged from 5.0% to 69.3% for different TKI types and symptoms. The key

symptom was fatigue, which is in line with previous clinician-reported adverse effects [30]. In an indirect comparison of imatinib, dasatinib, and nilotinib, statistically significant differences in symptom prevalence (any severity) were identified for 13 symptoms, generally favouring dasatinib and nilotinib over imatinib. Only one included study made a direct comparison between TKI types (the BFORE randomised trial), reporting on pain and anxiety/depression of any severity only [19]. Results for these symptoms were similar to other 3

The need to use patient-reported outcomes in CML research is emphasised when we compare the patient-reported symptom burden estimated in this study to the toxicity-burden estimated by studies that applied CTCAE, as reported by physicians (any severity). Our effect estimates are often at the upper end of or even above the range of the prevalence of all-grade toxicity assessed by CTCAE [7, 33], further supporting the idea that physician-assessed toxicity underestimates the symptoms patients experience. Fatigue and oedema, in particular, were reported at significantly lower rates (≥29% (fatigue) and 16% (oedema)) in CTCAE assessments for both dasatinib and nilotinib [7]. Similarly, for imatinib, muscle cramps were reported 23% less frequently [33], compared to our meta-analysed effect estimates.

Our study is the first to quantitatively analyse the prevalence of CML patient-reported symptoms during TKI therapy. Its strength lies in the detailed analysis and variation of patient-reported symptom burden provided by combining data from different studies. In contrast to CTCAE, there is no consensus on how patient-reported outcomes should be measured and reported. Seven out of 11 identified studies used validated instruments, with only four using leukaemia-specific instruments. Two of these four studies used the FACT-LEU, which is not sufficiently comprehensive, as it misses key TKI-related toxicity, such as muscle cramps [34]. However, the other validated CML-specific instrument used by two studies (EORTC QLQ-CML24) does not cover genitourinary symptoms, which appear to be prevalent based on one study using a generic instrument [23]. Unfortunately, this generic instrument lacks sufficient content validity [34]. Therefore, this review does not determine which instrument is best suited to assess TKI-related toxicity in CML.

As for the reporting of outcomes, we had to exclude many studies because of differences in reporting, e.g., 17 studies did not report per TKI type [35-51], while 19 studies reported other outcomes than prevalence [26, 52-69]. To some extent we can compare our findings to twelve of these studies, which

reported mean symptom scores [26, 52, 53, 55, 59, 62, 63, 65-69]. Eight of those twelve studies reported on multiple symptoms, identifying fatigue as the most severe symptom in dasatinib [52, 55, 69]; imatinib [52, 53, 55, 59, 62], nilotinib [26, 52, 53, 55, 65, 69]; as well as for ponatinib [69]. However, for bosutinib one study found that diarrhoea had the highest mean score [53]. This suggests that fatigue is both the most prevalent and the most severe symptom, though this might not apply to bosutinib, for which we had no prevalence data.

In addition, eight of those twelve studies made comparisons between different TKI types: six observational studies made unadjusted comparisons [52, 53, 55, 66, 67, 69], while two were randomised controlled trials [63, 68]. Comparisons were made between asciminib, bosutinib, dasatinib, imatinib, nilotinib and ponatinib. Overall, 22 different treatment groups were identified across these studies with six treatment groups including fewer than 30 patients and only two treatment groups including more than 100 patients. Findings from these eight studies were diverse: some reached statistical significance for certain comparisons, while others did not find significant differences for similar comparisons, or even reported contradictory significant differences. Notably, most studies used questionnaires that evaluated only a few separate symptoms. At present, the variety of questionnaires used and the variability in the reporting of outcomes compromise the comparability of patient-reported symptoms, making it challenging to draw consistent and clinically meaningful findings across studies [70]. Low patient numbers and unadjusted comparisons further contribute to this problem.

Two of those eight studies that compared mean severity scores but could not be included in our meta-analysis, still provide valuable supplementary data and merit discussion. The first is the randomised ASCEMBL trial, which compared asciminib to bosutinib [63]. Reporting adjusted mean differences in symptom scores using a mixed-effects model for repeated measurements, the study found that six out of 20 individual symptom items statistically favoured asciminib over bosutinib (nausea, lack of appetite, feeling drowsy, dry mouth, vomiting, and diarrhoea). However, the differences were small and did not reach the predefined clinical meaningful difference of 15%, except for diarrhoea, which worsened under bosutinib. The second study with supplementary data found that patients on ponatinib had significantly worse mean scores compared to patients on dasatinib and/or nilotinib for skin rash, muscle cramps, dry mouth and distress, disturbed sleep, malaise, swelling of extremities, and shortness of breath [69].

The research field is at the beginning stages of leveraging real-world data, and the availability and quality of such data are still limited [71]. As a starting point, it would be useful if more and larger studies reported on toxicity-symptoms, tabulating outcomes according to TKI type and response given, instead of reporting on quality of life expressed by mean scores. Even though we identified 47 symptoms on which 11 studies reported in this way, 53% (symptoms of any severity) to 70% (moderate-severe symptoms) of all our analyses are based on the results of only one study. Furthermore, data on bosutinib, ponatinib and asciminib were scarce or absent and are especially needed to inform clinical care.

Another limitation is the high heterogeneity of the meta-analysed effect sizes, reflecting either clinical heterogeneity, such as population or treatment differences, and/or methodological heterogeneity, such as different questionnaires. Previous studies have shown significant correlations between patient-reported symptoms and, for example, gender, treatment duration, age, comorbidity and comedication [20, 37, 72]. We identified age as a covariate that explained heterogeneity to a large extent for fatigue during nilotinib use. The prevalence of patient-reported fatigue of any severity decreased with increasing age, possibly due to a 'response shift', a psychological adaptation in which patients either change their internal standards for measuring a concept (what is 'fatigue' may shift during the course of a chronic illness); or redefine fatigue as they age [73]. However, the limited number of studies that provided (consistent) information on pre-determined covariates restricted exploratory analysis: we did not find a similar age-related effect on fatigue prevalence during imatinib or dasatinib use, for example.

In this meta-analysis frequent urination is found to be a high-prevalence symptom for CML patients treated by either of the three analysed TKIs, although it has not been described as a side effect of TKIs before. Besides it possibly being an effect of TKI treatment, feasible explanations could be that it is a consequence of oedema during TKI treatment, or due to comorbidity. Due to limited information on covariates, this could not be further specified. An individual-patient data meta-analysis of existing studies might shed more light on covariates. Future studies should incorporate consistent and standardised information on covariates, to help unravel the impact they have on the patient symptom-experience in the real-world.

Conclusion

Low-grade patient-reported symptom burden during TKI usage is high, with significant differences between TKI types for a third of reported symptoms. Though evidence is mainly indirect and unadjusted for covariates, this is the most in-depth overview of patients' experiences available to the best of our knowledge. These findings are a prerequisite for shared decision-making, when discussing treatment choices with patients. Future real-world studies should focus on direct comparisons between different TKI types, adjusted for covariates, including asciminib, bosutinib, and ponatinib.

References

1. Yeung, D.T., N. Shanmuganathan, and T.P. Hughes, *Asciminib: a new therapeutic option in chronic-phase CML with treatment failure.* Blood, 2022. **139**(24): p. 3474-3479.
2. Sasaki, K., et al., *Relative survival in patients with chronic-phase chronic myeloid leukaemia in the tyrosine-kinase inhibitor era: analysis of patient data from six prospective clinical trials.* Lancet Haematol, 2015. **2**(5): p. e186-93.
3. Bower, H., et al., *Life Expectancy of Patients With Chronic Myeloid Leukemia Approaches the Life Expectancy of the General Population.* J Clin Oncol, 2016. **34**(24): p. 2851-7.
4. Cella, D., C.J. Nowinski, and O. Frankfurt, *The impact of symptom burden on patient quality of life in chronic myeloid leukemia.* Oncology, 2014. **87**(3): p. 133-47.
5. Jiang, Q., et al., *Variables associated with patient-reported outcomes in persons with chronic myeloid leukemia receiving tyrosine kinase-inhibitor therapy.* J Cancer Res Clin Oncol, 2017. **143**(6): p. 1013-1022.
6. Eliasson, L., et al., *Exploring chronic myeloid leukemia patients' reasons for not adhering to the oral anticancer drug imatinib as prescribed.* Leuk Res, 2011. **35**(5): p. 626-30.
7. Caldemeyer, L., et al., *Long-Term Side Effects of Tyrosine Kinase Inhibitors in Chronic Myeloid Leukemia.* Curr Hematol Malig Rep, 2016. **11**(2): p. 71-9.
8. Hochhaus, A., et al., *European LeukemiaNet 2020 recommendations for treating chronic myeloid leukemia.* Leukemia, 2020. **34**(4): p. 966-984.
9. Efficace, F., et al., *Patient- versus physician-reporting of symptoms and health status in chronic myeloid leukemia.* Haematologica, 2014. **99**(4): p. 788-93.
10. Geelen, I.G.P., et al., *Treatment outcome in a population-based, 'real-world' cohort of patients with chronic myeloid leukemia.* Haematologica, 2017. **102**(11): p. 1842-1849.
11. Page, M.J., et al., *The PRISMA 2020 statement: an updated guideline for reporting systematic reviews.* Rev Esp Cardiol (Engl Ed), 2021. **74**(9): p. 790-799.
12. Modesti, P.A., et al., *Panethnic Differences in Blood Pressure in Europe: A Systematic Review and Meta-Analysis.* PLoS One, 2016. **11**(1): p. e0147601.
13. Grant, M.C., et al., *The prevalence of symptoms in 24,410 adults infected by the novel coronavirus (SARS-CoV-2; COVID-19): A systematic review and meta-analysis of 148 studies from 9 countries.* PLoS One, 2020. **15**(6): p. e0234765.
14. Migliavaca, C.B., et al., *Quality assessment of prevalence studies: a systematic review.* J Clin Epidemiol, 2020. **127**: p. 59-68.
15. Herzog, R., et al., *Are healthcare workers' intentions to vaccinate related to their knowledge, beliefs and attitudes? A systematic review.* BMC Public Health, 2013. **13**: p. 154.
16. Ioannidis, J.P., N.A. Patsopoulos, and E. Evangelou, *Uncertainty in heterogeneity estimates in meta-analyses.* BMJ, 2007. **335**(7626): p. 914-6.
17. Borenstein, M., et al., *Introduction to Meta-Analaysis*. Vol. 1. 2009: John Wiley & Sons.
18. Bostan, H., et al., *Quality of Life and Symptom Burden With First- and Second-generation Tyrosine Kinase Inhibitors in Patients With Chronic-phase Chronic Myeloid Leukemia.* Clin Lymphoma Myeloma Leuk, 2020. **20**(12): p. 836-842.
19. Cortes, J.E., et al., *Patient-reported outcomes in the phase 3 BFORE trial of bosutinib versus imatinib for newly diagnosed chronic phase chronic myeloid leukemia.* J Cancer Res Clin Oncol, 2019. **145**(6): p. 1589-1599.

20. Efficace, F., et al., *Health-related quality of life in chronic myeloid leukemia patients receiving long-term therapy with imatinib compared with the general population.* Blood, 2011. **118**(17): p. 4554-60.
21. Efficace, F., et al., *Health-related quality of life of newly diagnosed chronic myeloid leukemia patients treated with first-line dasatinib versus imatinib therapy.* Leukemia, 2020. **34**(2): p. 488-498.
22. Kekale, M., M. Peltoniemi, and M. Airaksinen, *Patient-reported adverse drug reactions and their influence on adherence and quality of life of chronic myeloid leukemia patients on per oral tyrosine kinase inhibitor treatment.* Patient Prefer Adherence, 2015. **9**: p. 1733-40.
23. Yu, L., et al., *Variables associated with patient-reported symptoms in persons with chronic phase chronic myeloid leukemia receiving tyrosine kinase inhibitor therapy.* Medicine (Baltimore), 2019. **98**(48): p. e18079.
24. Kapoor, J., et al., *Factors influencing adherence to imatinib in Indian chronic myeloid leukemia patients: A cross-sectional study.* Mediterranean Journal of Hematology and Infectious Diseases, 2015. **7(1) (e2015013)**.
25. Boons, C., et al., *Response and Adherence to Nilotinib in Daily practice (RAND study): an in-depth observational study of chronic myeloid leukemia patients treated with nilotinib.* Eur J Clin Pharmacol, 2020. **76**(9): p. 1213-1226.
26. Efficace, F., et al., *Health-related quality of life in patients with chronic myeloid leukemia receiving first-line therapy with nilotinib.* Cancer, 2018. **124**(10): p. 2228-2237.
27. Huguet, F., et al., *Nilotinib efficacy, safety, adherence and impact on quality of life in newly diagnosed patients with chronic myeloid leukaemia in chronic phase: a prospective observational study in daily clinical practice.* Br J Haematol, 2019. **187**(5): p. 615-626.
28. Kantarjian, H.M., et al., *Long-term patient-reported outcomes from an open-label safety and efficacy study of bosutinib in Philadelphia chromosome-positive chronic myeloid leukemia patients resistant or intolerant to prior therapy.* Cancer, 2018. **124**(3): p. 587-595.
29. Nguyen, C.T.T., et al., *Quality of life among chronic myeloid leukemia patients in the second-line treatment with nilotinib and influential factors.* Qual Life Res, 2022. **31**(3): p. 733-743.
30. Efficace, F., et al., *Chronic fatigue is the most important factor limiting health-related quality of life of chronic myeloid leukemia patients treated with imatinib.* Leukemia, 2013. **27**(7): p. 1511-9.
31. Thanarajasingam, G., et al., *Reaching beyond maximum grade: progress and future directions for modernising the assessment and reporting of adverse events in haematological malignancies.* Lancet Haematol, 2022. **9**(5): p. e374-e384.
32. Dahlen, T., et al., *Adverse outcomes in chronic myeloid leukemia patients treated with tyrosine kinase inhibitors: Follow-up of patients diagnosed 2002-2017 in a complete coverage and nationwide agnostic register study.* Am J Hematol, 2022. **97**(4): p. 421-430.
33. Adattini, J.A., et al., *Real-world efficacy and safety outcomes of imatinib treatment in patients with chronic myeloid leukemia: An Australian experience.* Pharmacol Res Perspect, 2022. **10**(5): p. e01005.
34. Smit, Y., et al., *Measuring chronic myeloid leukaemia TKI-related toxic effects in the real world: a systematic review and critical assessment of content validity of patient-reported outcome measures.* Lancet Haematol, 2023. **10**(10): p. e849-e859.

35. Atallah, E., et al., *Assessment of Outcomes After Stopping Tyrosine Kinase Inhibitors Among Patients With Chronic Myeloid Leukemia: A Nonrandomized Clinical Trial.* JAMA Oncol, 2021. **7**(1): p. 42-50.
36. Efficace, F., et al., *International development of an EORTC questionnaire for assessing health-related quality of life in chronic myeloid leukemia patients: the EORTC QLQ-CML24.* Qual Life Res, 2014. **23**(3): p. 825-36.
37. Efficace, F., et al., *Validation and reference values of the EORTC QLQ-CML24 questionnaire to assess health-related quality of life in patients with chronic myeloid leukemia.* Leuk Lymphoma, 2021. **62**(3): p. 669-678.
38. Jiang, Q., et al., *Young age and high cost are associated with future preference for stopping tyrosine kinase inhibitor therapy in Chinese with chronic myeloid leukemia.* J Cancer Res Clin Oncol, 2016. **142**(7): p. 1539-47.
39. Menon, H., *Non-adherence to CML therapy and its clinical implications in India.* Natl Med J India, 2017. **30**(3): p. 142-147.
40. Oswald, L.B., et al., *Correlates of fatigue severity in patients with chronic myeloid leukemia treated with targeted therapy.* Support Care Cancer, 2022. **30**(1): p. 87-94.
41. Phillips, K.M., et al., *Quality of life outcomes in patients with chronic myeloid leukemia treated with tyrosine kinase inhibitors: a controlled comparison.* Support Care Cancer, 2013. **21**(4): p. 1097-103.
42. Piekarska, A., et al., *Quality of life in CML patients after HCT pretreated with second generation tyrosine kinase inhibitors.* Acta Haematologica Polonica, 2015. **46(Supplement 1)**: p. 193-194.
43. Sharf, G., et al., *Treatment-free remission in chronic myeloid leukemia: the patient perspective and areas of unmet needs.* Leukemia, 2020. **34**(8): p. 2102-2112.
44. Shi, D., et al., *Variables associated with self-reported anxiety and depression symptoms in patients with chronic myeloid leukemia receiving tyrosine kinase inhibitor therapy.* Leuk Lymphoma, 2021. **62**(3): p. 640-648.
45. Tsai, Y.F., et al., *Side effects and medication adherence of tyrosine kinase inhibitors for patients with chronic myeloid leukemia in Taiwan.* Medicine (Baltimore), 2018. **97**(26): p. e11322.
46. Zheng, F., et al., *Health-related quality of life in children with chronic myeloid leukemia in the chronic phase.* J Cancer Res Clin Oncol, 2022. **148**(2): p. 341-350.
47. Foulon, S., et al., *Health state utility and quality of life measures in patients with chronic myeloid leukemia in France.* Qual Life Res, 2021. **30**(7): p. 2021-2032.
48. Chen, H., et al., *Patient Versus Physician Perspective in the Management of Chronic Myeloid Leukemia During Treatment with Tyrosine Kinase Inhibitors.* Oncology and Therapy, 2024. **12(1)**: p. 131-145.
49. Gibek, K., T. Sacha, and K. Cyranka, *Side effects of treatment with tyrosine kinase inhibitors in patients with chronic myeloid leukemia and the occurrence of anxiety symptoms.* Psychiatr Pol, 2024. **58**(1): p. 25-38.
50. Tadesse, F., et al., *Health-Related Quality of Life and Financial Burden in Ethiopian Patients With Chronic Myeloid Leukemia Receiving Tyrosine Kinase Inhibitors: A Cross-Sectional Study.* JCO Glob Oncol, 2024. **10**: p. e2400281.
51. Gibek, K., *Side effects of treatment with tyrosine kinase inhibitors in patients with chronic myeloid leukaemia and the occurrence of depressive symptoms.* Wspolczesna Onkologia, 2023. **27(4)**: p. 277-283.

52. Abulafia, A.S., et al., *Health-related quality of life in patients with chronic myeloid leukemia treated with first-versus second-generation tyrosine kinase inhibitors.* Journal of Clinical Medicine, 2020. **9(11)**: p. 1-9.
53. Ali, H., Z. Mohamed, and D. Abdullah, *Health-related quality of life in chronic myeloid leukaemia patients receiving long-term therapy with different tyrosine kinase inhibitors in Kurdistan region.* Georgian Med News, 2024(354): p. 173-180.
54. Aziz, Z., et al., *Assessment of quality of life with imatinib mesylate as first-line treatment in chronic phase-chronic myeloid leukemia.* Leuk Lymphoma, 2011. **52**(6): p. 1017-23.
55. Cho, H.J., et al., *Real-world experience in the management of chronic myeloid leukemia patients focused on tyrosine kinase inhibitors intolerance and health-related quality of life.* Acta Haematol, 2025: p. 1-14.
56. Hahn, E.A. and G.A. Glendenning, *Quality of life on imatinib.* Semin Hematol, 2003. **40**(2 Suppl 2): p. 31-6.
57. Hahn, E.A., et al., *Quality of life in patients with newly diagnosed chronic phase chronic myeloid leukemia on imatinib versus interferon alfa plus low-dose cytarabine: results from the IRIS Study.* J Clin Oncol, 2003. **21**(11): p. 2138-46.
58. Isfort, S., et al., *Step-in dosing of bosutinib in pts with chronic phase chronic myeloid leukemia (CML) after second-generation tyrosine kinase inhibitor (TKI) therapy: results of the Bosutinib Dose Optimization (BODO) Study.* Ann Hematol, 2023. **102**(10): p. 2741-2752.
59. Jain, P., et al., *Comparative study for the efficacy, safety and quality of life in patients of chronic myeloid leukemia treated with Imatinib or Hydroxyurea.* J Res Pharm Pract, 2013. **2**(4): p. 156-61.
60. Minami, Y., et al., *Asciminib in Patients With CML-CP Previously Treated With ≥ 2 Tyrosine Kinase Inhibitors: 96-Week Results From the Japanese Subgroup Analysis of the ASCEMBL Study.* Int J Hematol, 2024. **120**(3): p. 305-313.
61. Mo, X.D., et al., *Health-related quality of life of patients with newly diagnosed chronic myeloid leukemia treated with allogeneic hematopoietic SCT versus imatinib.* Bone Marrow Transplant, 2014. **49**(4): p. 576-80.
62. Rashid, A.A. and Z. Albasry, *Assessing the quality of life in imatinib-treated chronic myeloid leukemia patients: A cross sectional descriptive study.* International Journal of Pharmaceutical Research, 2020. **12(Supplementry 2)**: p. 2784-2791.
63. Réa, D., et al., *Health-related quality of life of patients with resistant/intolerant chronic phase chronic myeloid leukemia treated with asciminib or bosutinib in the phase 3 ASCEMBL trial.* Leukemia, 2023. **37**(5): p. 1060-1067.
64. Rosti, G., et al., *Impact of age and comorbidities on the efficacy and tolerability of bosutinib in previously treated patients with chronic myeloid leukemia: results from the phase 4 BYOND study.* Leukemia, 2024. **38**(1): p. 126-135.
65. Sacha, T., et al., *Quality of Life and Adherence to Therapy in Patients With Chronic Myeloid Leukemia Treated With Nilotinib as a Second-Line Therapy: A Multicenter Prospective Observational Study.* Clin Lymphoma Myeloma Leuk, 2017. **17**(5): p. 283-295.
66. Uyanik, M.S., et al., *The impact of new generic formulations of imatinib mesylate on general quality of life in chronic phase chronic myeloid leukaemia patients.* Eur J Cancer Care (Engl), 2017. **26**(2).
67. Williams, L.A., et al., *Measuring the symptom burden associated with the treatment of chronic myeloid leukemia.* Blood, 2013. **122**(5): p. 641-7.

68. Yu, L., et al., *Achieving optimal response at 12 months is associated with a better health-related quality of life in patients with chronic myeloid leukemia: a prospective, longitudinal, single center study.* BMC Cancer, 2018. **18**(1): p. 782.
69. Zulbaran-Rojas, A., et al., *A prospective analysis of symptom burden for patients with chronic myeloid leukemia in chronic phase treated with frontline second- and third-generation tyrosine kinase inhibitors.* Cancer Med, 2018. **7**(11): p. 5457-5469.
70. Schoenbeck, K.L. and K.E. Flynn, *Health-Related Quality of Life of Patients with Chronic Myeloid Leukemia as Measured by Patient-Reported Outcomes: Current State and Future Directions.* Curr Hematol Malig Rep, 2021. **16**(6): p. 491-499.
71. Hermans, S.J.F., et al., *Externally Controlled Studies Using Real-World Data in Patients With Hematological Cancers: A Systematic Review.* JAMA Oncol, 2024. **10**(10): p. 1426-1436.
72. Janssen, L., et al., *Fatigue in chronic myeloid leukemia patients on tyrosine kinase inhibitor therapy: predictors and the relationship with physical activity.* Haematologica, 2021. **106**(7): p. 1876-1882.
73. Sprangers, M.A.G. and C.E. Schwartz, *Integrating response shift into health-related quality-of-life research: a theoretical model*, in *Adaptation to changing health: response shift in quality-of-life research*, M.A.G. Sprangers and C.E. Schwartz, Editors. 2000, American Psychological Association. p. 11-23.

Supplementary material

Search strategy

The search strategy included terms and synonyms for 'chronic myeloid leukaemia', 'protein kinase inhibitors' and different TKI brand names. Case reports, editorials and phase I clinical trials were excluded in the search string and the language of the articles was restricted to English. Phase I clinical trials were excluded as these are dose finding studies: they usually do not evaluate patient-reported symptoms, and when they do, they would only involve small numbers of patients over a very short period of time.

	Search string	Citations January 2022	Citations July 2023	Citations February 2025
PubMed				
#1	"Leukemia, Myelogenous, Chronic, BCR-ABL Positive"[MeSH] OR "chronic myelogenous leukemia" OR "chronic myeloid leukemia" OR "Ph1 positive chronic myelogenous" OR "Ph1-positive chronic myelogenous" OR "Ph1 positive chronic myelogenous" OR "Ph1-positive chronic myelogenous" OR "Philadelphia positive chronic myeloid leukemia" OR "Philadelphia-positive chronic myeloid leukemia" OR "Chronic myelocytic leukemia" OR "CML"	35,356	37,107	38,951
#2	"Protein Kinase Inhibitors"[MeSH] OR "protein kinase inhibitor" OR "protein kinase inhibitors" OR "TKI" OR "TKIs" OR "TKI's" OR "imatinib" OR "imatinib mesylate"[MeSH] OR "Glivec" OR "nilotinib" OR "Tasigna" OR "dasatinib" OR "sprycel" OR "bosutinib" OR "bosulif" OR "ponatinib" OR "iclusig" OR "asciminib"	134,109	150,783	95,356
#3	#1 AND #2	10,173	10,966	11,659
#4	"case reports"[Publication Type] OR "clinical trial, phase i"[Publication Type] OR "editorial"[Publication Type]			
#5	#3 NOT #4	8,299	8,954	9,521
#6	humans[Filter] AND english[Filter]			
#7	#5 AND #6	**6,562**	7,104	7,471
#8	Filters: from 2021 - 2023(search July 2023 only)		**806**	
#9	Filters: from 2023 – 2025 (search February 2025 only)			**630**

Table Continued

	Search string	Citations January 2022	Citations July 2023	Citations February 2025
EMBASE OVID				
1	chronic myeloid leukemia.mp. or exp chronic myeloid leukemia/	50,577	54,279	57,897
2	cml.mp.	31,136	33,384	35,486
3	1 or 2	58,604	62,806	66,930
4	tyrosine kinase inhibitor.mp. or exp protein tyrosine kinase inhibitor/	340,117	411,877	466,369
5	tki.mp.	23,153	27,470	31,374
6	(imatinib or glivec).mp. or exp imatinib/	47,035	50,929	54,170
7	(nilotinib or tasigna).mp. or exp nilotinib/	10,425	11,733	13,041
8	(dasatinib or sprycel). mp. or exp dasatinib/	15,937	18,282	20,673
9	(bosutinib or bosulif). mp. or exp bosutinib/	3,070	3,697	4,272
10	(ponatinib or iclusig). mp. or exp ponatinib/	3,611	4,519	5,455
11	asciminib.mp. or exp asciminib/ [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword heading word, floating subheading word, candidate term word]	204	422	748
12	4 or 5 or 6 or 7 or 8 or 9 or 10 or 11	343,570	415,833	470,885
13	3 and 12	24,939	27,145	29,177
14	exp note/ or editorial/	1,510,164	1,637,632	1,743,631
15	13 not 14	23,692	25,827	27,785
16	limit 15 to (human and english language and exclude medline journals)	**2,507**		
16	Limit 15 to (human and English language and " remove medline records" and yr="2021 -Current" (July 2023 search only)		**1,342**	
16	Limit 15 to (human and English language and "remove medline records" and yr="2023 -Current" (Feb 2025 search only)			**1,001**
	PubMed and Embase combined	**9,069**	**2,148**	**1,631**
	Identified through reference tracking	1	0	1
	Subtotal	9,070	2,148	1,632
	De-duplication between PubMed and Embase records	8,988	2,039	1,632

Table Continued

Search string	Citations January 2022	Citations July 2023	Citations February 2025
Duplicates between January 2022 and July 2023 searches removed (July 2023 search only)		1,465	
Duplicates between July 2023 and Feb 2025 searches removed (Feb 2025 search only)			1,302
Screened on title and abstract	**8,988**	**1,465**	**1,302**
Excluded on title and/abstract	8,566	1,449	1,263
Screened full text	**422**	**16**	**39**
Excluded after full text screening:	412	15	39
No data on specific symptoms	*21*	*1*	*0*
No (original) patient-reported data	*365*	*10*	*25*
No data per TKI	*11*	*2*	*4*
No prevalence data	*13*	*1*	*5*
Patient reported symptoms after/during TKI withdrawal	*1*	*1*	*4*
Only severe symptoms reported	*1*	*0*	*0*
Only data on top 3 symptoms per patient	*0*	*0*	*1*
Included after full text screening:	**10**	**1**	**0**
Asciminib	*0*	*0*	*0*
Bosutinib	*2*	*0*	*0*
Dasatinib	*4*	*0*	*0*
Imatinib	*7*	*0*	*0*
Nilotinib	*5*	*1*	*0*
Ponatinib	*0*	*0*	*0*

Critical appraisal

Critical appraisal of included studies, according to adapted Newcastle Ottawa Scale

Study	Selection		
	Representativeness sample (max 1 star)	Sample size (max 1 star)	Non-respondents (max 1 star)
Boons 2020		*	*
Bostan 2020		*	*
Cortes 2019	*	*	*
Efficace 2011	*	*	*
Efficace 2020	*	*	*
Huguet 2019		*	
Kantarjian 2018		*	
Kapoor 2015		*	*
Kekale 2015		*	*
Nguyen 2022	*	*	
Yu 2019	*	*	*

Abbreviation: n.a.: not applicable

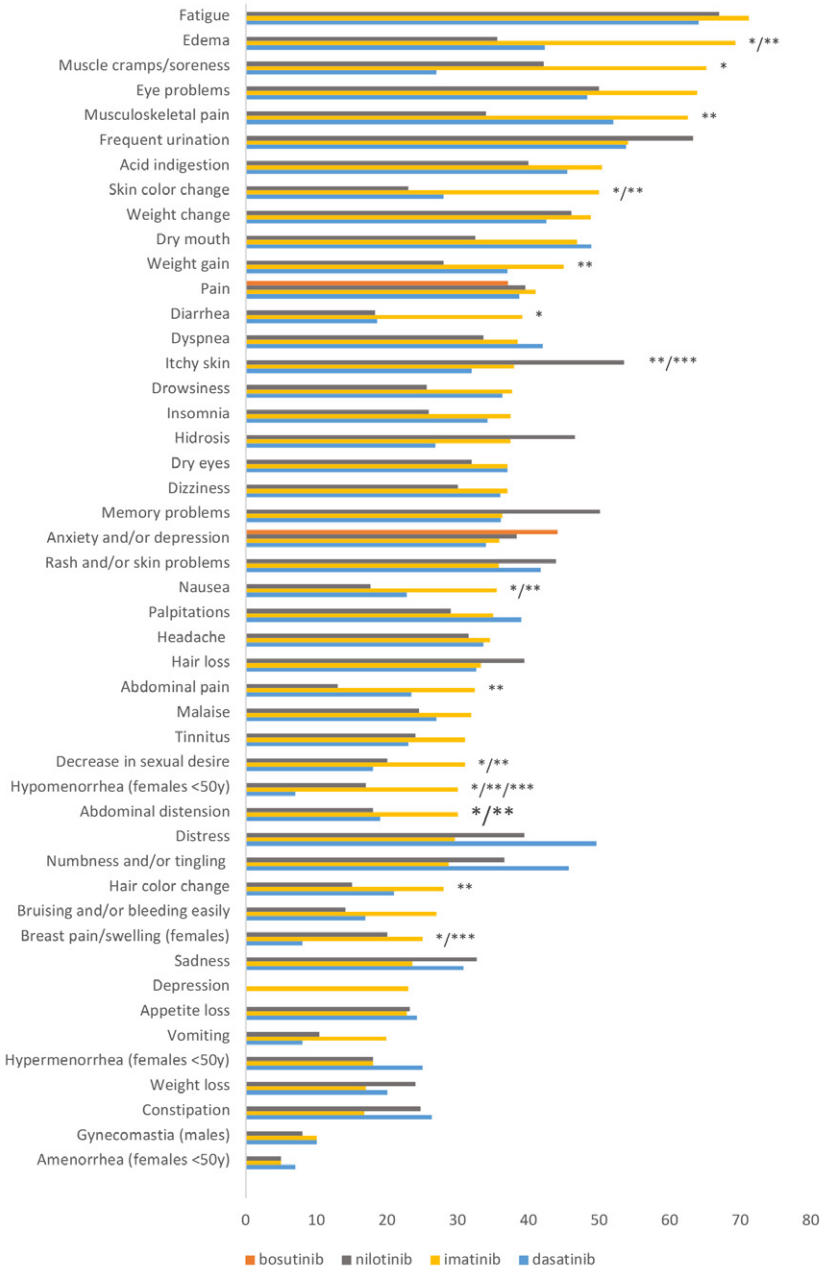
Symptom definitions

In so far as symptom definitions are not obvious, included terms are described here.

General term	Included terms
Bruising/bleeding easily	Bruising easily; easy bruising/bleeding; haemorrhagic tendency of skin
Oedema	Swelling of ankles, legs or around eyes; swelling of hands, legs, feet, or around eyes; periorbital and lower limb oedema; oedema
Rash and/or skin problems	Rash/skin changes; rash; skin problems
Pain	Pain; pain and discomfort

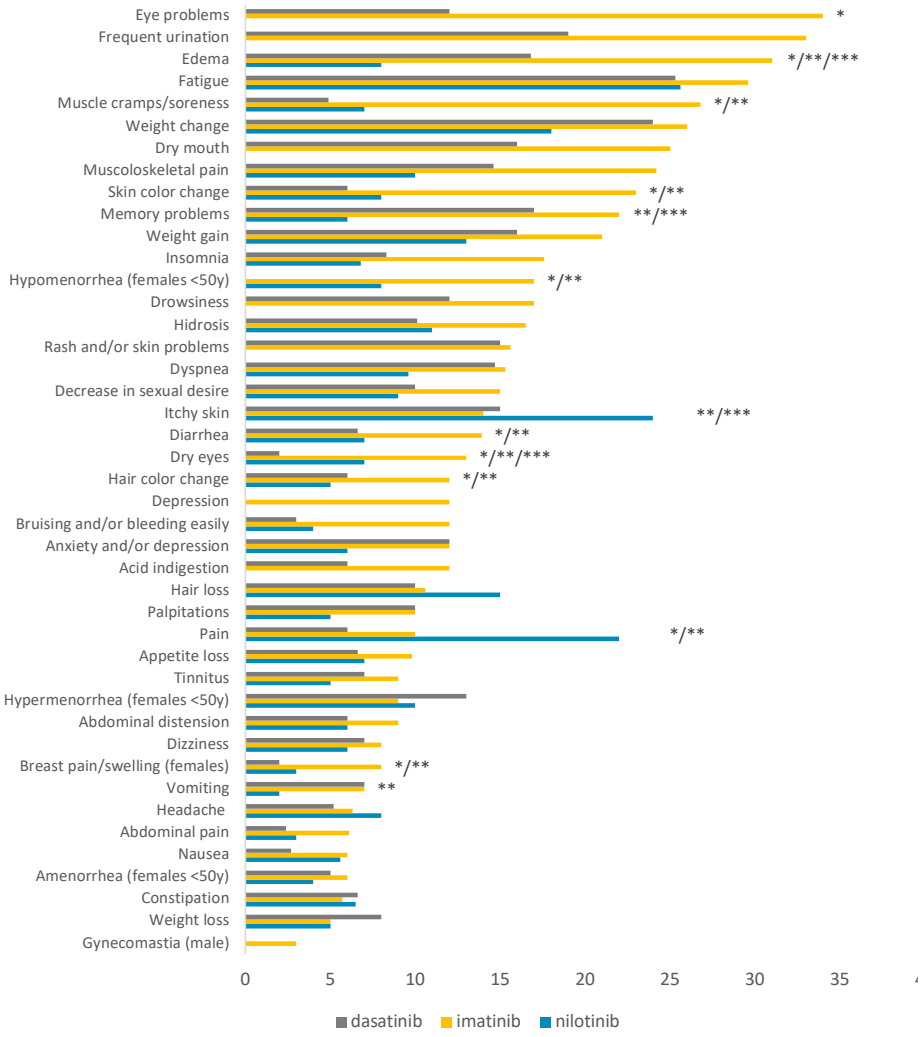
	Comparability	Outcomes		Total
Ascertainment of exposure (max 2 stars)	Control confounding factors	Assessment of outcome (max 2 stars)	Statistical test	0-7 stars
*	n.a.	*	n.a.	4
**	n.a.	*	n.a.	5
**	n.a.	*	n.a.	6
**	n.a.	*	n.a.	6
**	n.a.	*	n.a.	6
**	n.a.	*	n.a.	4
**	n.a.	*	n.a.	4
**	n.a.	*	n.a.	5
*	n.a.	*	n.a.	4
**	n.a.	*	n.a.	5
*	n.a.	*	n.a.	5

Figure 1 Meta-analysed effect estimates (weighted average) of the prevalence of 47 patient-reported symptoms of any severity, reported by 2,987 patients across 11 studies on bosutinib, dasatinib, imatinib, and nilotinib



*Significant difference between dasatinib and imatinib; **Significant difference between imatinib and nilotinib; *** Significant difference between dasatinib and nilotinib

Figure 2 Meta-analysed effect estimates (weighted average) of the prevalence of 43 patient-reported symptoms of moderate to severe intensity, reported by 1,973 patients across five studies on dasatinib, imatinib and nilotinib



*Significant difference between dasatinib and imatinib; **Significant difference between imatinib and nilotinib; *** Significant difference between dasatinib and nilotinib

Figure 3 Forest plot of the meta-analyses for fatigue (any severity)

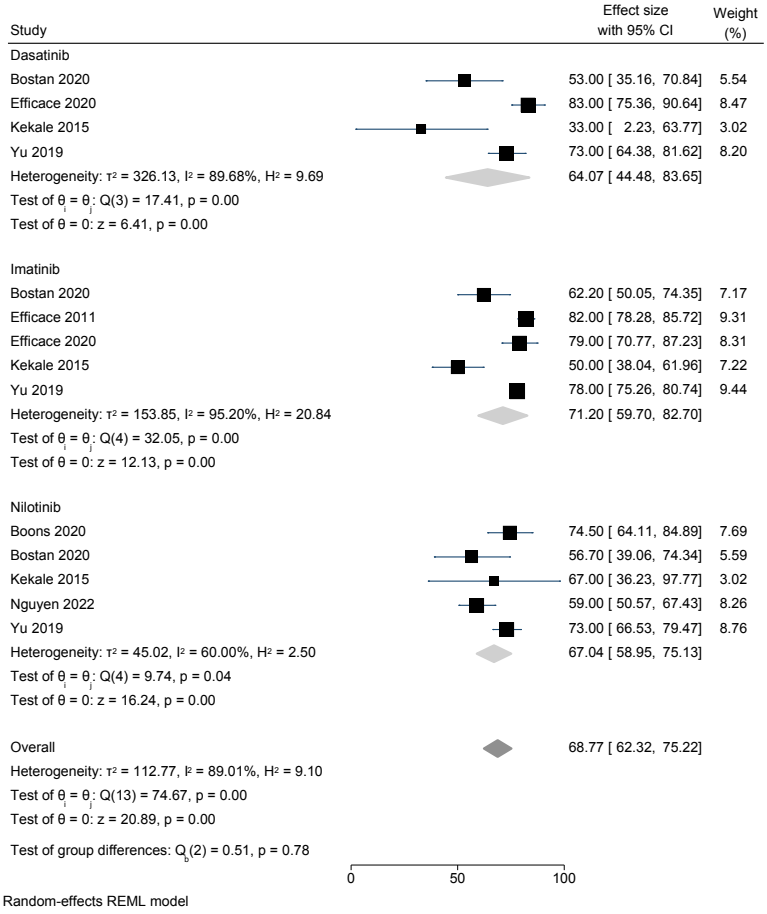


Figure 4 Forest plot of the meta-analyses for oedema (any severity)

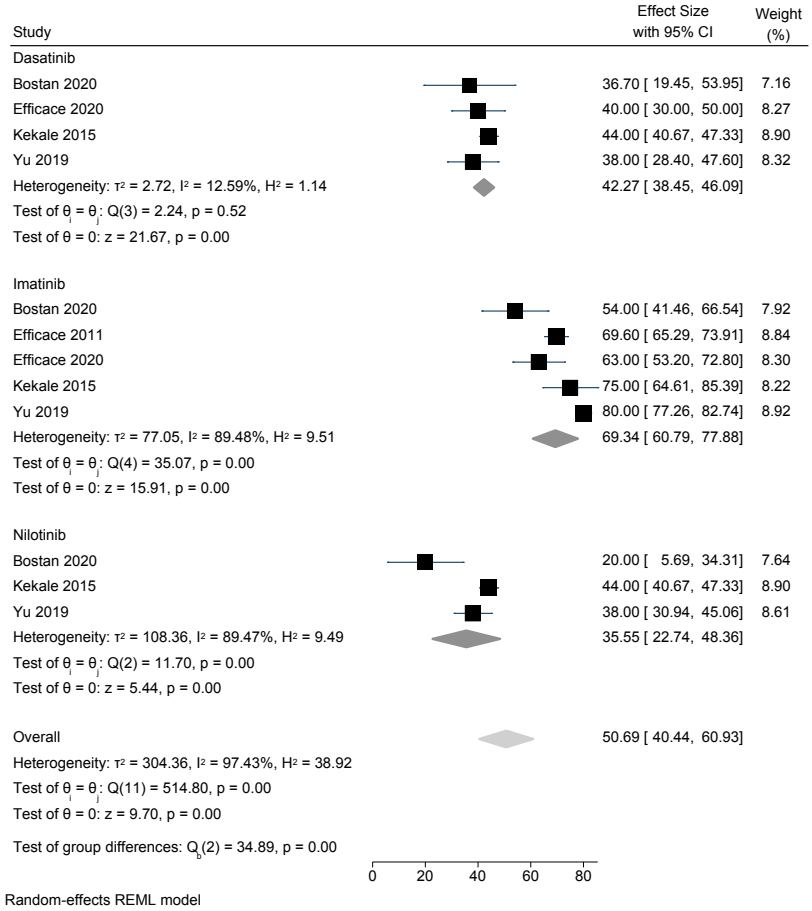


Figure 5 Forest plot of the meta-analyses for muscle soreness and cramps (any severity)

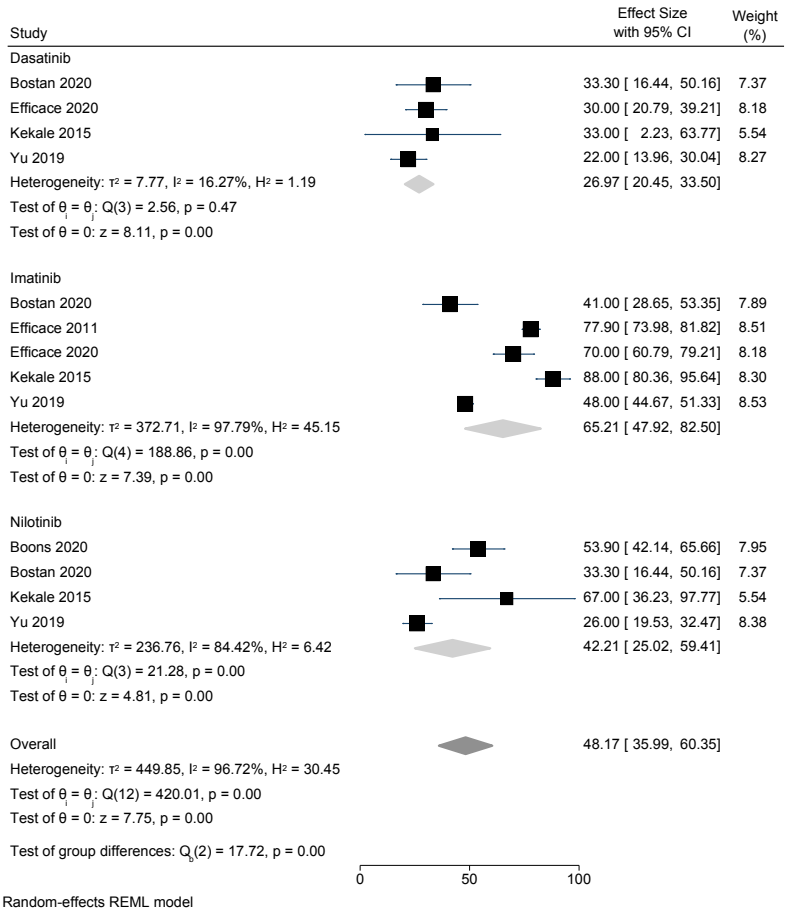


Figure 6 Forest plot of the meta-analyses for eye problems (any severity)

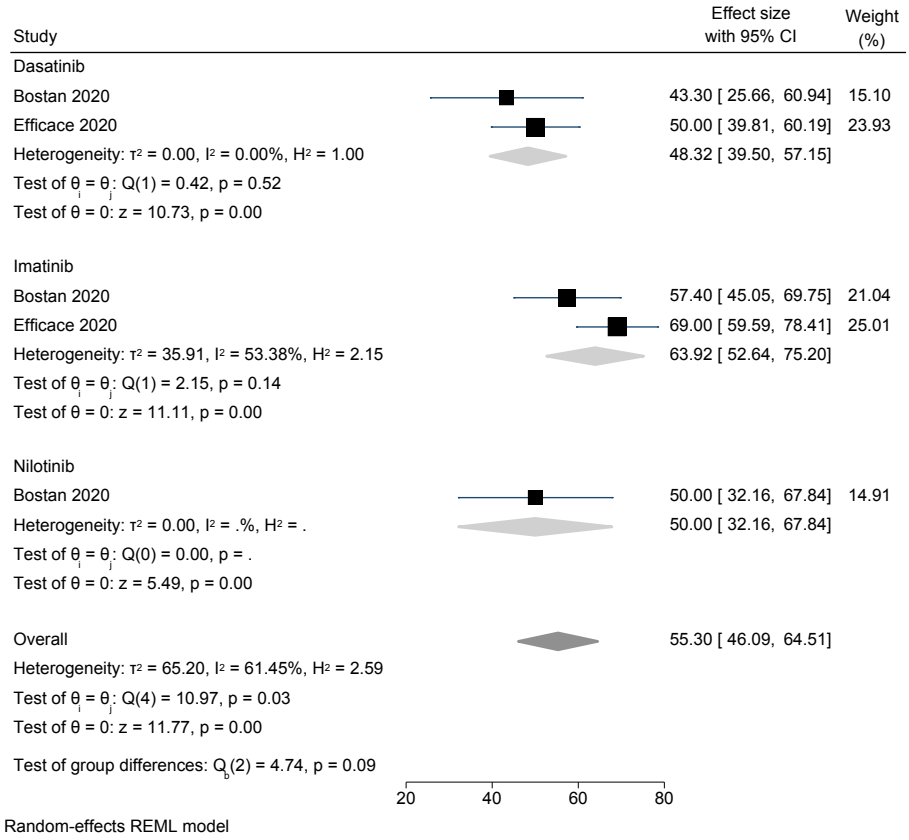


Figure 7 Forest plot of the meta-analyses for musculoskeletal pain (any severity)

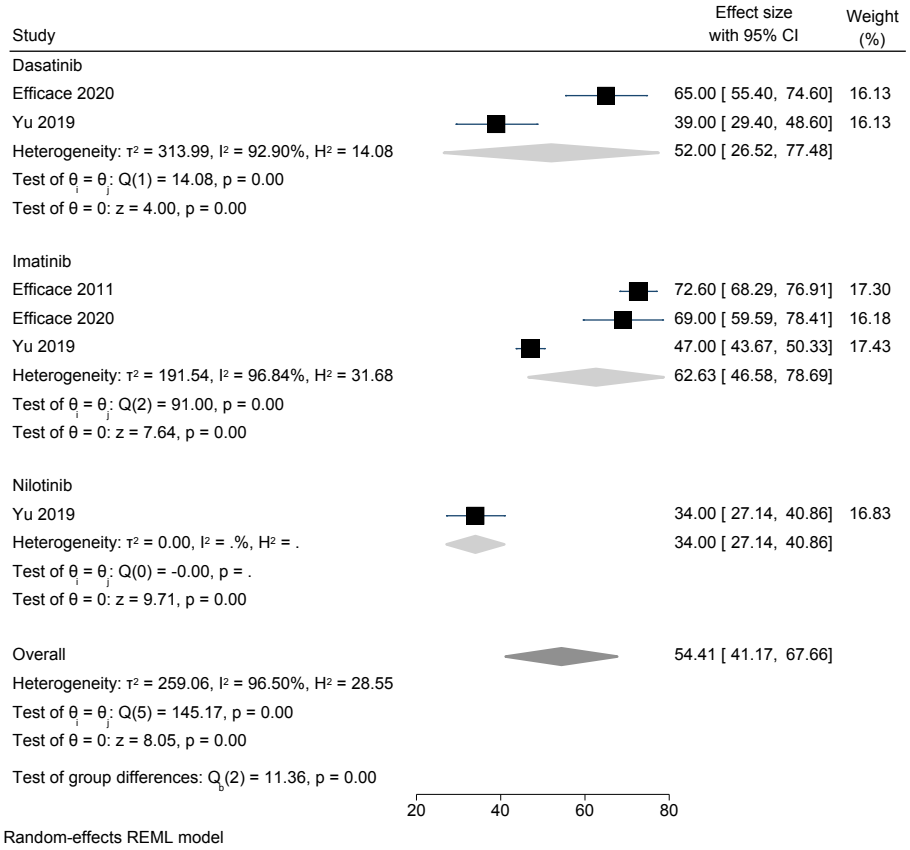


Figure 8 Forest plot of the meta-analyses for frequent urination (any severity)

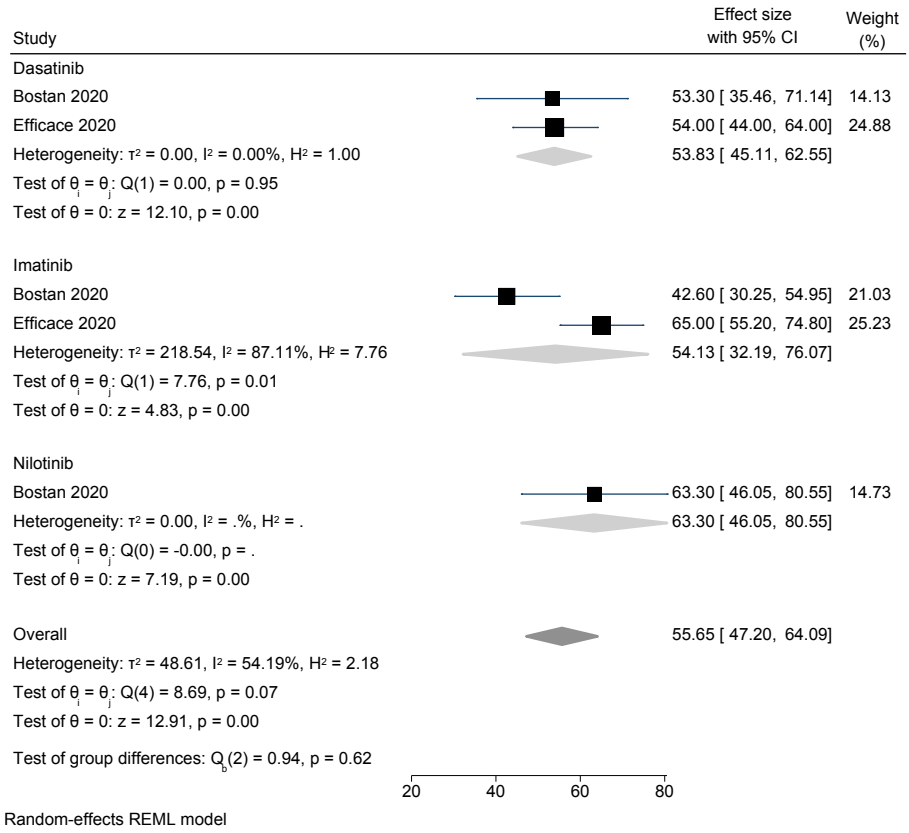


Figure 9 Forest plot of the meta-analyses for acid indigestion (any severity)

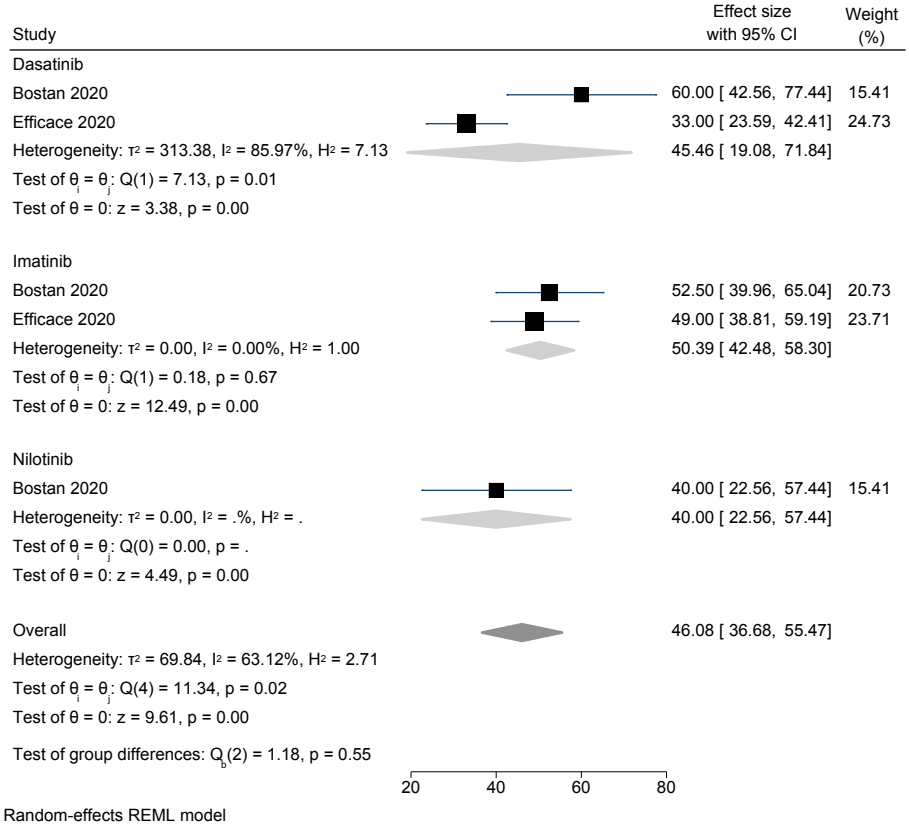


Figure 10 Forest plot of the meta-analyses for skin colour change (any severity)

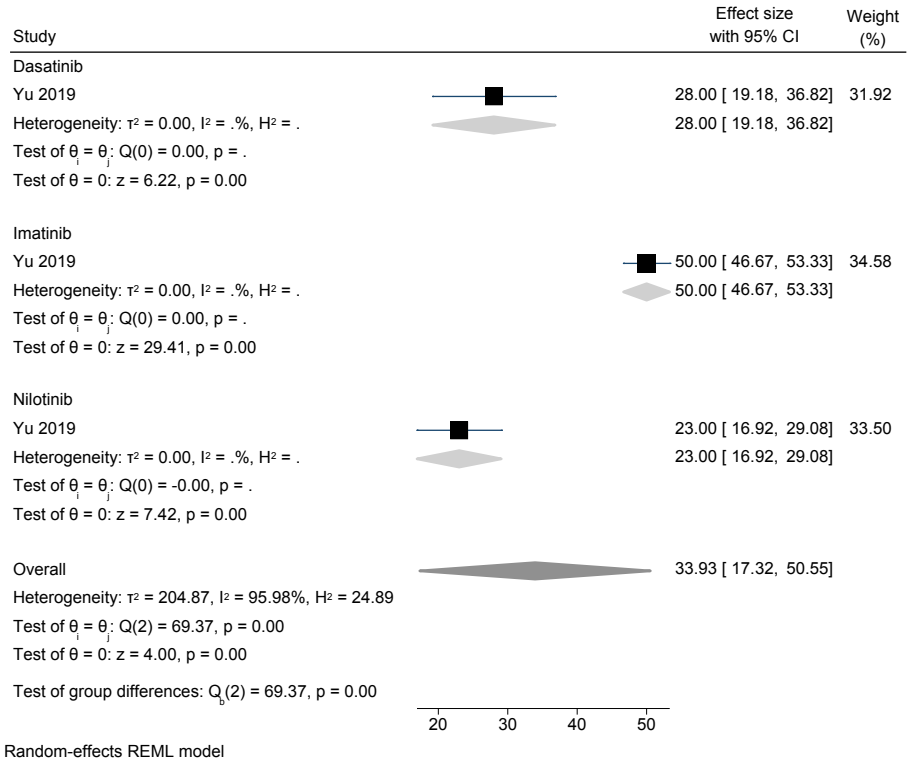


Figure 11 Forest plot of the meta-analyses for weight change (any severity)

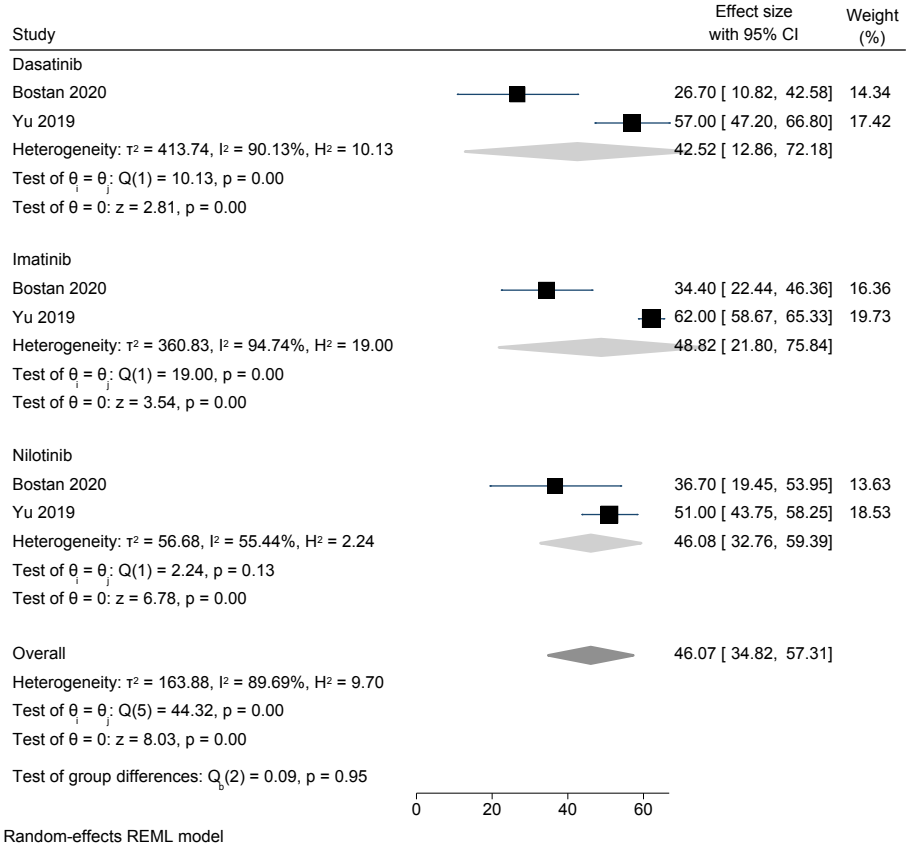


Figure 12 Forest plot of the meta-analyses for dry mouth (any severity)

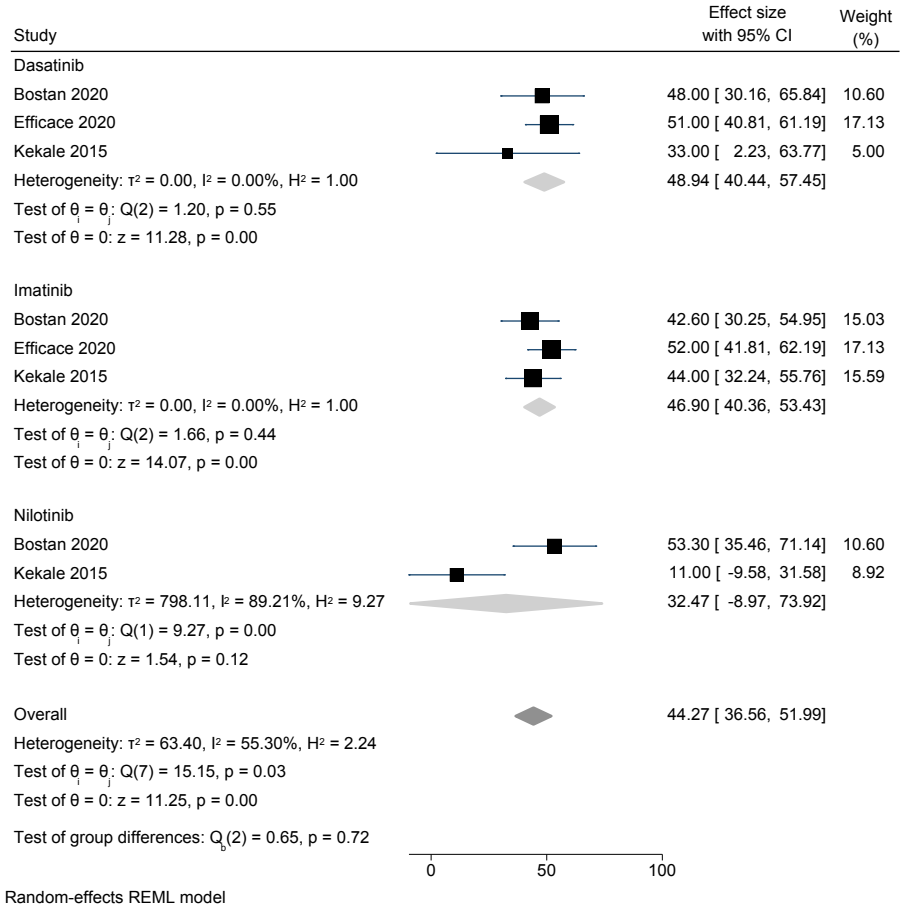


Figure 13 Forest plot of the meta-analyses for weight gain (any severity)

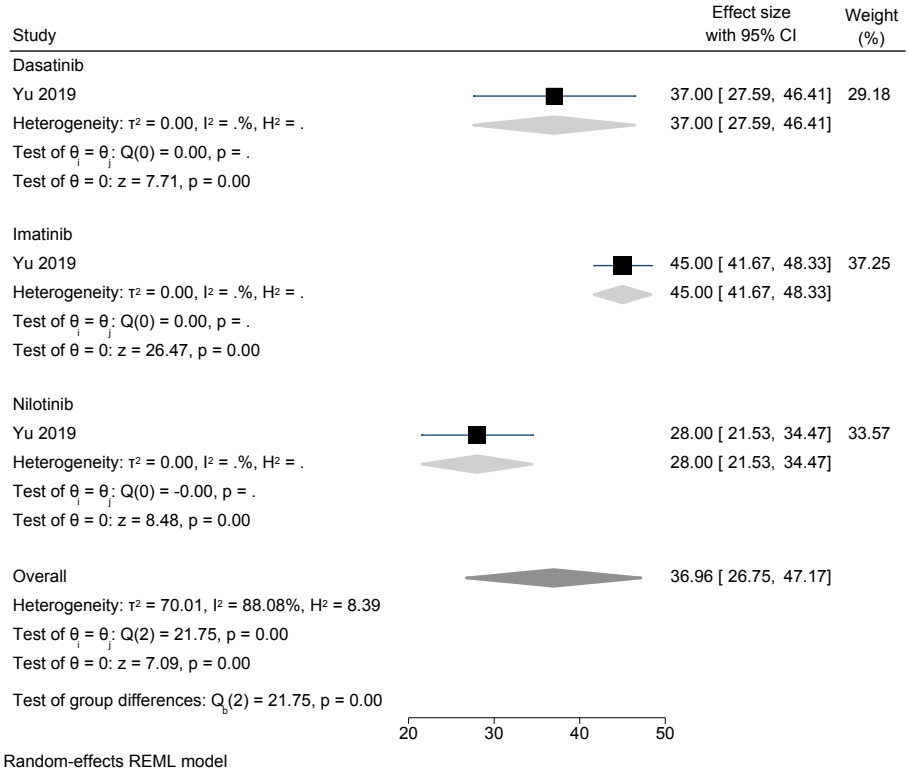


Figure 14 Forest plot of the meta-analyses for pain (any severity)

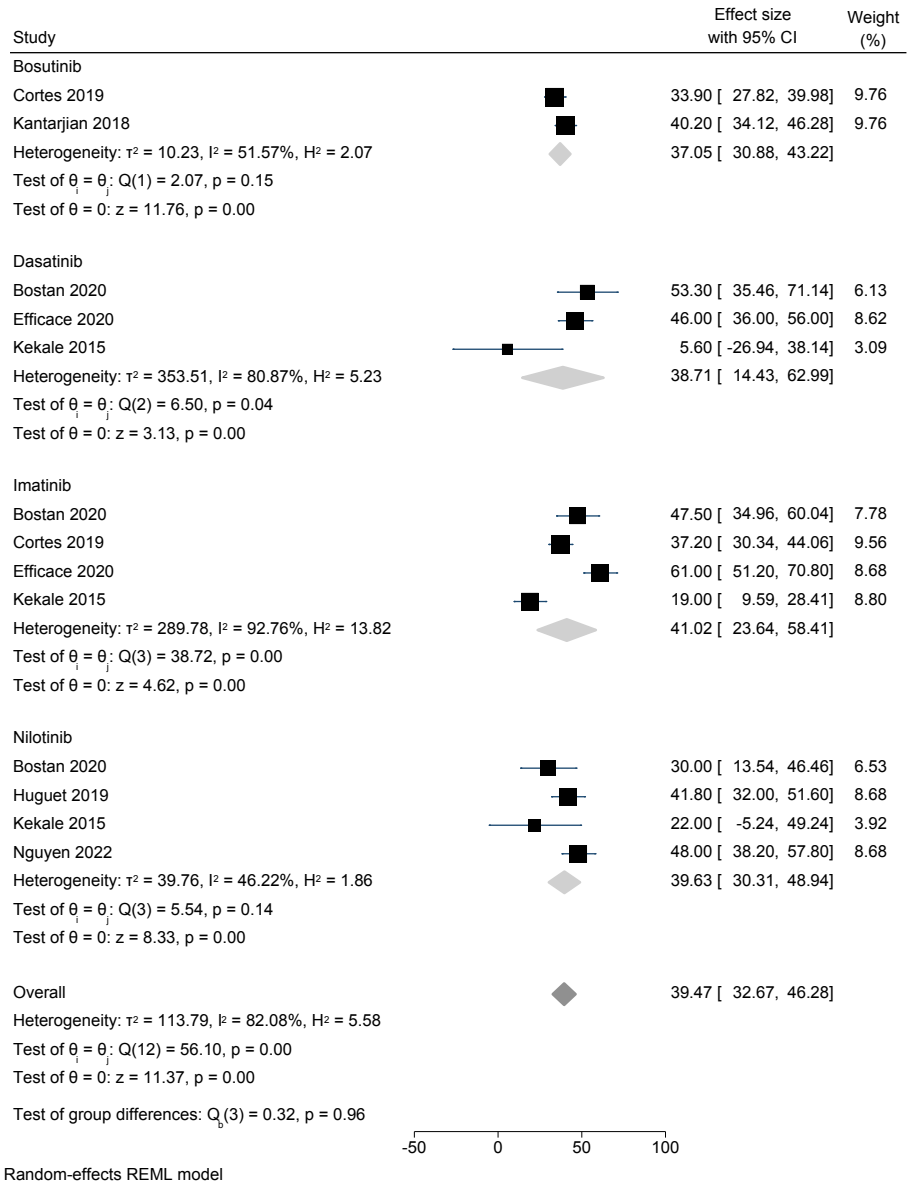


Figure 15 Forest plot of the meta-analyses for diarrhoea (any severity)

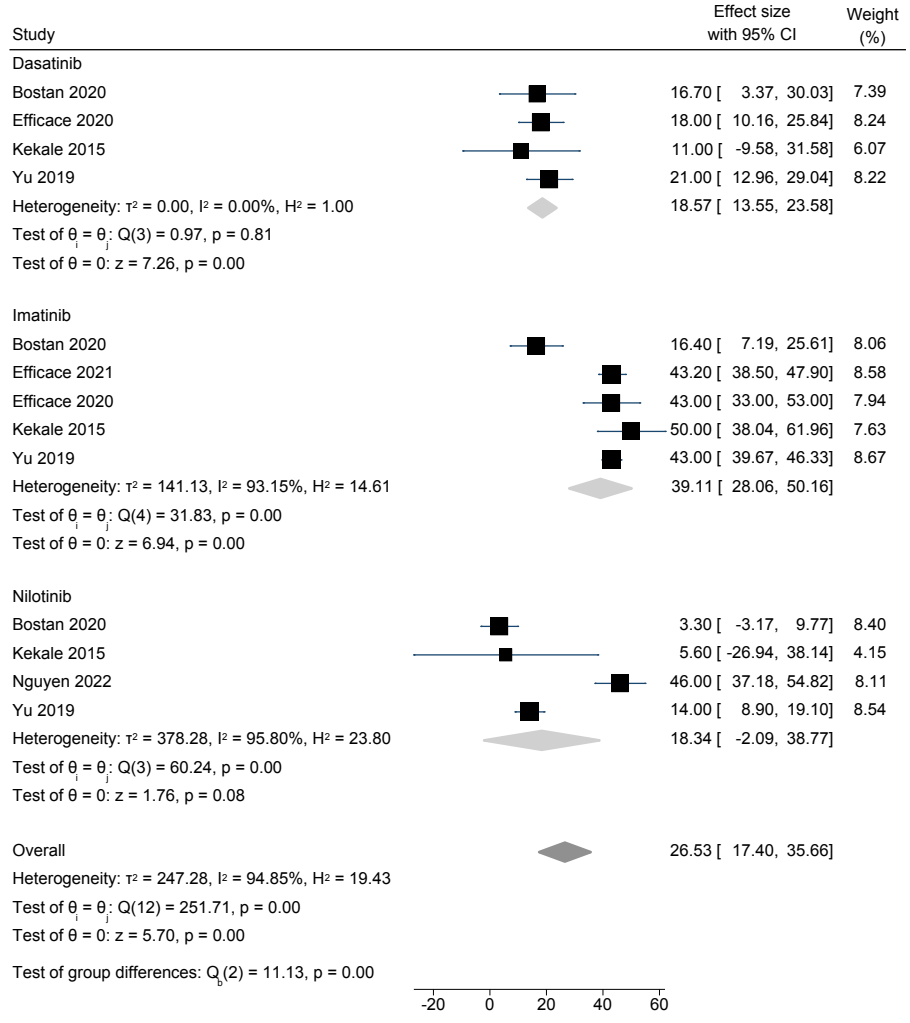


Figure 16 Forest plot of the meta-analyses for dyspnoea (any severity)

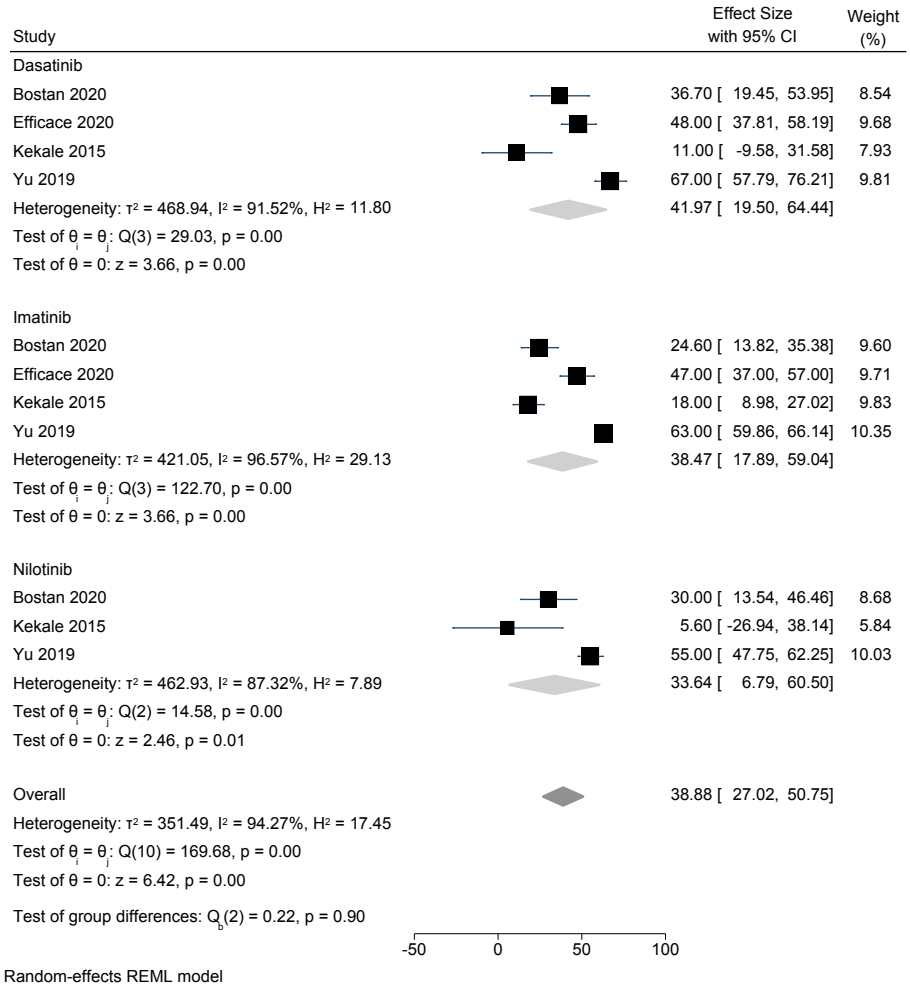


Figure 17 Forest plot of the meta-analyses for itchy skin (any severity)

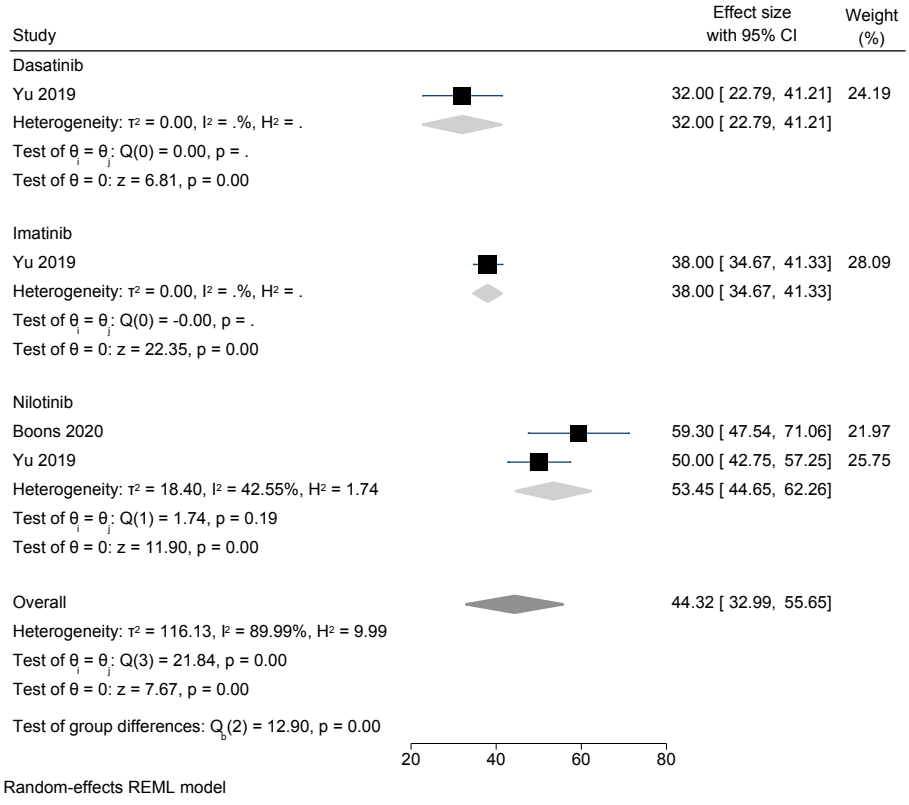


Figure 18 Forest plot of the meta-analyses for drowsiness (any severity)

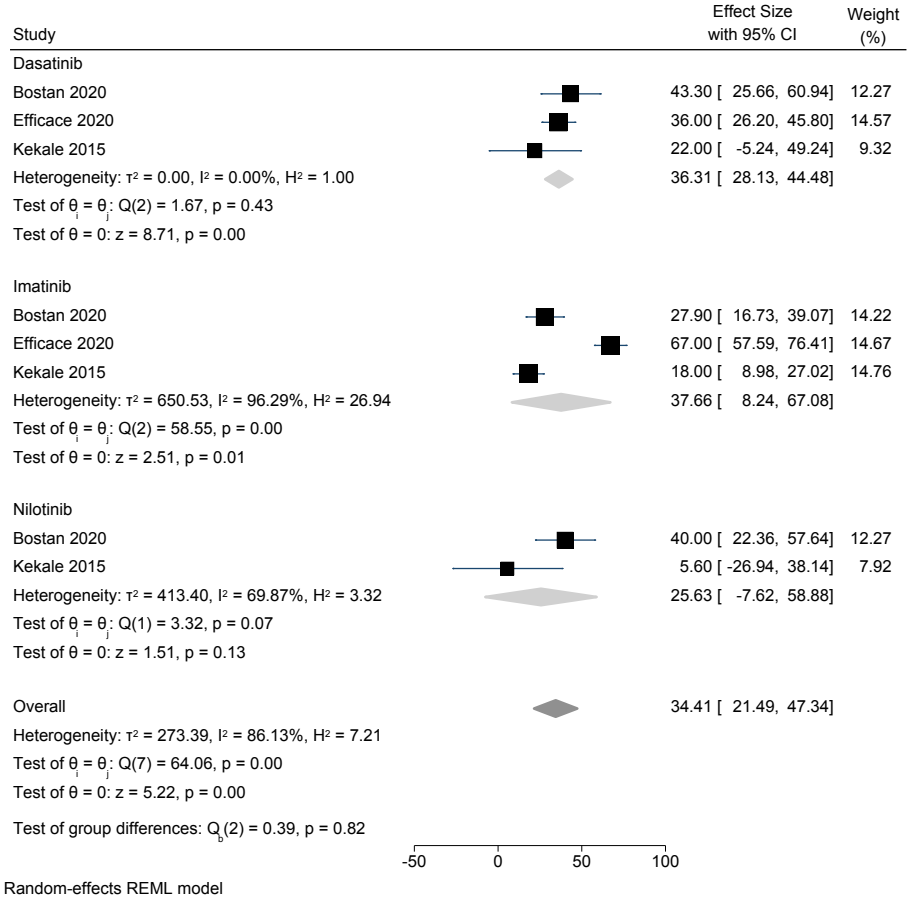


Figure 19 Forest plot of the meta-analyses for insomnia (any severity)

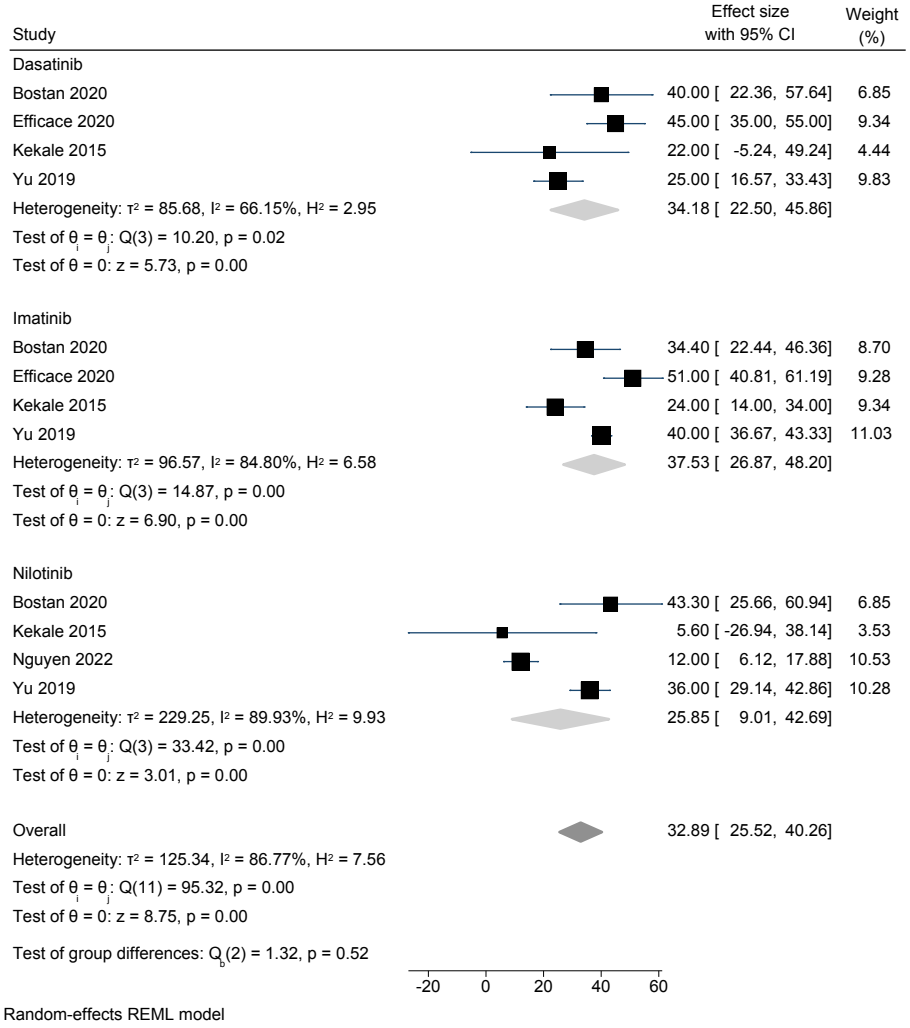


Figure 20 Forest plot of the meta-analyses for hidrosis (any severity)

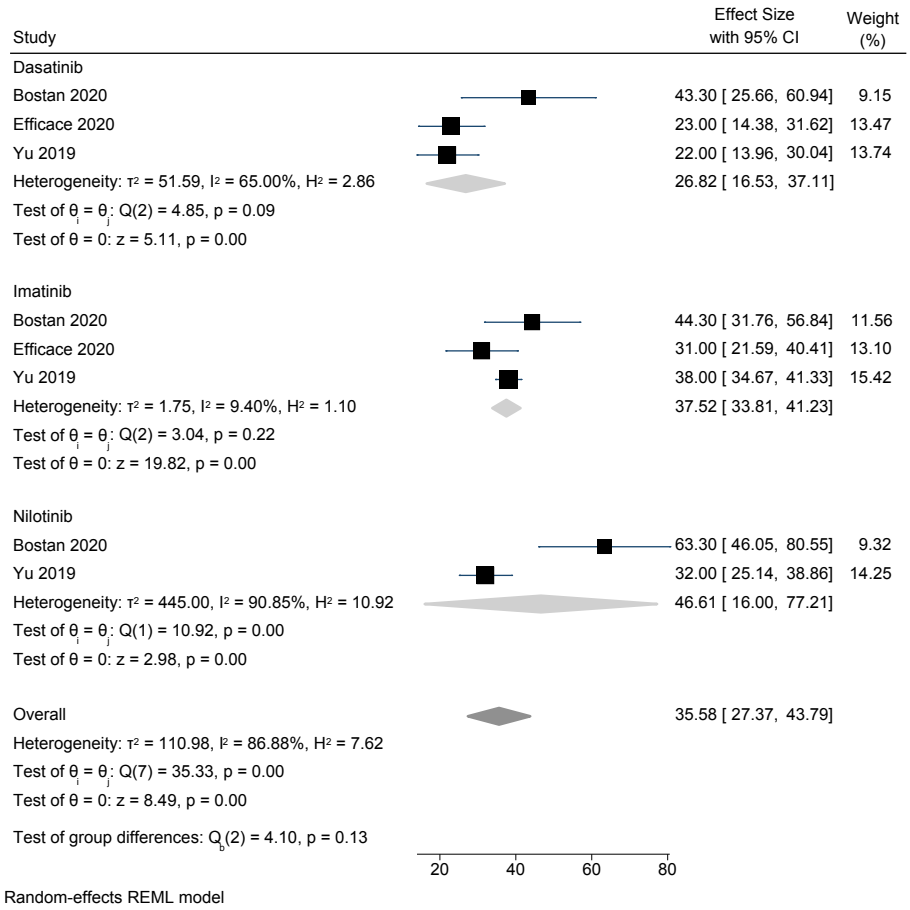


Figure 21 Forest plot of the meta-analyses for dry eyes (any severity)

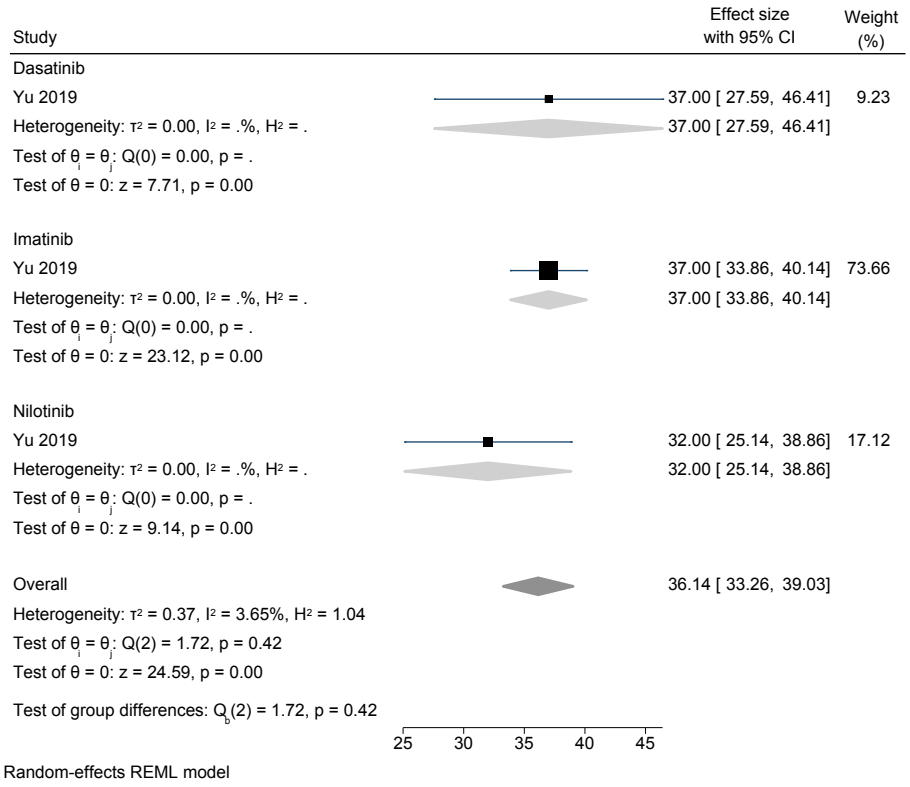


Figure 22 Forest plot of the meta-analyses for dizziness (any severity)

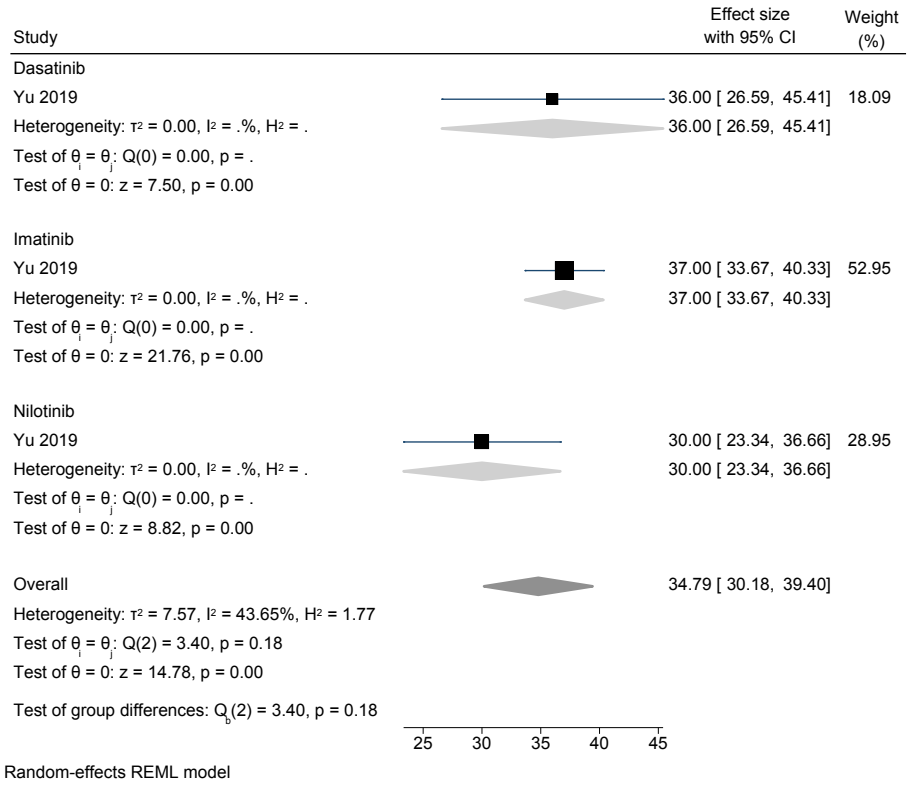


Figure 23 Forest plot of the meta-analyses for memory problems (any severity)

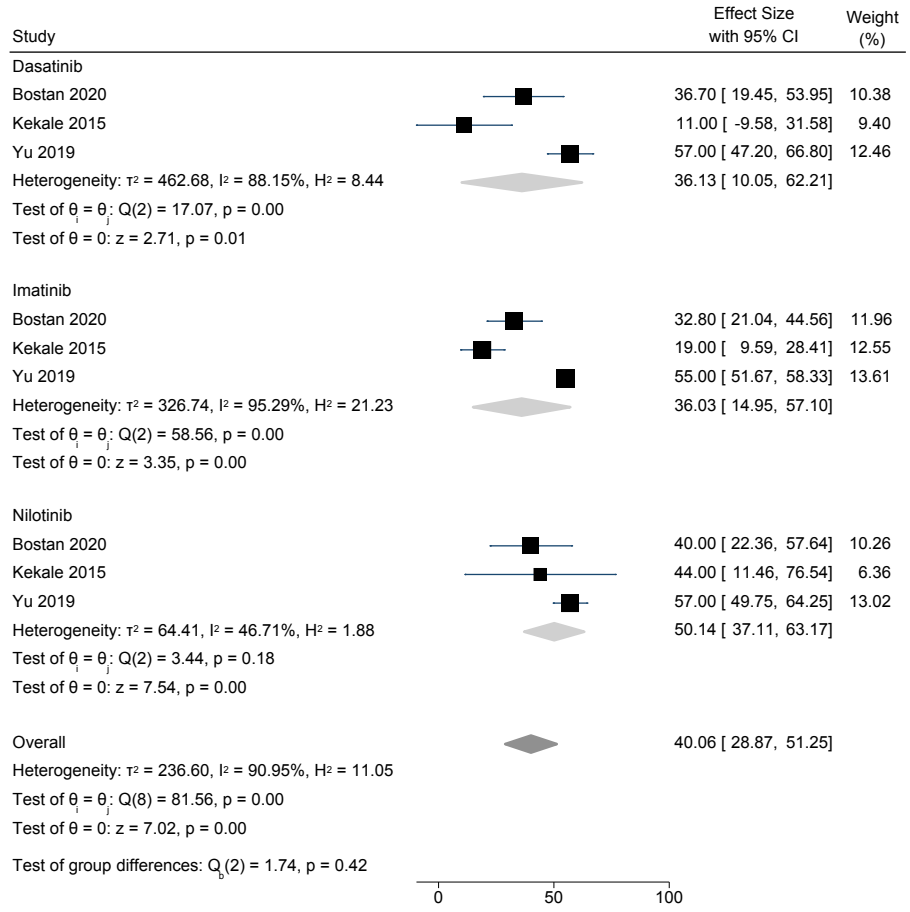


Figure 24 Forest plot of the meta-analyses for anxiety and depression (any severity)

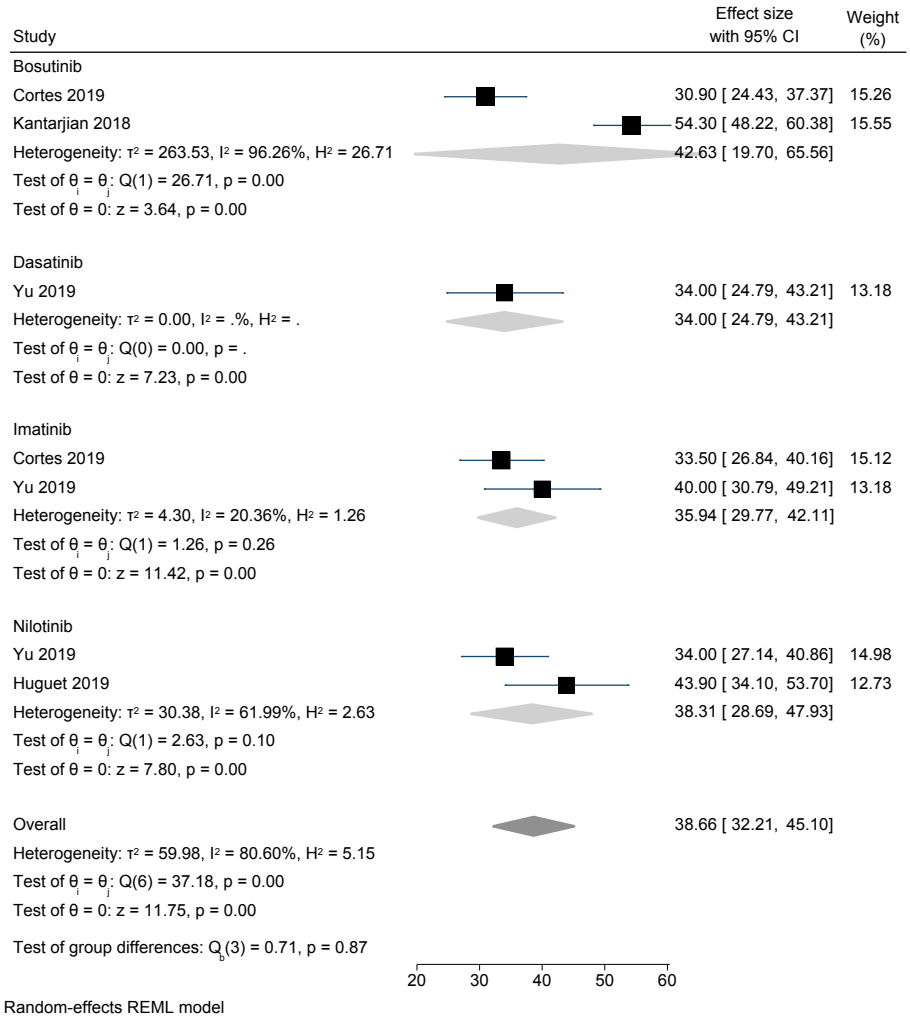


Figure 25 Forest plot of the meta-analyses for rash and/or skin problems (any severity)

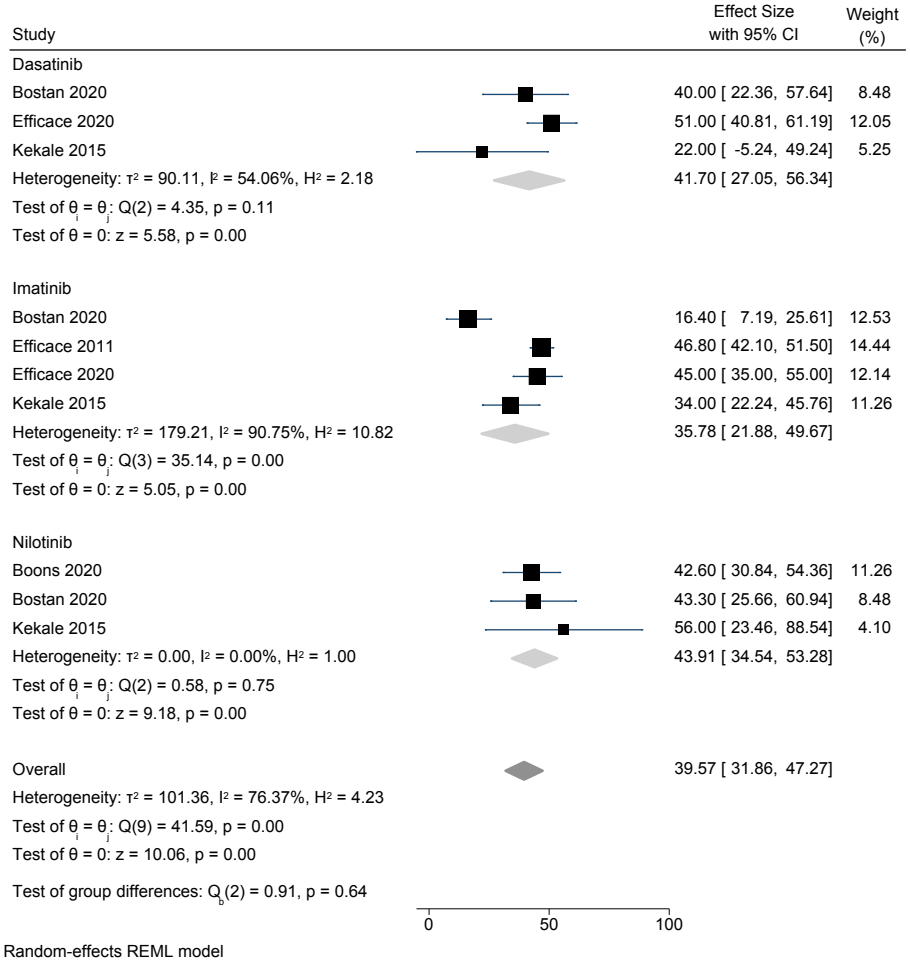


Figure 26 Forest plot of the meta-analyses for nausea (any severity)

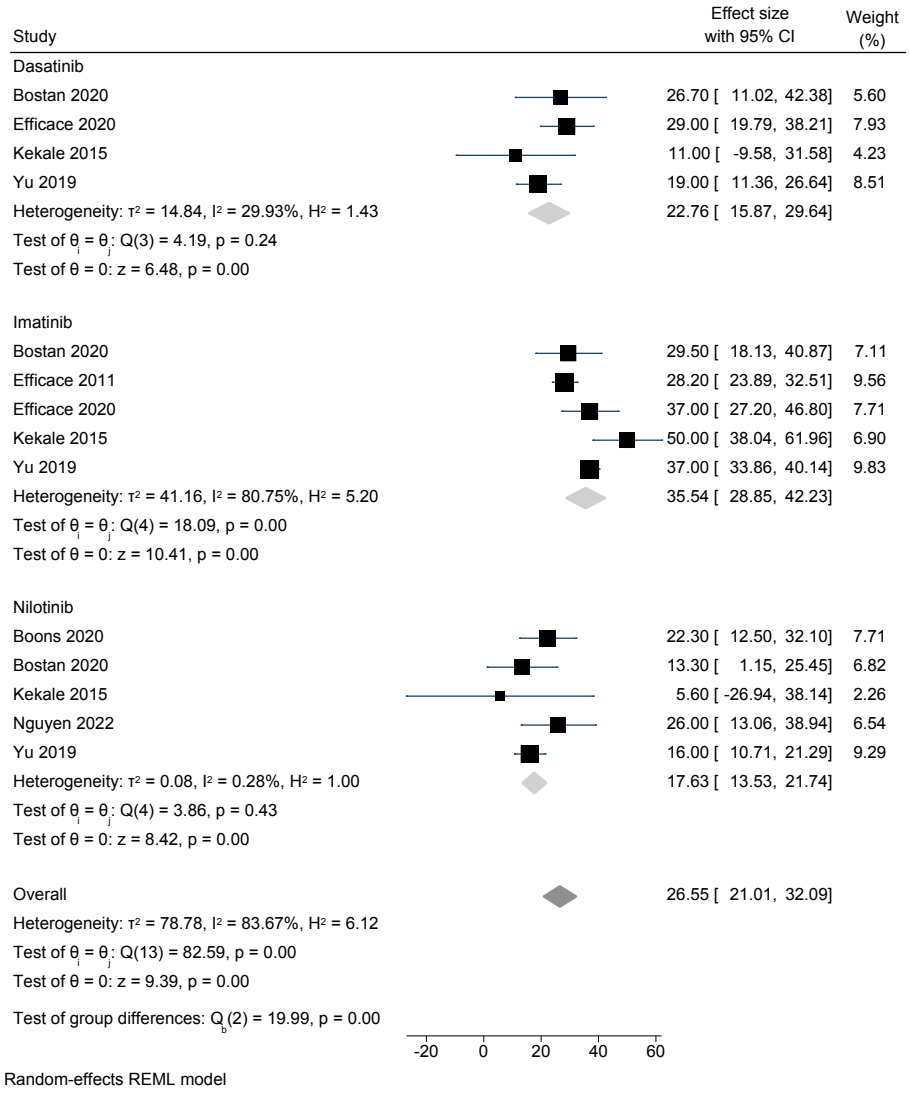


Figure 27 Forest plot of the meta-analyses for palpitations (any severity)

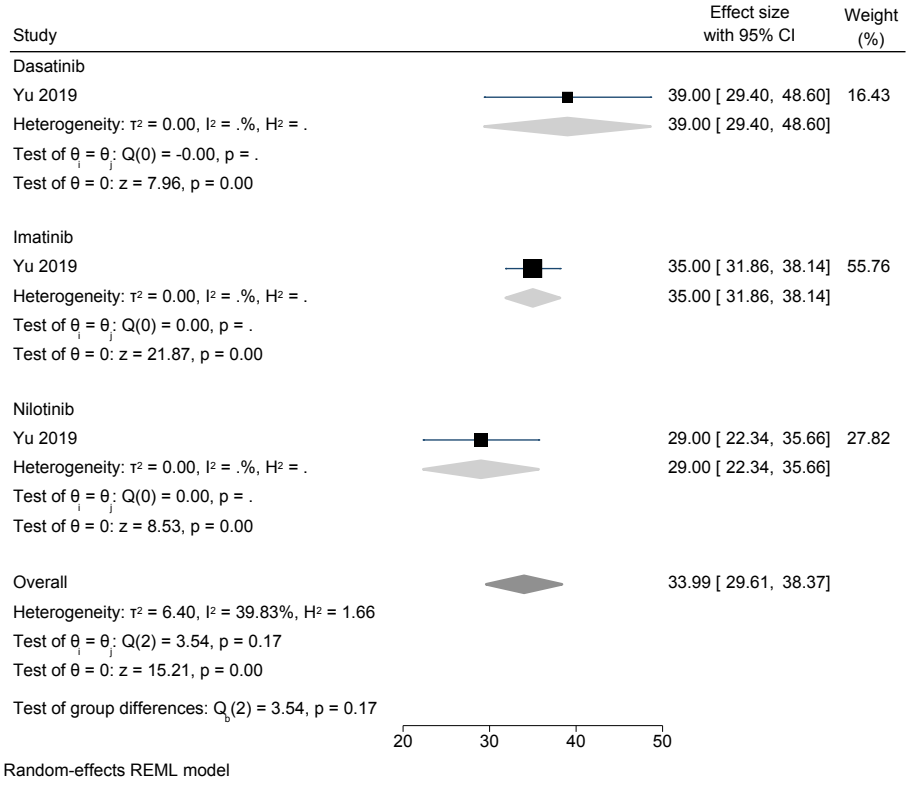


Figure 28 Forest plot of the meta-analyses for headache (any severity)

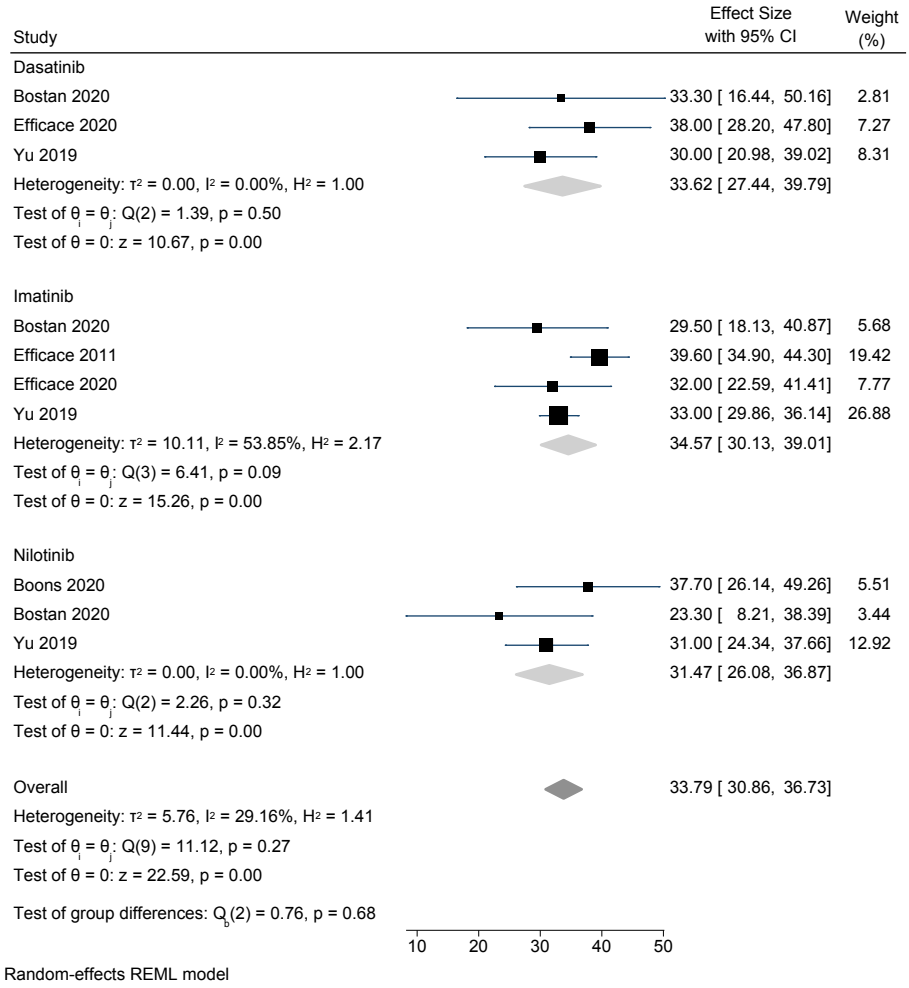


Figure 29 Forest plot of the meta-analyses for hair loss (any severity)

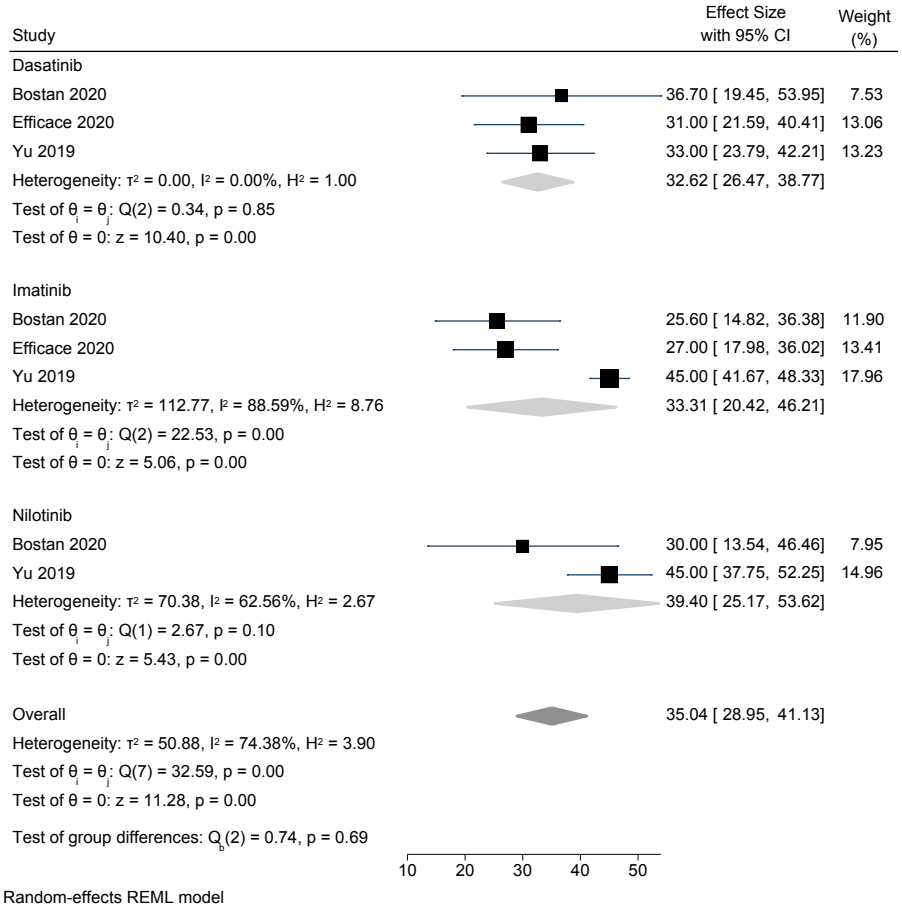


Figure 30 Forest plot of the meta-analyses for abdominal pain (any severity)

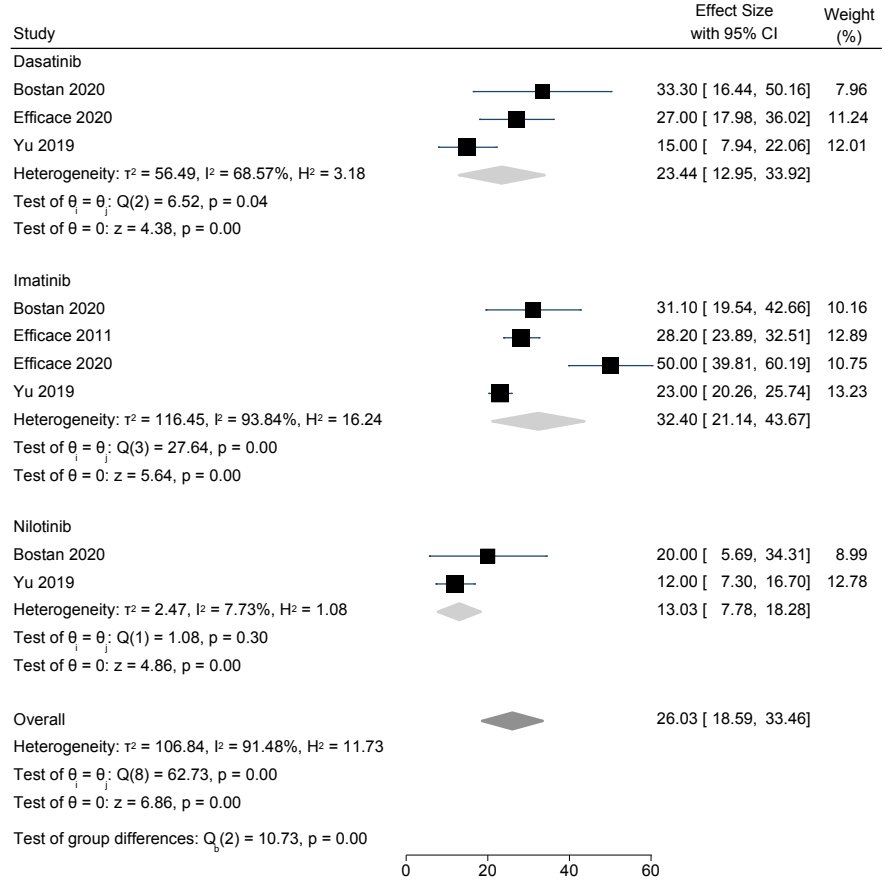


Figure 31 Forest plot of the meta-analyses for malaise (any severity)

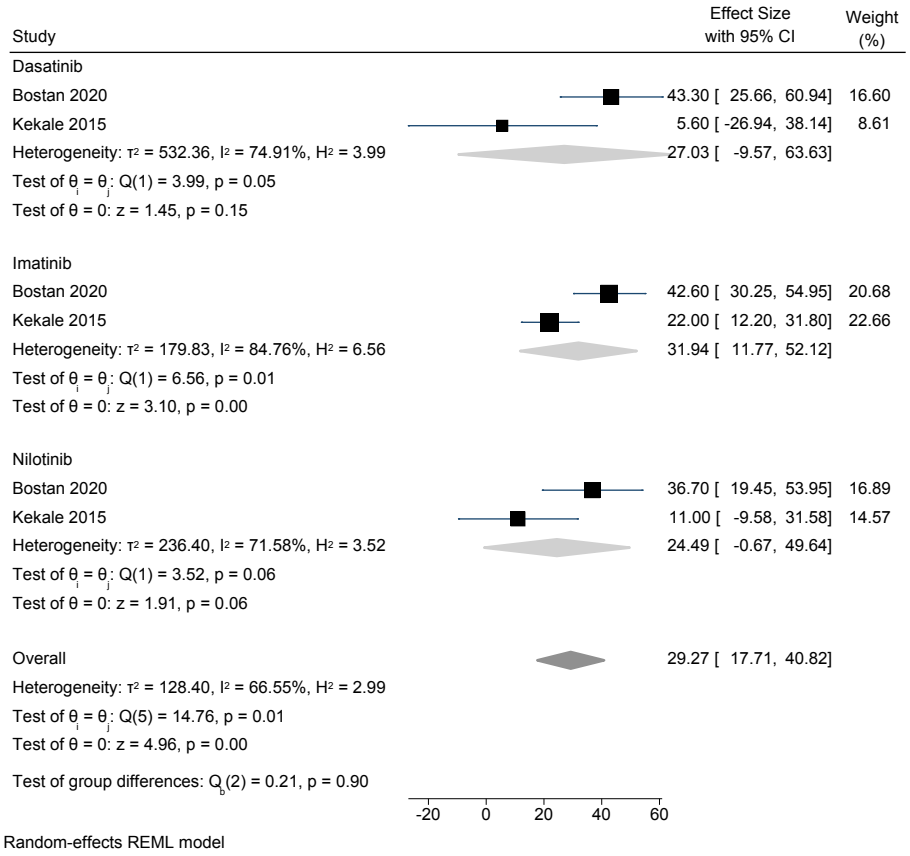


Figure 32 Forest plot of the meta-analyses for tinnitus (any severity)

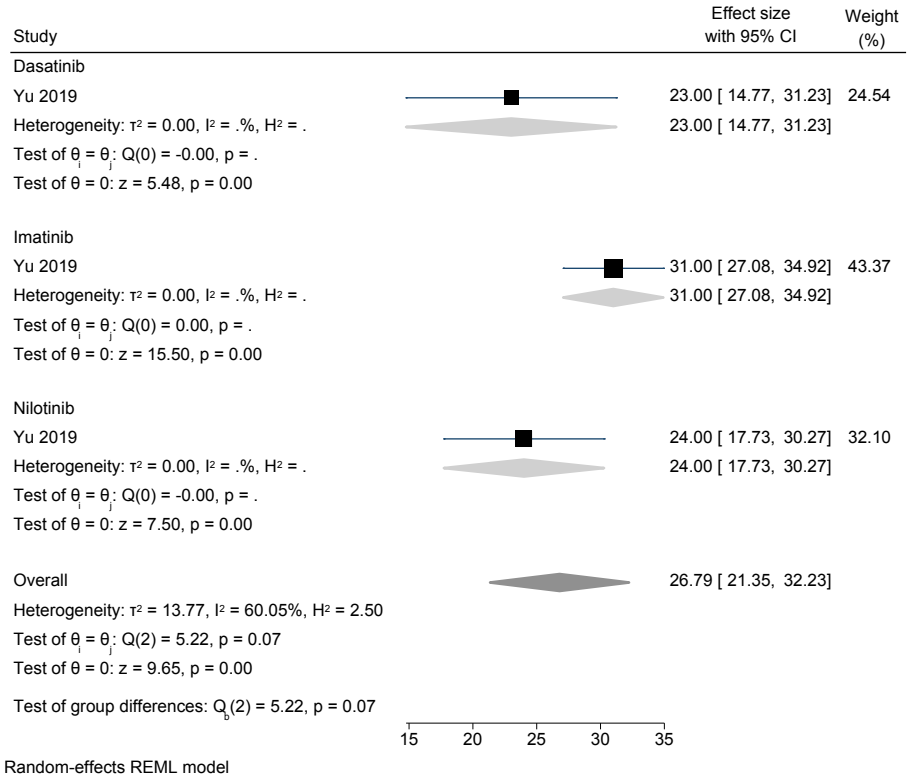


Figure 33 Forest plot of the meta-analyses for a decrease in sexual desire (any severity)

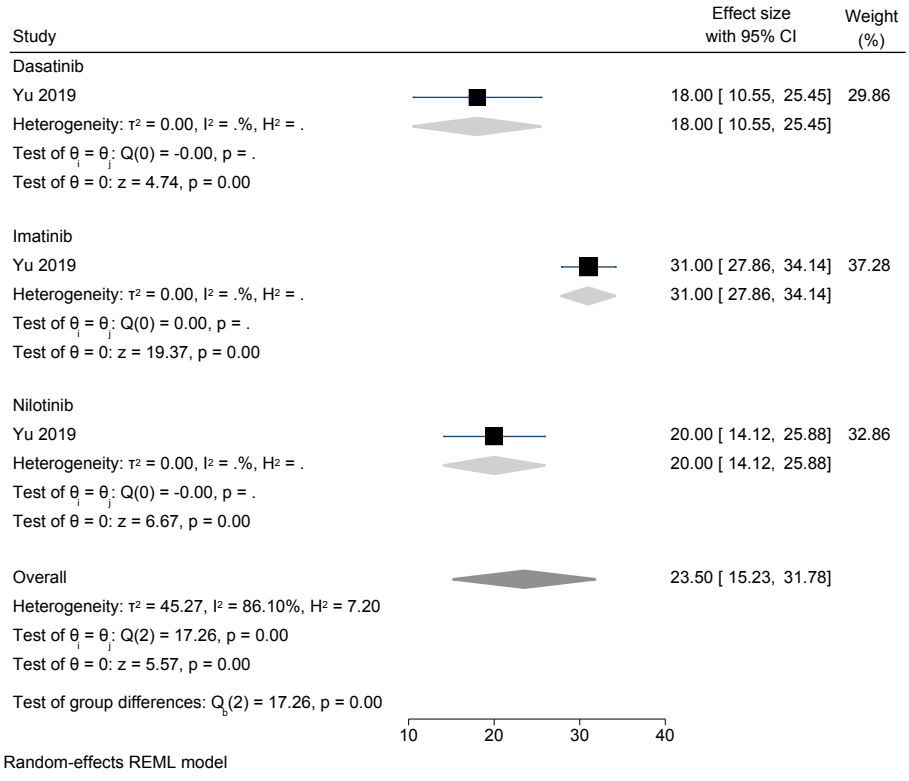


Figure 34 Forest plot of the meta-analyses for hypomenorrhea in females < 50 years of age (any severity)

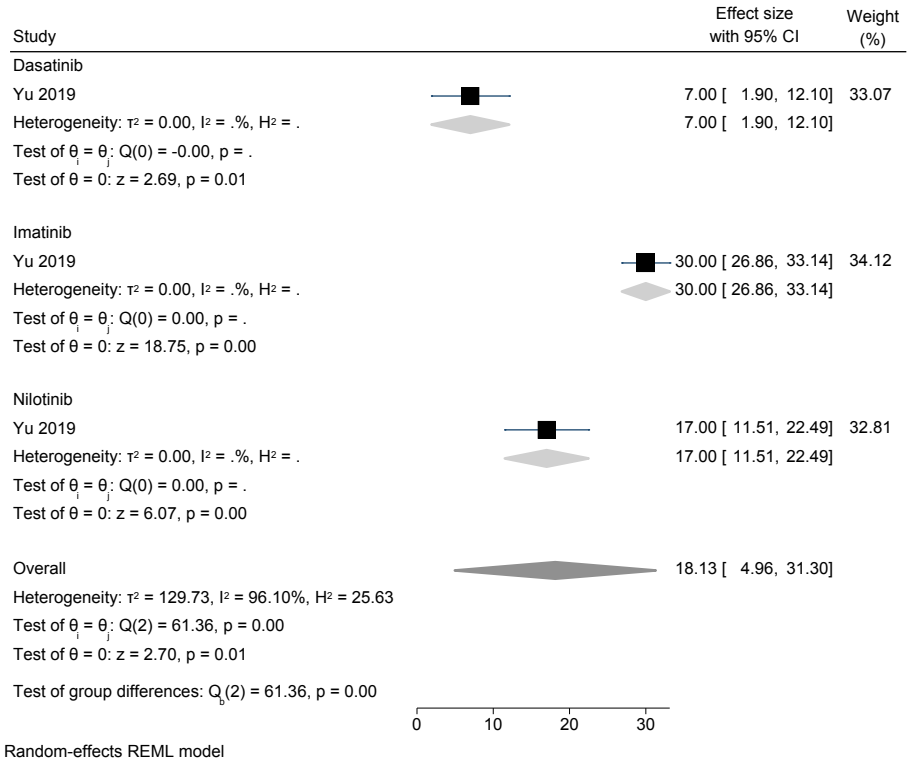


Figure 35 Forest plot of the meta-analyses for abdominal distension (any severity)

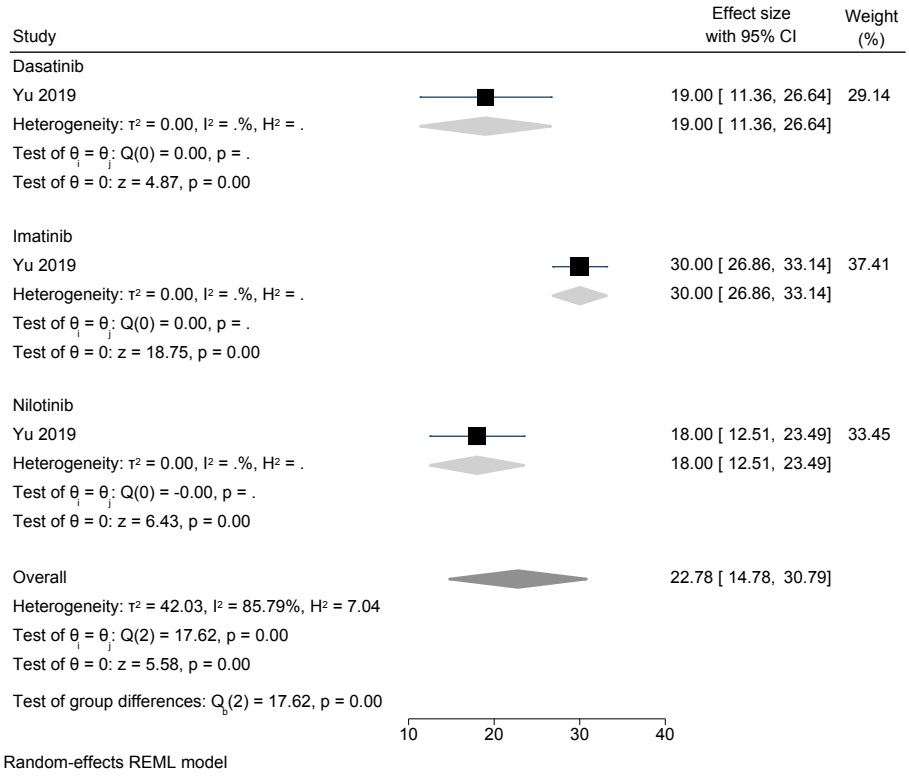


Figure 36 Forest plot of the meta-analyses for distress (any severity)

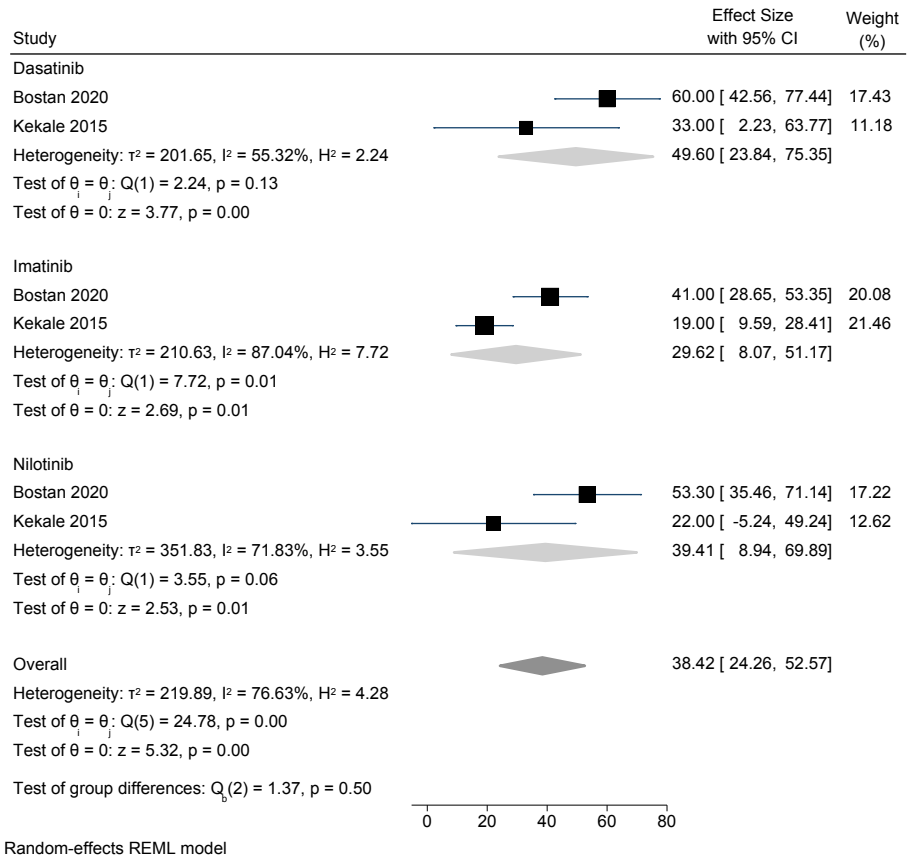


Figure 37 Forest plot of the meta-analyses for numbness or tingling (any severity)

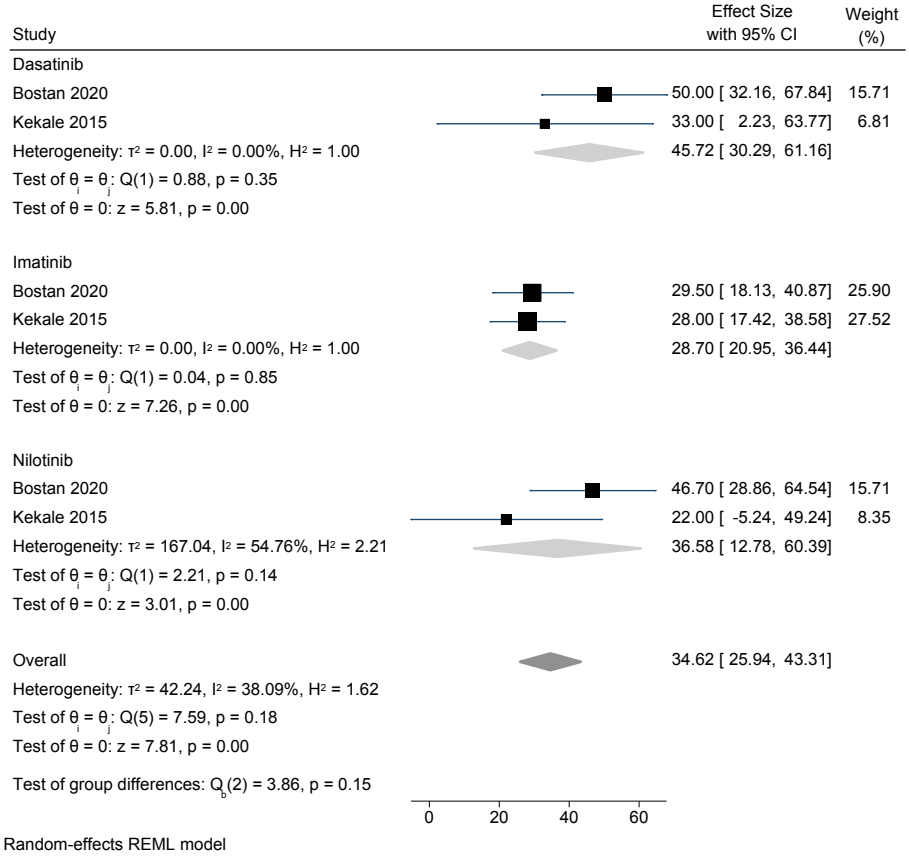


Figure 38 Forest plot of the meta-analyses for hair colour change (any severity)

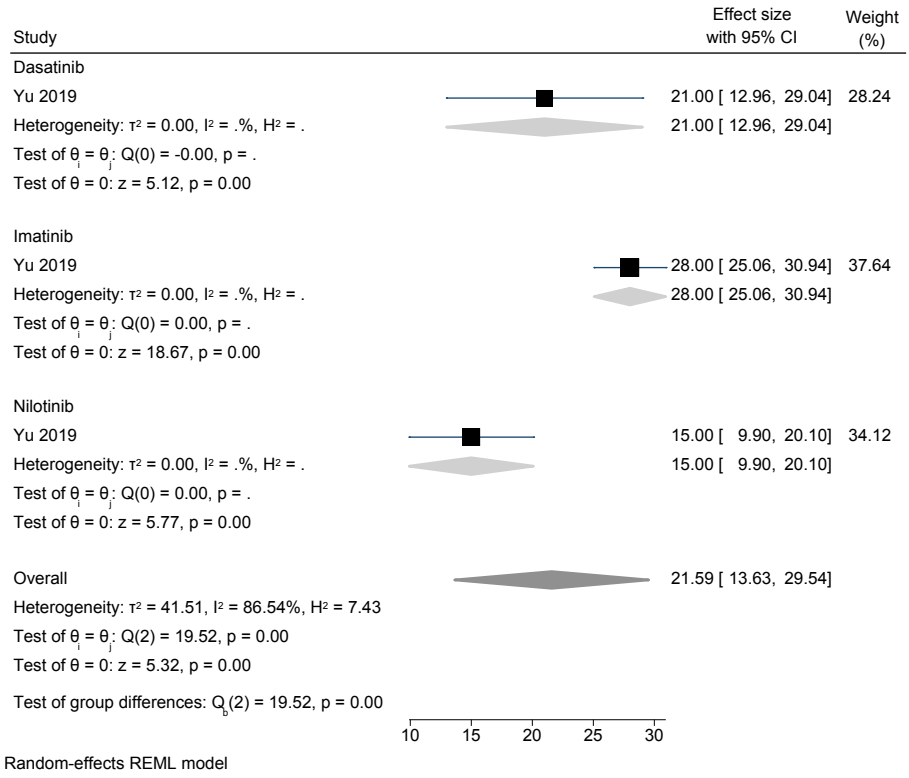


Figure 39 Forest plot of the meta-analyses for bruising and/or bleeding easily (any severity)

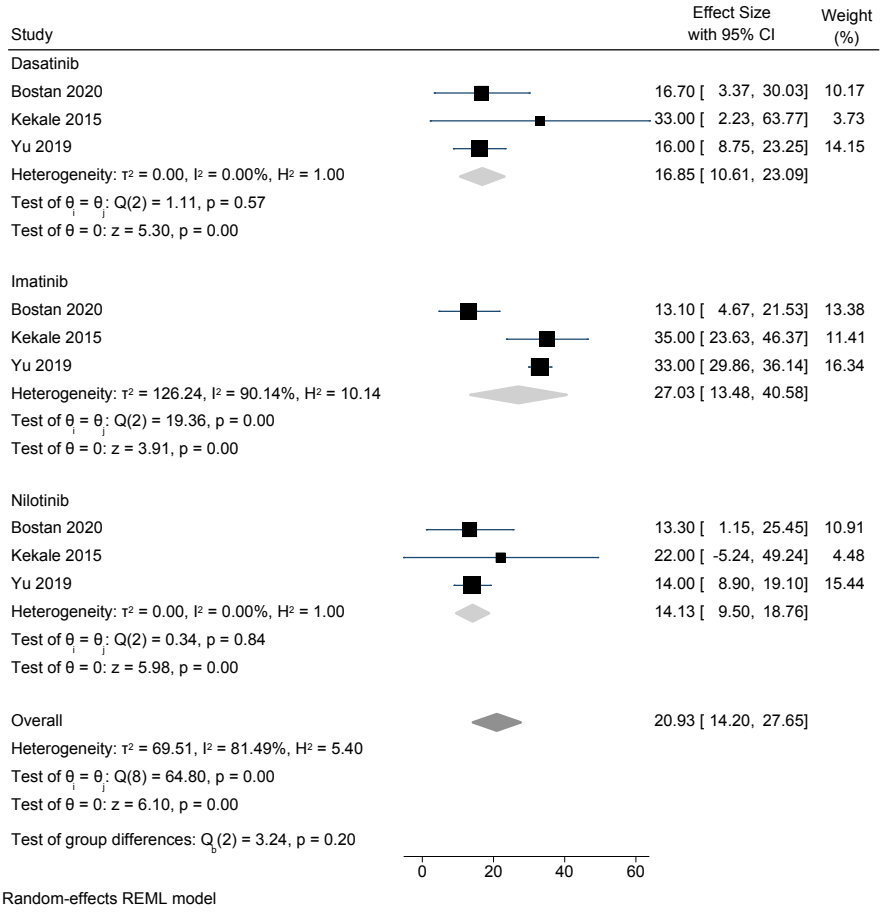


Figure 40 Forest plot of the meta-analyses for breast distension pain in females (any severity)

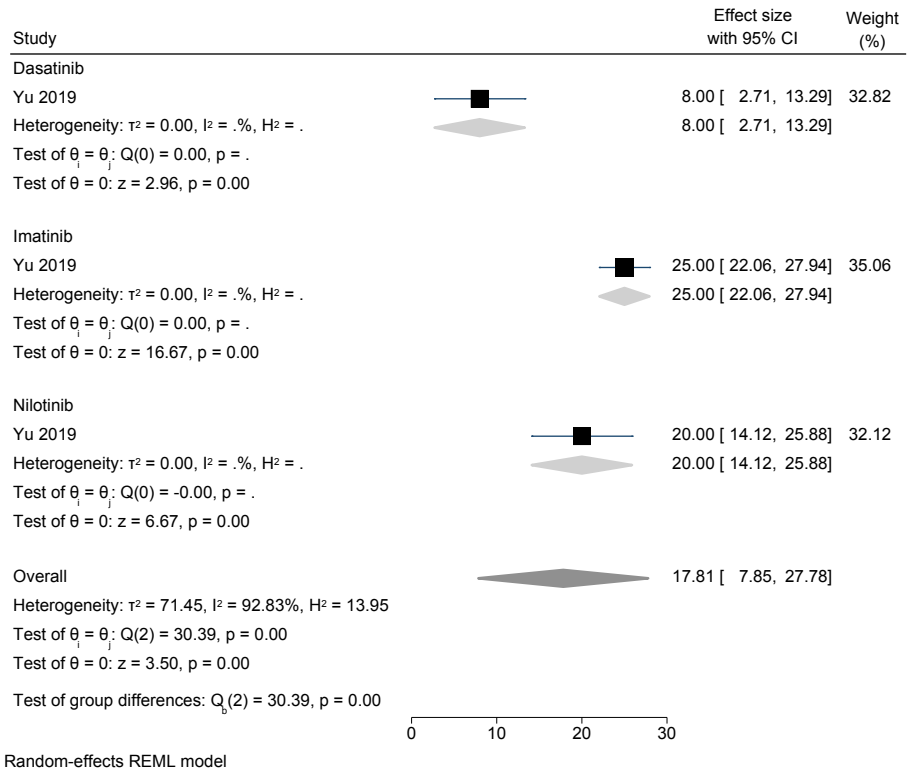


Figure 41 Forest plot of the meta-analyses for sadness (any severity)

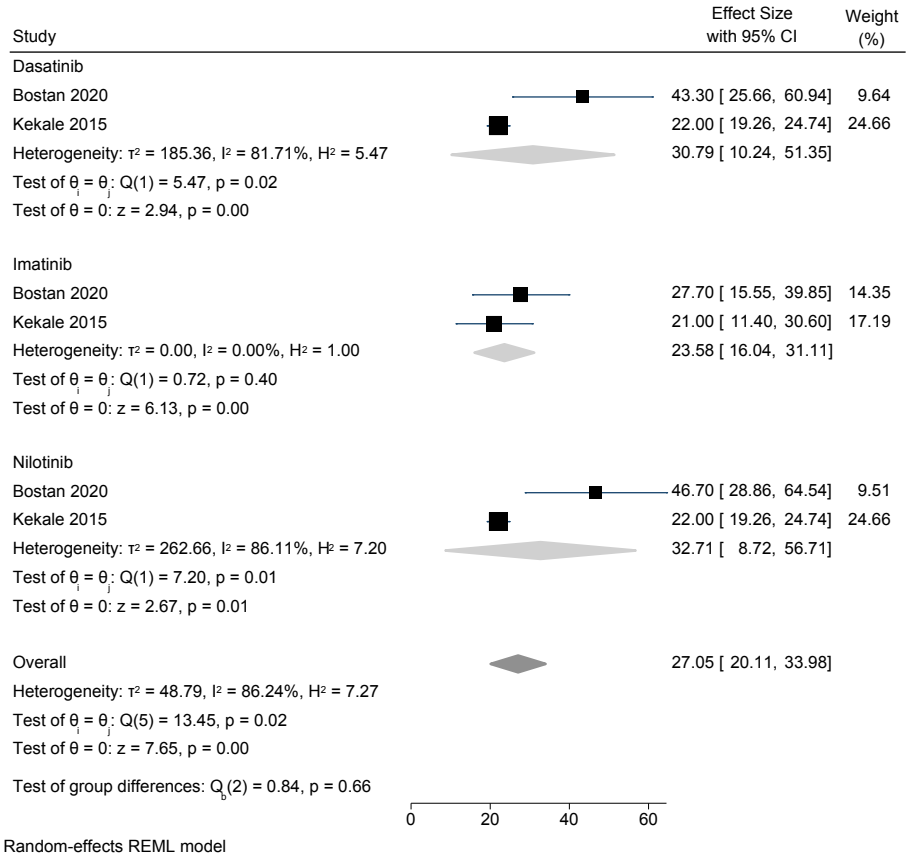


Figure 42 Forest plot of the meta-analyses for depression (any severity)

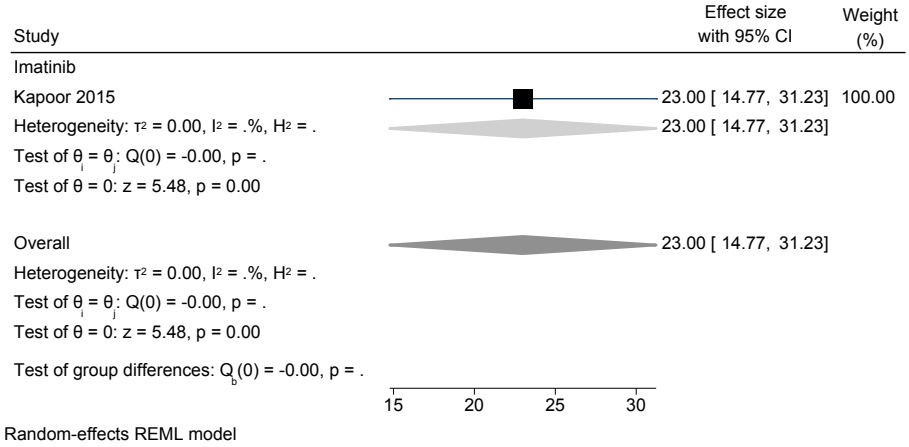


Figure 43 Forest plot of the meta-analyses for appetite loss (any severity)

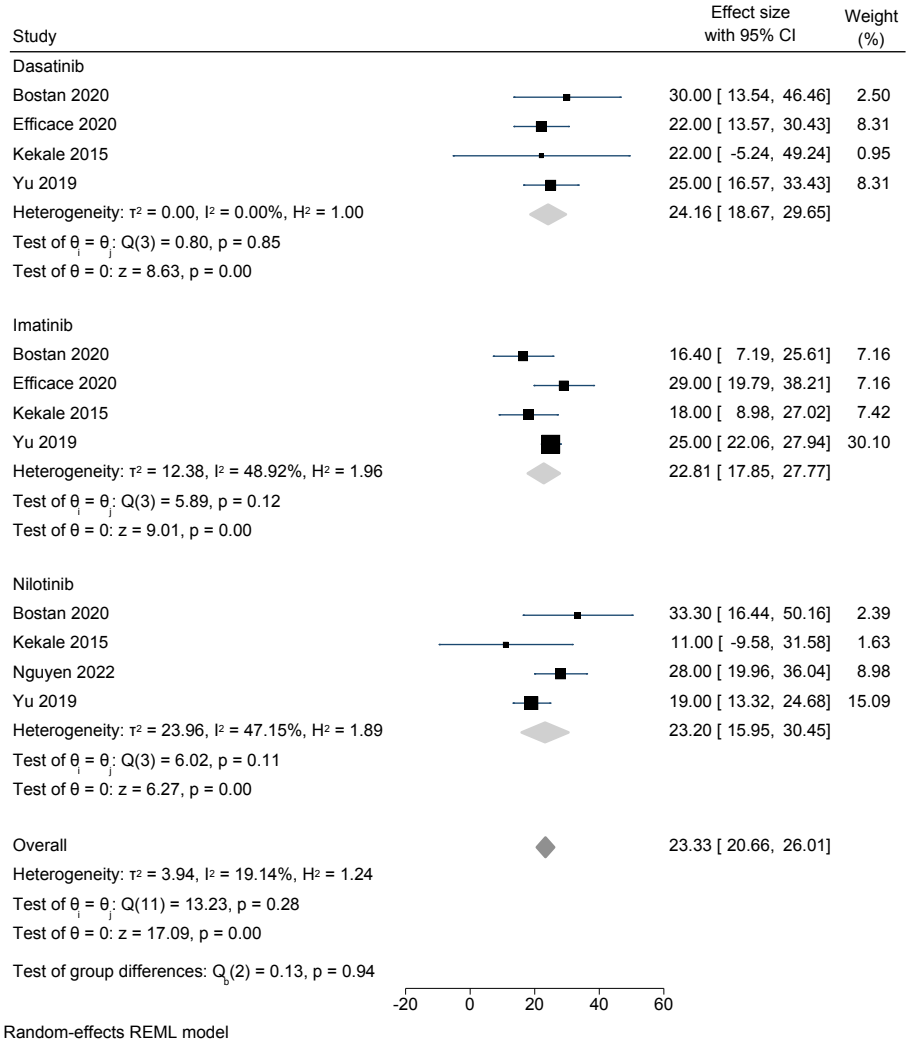


Figure 44 Forest plot of the meta-analyses for vomiting (any severity)

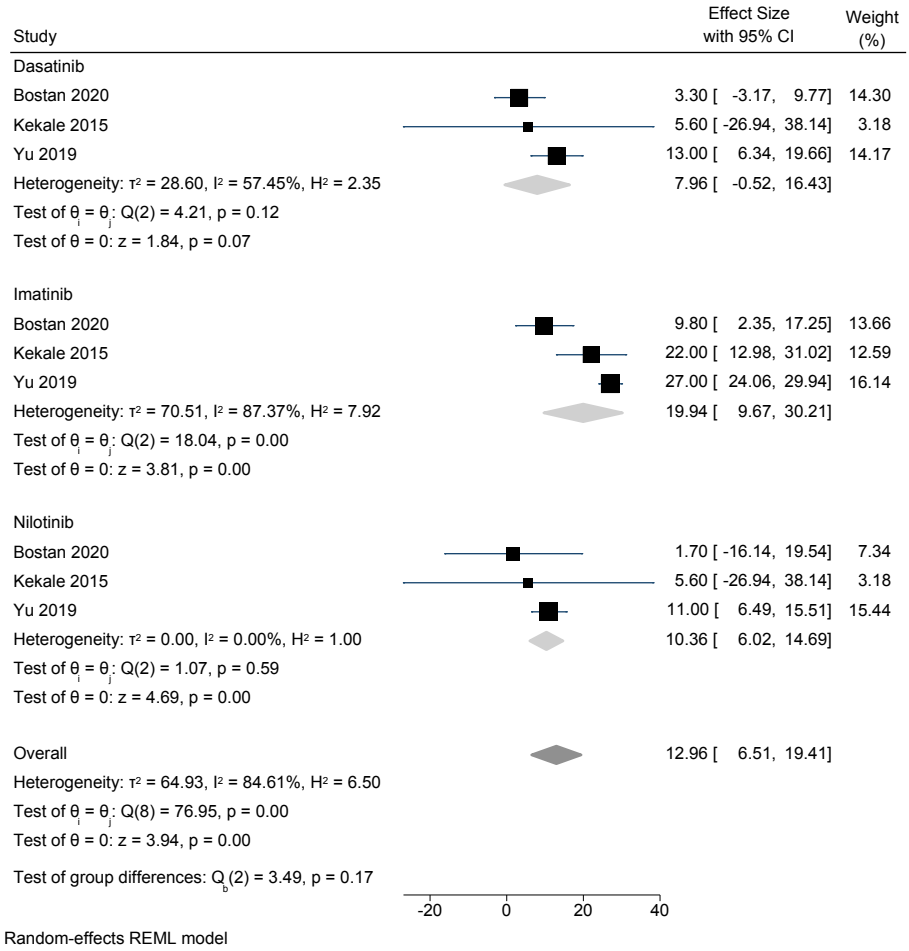


Figure 45 Forest plot of the meta-analyses for hypermenorrhoea in females aged <50 years (any severity)

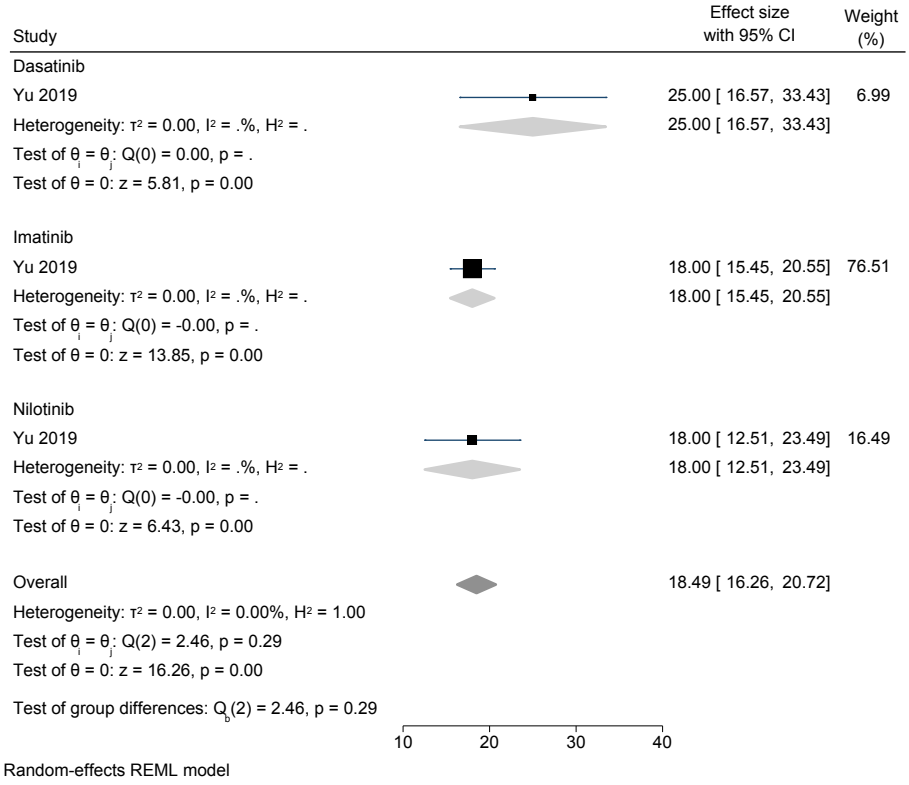


Figure 46 Forest plot of the meta-analyses for weight loss (any severity)

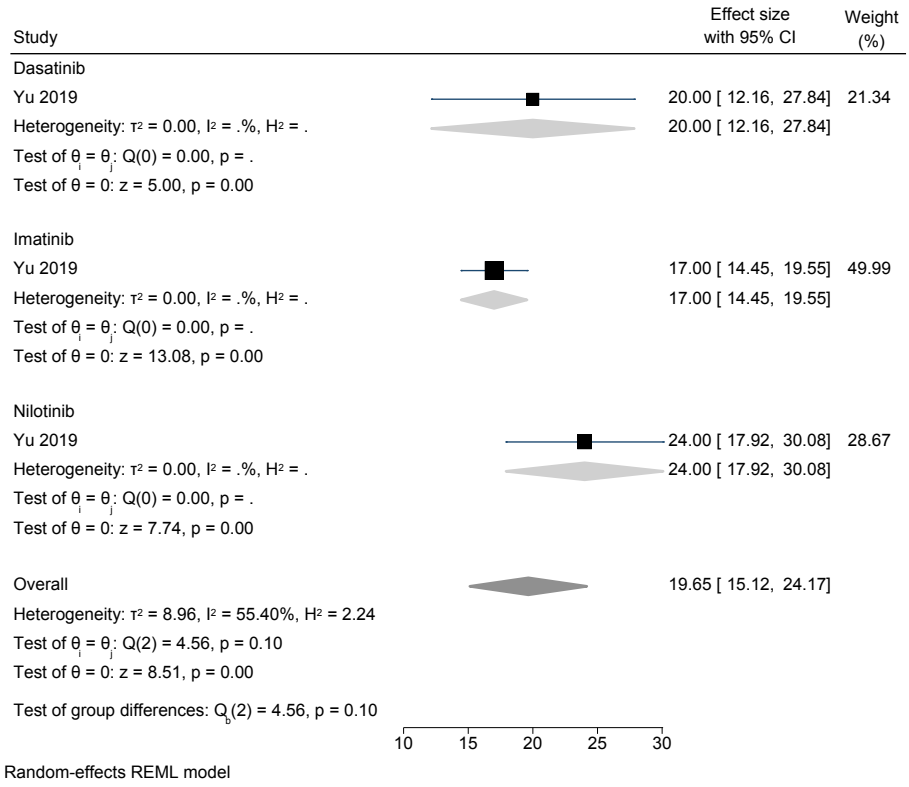


Figure 47 Forest plot of the meta-analyses for constipation (any severity)

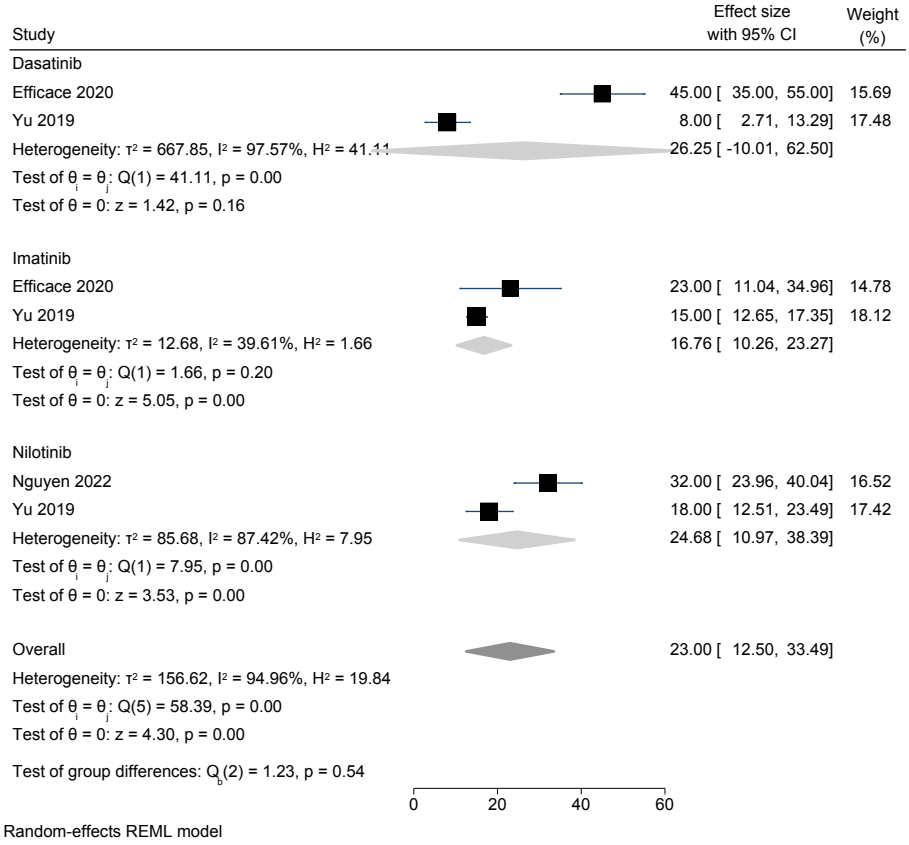


Figure 48 Forest plot of the meta-analyses for gynecomastia in males (any severity)

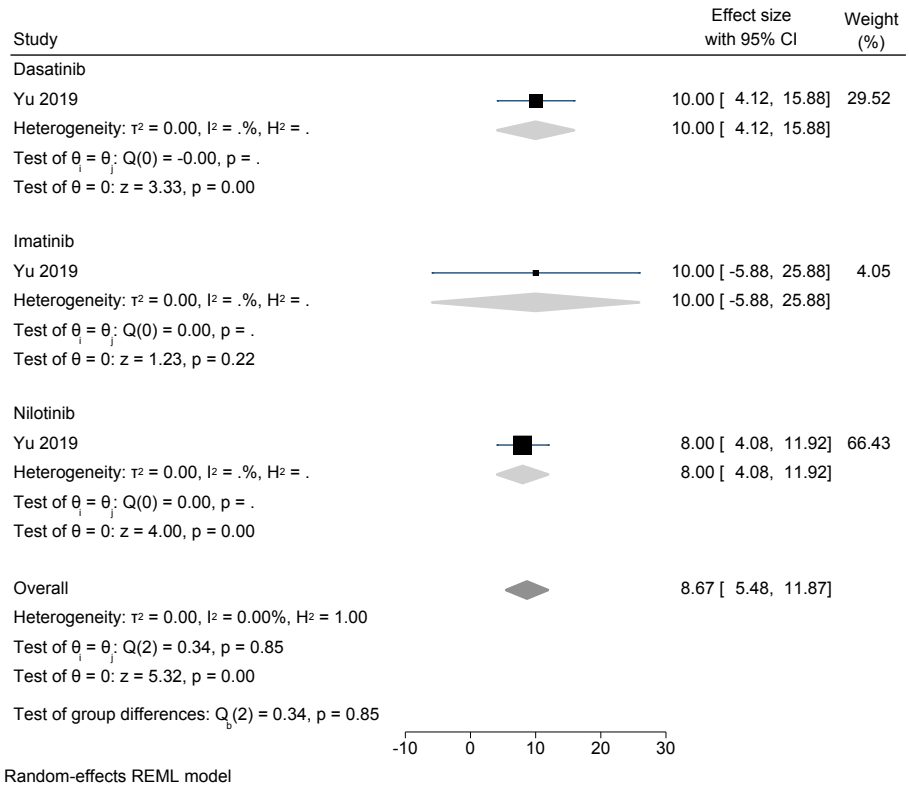
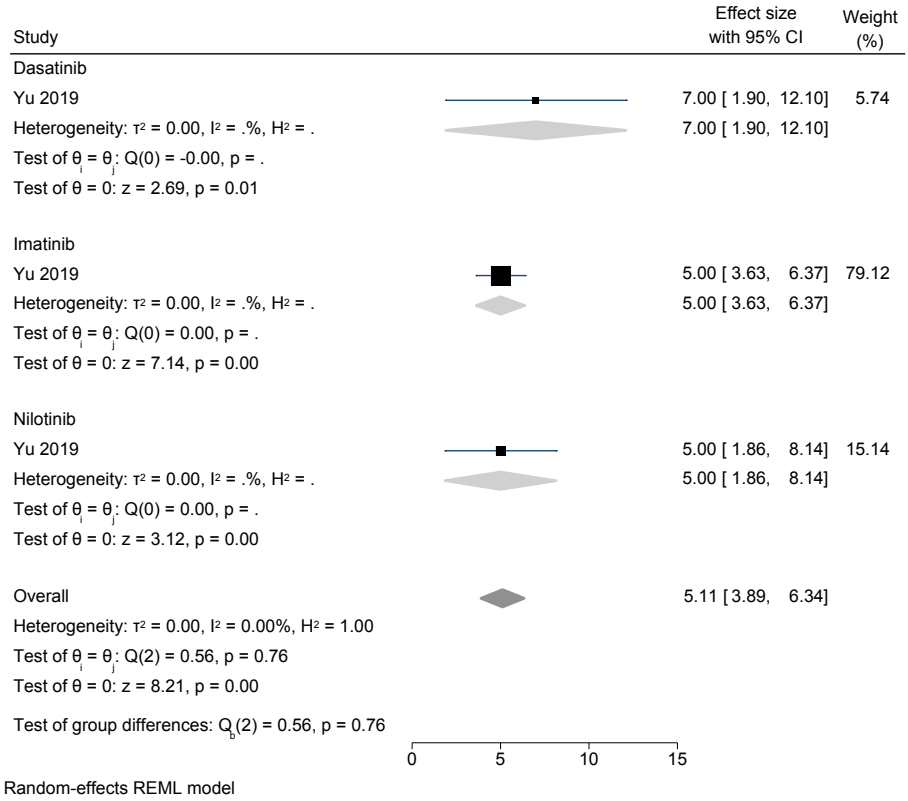


Figure 49 Forest plot of the meta-analyses for amenorrhea in females aged < 50 years (any severity)



Chapter 3

Measuring chronic myeloid leukaemia TKI-related toxic effects in the real world: a systematic review and critical assessment of content validity of patient-reported outcome measures

Yolba Smit, Sanne Metsemakers, Jeroen Janssen, Eduardus Posthuma, Iris Walraven, Rosella Hermens, Nicole Blijlevens

Lancet Haematology. 2023, 10(10);e849-59.

Abstract

Insight into real-world treatment-related toxic effects reported by patients has the potential to improve care, benchmark trials, and fill knowledge gaps, especially in patients with chronic myeloid leukaemia, which is treated in the majority of patients continually with tyrosine kinase inhibitors (TKIs). The aim of our systematic review was to investigate the content validity of instruments that elicit TKI-related toxic effects reported by patients with chronic myeloid leukaemia in the real world. We searched PubMed and Embase from Jan 1, 2017 to Oct 21, 2022. Studies on instruments used in or developed for patients with chronic myeloid leukaemia that assess a patient's symptoms were eligible. Content validity was assessed according to the Consensus-Based Standards for the Selection of Health Measurement Instruments (COSMIN): none of the six identified instruments were rated as sufficient. Five instruments (European Organisation for Research and Treatment of Cancer Core Quality of Life Questionnaire for chronic myeloid leukaemia with 24 items [EORTC QLQ-CML24], EORTC symptom set, Functional Assessment of Cancer Therapy-Leukaemia [FACT-LEU], haematological malignancies patient-reported outcomes [HM-PRO], and MD Anderson Symptom Inventory for chronic myeloid leukaemia [MDASI-CML]) were rated as inconsistent due to not being evaluated by professionals post-development, having very few patients with chronic myeloid leukaemia involved, or missing key symptoms. Moderate-quality to very low-quality evidence underpinned these ratings. The two EORTC instruments were the only ones not to miss key toxic effects (e.g., muscle cramps). However, their relevance was rated as inconsistent: the QLQ-CML24 includes questions on health-related quality-of-life, whereas the symptom set includes items sourced from solid cancer treatments. This Review shows the need for an instrument with sufficient content validity to measure toxic effects from TKI treatment in patients with chronic myeloid leukaemia. Until then, stakeholders can make an informed choice from currently used instruments with our assessment.

Introduction

The *Lancet Haematology Commission* called for action to modernise the assessment and reporting of adverse events in haematological malignancies [1, 2]. In particular, reporting of toxic effects in the real world is a priority [2]. Bringing in the patient's perspective through the administration of patient-reported outcome measures gives information on treatment tolerability in the real world, and supports decision making on dose modifications or treatment discontinuation [1]. Treatment decisions are especially important in chronic haematological malignancies, such as chronic myeloid leukaemia (CML), for which life-long oral therapy with tyrosine kinase inhibitors (TKIs) are the standard of care for the majority of patients. Although lifesaving, TKI therapy invokes considerable and enduring toxic effects. Compared with the general population or study controls, patients with CML on TKIs reported worse depression, dyspnoea, fatigue, pain, physical health-related quality of life, and composite symptom burden scores (e.g., nausea, diarrhoea, itching, skin changes and swelling of arms or legs) [3, 4]. In an observational study of 448 patients, around a third of patients on imatinib reported quite a bit or very much fatigue, 31% reported muscle cramps, 25% reported musculoskeletal pain and 25% reported oedema [5]. This high symptom burden is reflected in lower health-related quality of life scores compared with the general population, with marked impairments in role functioning in younger patients [5] and low grade adverse events negatively effecting quality of life [6]. Thus, with survival safeguarded in the vast majority of patients, the focus in CML care has shifted towards preventing and managing symptoms and complications of targeted therapy [7].

In CML in particular, patient-reported outcome measure data may help to select either the most appropriate TKI to be used as frontline therapy for different patient groups, or which TKI to switch to in case of intolerability. In addition, patient-reported outcome measure data might help with understanding the relationship between symptoms in patients with TKI use and symptoms in patients whom proceed to a TKI treatment-free interval. At present, around a third of patients in high-income countries will be eligible to proceed to a TKI treatment free interval [8]. Sixty percent of those will have withdrawal symptoms, especially musculoskeletal pain [9]. Therefore, monitoring of patient-reported symptoms also has value in patients in treatment-free remission, since there is little information on the course of symptoms after stopping TKIs, or on possible long term toxic effects of TKIs. However, at least

some symptoms decline over time after stopping treatment [10], which would make patient-reported outcome measure use increasingly irrelevant.

Unfortunately, providing a comprehensive overview of the patient-reported symptom burden related to TKIs at an individual and group level is not easy at present because data from patient-reported outcome measures are scarce, especially so for second-line and third-line TKIs. The fact that 94% of leukaemia trials did not include any patient-reported outcome measures as an outcome measure [1] illustrated that patient-reported outcome measures have not been considered important outcome data. To our knowledge, no data from a randomised trial exists that compare TKIs with each other in terms of specific TKI toxic effects reported by patients [11]. In most CML trials, side effects were reported by physicians to assess safety, which was done in a manner developed for short-course chemotherapy instead of long-term daily oral treatment [1]. Such reporting does not reflect a patient's symptom burden, as physicians tend to underestimate the incidence and severity of symptoms compared with a patient self-reporting symptoms [12-15].

The *Lancet Commission* indicated that improvement and expansion of reporting systems for patient self-reporting of adverse events is a solution that could bring improvement to adverse event assessment in the next few years [2]. The Commission's overarching vision shows the direction forward; however, what is the next practical step? Real-world longitudinal data on toxic effects reported by CML patients, including data before treatment initiation, after TKI withdrawal, and during dose switching, are needed, both for individual clinical care and for quality improvement at the aggregated level. Such data might help answer unresolved questions on the best choice of TKI in relation to expected toxic effects in particular patients, the effect of TKI switching on toxic effects, and long-term enduring toxic effects that persist after stopping or switching. These are important questions in the ageing population with CML, which has a substantial burden of comorbidity and comedication.

To gather reliable patient-reported toxic effect data related to the treatment with TKIs, clinicians, researchers and policy makers need to be informed which patient-reported outcome measures are available. In addition, stakeholders need to be able to assess the quality of the patient-reported outcome measurement instruments and the feasibility characteristics they have to suit their particular context. To address this need, we did a systematic review to identify patient-reported outcome measurement instruments that can be used

to assess the treatment-related toxic effects that CML patients have, critically assessed the content validity of these instruments, and described their feasibility characteristics.

Methods

This systematic review is reported following the PRISMA 2020 item checklist's items [16]. We did not prepare a protocol, nor was this review registered.

Search strategy and selection criteria and data extraction

We systematically searched PubMed and Embase on April 22, 2021 and updated this search on October 21, 2022 (full search strategy in Supplementary material). Key search terms included "chronic myeloid leukaemia", "haematologic malignancy", "quality of life", "patient-reported outcome measure", "symptom assessment", and "adverse effect". We restricted results to studies published from January 1, 2017 onwards, as a 2019 comprehensive systematic review included studies published up to December 31, 2016 [17]. We used this previous review to select studies to complement our search. In addition, 11 websites dedicated to patient-reported outcome measure instruments were examined for additional instruments (see Supplementary material) and references of systematic reviews, key publications [17-24], and included studies were searched for additional studies. Inclusion criteria were: studies written in Dutch, English, French, German or Spanish; original full-text research on self-report instruments which assess a patient's symptoms; the instrument is available for assessment; and the instrument has been developed for or validated in adult patients in the chronic phase of CML who are using TKIs. Instruments that focus on one category of symptoms (e.g., depression, sexual health) were excluded. Two researchers (YS and SM) selected studies independently.

Two researchers (YS and SM) selected studies independently. All disagreement was resolved through discussion. One researcher (YS) performed the data collection. Data were extracted on construct measured, total number of items, number of items related to symptoms, development process (e.g., goal, type and number of patients and health care professionals included, and countries participating), scales, subscales and single items included, response options and scoring range, languages available, licensing, and costs.

Overall content validity rating

We evaluated the quality of the content validity of identified patient-reported outcome measures with the Consensus-Based Standards for the Selection of Health Measurement Instruments (COSMIN) Risk of Bias checklist in its Excel template provided by the COSMIN group [25-27]. COSMIN evaluates patient-reported outcome measure development and content validity, plus eight other measurement properties concerned with validity, responsiveness and reliability. We focused on content validity exclusively as its three components - relevance, comprehensiveness and comprehensibility - form the basis of a reliable and valid patient-reported outcome measure. As per COSMIN recommendations, we rated each patient-reported outcome measure for the construct, population and context of use of our interest (i.e., treatment toxic effects, patients with CML in the chronic stage on TKIs, and clinical and research use [25]. We did not evaluate the instrument itself, but rather its application. We based our evaluation on information from the identified articles, the patient-reported outcome measures themselves, and their manuals. Authors were not contacted for further information.

COSMIN defines content validity as follows: all items are relevant for the construct of interest (i.e., relevance), no key aspects of the construct are missing (i.e., comprehensiveness), and the items are understood by patients as intended (i.e., comprehensibility) [25]. We rated the overall content validity of a patient-reported outcome measure in three steps (Figure 1) [25, 26]. All quality evaluations were done independently by two reviewers (YS and SM) and any differences were resolved in face-to-face meetings until a consensus was reached. One author (NB) collaborated in the validation of one identified patient-reported outcome measure (EORTC QLQ-CML24). They had no role in the selection process nor quality assessment.

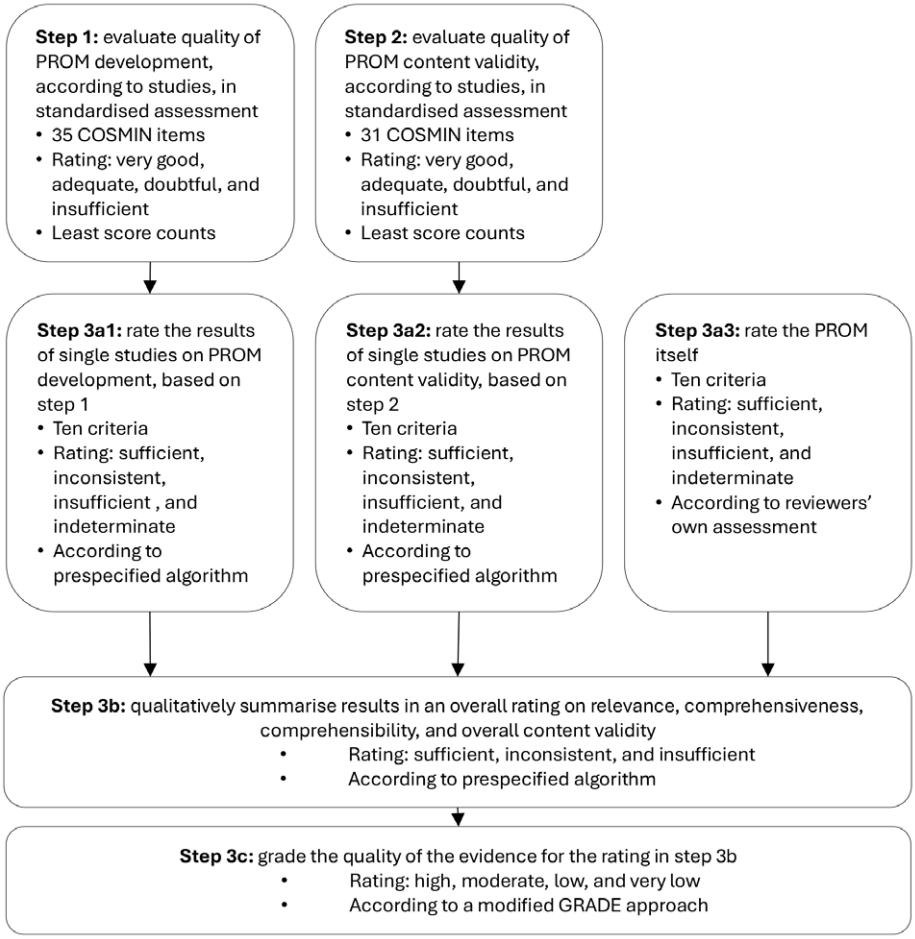


Figure 1 Steps in rating the overall content validity of a PROM according to COSMIN

Abbreviations: COSMIN: Consensus-Based Standards for the Selection of Health Measurement Instruments; GRADE: grading of recommendations, assessment, development, and evaluations; PROM: patient-reported outcome measure

Results

We reviewed 10,623 citations and 11 websites [17-24] (Figure 2), from which we identified six instruments that in part assessed symptoms in patients in the chronic phase of CML: the European Organisation for Research and Treatment of Cancer Core Quality of Life Questionnaire for CML (EORTC QLQ-CML24) [19, 28, 29], the EORTC Symptom Set [30]; the Functional Assessment of Cancer Therapy-Leukaemia (FACT-LEU) [31, 32], a generic Chinese language

patient-reported outcome measure [33], the haematological malignancies patient-reported outcomes (HM-PRO) [34-39]; and the MD Anderson Symptom Inventory for CML (MDASI-CML) [40, 41]. The Comprehensive Symptom Profile in patients with CML was only reported in abstracts and, therefore, was excluded [42, 43]. Reasons for full text exclusion are listed in the Supplementary material. The main characteristics of the six identified patient-reported outcome measures are described in Table 1. Together, the six instruments cover 90 symptom clusters (Table 2).

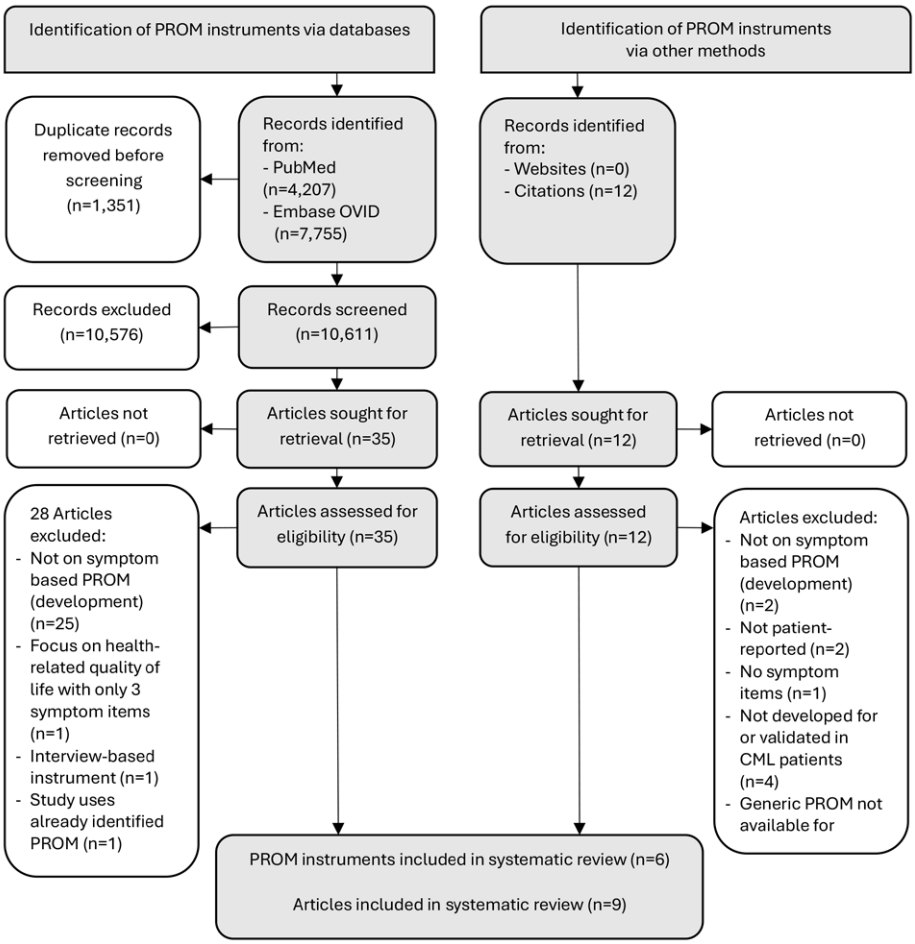


Figure 2 Study selection. Six PROM instruments that measure symptoms and have been developed for, or have been used in, patients with chronic myeloid leukaemia were identified. PROM: patient-reported outcome

In addition, the FACT-LEU asks whether patients have had side effects, and the HM-PRO offers an additional blank space to list any other symptom that concerns patients or any other comment they would like to make. Forty-seven items overlap across instruments, with some instruments asking more detailed questions than others, and 43 items are unique to only one of the patient-reported outcome measures. The EORTC QLQ-CML24 asks for details of skin problems in general, but not for the seven skin problems that are more specific (e.g., skin colour change, itchy skin, rash, dry, flaking or cracked skin, sore or painful skin, bruises, and haemorrhagic tendency of skin), which some other instruments address. Similarly, other instruments contain more detailed items for eye problems, mouth problems, joint or muscle pain, and swelling, compared with the EORTC QLQ-CML24. In the EORTC Symptom Set, hair loss and nausea are not measured, but this is the only questionnaire to address aches or pains in joints and has the most detailed questions on skin and eye problems. In the FACT-LEU, eye, mouth, joint and muscle problems are not assessed. The generic Chinese language questionnaire dedicates six items to sexual health and is the only questionnaire with items on hair colour change, hydrothorax and/or ascites, and weight gain. However, it does not address problems of the mouth, taste, or indigestion. As for the HM-PRO, eye, joint and muscle problems are missing. Lastly, in the MDASI, eye problems and swelling are not included.

Step 1: evaluation of development quality as described in studies

None of the patient-reported outcome measures were rated very good or adequate for development quality (step 1). Three patient-reported outcome measures were rated doubtful (EORTC QLQ-CML24, EORTC Symptom Set, and the MDASI-CML) because the origin of the construct (EORTC QLQ-CML24 and MDASI-CML), the context of use (EORTC Symptom Set), whether interviewers were trained or had experience (EORTC QLQ-CML24, EORTC Symptom Set, and the MDASI-CML), the approach used to analyse the data (MDASI-CML), or whether at least part of the data was coded independently (EORTC QLQ-CML24, EORT Symptom Set, MDASI-CML) were unclear. The three other patient-reported outcome measures (FACT-Leu, generic Chinese language questionnaire, and HM-PRO) were rated inadequate, because the patient-reported outcome measure was not developed in a sample representing the target population (generic Chinese language questionnaire), no cognitive interview study was done in the target population (the generic Chinese language questionnaire and HM-PRO), or comprehensibility or comprehensiveness were not assessed (all three questionnaires). A table with the ratings for each study is provided in the Supplementary material.

Table 1 Main characteristics of the six identified patient-reported outcome measures

	Construct measured (number of symptom items/total number of items)	**Target population (recall period)**	**Development method patient-reported outcome measure**
EORTC QLQ-CML24 (including the accompanying EORTC QLQ-C30) [19, 28, 29]	Health-related quality of life (25/54)	CML patients (past week)	Cross-culturally developed for research, involving a patient advocate group, 655 individual CML patients on TKIs, and 59 health care professionals experienced in CML (from Europe, United States, or Asia)
EORTC Symptom Set [30]	Symptom burden (61/61)	People with cancer treated with targeted therapy (past week)	Developed for research. Focus on symptoms during targeted therapy for CML, breast cancer, and gastrointestinal stromal tumours. 38 CML items developed from QLQ-CML24 data set, but not incorporated in the QLQ-CML24

	Scales and single items (number of items)	**Response options and range of scores or scoring**	• **Number of languages available** • **Licensing** • **Costs**
	EORTC QLQ-C30 • Global health status/ quality of life (2) • Physical functioning (5) • Role functioning (2) • Emotional functioning (4) • Cognitive functioning (2) • Social functioning (2) • Fatigue (3) • Nausea and vomiting (2) • Pain (2) • Single symptom items (5) • Financial difficulties (1) EORTC QLQ-CML24 • Symptom burden (13) • Impact on worry/mood (4) • Impact on daily life (3) • Body image problems (1) • Satisfaction with care and information (2) • Satisfaction with social life (1)	Either as: • Raw scores (single items only): o Not at all o A little o Quite a bit o Very much • Dichotomised scores: o E.g., 'not at all' vs. 'any extent' • Transformed scores: o 0 – 100 o Higher scores reflect more symptoms for symptom items	• 50 • Licencing through EORTC website • Free for academic use
	Single symptom items across 12 categories (skin, swelling, musculoskeletal, mouth, nails, heart and breathing, fatigue and energy, eyes, ear, nose, and throat, emotional functioning, digestion, and general) (61)	Each item should be scored separately rather than used to create subscale scores. Either as (see EORTC QLQ-CML24 above): • Raw scores • Dichotomised scores • Transformed scores	• 7 • Licensing through EORTC website • Free for academic use

Table 1 Continued

	Construct measured (number of symptom items/total number of items)	**Target population (recall period)**	**Development method patient-reported outcome measure**
FACT-LEU [31, 32]	Health-related quality of life (22/44)	Patients with chronic and acute leukaemia (past week)	Developed for research. Cross-culturally developed involving ten chronic leukaemia and 19 acute leukaemia patients, and 16 health care professionals experienced in leukaemia
Generic Chinese questionnaire [33]	Quality of life and symptom burden (37/59)	CML patients (not specified)	Generic Chinese language questionnaire developed for a specific study. Unclear how the development took place

	Scales and single items (number of items)	**Response options and range of scores or scoring**	• **Number of languages available** • **Licensing** • **Costs**
	Five subscales: • Physical wellbeing (7) • Social and family wellbeing (7) • Emotional wellbeing (6) • Functional wellbeing (7) • Additional concerns (17)	• Raw scores o Not at all o A little bit o Somewhat o Quite a bit o Very much The direct interpretation of raw scores from single items is hampered because some items have a negative direction, whereas other items have a positive direction • Sum scores Subscale scores range from 0-24, -28 or 0-68, depending on the number of items in a subscale. Total score ranges from 0 to 176. Higher scores indicate better quality of life	• 52 • Licensing through FACIT website • Free for investigator-initiated research, students, or clinical use
	• Patient factors, (co) medication, disease response, satisfaction with CML control, influence on work, costs of treatment • Single symptom items (37)	• Raw scores Degree o Mild o Moderate o Severe Discontinuation therapy due to symptoms o Yes o No Needing treatment o Yes o No	• 1 • Not described • Not described

Table 1 Continued

	Construct measured (number of symptom items/total number of items)	**Target population (recall period)**	**Development method patient-reported outcome measure**
HM-PRO [34-39]	Quality of life and symptom burden (18 plus blank space/42 plus 2 blank spaces)	Patients with haematological malignancy (quality of life: today; symptom burden: 3 days)	Developed for clinical practice in all haematological malignancies, on the basis of interviews with 129 patients from the United Kingdom, of whom 12 had CML
MDASI-CM (including the accompanying generic MDASI) [40, 41]	Symptom burden and symptom interference (20/26)	CML patients (24 hours)	Developed for clinical practice and research, with 38 CML patients (mainly in chronic phase), four physicians, four nurses, and four family caregivers, all from the United States

Abbreviations: EORTC: European Organization for Research and Treatment of Cancer; EORTC QLQ-CML24: EORTC Core Quality of Life Questionnaire for chronic myeloid leukaemia with 24 items; EORTC QLQ-C30: EORTC Core Quality of Life Questionnaire for cancer patients with 30 items; FACT-LEU: Functional Assessment of Cancer Therapy-Leukaemia; HM-PRO: haematological malignancies patient-reported outcomes; MDASI-CML: MD Anderson Symptom Inventory for chronic myeloid leukaemia

	Scales and single items (number of items)	**Response options and range of scores or scoring**	• **Number of languages available** • **Licensing** • **Costs**
	Part A 4 subscales: • Physical behaviour (7) • Social wellbeing (3) • Emotional behaviour (11) • Eating and drinking (3) Part B: • Symptom items (18)	• Raw scores Part A o Not at all o A little o A lot o Not applicable Part B o Not at all o Mild o Severe • Sum scores Subscale scores range from 0-6, 0-14 or 0-22, 0-36, or 0-48, depending on the number of items in a subscale. The higher the total score, the greater the effect on a patient's quality of life	• 11 • Licensing through HM-PRO website • Currently no costs for use in academic research and routine daily practice
	Two subscales: • Symptoms (20) • Symptom interference (6)	• Raw scores 0 being not present to 10 being as bad as you can imagine • Mean subscale scores as the arithmetic mean of items in the subscale	• 19 • Licensing through MD Anderson website • Fees vary depending on the type and extent of use, the setting and who is sponsoring

Table 2 Symptom items included in the six identified patient-reported outcome measures

Items	EORTC QLQ-CML24	EORTC Symptom Set	FACT-LEU	Generic Chinese language questionnaire	HM-PRO	MDASI-CML
Skin problems	√	..	..	..	√	√
Skin colour change	..	√	..	√	..	..
Itchy skin	..	√	..	√	..	..
Skin rash	..	√	..	..	..	..
Dry, flaking or cracked skin	..	√	..	..	..	..
Sore or painful skin	..	√	..	..	..	..
Bruises	..	√	√	..	..	√
Haemorrhagic tendency of skin	..	..	..	√	..	..
Nails break easily	..	√	..	..	..	..
Hair loss	√	..	..	√	√	..
Hair colour change	..	..	..	√	..	..
Eye problems	√	..	..	..	..	..
Watery eyes	..	√	..	..	..	..
Conjunctivitis	..	..	..	√	..	..
Conjunctival haemorrhage	..	..	..	√	..	..
Dry eyes	..	√	..	√	..	..
Burning eyes	..	√	..	..	..	..
Red eyes	..	√	..	..	..	..
Itchy eyes	..	√	..	..	..	..
Light sensitivity	..	√	..	..	..	..
Blurred vision	..	√	..	..	..	..
Nose bleeds	..	√	..	..	..	..
Other nose problems (smell, sneezing)	..	√	..	..	..	..
Hearing problems	..	√	..	..	..	..
Tinnitus	..		..	√	..	..
Dry mouth	√	√	..	..	..	√
Taste	..	√	..	..	√	..
Pain or soreness mouth	..	√	..	..	..	..
Sore throat	..	√	..	..	..	..
Changes to voice	..	√	..	..	..	..
Cough	..	√	..	..	√	√
Difficulty breathing	√	√	..	..	..	..
Chest pain	..	√	..	..	√	..
Chest distress and shortness of breath	..	..	..	√	..	..

Table 2 Continued

Items	EORTC QLQ-CML24	EORTC Symptom Set	FACT-LEU	Generic Chinese language questionnaire	HM-PRO	MDASI-CML
Palpitations	..	√	..	√	..	..
Hydrothorax and/or ascites	..	..	..	√	..	..
Acid indigestion/ heartburn	√	√	..		..	..
Constipation	√	√	..	√	√	..
Diarrhoea	√	√	..	√	√	√
Toilet urgency	..	√	..	..	..	..
Nausea	√	..	√	√	√	√
Vomiting	√	√	..	√	..	√
Lack of appetite	√	√	√	√	√	√
Abdominal pains or cramps	√	√	..	√	√	..
Painful bowel movements	..	√	..	…	..	..
Abdominal distension	..	√	..	√	..	..
Flatulence	..	√	..	..	..	..
Urinating frequently	√	..	..	..	..	..
Joint or muscle pain	√	√	..	√	..	..
Aches or pains in joints	..	√	..	..	..	..
Aches or pains in bones	..	√	..	..	..	..
Muscle aches, pains or cramps	..	√	..	..	..	..
Muscle cramps	√	..	..	√	..	√
Swelling	√	..	..	..	..	√
Swelling legs and ankles	..	√	..	..	..	..
Swelling face/ around eyes	..	√	..	..	..	..
Periorbital and lower limb oedema	..	..	..	√	..	..
Lumps	..	..	√	..	√	..
Pain	√ (2Q)	..	√ (2 Q)	..	..	√
Body pain	..	..	..	..	√	..
Back pain	..	√	..	..	√	..
Headaches	√	√	..	√	√	√
Dizziness	..	√	..	√	…	..
Tiredness	√ (3 Q)	√	√	√	√	√
Lack of energy	..	√	√	..	..	..

Table 2 Continued

Items	EORTC QLQ-CML24	EORTC Symptom Set	FACT-LEU	Generic Chinese language questionnaire	HM-PRO	MDASI-CML
Drowsiness	√	..	..	..	..	√
Disturbed sleep	√	√	..	√	√	√
Weakness	√	√	√	..	..	..
Distress	√	√ (3 Q)	√ (5Q)	..	√	√ (2Q)
Depression/anxiety	..	..	..	√	√	..
Emotional ups and downs	..	..	√	..	..	...
Memory problems	√	√	..	√	..	√
Concentration	√	√	..	..	√	..
Fever/chills	..	√	√ (2 Q)	..	√	..
Infection	..	..	√	..	√	..
Malaise	..	√	..	..	√	√
Numbness/tingling	..	√	..	..	..	√
Pale/cold fingers/toes	..	√	..	..	..	..
Sweating	√	√	√	√	√	..
Hot flushes	..	√	..	..	..	..
Weight loss	..	..	√	√	..	..
Weight gain	..	..	..	√	..	..
Bleeding easily	..	..	√	..	..	..
Bleeding gums	..	√	..	..	..	..
(Male) swollen breast	..	..	..	√	..	..
(Female) swollen breast	..	..	..	√	..	..
(Female ≤ 50 years) hypermenorrhoea	..	..	..	√	..	..
(Female ≤ 50 years) hypomenorrhea	..	..	..	√	..	..
(Female ≤ 50 years) amenorrhea	..	..	..	√	..	..
Decrease in sexual desire	..	..	..	√	..	..
Bothered by side effects	..	..	√	..	..	..
Blank space	..	..	..	..	√	..

√ Item included in patient-reported outcome instrument; Q: questions. Abbreviations: EORTC: European Organization for Research and Treatment of Cancer; EORTC QLQ-CML24: EORTC Core Quality of Life Questionnaire for chronic myeloid leukaemia with 24 items; FACT-LEU: Functional Assessment of Cancer Therapy-Leukaemia; HM-PRO: haematological malignancies patient-reported outcomes; MDASI-CML: MD Anderson Symptom Inventory for chronic myeloid leukaemia

Step 2: evaluation of content validity as described in studies

On the basis of content validity studies, none of the patient-reported outcome measures were rated very good or adequate for their content validity. HM-PRO and MDASI-CML were rated doubtful because whether interviewers were trained or had experience (MDASI-CML), a topic guide was used (MDASI-CML), all interviews were recorded and transcribed ad verbatim (MDASI-CML), two researchers were involved in the analysis (MDASI-CML), or professionals from all required disciplines were included (MDASI-CML) was unclear, because the approach used to analyse the data was unclear (MDASI-CML), or because less than 30 professionals or very few CML patients were involved in a quantitative survey to assess content validity (HM-PRO and MDASI-CML). The other four patient-reported outcome measures were rated inadequate because professionals, different from the professionals involved in the development phase, were not involved in assessing the relevance and comprehensiveness of patient-reported measurement items. A table with the ratings per study is provided in the Supplementary material.

Step 3: rating of overall content validity

Overall content validity was based on steps 1 and 2 plus reviewers' assessment, with quality appraisal of the evidence. Five patient-reported outcome measures were rated as having inconsistent overall content validity (Table 3). Although many items were rated sufficient, insufficient ratings that led to overall inconsistent ratings were given because not all items were relevant for symptom burden (EORTC QLQ-CML24, FACT-LEU, and MDASI-CML), comprehensiveness was not assessed in different groups from the patient-reported outcome measurement development group (EORTC QLQ-CML24, EORTC Symptom Set, and FACT-LEU), comprehensiveness was assessed as insufficient by the reviewers (HM-PRO and MDASI-CML), or not all items were relevant to CML patients (EORTC Symptom Set). Reviewers' assessment of the patient-reported outcome measures was similar to, or more positive than, ratings based on studies from steps 1 and 2, except for their assessment of HM-PRO and MDASI-CML. For these two instruments reviewers assessed the comprehensiveness as insufficient because key symptoms for CML patients (e.g., eye problems or muscle cramps) were missing.

The quality of the evidence of the rating of the five instruments was moderate (MDASI-CML) to low (EORTC QLQ-CML-24, EORTC Symptom Set, FACT-LEU, and HM-PRO). We downgraded the quality of the evidence because at least one content validity study of doubtful quality indicated moderate quality evidence

(MDASI-CML), and we further downgraded for indirectness because quality of life, not symptom burden was the focus of the patient-reported outcome measure or the target population differed from CML patients (EORTC QLQ-CML-24, EORTC Symptom Set, FACT-LEU, and HM-PRO). The generic Chinese language patient-reported outcome measure was rated as having insufficient overall content validity with very low quality evidence, downgraded because the development study was inadequate and content validity studies were missing.

Table 3 Overall content validity of identified patient-reported outcome measures, and quality appraisal according to the COSMIN-modified GRADE approach

	Relevance	**Comprehen-siveness**	**Comprehen-sibility**	**Overall content validity rating (quality of the evidence)**
EORTC QLQ-CML24	Inconsistent	Inconsistent	Sufficient	Inconsistent (low)
EORTC Symptom Set	Inconsistent	Inconsistent	Sufficient	Inconsistent (low)
FACT-LEU	Inconsistent	Insufficient	Inconsistent	Inconsistent (low)
Generic Chinese questionnaire	Insufficient	Insufficient	Insufficient	Insufficient (very low)
HM-PRO	Inconsistent	Inconsistent	Sufficient	Inconsistent (low)
MDASI-CML	Sufficient	Insufficient	Sufficient	Inconsistent (moderate)

Quality of the evidence is our confidence (i.e., high, moderate, low or very low) in the validity rating. Abbreviations: EORTC: European Organization for Research and Treatment of Cancer; EORTC QLQ-CML24: EORTC Core Quality of Life Questionnaire for chronic myeloid leukaemia with 24 items; FACT-LEU: Functional Assessment of Cancer Therapy-Leukaemia; HM-PRO: haematological malignancies patient-reported outcomes; MDASI-CML: MD Anderson Symptom Inventory for chronic myeloid leukaemia.

Discussion

When choosing an appropriate patient-reported outcome measure to evaluate treatment toxic effects, both content validity and feasibility for the context of use should be considered. Therefore, we set out to give an overview of available instruments, their content validity and feasibility characteristics. Regrettably, none of the six identified patient-reported outcome measures were rated as sufficient for overall content validity. A major shortcoming was the absence of comprehensiveness evaluation in a patient or clinician sample independent from the development phase. Five instruments were rated inconsistent but have potential. However, of these five patient-reported outcome measures, reviewers assessed the two EORTC instruments as the

only instruments sufficient for comprehensiveness, as the other three patient-reported outcome measures did not include eye problems and muscle cramps, which we consider key symptoms of TKI side effects. Unfortunately, both EORTC instruments were rated inconsistent for relevance, as the QLQ-CML24 includes items specific to quality of life, whereas the Symptom Set includes items sourced from patients with breast and gastrointestinal stromal tumours cancers on targeted therapy.

In previous studies, patients with chronic haematology cancer reported that questions from the EORTC QLQ-C30 and Outcomes and Experiences Questionnaire did not resonate with their personal experience [44, 45], with some patients feeling they were forced to complete irrelevant items to complete the questionnaire. Such negative feelings towards more generic instruments, mostly aimed towards quality of life in more impaired patients, may hamper patient-reported outcome measure response in real-world settings. The Symptom Set, on the other hand, has more detailed questions on side effects of importance to CML patients, such as skin and eye problems. However, the many items that were not sourced from CML patients might also negatively influence a patient's response. Unfortunately, no evidence is available as to which patient-reported outcome measure CML patients prefer.

Further guidance in the selection of the most appropriate patient-reported outcome measure cannot be found in instruments that are typically used in CML studies. A mere 6% of leukaemia trials included any patient-reported outcome measure as an outcome measure according to an assessment published in 2018 [1]. Landmark trials that evaluated patient-reported outcome measures are few and far between, reporting on few symptoms [46] or only reporting only on subscales for quality of life but not for separate symptoms [47]. Moreover, to the best of our knowledge, at present no single patient-reported outcome measure is recommended by standard initiatives, such as professional societies or research organisations.

Our search and selection process was quite rigorous: two researchers independently searched multiple sources in different languages. In addition to our search, to cover the time before the search start date, we used a published systematic review [17]. Although this approach is pragmatic, we feel confident that we have not missed patient-reported outcome measures with an extensive development trajectory in CML patients, as such labour-intensive trajectories would be likely to have resulted in additional publications. Furthermore,

our appraisal of identified instruments according to the rigorous COSMIN methodology is also a strength. COSMIN is strict and based, in part, on the least score counts method. In addition, COSMIN does not distinguish poor quality from poor reporting, and might, therefore, be unnecessarily strict in some cases. However, this harshness does reveal the weaknesses of the development and validation of assessed instruments, and thus might help in improving them.

In our process, we focused on CML-specific patient-reported outcome measures, but more generic instruments, such as the National Cancer Institute's patient-reported outcome measurement system companion to the Common Terminology Criteria for Adverse Events (PRO-CTCAE), might also be of use [48]. PRO-CTCAE has gained much attention since remote self-reporting with feedback to clinicians helped to improve health-related quality of life and survival, and reduce emergency room visits and hospitalisations in patients with a solid cancer [49, 50]. Although positive results have also been reported with EORTC instruments [51], most research uses PRO-CTCAE symptom sets at present. The PRO-CTCAE is especially aimed towards capturing treatment toxic effects and consists of an item library with 124 items, representing 78 symptomatic toxic effects. However, the EORTC Item Library is ten times larger with over 1,000 items available [29], allowing for rapid expansion of items, whereas the PRO-CTCAE does not have many TKI-specific toxic effect symptoms. When we consider the 38 unique symptoms from the two EORTC instruments (taking only those symptoms from the EORTC Symptom Set that were drawn from CML patients), around a third of items is unavailable from the PRO-CTCAE item library, including key TKI toxic effects, such as eye problems and muscle cramps.

Symptom items from both the EORTC instruments were considered the most comprehensive and, therefore, seem to be a good base to start development of a future patient-reported outcome measure, specific for toxic effects in CML patients. Patients and patient organisations need not only be engaged in the further development of patient-reported outcome measures, but also in their use in daily life to self-report toxic effects [2] beyond the current practice of top-down toxic effect self-reporting in clinical studies. Existing frameworks for patient-reported outcome measure development and implementation offer little guidance on how and when to involve patients and patient advocates [52, 53], although some more elaborate guidance is targeted at patients advocates [54]. In our opinion, involvement starts with collaborating

with patient advocates on the goal of the patient-reported outcome measure and interviewing individual patients on their needs and preferences. The involvement of diverse patients, including patients with low health-literacy, expert patients [55], and patient advocates, in participatory efforts based on design thinking methodology [56], action research [57, 58], or implementation research methodology [59, 60] is best practice. In addition, key features to enhance the feasibility and acceptability of electronic patient-reported outcome measure administration for patients, including haematology patients, have been described in the literature and include reminders to complete surveys, free text options for other symptoms, the ability to monitor and patients' own data, and password-free log in, among others [61, 62]. However, in practice, patients' involvement in patient-reported outcome measure development, especially beyond item development and comprehensibility, still needs strengthening to cater fully to patients' needs for self-management of toxic effects, such as insight into their symptoms over time and tailored self-management advice [63, 64].

A patient-reported outcome measure with sufficient content validity and patient adherence is needed to monitor and manage symptoms during TKI treatment and -withdrawal [9, 11]. Evidence from solid tumours, encompassing reduced symptom burden, improved adherence, physical functioning and quality of life, and a potential increase in survival, has led to electronic symptom monitoring being a level 1 grade A recommendation in the European Society for Medical Oncology Clinical Practice Guidelines [65]. Clinical decision making in CML, on the basis of captured symptoms, might be complex due to population and treatment diversity, coupled with a need for disease control. However, TKI-related toxic effect monitoring is likely to enhance clinical decision making, as has happened in solid tumours, through actionable clinical responses based on best practices and patient involvement through self-management [65]. In addition, there are knowledge gaps in the practical application of current clinical CML guidance that can be evaluated with aggregated TKI toxic effect data, which might subsequently lead to practical recommendations. For example, up to 30% of patients switch TKI due to intolerability of the initial TKI [11], but whether switching decreases symptoms, and, if so, to what extent and in which cases, is unknown at present.

This systematic review identified six patient-reported outcome measures that elicit TKI toxic effects in chronic phase CML patients. None were rated sufficient for overall content validity according to the rigorous COSMIN

method. Two EORTC patient-reported outcome measures scored best on comprehensiveness for CML treatment-specific toxic effects, as assessed by reviewers, but were rated inconsistent on relevance for TKI-specific toxic effects in CML patients. The assessment and transparent description of content validity and feasibility features in this review enables clinicians, researchers and policy makers to make an informed choice from available instruments.

Future research is needed to establish a patient-reported outcome measure for toxic effects chronic phase CML patients treated with TKIs that would be rated as sufficient for overall content validity, either through further validating and developing an already existing patient-reported outcome measure, or by composing a new patient-reported outcome measure. A future patient-reported outcome measure might best be constructed out of the CML treatment-specific items of the two EORTC instruments. In addition, the psychometric properties of individual symptom items need to be evaluated, especially the responsiveness of items to starting, stopping or altering the dose of TKIs. Finally, we need knowledge on the preferences of CML patients for the practical use of self-reported treatment toxic effects in the real-world, to be able to implement suitable instruments successfully.

References

1. Thanarajasingam, G., et al., *Beyond maximum grade: modernising the assessment and reporting of adverse events in haematological malignancies.* Lancet Haematol, 2018. **5**(11): p. e563-e598.
2. Thanarajasingam, G., et al., *Reaching beyond maximum grade: progress and future directions for modernising the assessment and reporting of adverse events in haematological malignancies.* Lancet Haematol, 2022. **9**(5): p. e374-e384.
3. Foulon, S., et al., *Health state utility and quality of life measures in patients with chronic myeloid leukemia in France.* Qual Life Res, 2021. **30**(7): p. 2021-2032.
4. Phillips, K.M., et al., *Quality of life outcomes in patients with chronic myeloid leukemia treated with tyrosine kinase inhibitors: a controlled comparison.* Support Care Cancer, 2013. **21**(4): p. 1097-103.
5. Efficace, F., et al., *Health-related quality of life in chronic myeloid leukemia patients receiving long-term therapy with imatinib compared with the general population.* Blood, 2011. **118**(17): p. 4554-60.
6. Guerin, A., et al., *Impact of low-grade adverse events on health-related quality of life in adult patients receiving imatinib or nilotinib for newly diagnosed Philadelphia chromosome positive chronic myelogenous leukemia in chronic phase.* Curr Med Res Opin, 2014. **30**(11): p. 2317-28.
7. Steegmann, J.L., et al., *European LeukemiaNet recommendations for the management and avoidance of adverse events of treatment in chronic myeloid leukaemia.* Leukemia, 2016. **30**(8): p. 1648-71.
8. Geelen, I.G.P., et al., *Treatment outcome in a population-based, 'real-world' cohort of patients with chronic myeloid leukemia.* Haematologica, 2017. **102**(11): p. 1842-1849.
9. Sharf, G., et al., *Treatment-free remission in chronic myeloid leukemia: the patient perspective and areas of unmet needs.* Leukemia, 2020. **34**(8): p. 2102-2112.
10. Ross, D.M., et al., *Durable treatment-free remission in patients with chronic myeloid leukemia in chronic phase following frontline nilotinib: 96-week update of the ENESTfreedom study.* J Cancer Res Clin Oncol, 2018. **144**(5): p. 945-954.
11. Hochhaus, A., et al., *European LeukemiaNet 2020 recommendations for treating chronic myeloid leukemia.* Leukemia, 2020. **34**(4): p. 966-984.
12. Basch, E., et al., *Patient versus clinician symptom reporting using the National Cancer Institute Common Terminology Criteria for Adverse Events: results of a questionnaire-based study.* Lancet Oncol, 2006. **7**(11): p. 903-9.
13. Fromme, E.K., et al., *How accurate is clinician reporting of chemotherapy adverse effects? A comparison with patient-reported symptoms from the Quality-of-Life Questionnaire C30.* J Clin Oncol, 2004. **22**(17): p. 3485-90.
14. Pakhomov, S.V., et al., *Agreement between patient-reported symptoms and their documentation in the medical record.* Am J Manag Care, 2008. **14**(8): p. 530-9.
15. Efficace, F., et al., *Patient- versus physician-reporting of symptoms and health status in chronic myeloid leukemia.* Haematologica, 2014. **99**(4): p. 788-93.
16. Page, M.J., et al., *The PRISMA 2020 statement: an updated guideline for reporting systematic reviews.* Bmj, 2021. **372**: p. n71.

17. Goswami, P., Y. Khatib, and S. Salek, *Haematological malignancy: Are we measuring what is important to patients? A systematic review of quality-of-life instruments.* Eur J Haematol, 2019. **102**(4): p. 279-311.
18. Salas, M., et al., *Validated instruments of quality of life (QOL) in patients with acute myeloid leukemia (AML) and other cancers.* Front Pharmacol, 2020. **11**: p. 1109.
19. Efficace, F., et al., *International development of an EORTC questionnaire for assessing health-related quality of life in chronic myeloid leukemia patients: the EORTC QLQ-CML24.* Qual Life Res, 2014. **23**(3): p. 825-36.
20. Cella, D., C.J. Nowinski, and O. Frankfurt, *The impact of symptom burden on patient quality of life in chronic myeloid leukemia.* Oncology, 2014. **87**(3): p. 133-47.
21. Efficace, F., et al., *Optimizing health-related quality of life in patients with chronic myeloid leukemia treated with tyrosine kinase inhibitors.* Expert Review of Hematology., 2021.
22. Efficace, F., et al., *Tyrosine-kinase inhibitors and patient-reported outcomes in chronic myeloid leukemia: a systematic review.* Leuk Res, 2013. **37**(2): p. 206-13.
23. Efficace, F., et al., *Health-related quality of life assessment and reported outcomes in leukaemia randomised controlled trials - a systematic review to evaluate the added value in supporting clinical decision making.* Eur J Cancer, 2008. **44**(11): p. 1497-506.
24. Schoenbeck, K.L. and K.E. Flynn, *Health-related quality of life of patients with chronic myeloid leukemia as measured by patient-reported outcomes: current state and future directions.* Curr Hematol Malig Rep, 2021. **16**(6): p. 491-499.
25. Terwee, C.B., et al., *COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study.* Qual Life Res, 2018. **27**(5): p. 1159-1170.
26. Mokkink, L.B., et al., *COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures.* Qual Life Res, 2018. **27**(5): p. 1171-1179.
27. Prinsen, C.A.C., et al., *COSMIN guideline for systematic reviews of patient-reported outcome measures.* Qual Life Res, 2018. **27**(5): p. 1147-1157.
28. Efficace, F., et al., *Validation and reference values of the EORTC QLQ-CML24 questionnaire to assess health-related quality of life in patients with chronic myeloid leukemia.* Leuk Lymphoma, 2021. **62**(3): p. 669-678.
29. *EORTC Item Library*. [cited 2022 October 3]; Available from: https://www.eortc.be/itemlibrary/#home/tab/official.
30. Sodergren, S.C., et al., *Developing symptom lists for people with cancer treated with targeted therapies.* Targeted Oncology, 2021. **16**(1): p. 95-107.
31. Cella, D., et al., *Measuring health-related quality of life in leukemia: the Functional Assessment of Cancer Therapy--Leukemia (FACT-Leu) questionnaire.* Value Health, 2012. **15**(8): p. 1051-8.
32. *FACT-Leu Functional Assessment of Cancer Therapy - Leukemia*. [cited 2022 October 3]; Available from: https://www.facit.org/measures/FACT-Leu.
33. Yu, L., et al., *Variables associated with patient-reported symptoms in persons with chronic phase chronic myeloid leukemia receiving tyrosine kinase inhibitor therapy.* Medicine (Baltimore), 2019. **98**(48): p. e18079.
34. Goswami, P., et al., *Reliability of a novel hematological malignancy specific patient-reported outcome measure: HM-PRO.* Front Pharmacol, 2020. **11**: p. 571066.
35. Goswami, P., et al., *Quality-of-life issues and symptoms reported by patients living with haematological malignancy: a qualitative study.* Ther Adv Hematol, 2020. **11**: p. 2040620720955002.

36. Goswami, P., et al., *Paper and electronic versions of HM-PRO, a novel patient-reported outcome measure for hematology: an equivalence study.* J Comp Eff Res, 2019. **8**(7): p. 523-533.
37. Goswami, P., et al., *Hematological malignancy specific patient-reported outcome measure (HM-PRO): construct validity study.* Frontiers in Pharmacology, 2020. **11 (aricle 1308)**.
38. Goswami, P., et al., *Development of a novel hematological malignancy specific patient-reported outcome measure (HM-PRO): content validity.* Frontiers in Pharmacology, 2020. **11 (article 209)**.
39. *HM-PRO*. [cited 2022 October 3]; Available from: https://hmpro.co.uk/.
40. Williams, L.A., et al., *Measuring the symptom burden associated with the treatment of chronic myeloid leukemia.* Blood, 2013. **122**(5): p. 641-7.
41. *MDASI Chronic Myeloid Leukemia Module*. [cited 2022 October 3]; Available from: https://www.mdanderson.org/research/departments-labs-institutes/departments-divisions/symptom-research/symptom-assessment-tools/md-anderson-symptom-inventory-chronic-myeloid-leukemia-module.html.
42. Nikitina, T., et al. *Comprehensive symptom profile in patients with chronic myeloid leukemia: practicability and sensitivity of the new symptom assessment tool CSP LEUK-CML*. 2014 2021 May 6]; E-poster]. Available from: https://library.ehaweb.org/eha/2014/19th/54072/tatyana.ionova.comprehensive.symptom.profile.in.patients.with.chronic.myeloid.html.
43. Zhang, S., et al., *Targeted literature review of patient reported outcomes (PROS) in chronic myeloid leukemia (CML) patients receiving second and later lines of treatment.* Blood, 2020. **136 (SUPPL 1)**: p. 26-27.
44. Thestrup Hansen, S., et al., *"I am sure that they use my PROM data for something important." A qualitative study about patients' experiences from a hematologic outpatient clinic.* Cancer Nurs, 2020. **43**(5): p. E273-E282.
45. Thestrup Hansen, S., et al., *User experiences on implementation of patient reported outcome measures (PROMs)in a Haematological outpatient clinic.* J Patient Rep Outcomes, 2020. **4**(1): p. 87.
46. Hahn, E.A., et al., *Quality of life in patients with newly diagnosed chronic phase chronic myeloid leukemia on imatinib versus interferon alfa plus low-dose cytarabine: results from the IRIS Study.* J Clin Oncol, 2003. **21**(11): p. 2138-46.
47. Cortes, J.E., et al., *Patient-reported outcomes in the phase 3 BFORE trial of bosutinib versus imatinib for newly diagnosed chronic phase chronic myeloid leukemia.* J Cancer Res Clin Oncol, 2019. **145**(6): p. 1589-1599.
48. Basch, E., et al., *Development of the National Cancer Institute's patient-reported outcomes version of the common terminology criteria for adverse events (PRO-CTCAE).* J Natl Cancer Inst, 2014. **106**(9).
49. Basch, E., et al., *Symptom monitoring with patient-reported outcomes during routine cancer treatment: a randomized controlled trial.* J Clin Oncol, 2016. **34**(6): p. 557-65.
50. Basch, E., et al., *Overall survival results of a trial assessing patient-reported outcomes for symptom monitoring during routine cancer treatment.* JAMA, 2017. **318**(2): p. 197-198.
51. Gibbons, C., et al., *Routine provision of feedback from patient-reported outcome measurements to healthcare providers and patients in clinical practice.* Cochrane Database Syst Rev, 2021. **10**: p. CD011589.
52. van der Wees, P.J., et al., *Development of a framework with tools to support the selection and implementation of patient-reported outcome measures.* J Patient Rep Outcomes, 2019. **3**(1): p. 75.

53. Aaronson, N., et al., *User's guide to implementing patient-reported outcomes assessment in clinical practice*. 2015, International Society for Quality of Life Research (ISOQOL).
54. van der Wees, P. and V. Ildikó, *Patient participation in patient reported outcome measures (PROMs) [In Dutch]*. 2015, Netherlands Patient Federation
55. Anampa-Guzman, A., et al., *The rise of the expert patient in cancer: from backseat passenger to co-navigator.* JCO Oncol Pract, 2022. **18**(8): p. 578-583.
56. Ector, G.I., et al., *The development of a web-Based, patient-centered intervention for patients with chronic myeloid leukemia (CMyLife): design thinking development approach.* J Med Internet Res, 2020. **22**(5): p. e15895.
57. Baum, F., C. MacDougall, and D. Smith, *Participatory action research.* J Epidemiol Community Health, 2006. **60**(10): p. 854-7.
58. Verweij, L., et al., *Participatory action research: method to improve, implement and evaluate a real-life digital care platform (under review).*
59. Stover, A.M., et al., *Using an implementation science approach to implement and evaluate patient-reported outcome measures (PROM) initiatives in routine care settings.* Qual Life Res, 2021. **30**(11): p. 3015-3033.
60. Manalili, K., M.J. Santana, and I.P.P.i.c.p.i.s.w. group, *Using implementation science to inform the integration of electronic patient-reported experience measures (ePREMs) into healthcare quality improvement: description of a theory-based application in primary care.* Qual Life Res, 2021. **30**(11): p. 3073-3084.
61. Rocque, G.B., et al., *Adaptation of remote symptom monitoring using electronic patient-reported outcomes for implementation in real-world settings.* JCO Oncol Pract, 2022. **18**(12): p. e1943-e1952.
62. Warrington, L., et al., *Electronic systems for patients to report and manage side effects of cancer treatment: systematic review.* J Med Internet Res, 2019. **21**(1): p. e10875.
63. Lai-Kwon, J., et al., *The feasibility, acceptability, and effectiveness of electronic patient-reported outcome symptom monitoring for immune checkpoint inhibitor toxicities: a systematic review.* JCO Clin Cancer Inform, 2023. **7**: p. e2200185.
64. Wiering, B., D. de Boer, and D. Delnoij, *Patient involvement in the development of patient-reported outcome measures: a scoping review.* Health Expect, 2017. **20**(1): p. 11-23.
65. Di Maio, M., et al., *The role of patient-reported outcome measures in the continuum of cancer clinical care: ESMO Clinical Practice Guideline.* Ann Oncol, 2022. **33**(9): p. 878-892.

Supplementary material

Full Systematic search for available instruments that measure symptoms in chronic phase CML patients

		Results April 2021	Results October 2022
PubMed			
1	"Leukemia, Myelogenous, Chronic, BCR-ABL Positive"[Mesh] OR "chronic myelogenous leukemia" OR "chronic myeloid leukemia" OR "Ph1 positive chronic myelogenous" OR "Ph1-positive chronic myelogenous" OR "Ph1 positive chronic myelogenous" OR "Ph1-positive chronic myelogenous" OR "chronic myelogenous Ph1 positive" OR "chronic myelogenous Ph1-positive" OR "chronic myeloid Ph1 positive" OR "chronic myeloid Ph1-positive" OR "Philadelphia positive chronic myeloid leukemia" OR "Philadelphia-positive chronic myeloid leukemia" OR "Hematologic Neoplasms"[Mesh] OR "haematological malignancy" OR "hematological malignancy" OR "haematological malignancies" OR "hematological malignancies" OR "haematologic malignancy" OR "hematologic malignancy" OR "haematologic malignancies" OR "hematologic malignancies" OR "haematological neoplasm" OR "hematological neoplasm" OR "haematological neoplasms" OR "hematological neoplasms" OR "haematologic neoplasm" OR "hematologic neoplasm" OR "haematologic neoplasms" OR "hematologic neoplasms" OR "haematopoietic malignancy" OR "hematopoietic malignancy" OR "haematopoietic malignancies" OR "hematopoietic malignancies" OR "haematopoietic neoplasm" OR "hematopoietic neoplasm" OR "haematopoietic neoplasms" OR "hematopoietic neoplasms"	74,089	81,591
2	"Life Quality" OR "Quality of life" OR "Health-Related Quality Of Life" OR "Health Related Quality Of Life" OR "HRQoL" OR "HR-QoL" OR "QoL" OR "Quality of Life"[Mesh] OR "Patient Reported Outcome Measures"[Mesh] OR "Patient reported outcome measure" OR "patient reported outcome" OR "patient reported outcomes" OR "patient-reported outcome" OR "patient-reported outcomes" OR "Symptom Assessment"[Mesh] OR "symptom assessment" OR "symptom assessments" OR "symptom evaluation" OR "symptoms evaluations" OR "symptom monitoring" OR "symptoms monitoring" OR "symptom tracking" OR "symptoms tracking" OR "outcomes assessment" OR "outcome assessment" OR "Drug-Related Side Effects and Adverse Reactions"[Mesh] OR "adverse effect" OR "adverse effects" OR "side effect" OR "side effects" OR "adverse reaction" OR "adverse reactions" OR "adverse event" OR "adverse events"	2,681,373	2,952,210

Table Continued

		Results April 2021	Results October 2022
3	1 AND 2	13,947	15,788
4	Filters: Humans, Dutch, English, French, German, Spanish	12,277	14,056
5	Filters: from 2017 – 2021 (April 2021 search only)	2,677	
6	Filters: from 2021 – 2022 (October 2022 search only)		1,530
Embase OVID			
1	chronic myeloid leukemia.mp. or exp chronic myeloid leukemia/ or hematologic malignancy/ or haematological malignancy.mp. or hematological malignancy.mp. or haematological malignancies.mp. or hematological malignancies.mp. or haematologic malignancy.mp. or hematologic malignancy.mp. or haematologic malignancies.mp. or hematologic malignancies.mp. or haematological neoplasm.mp. or hematological neoplasm.mp. or haematological neoplasms.mp. or hematological neoplasms.mp. or haematologic neoplasm.mp. or hematologic neoplasm.mp. or haematologic neoplasms.mp. or hematologic neoplasms.mp. or haematopoietic malignancy.mp. or hematopoietic malignancy.mp. or haematopoietic malignancies.mp. or hematopoietic malignancies.mp. or haematopoietic neoplasm.mp. or hematopoietic neoplasm.mp. or haematopoietic neoplasms.mp. or hematopoietic neoplasms.mp.	109,642	123,982
2	"quality of life"/ or quality of life.mp. or life quality.mp. or HRQoL.mp. or HR-QoL.mp. or QoL.mp. or patient-reported outcome/ or patient reported outcome*.mp. or patient-reported outcome*.mp. or symptom assessment/ or symptom assessment*.mp. or symptoms assessment*.mp. or symptom evaluation*.mp. or symptoms evaluation*.mp. or symptom monitoring.mp. or symptoms monitoring.mp. or symptom tracking.mp. or symptoms tracking.mp. or outcomes assessment.mp. or outcome assessment.mp. or side effect.mp. or side effect/ or adverse event/ or adverse effect*.mp. or side effect*.mp. or adverse reaction*.mp. or adverse event*.mp.	2,752,165	3,127,986
3	1 and 2	23,901	28,260
4	limit 3 to (human and (dutch or English or german or french or spanish))	22,511	26,761
5	limit 4 to yr="2017 -Current (April 2021 search only)	5,972	
5	limit 4 to yr="2021 -Current (October 2022 search only)		4,008
6	limit 5 to "remove medline records" (October 2022 search only)		2,042
	Duplicates removed	75	184
	Final results Embase	5,897	1,858
	Pubmed and Embase combined	**8,574**	**3,388**

Table Continued

	Results April 2021	Results October 2022
Duplicates between Pubmed and Embase removed	1,196	11
Final results Pubmed and Embase combined	7,378	3,377
Duplicates with April 2021 search removed (October 2022 search only)		3,233
Excluded on title and/or abstract	7,360	3,216
Included on title and/or abstract	18	17
Identified through reference tracking as needed to examine full text	12	0
Total articles examined full text	30	17
Articles excluded full text	21	17
Articles included	9	0

Websites searched for instruments that measure symptoms of chronic phase CML patients

Organisation/ initiative	URL	Access data	New results
COMET	www.comet-initiative.org/	May 27, 2021	0
COSMIN	Database.cosmin.nl	May 27, 2021	0
EORTC	qol.eortc.org/questionnaires/	May 27, 2021	0
eProvide	eprovide.mapi-trust.org/	May 27, 2021	0
Health Measures, including PROMIS	www.healthmeasures.net/	May 27, 2021	0
ICHOM	www.ichom.org/	May 27, 2021	0
ISOQOL	www.isoqol.org	May 27, 2021	0
ISQOLS	isqols.org/	May 27, 2021	0
PoCoG	www.pocog.org.au	May 27, 2021	0
PRO-CTCAE	Healthcaredelivery.cancer.gov/pro-ctcae	May 27, 2021	0
WHO ICF	www.icf-core-sets.org	May 27, 2021	0

Reasons for the exclusion of articles assessed full text

	Reference	Reason for full text exclusion
1	Ahn, S. Y., et al. (2022). "Safety and efficacy of nilotinib in adult patients with chronic myeloid leukemia: a post-marketing surveillance study in Korea." Blood Research 57(2): 144-151.	Not on a symptom burden instrument
2	Al Hadidi, S., et al. (2021). "Assessment and reporting of quality-of-life measures in pivotal clinical trials of hematological malignancies." Blood Adv 5(22): 4630-4633.	Review, not a study undertaken to develop an instrument
3	Atallah E, Schiffer CA, Radich JP, Weinfurt KP, Zhang MJ, Pinilla-Ibarz J, et al. Assessment of Outcomes After Stopping Tyrosine Kinase Inhibitors Among Patients With Chronic Myeloid Leukemia: A Nonrandomized Clinical Trial. JAMA Oncol. 2021;7(1):42-50.	PROMIS: not developed for, nor validated in CML patients; patients whom stopped TKIs, only 5 symptoms monitored
4	Bacik J, et al. The functional assessment of cancer therapy-BRM (FACT-BRM): a new tool for the assessment of quality of life in patients treated with biologic response modifiers. Qual Life Res. 2004;13(1):137-54.	FACT-BMR: not developed for, nor validated in CML patients
5	Balitsky, A. K., et al. (2022). "Important questions for the malignant hematologist to consider when designing or evaluating a study with patient-reported outcome measures (PROMs)." Eur J Haematol 109(1): 3-9.	Review, not a study undertaken to develop an instrument
6	Bertero, C., B.E. Eriksson, and A.C. Ek, A substantive theory of quality of life of adults with chronic leukaemia. Int J Nurs Stud, 1997. 34(1): p. 9-16.	Qualitative study, no development of symptom burden instrument
7	Boons C, Timmers L, Janssen J, Westerweel PE, Blijlevens NMA, Smit WM, et al. Response and Adherence to Nilotinib in Daily practice (RAND study): an in-depth observational study of chronic myeloid leukemia patients treated with nilotinib. Eur J Clin Pharmacol. 2020;76(9):1213-26.	Generic questionnaire, not available for assessment
8	De Marchi, F., et al., How could patient reported outcomes improve patient management in chronic myeloid leukemia? Expert Rev Hematol, 2017. 10(1): p. 9-14.	Review, not a study undertaken to develop an instrument
9	Eeltink, C., et al. (2022). "Sexual problems in patients with hematological diseases: a systematic literature review." Support Care Cancer 30(6): 4603-4616.	Review, not a study undertaken to develop an instrument
10	Efficace, F. and F. Cottone, Time for patient reported outcomes assessment in routine hematology practice: the case of chronic myeloid leukemia. Expert Rev Hematol, 2019. 12(1): p. 1-3.	Review, not a study undertaken to develop an instrument
11	Efficace, F., et al., Optimizing health-related quality of life in patients with chronic myeloid leukemia treated with tyrosine kinase inhibitors. Expert Review of Hematology., 2021.	Systematic review, not a study undertaken to develop an instrument
12	Efficace, F., et al. (2021). "Patient-Reported Outcomes as Independent Prognostic Factors for Survival in Oncology: Systematic Review and Meta-Analysis." Value in Health 24(2): 250-267.	Review, not a study undertaken to develop an instrument

Table Continued

	Reference	Reason for full text exclusion
13	Efficace, F., et al. (2022). "Physicians' Perceptions of Clinical Utility of a Digital Health Tool for Electronic Patient-Reported Outcome Monitoring in Real-Life Hematology Practice. Evidence From the GIMEMA-ALLIANCE Platform." Frontiers in Oncology 12 (no pagination).	Not on a symptom burden instrument
14	Etienne, G., et al. (2021). "Ponatinib long-term follow-up of efficacy and safety in CP-CML patients in real world settings in France: The POST-PACE study." Leuk Res 104: 106541.	Not on a symptom burden instrument
15	Flores, B., S. Klaar, and D.J. O'Connor, Changing views on adverse event reporting. Lancet Haematol, 2018. 5(11): p. e506-e507.	Not a study undertaken to develop an instrument
16	Fracchiolla, N.S., et al., FarmaREL: An Italian pharmacovigilance project to monitor and evaluate adverse drug reactions in haematologic patients. Hematol Oncol, 2018. 36(1): p. 299-306.	Not on a patient-reported outcome measure
17	Hahn EA, Glendenning GA, Sorensen MV, Hudgens SA, Druker BJ, Guilhot F, et al. Quality of life in patients with newly diagnosed chronic phase chronic myeloid leukemia on imatinib versus interferon alfa plus low-dose cytarabine: results from the IRIS Study. J Clin Oncol. 2003;21(11):2138-46.	FACT-BMR: not developed for, nor validated in CML patients
18	Henderson, J.R., 2nd, et al., Patient-Reported Health-Related Quality-of-Life Assessment at the Point-of-Care with Adolescents and Young Adults with Cancer. J Adolesc Young Adult Oncol, 2018. 7(1): p. 97-102.	10-item PROMIS Global: Focus on health-related quality of life, not symptom burden. Only 3 items refer to symptom burden (pain, fatigue, and emotional problems)
19	Martínez-López, J., et al. (2021). "The safety and efficacy of dasatinib plus nivolumab in patients with previously treated chronic myeloid leukemia: results from a phase 1b dose-escalation study." Leuk Lymphoma 62(8): 2040-2043.	Not on a symptom burden instrument
20	Moulin, S.M., et al., The role of clinical pharmacists in treatment adherence: fast impact in suppression of chronic myeloid leukemia development and symptoms. Support Care Cancer, 2017. 25(3): p. 951-955.	Not on a symptom burden instrument
21	Nguyen, C. T. T., et al. (2022). "Quality of life among chronic myeloid leukemia patients in the second-line treatment with nilotinib and influential factors." Qual Life Res 31(3): 733-743.	Interview-based instrument
22	Osborne, T.R., et al., Improving the assessment of quality of life in the clinical care of myeloma patients: the development and validation of the Myeloma Patient Outcome Scale (MyPOS). BMC Cancer, 2015. 15: p. 280.	MyPoS: developed for multiple myeloma patients
23	Papaila, A., et al. (2021). "Adverse outcomes in patients with chronic myeloid leukaemia treated with tyrosine kinase inhibitors." Heart 107(20): 1680.	Not on a symptom burden instrument

Table Continued

	Reference	Reason for full text exclusion
24	Saultz, J. and R. T. Maziarz (2021). "The evolving need for patient reported outcome data in determining choice of immune therapy treatment." Med (N Y) 2(8): 907-911.	Not on a symptom burden instrument
25	Schoenbeck, K. L., et al. (2022). "Patient-Reported Functional Outcomes in Patients With Chronic Myeloid Leukemia After Stopping Tyrosine Kinase Inhibitors." J Natl Cancer Inst 114(1): 160-164.	On functional outcomes after stopping TKIs
26	Schoenbeck, K. L. and K. E. Flynn (2021). "Health-Related Quality of Life of Patients with Chronic Myeloid Leukemia as Measured by Patient-Reported Outcomes: Current State and Future Directions." Curr Hematol Malig Rep 16(6): 491-499.	Review, not a study undertaken to develop an instrument
27	Senf, B., et al., Quality of life and distress assessed with self and external assessment screening tools in patients with hematologic malignancies attending treatment in an acute hospital. Qual Life Res, 2020. 29(12): p. 3375-3385.	No CML patients included, not on a symptom burden instrument
28	Singh, A. K., et al. (2022). "Imatinib and Patient-Related Outcomes in Chronic Myeloid Leukemia: A Single Centric Experience." SN Comprehensive Clinical Medicine 4(1) (no pagination).	Not on a symptom burden instrument
29	Stalfelt, A.M., Quality of life during induction treatment of acute myeloid leukaemia. A comparison of three intensive chemotherapy regimens using three instruments for quality of life assessment. Acta Oncol, 1994. 33(5): p. 477-85.	See Stalfelt 1993
30	Stalfelt, A.M. and B. Wadman, Assessing quality of life in leukemia: presentation of an instrument for assessing quality of life in patients with blood malignancies. Qual Assur Health Care, 1993. 5(3): p. 201-11.	To be used as a structured interview instrument. LIP (Life Ingredient Profile) instrument: developed based on quality of life literature and post hoc validation in haematological patients, but not CML patients.
31	Thestrup Hansen, S., et al., I Am Sure That They Use My PROM Data for Something Important. A Qualitative Study about Patients' Experiences from a Hematologic Outpatient Clinic. Cancer Nursing, 2020. 43(5): p. E273-E282.	Qualitative study, no development of symptom burden instrument
32	Tiribelli, M., G. Binotto, and M. Bonifacio, The significance of early warning in chronic myeloid leukemia. Expert Review of Hematology, 2018. 11(4): p. 265-266.	Not on a symptom burden instrument
33	Tuchler, H., et al., A short multilingual quality of life questionnaire--practicability, reliability and interlingual homogeneity. Qual Life Res, 1992. 1(2): p. 107-17.	QAHL: no items on symptom burden
34	Turkina, A., et al., TARGET: a survey of real-world management of chronic myeloid leukaemia across 33 countries. Br J Haematol, 2020. 190(6): p. 869-876.	Not on a symptom burden instrument

Table Continued

	Reference	Reason for full text exclusion
35	Yang, Y. P., et al. (2022). "Effects of physical exercise on the quality-of-life of patients with haematological malignancies and thrombocytopenia: A systematic review and meta-analysis." World Journal of Clinical Cases 10(10): 3143-3155.	Review, not a study undertaken to develop an instrument
36	Zhang, N., et al. (2022). "Illness uncertainty, self-perceived burden and quality of life in patients with chronic myeloid leukaemia: A cross-sectional study." J Clin Nurs 31(19-20): 2935-2942.	Not on a symptom burden instrument
37	Zhang, Q., et al. (2022). "Chronic disease management practices and factors associated with health-related quality-of-life for persons with chronic myeloid leukemia receiving tyrosine kinase inhibitor therapy." Ann Palliat Med 11(4): 1336-1350.	Using an already identified instrument
38	Zysberg, L., S. Hai, and N. Dally, A New Take on a Resource-Based Model of Quality of Life in Hemato-Oncological Patients: Demographic, Personal, and Social Factors. J Clin Psychol Med Settings, 2019. 26(4): p. 430-439.	Not on symptom burden

Step 1: evaluation of development quality of identified PROMs as described in studies, according to COSMIN risk of bias table

	Ref	PROM design				
		General design requirements				
		Clear construct	Clear origin of construct	Clear target population for which the PROM was developed	Clear context of use	PROM developed in sample representing the target population
EORTC QLQ-CML24	[1]	V	D	V	V	V
EORTC Symptom Set	[2]	V	V	V	D	V
FACT-LEU	[3]	V	D	V	D	A
Generic Chinese	[4]	V	D	V	V	I
HM-PRO	[5]		V	V		A
	[6]	V		V	V	
MDASI-CML	[7]	V	D	V	D	V

Abbreviations: A: adequate; CI: cognitive interview; D; doubtful; I; inadequate; PROM: patient-reported outcome measurement; V: very good. [1] When the PROM was not developed in a sample representing the target population, the concept elicitation was not further rated. [2] Empty cells indicate that a cognitive interview study (or part of it) was not performed.

Step 2: evaluation of content validity of the identified PROMs as described in studies, according to COSMIN risk of bias

		Asking patients			Asking experts	
PROM	Reference	Relevance	Compre-hensiveness	Compre-hensibility	Relevance	Compre-hensiveness
EORTC QLQ-CML24	[1]	I	D	D	I	I
EORTC Symptom Set	[8]	D	D	D	I	I
FACT-LEU	[3]	D	D	D	I	I
Generic Chinese	[4]	I	I	I	I	I
HM-PRO	[6]	D	D	D	D	D
MDASI-CML	[7]	D	D	D	D	D

Abbreviations: A: adequate; CI: cognitive interview; D; doubtful; I: inadequate; PROM: patient-reported outcome measure; V: very good.

			Cognitive interview (CI) study[2]				TOTAL PROM DEVELOPMENT
	Concept elicitation[1]	Total PROM design	General design requirements	Compre-hensibility	Compre-hensiveness	Total CI study	
			CI study performed in sample representing the target population				
	D	D	V	D	D	D	D
	D	D	V	D	D	D	D
	I	I	V	I	D	I	I
		I	I	I	I	I	I
	A	A					
			I	D	I	I	I
	D	D	V	D	D	D	D

Supplementary references

1. Efficace, F., et al., *International development of an EORTC questionnaire for assessing health-related quality of life in chronic myeloid leukemia patients: the EORTC QLQ-CML24.* Qual Life Res, 2014. **23**(3): p. 825-36.
2. Sodergren, S.C., et al., *Developing symptom lists for people with cancer treated with targeted therapies.* Targeted Oncology, 2021. **16**(1): p. 95-107.
3. Cella, D., et al., *Measuring health-related quality of life in leukemia: the Functional Assessment of Cancer Therapy--Leukemia (FACT-Leu) questionnaire.* Value Health, 2012. **15**(8): p. 1051-8.
4. Yu, L., et al., *Variables associated with patient-reported symptoms in persons with chronic phase chronic myeloid leukemia receiving tyrosine kinase inhibitor therapy.* Medicine (Baltimore), 2019. **98**(48): p. e18079.
5. Goswami, P., et al., *Quality-of-life issues and symptoms reported by patients living with haematological malignancy: a qualitative study.* Ther Adv Hematol, 2020. **11**: p. 2040620720955002.
6. Goswami, P., et al., *Development of a novel hematological malignancy specific patient-reported outcome measure (HM-PRO): content validity.* Frontiers in Pharmacology, 2020. **11 (article 209)**.
7. Williams, L.A., et al., *Measuring the symptom burden associated with the treatment of chronic myeloid leukemia.* Blood, 2013. **122**(5): p. 641-7.
8. Efficace, F., et al., *Health-related quality of life in chronic myeloid leukemia patients receiving long-term therapy with imatinib compared with the general population.* Blood, 2011. **118**(17): p. 4554-60.

Chapter 4

Development, validation, and pilot-testing of a newly developed patient-reported instrument for monitoring real-world TKI-related toxicity in CML care

Yolba Smit, Jacolien Timmerman, Alexander Currie, Hatem Soltan, Inger Abma, Jeroen Janssen, Eduardus Posthuma, Rianne Fijten, Ewald Bronkhorst, Andre Dekker, Rosella Hermens, Nicole Blijlevens

Submitted

Research in context

Evidence before this study

To compare the toxicity symptoms of different tyrosine kinase inhibitors (TKIs), as reported by chronic myeloid leukaemia (CML) patients, we conducted a systematic review of PubMed and Embase, from inception to February 2025. We included studies that reported on individual symptoms, as composite scores like 'symptom burden' are less helpful for shared decision-making in cases of intolerability, which is often driven by specific symptoms.

We identified 16 studies—including two randomised trials—that evaluated symptom severity of current TKI treatments using validated patient-reported outcome measures in at least 30 patients per TKI. The findings were scattered, mostly unadjusted, often inconsistent, and sometimes conflicting. Few studies addressed newer TKIs.

A prior systematic review and critical appraisal also found that existing instruments lack sufficient content validity to capture toxicity as experienced by CML patients.

Added value of this study

We developed the new 77-item EORTC IL386, the most comprehensive validated instrument to date for measuring CML patients' experiences with TKIs, covering 50 symptoms and scales. It is also the first CML-specific tool designed for clinical use, with a workflow co-developed with patients, who prioritised comprehensiveness over brevity, and overwhelmingly preferred a personalised format with filter questions over completing the full instrument.

We found that the mean severity of one-third of symptoms/scales differed significantly—by up to 30%—among the three most common treatments: dasatinib, imatinib, and nilotinib. Notably, for most individual symptoms, no other studies provided data suitable for comparing mean severity across TKIs.

The importance of the comprehensiveness of our instrument is underscored by the high prevalence and severity of sexual problems in collected pilot-data. Sexual problems have not been captured by previous studies using validated instruments, nor are they adequately reflected in the TKI product information, and they are under evaluated in both CML and haematology more broadly.

Implications of all the available evidence

We offer a modernised approach of adverse event reporting which can realise the potential of personalised medicine with sex and gender equity, the ability to monitor tolerability, and support post-marketing surveillance. Shared decision-making on which TKI to use can be supported by the identified treatment differences. Future studies should confirm symptom-specific differences for which no other studies have provided data yet, including the newer-generation TKIs, and examine the impact of dose adjustments or TKI switching on patient-reported toxicity.

Abstract

Background

In chronic myeloid leukaemia (CML), treatment choices between tyrosine kinase inhibitors (TKIs) are based on efficacy and tolerability. However, no instrument adequately measures patient-reported toxicity to monitor tolerability. We developed, validated and pilot-tested a new instrument to capture patients' experiences in real-world settings, and explored pilot data.

Methods

A mixed-methods study, following COREQ and STROBE guidelines. Instrument development started with CML-specific items from existing instruments, followed by semi-structured sessions with purposefully-sampled CML patients, and convenience-sampled CML professionals. Interviews focused on comprehensibility, comprehensiveness, relevance, acceptability, workflow, and graphics. Two researchers analysed transcribed sessions independently. Patients completed the evolving instrument at four time points for quantitative validation and pilot-testing. Psychometric evaluation assessed filter question validity, redundancy, reliability, measurement error, responsiveness, scale construction, and symptom severity differences in pilot data and meta-analyses of published studies.

Findings

A new EORTC instrument was developed, covering 50 symptoms and scales. Twenty-four generic filter questions are available for computer adaptive testing, an option preferred by 87% of patients. For 23 issues no existing EORTC item was available, these need to be added later. Psychometric properties were fair-excellent for reliability, and responsiveness was indicated for half of 50 symptoms. Pilot data showed significant severity differences between dasatinib, imatinib, and nilotinib in a third of symptoms and scales, with differences up to 30%.

Interpretation

Identified differences between common TKIs can influence shared decision-making. Sexuality problems are excluded from existing instruments, yet are prevalent and severe.

Introduction

The *Lancet Haematology Commission* has called for modernizing adverse event assessment in haematologic malignancies—particularly by incorporating patient-reported outcomes in real-world settings [1-5]. Such data are especially relevant in conditions requiring lifelong oral treatment, where multiple options exist and decisions are shaped by both efficacy and tolerability. This applies to chronic myeloid leukaemia (CML), where six commonly used tyrosine kinase inhibitors (TKIs) are available, alongside the possibility of treatment-free remission for some patients.

Due to intolerability, one in four to five patients switch to a different TKI [6, 7]. Unfortunately, evidence on tolerability, experienced by patients, from randomised [8, 9] and non-randomised [10-23] studies is scattered, inconsistent and sometimes conflicting [24]. As a result, treatment switches are often based on expert knowledge, which tends to underestimate both the prevalence and severity of what patients experience [25]. Therefore, patient-reported toxicity data are essential to incorporate the patient's perspective, provide real-world insights into treatment tolerability, and guide decision-making regarding dose modifications or treatment switches. Yet, no current instrument sufficiently captures TKI-related toxicity [26]. We therefore developed and validated a new instrument to allow CML patients to monitor TKI toxicity in daily clinical care, complete with a workflow and graphical overview, while evaluating between-TKI differences.

Methods

This mixed-methods study applied both qualitative and quantitative methodology, reported according to COREQ [27] and STROBE [28] (overviews in the Supplementary material). Ethical approval was waived, as the study did not fall under the Medical Research Involving Human Subjects Act.

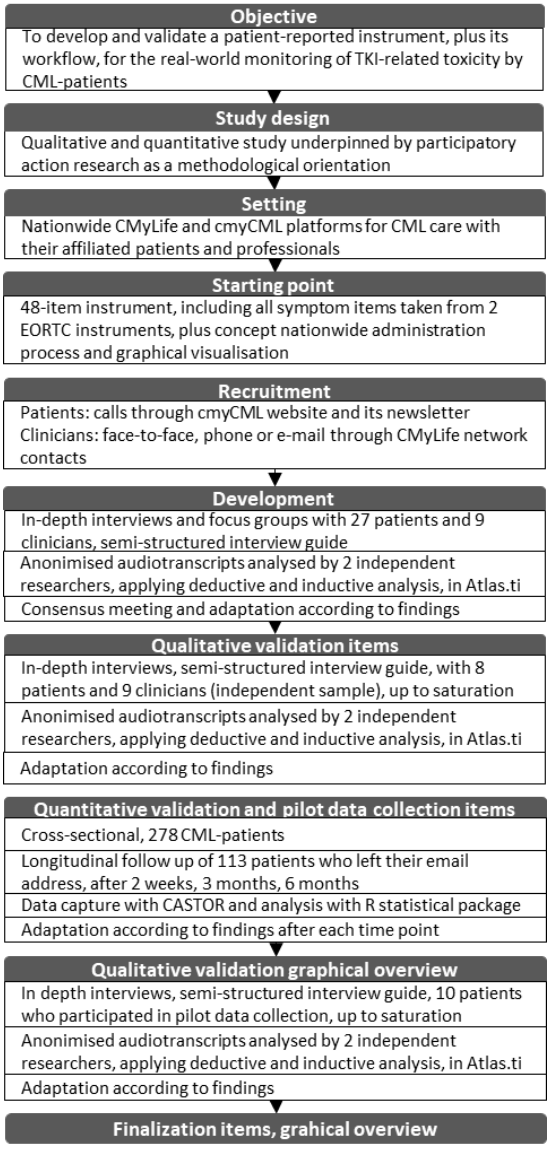


Figure 1 Study flow diagram

In brief, this study consisted of a development phase, followed by qualitative and quantitative validation phases, pilot data collection and analyses. Figure 1 summarises the study flow, for which we here describe the main points of the Methods. In the Supplementary material the Methods are described in detail. Participatory action research was the methodological orientation underpinning this study: it aims to improve daily practice through change, and wants to empower the research participants in improving daily practice [29].

Design requirements and setting

The instrument was primarily designed for real-world clinical settings, alongside research, to gain insights into the symptoms and complaints CML patients experience due to TKI use, enabling personalised advice, in line with an up-to-date guideline. The instrument was intended for electronic administration via the CMyLife platform, which helps Dutch CML patients manage their CML, interpret BCR::ABL1 results, and take action. Currently, it includes a patient-portal (www.cmycml.nl), a CML guideline app, and a clinical decision support system [30, 31].

Development and qualitative validation phases

The instrument started as a 48-item tool, combining symptom items from the EORTC QLQ-CML-24 and the EORTC Symptom Set [16, 32, 33] as these were the only instruments that captured key toxicities [26]. The 48-item tool was extended and changed in the development and validation phases, based upon in-depth interviews and focus groups – hereafter referred to as sessions - with CML patients and CML professionals. CML professionals included nurse specialists and haematologists working with CML patients. The EORTC has an Item Library with over 1,000 items that can be browsed and used to build custom-made item lists [34]. This Item Library and the concomitant contracting pipeline was used throughout this study, to add and exchange items when needed.

Quantitative validation phase

We advertised on www.cmycml.nl to recruit chronic-phase CML patients for an online survey on TKI-related toxicity symptoms and feedback (T1). Respondents were invited to participate in follow-up surveys at two weeks (T2), three months (T3), and six months (T4). Questionnaires were administered via Castor electronic data-capture software and adapted between the different timepoints, based on feedback received, and design and research needs. We evaluated feedback, the validity of filter items, redundancy, item-level reliability, measurement error, and responsiveness.

Pilot data

Pilot data was collected with the evolving instrument at T1-4. This data was used for scale construction, condensing multiple symptom-centred questions into a scale, in order to enhance the interpretability of the instrument for real-world research and to report pilot data more meaningfully. Furthermore, we evaluated symptom severity, differences between TKI treatments - based on both pilot data and systematically sourced existing data [24] in meta-analyses - and performed cluster analysis.

Development and validation graphical overview

In addition to instrument development and validation, we developed different graphical concepts to represent symptom severity over time, based on best practices [35-37]. We subsequently held sessions in a similar manner as for the instrument, with patients and professionals to elicit preferences and validate the adapted design.

Results

A total of at least 278 patients and 18 health care professionals participated in the development and validation phases, depicted in Figure 1 and described here below. Patient participants in the qualitative phases may overlap with those in the quantitative phase. Table 1 provides the patient characteristics per phase. Translated illustrative quotations and analyses details are provided in the Supplementary material.

Table 1 Characteristics of participants in each phase

	Development [$]	**Qualitative validation** [$]	**Quantitative validation** [$#]	**Qualitative validation graphical overview** [#]
Patients				
Number	27	8	278	10
Female	16 (59.3%)	6 (75.0%)	145 (52.1%)	
Median age in years (range)	60.5 y (33 to 77)	61 y (49 to 84)	63 (28 to 84)	59 (47 to 84)
Median time since diagnosis in years	5 y (0 to 16)	5 y (1 to 23)	6 (1 to 22)	7 (3 to 10)
University medical centre	9 (33.3%)	2 (25.0%)	Not available	Not available

Table 1 Continued

	Development [$]	**Qualitative validation** [$]	**Quantitative validation** [$#]	**Qualitative validation graphical overview** [#]
*Educational level**	Low 1 (3.7%)	Low 1 (12.5%)	Not available	Not available
	Middle 3 (11.1%)	Middle 2 (25.0%)		
	High 19 (70.4%)	High 5 (62.5%)		
	Unknown 4 (14.8%)	Unknown 0		
Treatment:	Not available	Not available		
asciminib			12 (4.5%)	1 (10.0%)
bosutinib			14 (5.2%)	1 (10.0%)
dasatinib			70 (26.2%)	2 (20.0%)
imatinib			84 (31.5%)	1 (10.0%)
nilotinib			39 (14.6%)	1 (10.0%)
ponatinib			8 (3.0%)	1 (10.0%)
Treatment-free remission			40 (15.0%)	3 (30.0%)
Nurse specialists			Not applicable	Not applicable
Number	5	2		
Female	5 (100%)	2 (100%)		
University medical centre	2 (40%)	1 (50%)		
Haematologists			Not applicable	Not applicable
Number	4	7		
Female	1 (25%)	5 (71%)		
University medical centre	2 (50%)	4 (57%)		

[$]Participants in the qualitative validation phase were an independent sample from the development phase. Patients in qualitative and quantitative phases may have overlapped. [#]Sampled from quantitative validation participants. *According to the definition of Statistics Netherlands.

Development phase

Twenty-seven patients, five nurse specialists (one in-training), and four haematologists participated. Patients were purposefully sampled from 59 applicants to our online recruitment, with no refusals. They took part in three online group sessions and three individual sessions (one by telephone) based on patient preference. Professionals – both nurse specialists and haematologists - participated in three individual sessions and three sessions

with two attendees, all online, to accommodate their schedules. Three nurse specialists declined due to time constraints, and one did not respond.

Participants identified issues across eight topics: Comprehensibility, Comprehensiveness, (Personal) Relevance, Frequency of administration, Order of items, Open text-box, Recall period, and Graphical concepts. In response to Comprehensibility, one item on 'pain in chest' was replaced by 'pain on chest' to better reflect common Dutch phrasing. In response to Comprehensiveness, thirteen items were added to the 48 items we started the development phase with. In addition, patients identified seven missing issues which had no corresponding items in the EORTC Item Library (Table 2).

Table 2 The final instruments[$] with category names, items pertaining to categories and scales, and missing issues

Category (scale)	EORTC number	Item
Fatigue (all items included in Fatigue scale)	IL281	Tired
		Need to rest
		Lacked energy
		Lethargic
		Weak
		Drowsy
Head and brain (all items included in Cognitive problems, except Headaches and Dizziness)	IL255	Headaches
		Dizziness
		Difficulty remembering
		Difficulty concentrating
		Trouble thinking clearly
Skin (all items included in Skin problems scale, except Bruised easily)	IL282	Skin colour change
		Itchy skin
		Skin rash
		Dry/flaking/cracked skin
		Sore/painful skin
		Bruised easily
Wound healing	IL283	Wounds healing problems
Hair and nails	IL284	Hair loss
		Nail problems
Eyes (all items included in Eye problems scale)	IL387	Watery eyes
		Dry eyes
		Burning eyes
		Discomfort bright light
		Blurred vision
		Red eyes

Table 2 Continued

Category (scale)	EORTC number	Item
Oedema	IL286	Swelling face/eyes
		Swelling body parts
Hearing	IL258	Hearing problems
Nose bleeds and nose problems	IL259	Nose bleeds
		Nose problems
Mouth and teeth	IL260	Dry mouth
		Pain/soreness mouth
		Teeth problems
		Bleeding gums
Food and taste	IL261	Taste change
		Appetite loss
Mood (all items included in Mood problems scale)	IL287	Depressed
		Irritable
		Tense
		Worried
		Afraid
		Difficulty sleeping
Heart	IL263	Palpitations
		Pain in chest
Breathing	IL264	Short of breath
		Coughed
Stomach and intestines	IL265	Feeling bloated
		Indigestion
		Nausea
		Vomited
		Diarrhoea
		Obstipation
		Abdominal pain/cramps
		Flatulence
Urination	IL266	Frequent urination
Muscles and joints (all items included in Muscle/joint cramps/pain scale, plus items Trouble walking pain and Pain)	IL288	Muscle cramps
		Aches/pains muscles/joints
		Muscle weakness
Fingers and toes	IL289	Tingling/numbness hands/feet
		Pale/cold fingers/toes
		Problems finger dexterity
Pain	IL290	Trouble walking pain
		Back pain
		Pain
Fever and sweating	IL291	Fever/chills
		Sweating excessively

Table 2 Continued

Category (scale)	EORTC number	Item
Warmth and cold	IL292	Problems tolerating heat/cold
		Hot flushes
Infection	IL293	Airway infections
		Other infections
Nipples and breasts	IL294	Sore/enlarged nipples/breasts
Sexuality (all items included in Sexuality problems scale, except gender specific items (Difficulty getting/ maintaining erection and Dry vagina))	IL295	Less interest in sex
		Less sexual enjoyment
		Pain during/after sex
		Difficulty getting/maintaining erection
		Dry vagina
Issues not available as EORTC item (in Dutch) at present		Itching/painful scalp
		Change in hair structure *
		Change in hair colour *
		Greasy hair
		Skin lets loose
		UV-sensitivity skin *
		Skin infections
		Petechiae
		Bleeding tendency
		Sensitivity to sound
		Tinnitus
		Change in thirst
		Receding gums
		Increased appetite
		Productive cough
		Slow heartbeat
		Rumbling, grunting intestines *
		Muscle stiffness
		Nerve pain
		Chronic bladder infections
		Change menstruation
		Absence of menstruation
		Fertility problems

Items in shaded cells are not included in any other currently available, validated questionnaire [1]
[$] Also available as one questionnaire (EORTC IL386) or as an add-on instrument to the QLQ-CML-24 (EORTC IL407) with 49 symptom items additional to the QLQ-CML24
* Not available in Dutch language

Regarding (Personal) Relevance, patients preferred not to classify items as irrelevant and valued comprehensiveness over brevity. Professionals considered issues outside of their expertise or unamendable to be less relevant, though no items were deemed irrelevant. All participants were interested in how to tailor the questionnaire to make it a more personalised experience. Examples mentioned were: a more ample use of filter questions, only administer items which previously scored either 'a little', 'quite a bit', or 'very much', and: enable to decline administration with a simple: 'I have no symptoms'.

Four perspectives emerged with regard to the Frequency of administration: link the administration to BCR::ABL1 monitoring, administer more frequently at the start of (a new) TKI, also administer during treatment-free remission, and enable on-demand administration on patients' own initiative. Concerning the Order of items, participants preferred to keep similar items together, start with the most prevalent symptoms, and move from physical items first to emotional items later on. In addition, an Open text box for entering any missing symptom or lifestyle factors that might relate to symptoms was requested by several patients One patient mentioned in the chat of an online group session that the Recall period of a week is rather short. No more information was available on the reason, or for which items the recall period was considered too short.

With regard to Graphical concepts, participants were asked to objectively interpret both bar and line concept graphs depicting the course of symptoms over time. Almost all interpretations were correct, both for bar as well as for line graphs. Patients preferred a line graph to a bar graph. For the symptom overview, participants preferred an overview with multiple line graphs, requested addition of TKI-type and dosage, preferred clustering of symptoms in logical groupings for better oversight, and wanted to be able to share their symptom overview with professionals.

Qualitative validation phase

Eight patients, two nurse specialists and seven haematologists, all of whom had not participated in the development phase, participated in the qualitative validation phase. Two lower educated [2] patients were recruited through nurse specialists, because very little lower educated patients applied via the online recruitment. Six patients were purposefully sampled from 59 applicants to our online recruitment call, with no refusals. All patients were interviewed individually: two by telephone at their request and the other six online. Professionals took part in seven individual sessions and one

session with two professionals, all online and so planned for convenience to their schedules. Participants raised new issues related to Comprehensibility and Comprehensiveness. Twelve items were added and one item, considered unclear because of the phrasing 'Have you had problems with coughing?', was replaced by 'Have you coughed?'. For eight issues that patients identified as missing, no corresponding items were available in the EORTC Item library (Table 2). No new issues emerged in the last two interviews after the sixth interview, indicating saturation.

Quantitative validation phase

Response, feedback

At T1, 278 patients responded, out of an unknown number of patients viewing the CMyCML website or its newsletter. Of these, 113 patients (40.6%) opted to complete subsequent surveys, with response rates of 78.8% (T2), 75.5% (T3) and 71.7% (T4). On average, patients took 8.0 minutes (SD 4.4) to complete 72 items in a single questionnaire (T1), compared to 9.4 minutes (SD 3.8 minutes) to complete 17 filter questions, leading to 17 questionnaires with a maximum of 77 items (T3). Most patients (86.8%) preferred completing multiple questionnaires with filter questions, 10.8% preferred a single questionnaire with all items, and 2.4% preferred answering only the items that had bothered them in the previous administration.

Overall, patients' attitude towards the questionnaire was positive; most patients found it easy to complete (63.2%) or were neutral (12.9%); 94.5% stated they had understood all items; 47.7% found that the questionnaire gave them insight into their complaints and symptoms; and only 1.4% found it stressful to fill in. 59.5% of patients found it better to divide the questionnaire into categories, to more easily find questions that would apply to them. At T1, a minority of patients (14.0%) said they would not use the questionnaire to get an overview of their symptoms or complaints, whereas 44.7% would, and 41.3% was neutral. Patients who would not use the questionnaire to get an overview, stated as reasons: not wanting to be occupied with their disease, not seeing added value, as symptoms were known to them, not having (a lot of) symptoms, and difficulty in interpreting their symptoms as TKI- or otherwise related. Patients preferred a median of 75 days (IQR: 148.8) between filling out surveys. At T2, 56% of patients indicated they preferred an open text-box at the end of the instrument, to either register other symptoms or complaints, give extra or more nuanced information, e.g., on co-morbidities, co-medication, social functioning, quality of life, or pose questions. Patients gave as new feedback that they had comorbidities or used

comedications that might affect their answers. At T3, all patients understood all items. No new feedback was received at T4.

Seven items were added and one item (about taste) was replaced for better comprehensibility, and no corresponding items were available in the EORTC Item Library for eight missing issues (Table 2, details in the Supplementary material).

Psychometric properties

We evaluated the **Validity of filter questions** for symptom categories (Methods in the Supplementary material), because of the wish for a more personalised experience, and a preference for symptom categories. Filter questions were considered the most feasible approach to achieve this. Two overarching EORTC-items (skin problems, eye problems) were invalid as filters for more detailed items, and would not suffice for the instrument as a whole. Therefore, custom generic filter questions were developed and tested for item categories during T3-T4, resulting in 24 filter questions. Of these, 22 were deemed valid at T4, while two (on 'Nose bleeds and nose problems', and 'Sexuality') were considered invalid, with adapted versions available for further field testing in the future (Supplementary material). Table 2 shows the final instrument, available as a single 77-item EORTC instrument or 24 separate EORTC instruments with 24 generic filter questions, plus 23 missing issues that need to be added later. **Redundancy:** no inter-item correlation exceeded 0.9 for any item combination (data not shown), so all items were retained. **Reliability** and **measurement error** were evaluated for the 70 items administered at T1 and T2. All items had at least fair-to-good reliability (weighted kappa ≥0.40), though the lower 95%CI was below 0.40 for 22 items. Fourteen items showed excellent reliability (weighted kappa ≥0.75). The percentage positive agreement between T1 and T2 ranged from 45.3% for Dry, flaky, or cracked skin to 93.3% for Sore/enlarged nipples/breasts (details in the Supplementary material). Indications for sufficient **responsiveness** were found for 21/50 symptoms/scales (see Scale construction under Pilot-data). The median change from T3 to T4 was 29.6% in self-reported improved patients, 7.7% in stable patients, and -18.5 in deteriorated patients (details in the Supplementary material).

Graphical overview

Ten patients participated; four invited patients did not respond, one declined. All were interviewed online, except for one who requested an in-person interview. No new issues were raised in the last two interviews after the eighth interview, indicating saturation. Issues brought forward (legend; category

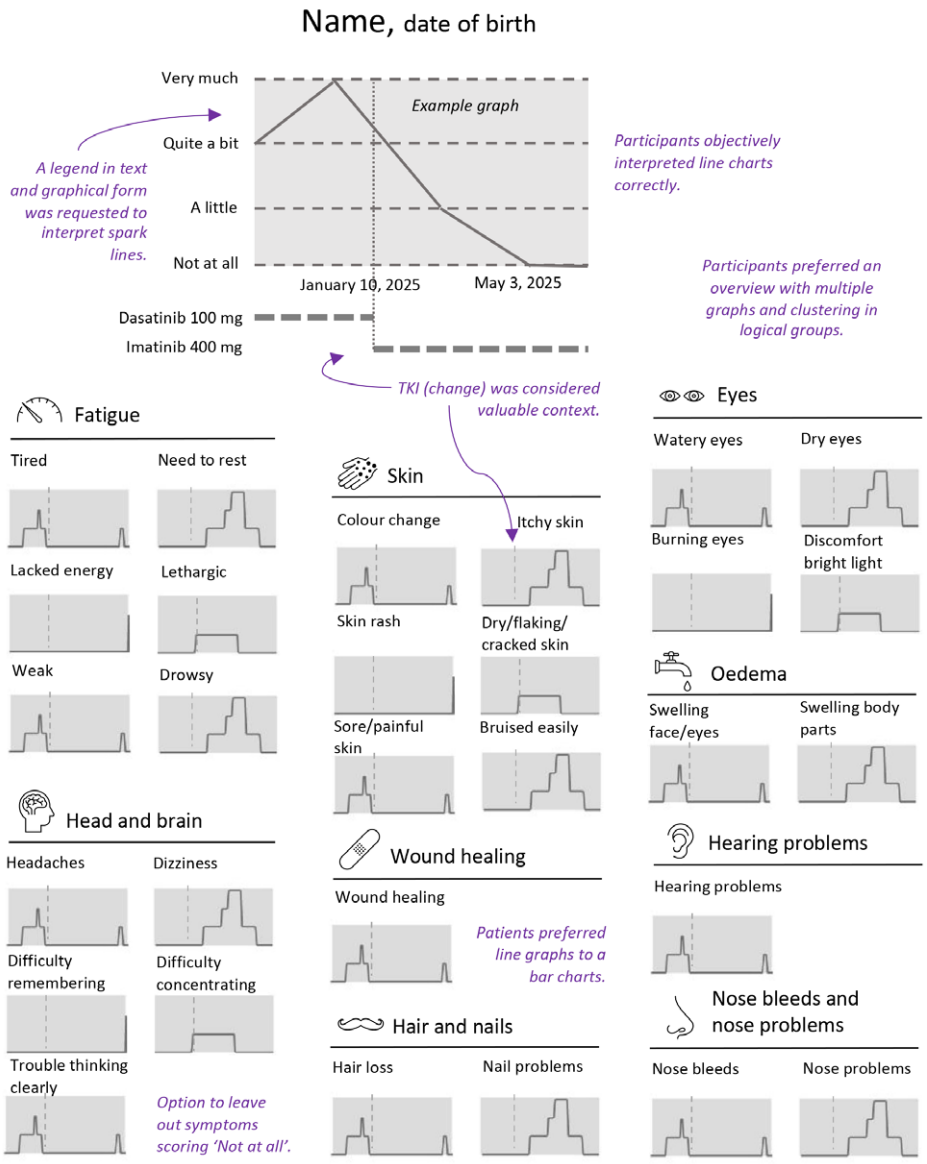


Figure 2 The final graphical overview, with issues participants brought forward in the development and validation phases in purple

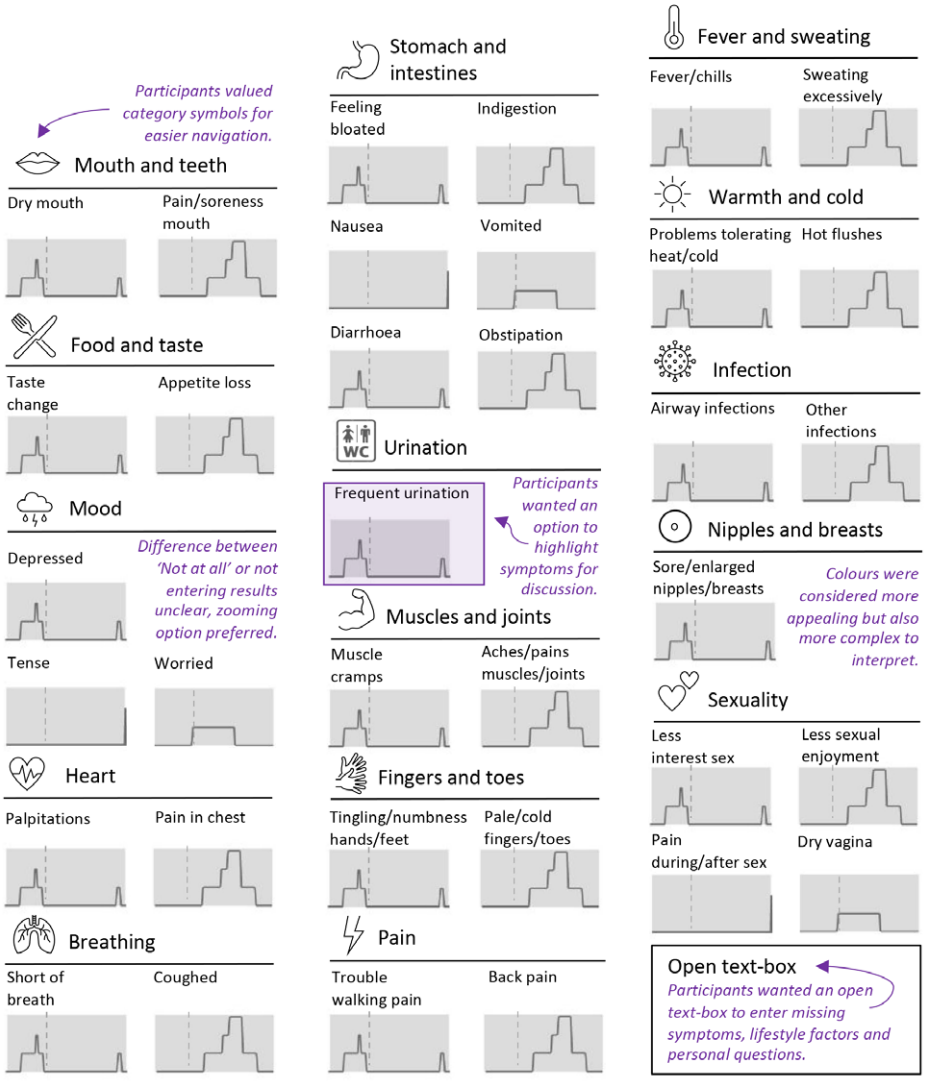


Figure 2 Continued

symbols; medication change; highlight symptoms; coloured design; zooming option, leave out symptoms that scored 'not at all') are depicted in Figure 2, the final design, and in detail in the Supplementary material.

Pilot-data

Scale construction and characteristics: 71 out of 77 items were retained for factor analysis, resulting in 7 scales with 33 items and meaningful interpretations (Fatigue, Mood problems, Muscle/joint pain/cramps, Skin problems, Eye problems, Cognitive problems, Sexuality problems, Table 3). These scales demonstrated good internal consistency (Cronbach's alpha ≥0.73). Therefore, the pilot data analyses included 7 scales and 43 single items, totalling 50 symptoms/scales. **(Comparative) analyses**: Fatigue was both the most prevalent and severe symptom across all TKIs (summarised in Table 4, detailed in the Supplementary material). Other common issues included Muscle/joint cramps/pain, Eye problems, Skin problems, Mood problems, Cognitive problems, and Sexuality problems, as well as Flatulence, Frequent urination, and Tingling/numbness of hands/feet. At T1, the adjusted mean severity of 13 of 44 evaluable symptoms/scales (29.5%) differed significantly across treatment options with ≥30 participants (Table 5). However, only Swelling face/eyes and Feeling bloated showed robust significant differences after correction for multiple comparisons. **Meta-analyses**: Eight studies [1-8] contributed data for meta-analyses of eight individual symptoms (detailed in the Supplementary material). For three symptoms - Nausea/vomiting, Dyspnoea and Appetite loss - indirect and unadjusted comparisons indicated statistically significant differences between TKI treatments, in addition to the differences observed with our pilot data alone. **Cluster analysis**: The clustering separated patients with less severe symptoms from those with more severe symptoms. Clusters differed significantly by gender, with a higher proportion of females in the higher severity cluster (p=0.03). No significant differences were found for other covariates, though there was a trend suggesting a more recent diagnosis was associated with the more severe symptom cluster (p=0.05).

Table 3 Constructed scales with included items and Cronbach's alpha

(Sub)scale	Included items	Number of included items	Cronbach's alpha
Fatigue	All fatigue category items	6	0.93 (0.91 – 0.94)
Mood problems	All mood category items	6	0.86 (0.83 – 0.88)
Muscle/joint pain/ cramps	All category items plus items on pain walking and pain in general	5	0.82 (0.79 – 0.85)
Skin problems	All skin category items except 'bruising'	5	0.77 (0.72 – 0.81)
Eye problems	All eye category items	5	0.73 (0.67 – 0.78)
Cognitive problems	All category items except headache and dizziness	3	0.88 (0.85 – 0.90)
Sexuality problems	All category items except gender-specific items	3	0.80 (0.75 – 0.83)

Table 4 Most prevalent and severe symptoms per current treatment

Current CML treatment	Prevalence		
	N of symptoms/scales which bothered ≥50% of patients	Top 5	Prevalence (95% confidence interval)
Any TKI	12	Fatigue (scale)	91.6 (87.2-94.9)
		Eye problems (scale)	86.6 (80.9-91.2)
		Muscle/joint cramps/pain (scale)	85.0 (78.9 (89.9)
		Skin problems (scale)	83.1 (76.9-88.1)
		Mood problems (scale)	80.5 (74.3 (85.8)
Asciminib	31	Muscle/joint cramps/pain (scale)	100 (63.1-100)
		Fatigue (scale)	91.7 (61.5-99.8)
		Skin problems (scale)	91.7 (61.5-99.8)
		Mood problems (scale)	90.0 (55.5-99.7)
		Cognitive problems (scale)	88.9 (51.8-99.7)
Bosutinib	24	Fatigue (scale)	100 (76.8-100)
		Flatulence	100 (69.2-100)
		Muscle/joint cramps/pain (scale)	81.8 (48.2-97.7)
		Cognitive problems (scale)	75.0 (34.9-96.8)
		Skin problems (scale)	75.0 (34.9-96.8)
		Eye problems (scale)	75.0 (34.9-96.8)
		Mood problems (scale)	75.0 (34.9-96.8)
Dasatinib	17	Fatigue (scale)	91.4 (82.3-96.8)
		Skin problems (scale)	85.7 (73.8-93.6)
		Mood problems (scale)	84.7 (73.0-92.8)
		Eye problems (scale)	81.8 (69.1-90.9)
		Muscle/joint cramps/pain (scale)	81.8 (69.1-90.9)
Imatinib	16	Muscle/joint cramps/pain (scale)	89.7 (79.9-95.8)
		Fatigue (scale)	89.3 (80.6-95.0)
		Eye problems (scale)	88.7 (79.0-95.0)
		Skin problems (scale)	78.9 (67.6-87.7)
		Mood problems (scale)	78.7 (67.7-87.3)
Nilotinib	11	Fatigue (scale)	94.9 (82.7-99.4)
		Eye problems (scale)	91.2 (76.3-98.1)
		Skin problems (scale)	82.9 (66.4-93.4)
		Mood problems (scale)	78.4 (61.8-90.2)
		Muscle/joint cramps/pain (scale)	75.0 (56.6-88.5)
Ponatinib	15	Fatigue (scale)	87.5 (47.3-99.7)
		Cognitive problems (scale)	85.7 (42.1-99.6)
		Eye problems (scale)	85.7 (42.1-99.6)
		Dry mouth	85.7 (42.1-99.6)
		Indigestion	83.3 (35.9-99.6)
		Frequent urination	83.3 (35.9-99.6)
		Tingling/numbness hands/feet	83.3 (35.9-99.6)
Treatment-free remission	13	Muscle/joint cramps/pain (scale)	93.8 (79.2-99.2)
		Fatigue (scale)	90.0 (76.3-97.2)
		Mood problems (scale)	83.3 (67.2-93.6)
		Skin problems (scale)	82.9 (66.4-93.4)
		Nose problems	75.0 (34.9-96.8)

Severity ranges from 0 ('not at all') to 100 ('a lot'), with a score of 33.3 indicating 'a little', and 66.6 'quite a bit'. No 95% confidence interval calculated if only one score available, or if all scores were the same.

	Severity	
N of symptoms/scales with mean symptom severity >33.3	**Top 5**	**Mean severity (95% confidence interval)**
2	Fatigue (scale)	41.5 (38.2-44.8)
	Flatulence	36.3 (31.5-41.0)
	Frequent urination	31.6 (26.9-36.4)
	Muscle/joint cramps/pain (scale)	28.6 (25.2-31.9)
	Dry vagina	27.9 (21.0-34.8)
10	Feeling bloated	45.8 (3.9-87.8)
	Short of breath	45.8 (16.3-75.4)
	Coughed	45.8 (16.3-75.4)
	Abdominal pain/cramps	41.6 (9.2-74.1)
	Fatigue (scale)	40.3 (23.8-56.7)
9	Flatulence	70.0 (49.1-90.9)
	Nose problems	46.6 (0-100)
	Abdominal pain/cramps	43.3 (9.5-77.1)
	Fatigue (scale)	42.8 (29.0-56.6)
	Feeling bloated	42.4 (15.7-69.1)
5	Fatigue (scale)	46.3 (39.8-52.9)
	Pale/cold fingers toes	40.0 (20.7-59.3)
	Frequent urination	35.7 (26.1-45.4)
	Tingling/numbness hands/feet	35.1 (25.3-45.0)
	Flatulence	34.5 (25.4-43.7)
3	Fatigue (scale)	39.0 (33.7-44.3)
	Swelling face/eyes	38.9 (31.3-46.4)
	Flatulence	36.7 (29.5-44.0)
	Muscle/joint cramps/pain (scale)	31.3 (25.7-36.9)
	Diarrhoea	30.9 (22.3-39.5)
2	Difficult erection	41.6 (19.3-64.0)
	Fatigue (scale)	38.5 (31.2-45.7)
	Dry vagina	33.3 (16.2-50.5)
	Flatulence	32.3 (21.1-43.5)
	Back pain	30.9 (10.3-51.6)
2	Fatigue (scale)	40.5 (20.2-60.9)
	Frequent urination	38.9 (12.5-65.2)
	Nose problems	33.3
	Dry mouth	33.3 (15.5-51.1)
	Feeling bloated	33.3 (2.0-64.6)
3	Fatigue (scale)	40.5 (32.3-48.8)
	Back pain	37.5 (6.1-68.8)
	Dry vagina	35.5 (15.3-55.8)
	Muscle/joint cramps/pain (scale)	30.9 (24.0-37.8)
	Nose problems	29.1 (11.3-47.0)

Table 5 Symptoms with a significant difference between current TKI treatments, used by ≥30 patients, from our pilot data (top 13 rows) and meta-analyses (bottom three rows) of eight systematically identified studies [1-8] and our pilot data

	Raw mean estimates		
Symptom	**Dasatinib**	**Imatinib**	**Nilotinib**
Headaches	23.7 (16.1-31.3)	12.6 (7.7-17.5)	7.4 (1.9-12.9)
Cognitive problems (scale)	30.5 (23.6-37.5)	21.1 (16.3-25.9)	16.5 (10.4-22.6)
Skin problems (scale)	28.5 (23.4-33.5)	19.5 (15.1-24.0)	24.2 (6.8-31.7)
Bruised easily	12.9 (6.9 -18.8)	21.7 (15.7-27.8)	9.5 (4.3-14.8)
Wounds healing problems	11.1 (5.0-17.2)	10.2 (5.7-14.7)	1.0 (0.0-2.9)
Hair loss	26.8 (19.1-34.5)	11.6 (6.6-16.5)	24.5 (13.7-35.3)
Swelling face/eyes	20.0 (11.4-28.6)	38.9 (31.3-46.4)	11.8 (4.3-19.3)
Mood problems (scale)	24.2(18.3-30.2)	24.3 (18.6-30.1)	16.4 (11.1-21.7)
Feeling bloated	30.9 (23.0-38.9)	22.4 (15.6-29.1)	14.1 (6.3-22.0)
Obstipation	22.6 (14.1-31.1)	8.7 (4.4-13.0)	18.7 (7.0-30.5)
Diarrhoea	16.7 (10.1-23.2)	30.9 (22.3-39.5)	7.3 (0-14.6)
Abdominal pain/cramps	24.8 (16.5-33.2)	21.7 (15.9 -27.5)	16.7 (7.0-26.3)
Hot flushes	17.0 (9.6-24.3)	7.3 (3.0-11.7)	13.5 (3.0-24.0)
Nausea/vomiting (meta-analysis)	7.3 (4.1-10.5)	12.7 (11.5-14.0)	4.8 (1.4-8.1)
Dyspnoea (meta-analysis)	21.2 (17.2-25.1)	23.8 (19.8-27.8)	12.4 (9.5-15.2)
Appetite loss (meta-analysis)	11.3 (7.7-14.9)	17.3 (12.0-22.7)	7.3 (4.0-10.5)

$: p=0.04; £: p<0.01; †: p=0.03; ¥: p=0.01; * Robust in Benjamini-Hochberg procedure for multiple comparisons; §: p=0.02; Ω: p<0.05. Symptom scores range from 0 ('not at all') to 100 ('a lot'), with a score of 33.3 indicating 'a little', and 66.6 'quite a bit'. Shaded areas indicate a covariant-adjusted (age, gender, years since diagnosis, years on current treatment, TKI-dosage) or meta-analysed significant difference between treatment options, with red indicating a worse symptom experience; green a better symptom experience; orange an intermediate symptom experience; and grey an adjusted (pilot data) or unadjusted (meta-analysis) estimate of significant differences.

	Estimated differences (standard error)		
	Dasatinib – imatinib	**Imatinib – nilotinib**	**Dasatinib - nilotinib**
	10.9 (5.3) $		19.6 (6.1) £
			13.3 (6.1) †
	10.5 (3.9) £		
		13.3 (6.0) †	
		10.9 (4.2) ¥	8.5 (3.7)†
	15.0 (5.5) £		
	-19.8 (6.7) £	* 29.7 (7.4) £	
			12.6 (5.1) §
	15.6 (6.2) ¥		* 24.5 (6.8) £
	15.4 (5.3) £		
	-15.6 (6.9) £	22.2 (8.3) †	12.7 (5.9) †
			15.8 (7.1) †
	8.2 (4.1) Ω		
	-5.4 (1.8) £	8.0 (1.8) £	
		11.4 (2.5) £	8.8 (2.5) £
		10.1 (3.2) £	

Discussion

In this study, we developed, validated and pilot-tested a new 77-item instrument to measure CML patient-reported toxicity in a real-world context, along with its workflow and graphical feedback. The instrument is available through the EORTC Item Library https://qol.eortc.org/item-library/ as a single EORTC questionnaire (IL386), as an add-on to the EORTC QLQ-CML24 (IL407), or as 24 separate EORTC questionnaires with 24 generic filter questions, the latter preferred by most patients. Compared to existing validated instruments [1], this new instrument contains 17 additional items. Notably, items on sexuality - absent in current validated instruments - ranked amongst the most prevalent and severe symptoms in our pilot data. This finding is not reflected in product information brochures [2-7], highlighting the need for further research into sexual health in CML and haematology more broadly [8, 9].

On 23 relevant issues no EORTC items were available in the EORTC Item Library - the most comprehensive library available - indicating a need for inclusion in the future. Of note, there are no available items on menstrual disorders, though patients indicated they wanted such items included, and the product information brochures of dasatinib, imatinib and nilotinib indicate menstrual disorders as a side effect [4, 5, 7]. Adaptations to the instrument will also be necessary as new TKIs or knowledge on existing TKIs emerge, given that few patients used asciminib, bosutinib, or ponatinib. Patients preferred an open text-box to provide additional information, offering an opportunity to capture emerging toxicities. Psychometric properties were fair to excellent for reliability, and responsiveness was indicated for nearly half of the symptoms/scales. Responsiveness for other symptoms/scales could not be evaluated, likely due to insufficient data in some cases, and possibly due to a long recall period or faulty anchor questions in others. Overall, validation studies for responsiveness are challenging to design and execute effectively, when no gold standard exists. Importantly, comprehensibility as a prerequisite for responsiveness has been validated in our study.

Pilot data indicated a high prevalence and severity of symptoms. Patients with more severe symptoms may have been more likely to participate, introducing selection bias. However, for eight symptoms/scales reported in the literature, the mean severity measured with EORTC items [10-16] was comparable for six symptoms. Only for fatigue during dasatinib and nilotinib, and insomnia during imatinib did we report more severe symptoms than some other studies (detailed in the Supplementary material). Notably, we were unable to meta-

analyse most items/scales with existing data, as either individual symptom data were not provided or not evaluated in published studies. The higher proportion of females in the high-severity cluster aligns with previous findings of more severe symptoms in female CML patients [15] and broader sex differences in reported adverse drug reactions [17]. Possibly, TKI dosages are relatively too high in women, whom have, on average, a lower body mass than men. However, other sex- and gender-related factors, such as physiology, lifestyle, behaviour, help-seeking, communication style, or adherence, may play a role. Ultimately, sex- and gender-specific TKI prescribing may be justified, as treatment response disparities are increasingly recognised in other fields as well [18]. Our new instrument may support more sex- and gender-equitable medical care.

In our pilot data and meta-analyses of systematically sourced studies, a third of evaluated symptoms/scales showed significant differences in severity between the three most common treatments, with differences of up to 30%. These findings can inform shared decision-making when switching TKIs. The significant differences aligned with expected pathophysiological mechanisms. For example, dasatinib, which showed the most severe symptoms in most cases, has many more off-tumour on-target effects than imatinib and nilotinib [19]. More severe bruising, wound healing issues and swelling of the face/eyes with imatinib may be attributed to its effects on heightened endothelial permeability, and impeded thrombus formation and collagen synthesis [20].

This raises the broader issue of whether a patient-reported symptom is causally related to TKI treatment. Using a validated instrument does not imply causality, just as symptoms or adverse events reported by professionals do not [21, 22]. Patients included in our pilot data also indicated that they found it difficult to say what caused their symptoms. Literature indicates that CML patients on TKIs report worse scores for depression, dyspnoea, fatigue, insomnia, cognitive issues, pain, physical health-related quality of life, and composite symptom burden (e.g., nausea, diarrhoea, itching, skin changes, and swelling) compared to the general population or study controls [16, 23, 24] However, for many symptoms in our study, and by TKI, this information is not available. A large nationwide cohort study with complete coverage found that, compared to matched controls, TKI-treated CML patients had an increased risk for 142 out of the 670 analysed disease categories, including ophthalmic, infectious, ear and brain diseases, with no category more common in controls [25]. Thus, previous evidence on symptoms experienced by CML patients remains limited.

A major strength of this study is the involvement of patients and professionals, not only in the development and validation of the instrument, but also in providing results and feedback on the workflow. While there is rigorous methodological guidance on selecting and assessing patient-reported outcome measures [26], to our knowledge, no such guidance exists on providing results and developing workflows.

A limitation of our study is the small number of patients in some subgroups of our pilot-data, which hindered comparisons with newer TKIs, limited information on different dosages, and prevented exploration of changes in treatment (e.g., switching between TKIs, stopping, or altering dosages). We may have lacked the power to detect significant differences between the three most common TKIs for some symptom-TKI combinations. Larger prospective, longitudinal studies, or studies such as the RODEO study [27], are needed to address how dose reductions affect symptom severity, in which patients, and by what extent. Such evidence, addressing these knowledge gaps, is crucial for personalising TKI strategies [28].

Conclusions

We developed a new validated 77-item EORTC instrument and workflow to monitor CML patient-reported toxicity symptoms, which our pilot-data showed to be more extensive than those captured by existing validated tools. Notably, one-third of the symptoms varied across treatment modalities, offering choice and valuable insights for shared decision-making. Developing and validating patient-reported instruments for real-world clinical use is an ongoing, iterative processes - especially with the introduction of new treatments. Future studies should aim to confirm these findings and explore newer-generation TKIs, dosage adjustments, and treatment switches to fully harness patient-reported toxicity data in clinical care.

References

1. Thanarajasingam, G., et al., *Beyond maximum grade: modernising the assessment and reporting of adverse events in haematological malignancies.* Lancet Haematol, 2018. **5**(11): p. e563-e598.
2. Thanarajasingam, G., et al., *Reaching beyond maximum grade: progress and future directions for modernising the assessment and reporting of adverse events in haematological malignancies.* Lancet Haematol, 2022. **9**(5): p. e374-e384.
3. Major, A., A.C. Dueck, and G. Thanarajasingam, *Beyond maximum grade: advancing the measurement and analysis of adverse events in malignant haematology trials in the modern era.* Lancet Haematol, 2025. **12**(6): p. e451-e462.
4. Bhatnagar, V., et al., *Beyond maximum grade: using patient-generated data to inform tolerability of treatments for haematological malignancies.* Lancet Haematol, 2025. **12**(6): p. e463-e469.
5. Brockelmann, P.J., et al., *Beyond maximum grade: tolerability of immunotherapies, cellular therapies, and targeted agents in haematological malignancies.* Lancet Haematol, 2025. **12**(6): p. e470-e481.
6. Geelen, I.G.P., et al., *Treatment outcome in a population-based, 'real-world' cohort of patients with chronic myeloid leukemia.* Haematologica, 2017. **102**(11): p. 1842-1849.
7. Busque, L., et al., *Switching TKIs during CML therapy is frequent, mostly driven by intolerance, and does not affect survival: a prospective Quebec registry study.* Blood Cancer J, 2025. **15**(1): p. 35.
8. Cortes, J.E., et al., *Patient-reported outcomes in the phase 3 BFORE trial of bosutinib versus imatinib for newly diagnosed chronic phase chronic myeloid leukemia.* J Cancer Res Clin Oncol, 2019. **145**(6): p. 1589-1599.
9. Rea, D., et al., *Health-related quality of life of patients with resistant/intolerant chronic phase chronic myeloid leukemia treated with asciminib or bosutinib in the phase 3 ASCEMBL trial.* Leukemia, 2023. **37**(5): p. 1060-1067.
10. Shacham Abulafia, A., et al., *Health-related quality of life in patients with chronic myeloid leukemia treated with first- versus second-generation tyrosine kinase inhibitors.* J Clin Med, 2020. **9**(11).
11. Ali, H., Z. Mohamed, and D. Abdullah, *Health-related quality of life in chronic myeloid leukaemia patients receiving long-term therapy with different tyrosine kinase inhibitors in Kurdistan region.* Georgian Med News, 2024(354): p. 173-180.
12. Bostan, H., et al., *Quality of life and symptom burden with first- and second-generation tyrosine kinase inhibitors in patients with chronic-phase chronic myeloid leukemia.* Clin Lymphoma Myeloma Leuk, 2020. **20**(12): p. 836-842.
13. Cho, H.J., et al., *Real-world experience in the management of chronic myeloid leukemia patients focused on tyrosine kinase inhibitors intolerance and health-related quality of life.* Acta Haematol, 2025: p. 1-14.
14. Efficace, F., et al., *Health-related quality of life in patients with chronic myeloid leukemia receiving first-line therapy with nilotinib.* Cancer, 2018. **124**(10): p. 2228-2237.
15. Efficace, F., et al., *Health-related quality of life of newly diagnosed chronic myeloid leukemia patients treated with first-line dasatinib versus imatinib therapy.* Leukemia, 2020. **34**(2): p. 488-498.

16. Efficace, F., et al., *Validation and reference values of the EORTC QLQ-CML24 questionnaire to assess health-related quality of life in patients with chronic myeloid leukemia.* Leuk Lymphoma, 2021. **62**(3): p. 669-678.
17. Foulon, S., et al., *Health state utility and quality of life measures in patients with chronic myeloid leukemia in France.* Qual Life Res, 2021. **30**(7): p. 2021-2032.
18. Nguyen, C.T.T., et al., *Quality of life among chronic myeloid leukemia patients in the second-line treatment with nilotinib and influential factors.* Qual Life Res, 2022. **31**(3): p. 733-743.
19. Rashid, A. and Z. Albasry, *Assessing the quality of life in imatinib-treated chronic myeloid leukemia patients: a cross sectional descriptive study.* Int J Pharmaceut Res, 2020. **12**(2): p. 2784-91.
20. Sacha, T., et al., *Quality of life and adherence to therapy in patients with chronic myeloid leukemia treated with nilotinib as a second-line therapy: a multicenter prospective observational study.* Clin Lymphoma Myeloma Leuk, 2017. **17**(5): p. 283-295.
21. Uyanik, M.S., et al., *The impact of new generic formulations of imatinib mesylate on general quality of life in chronic phase chronic myeloid leukaemia patients.* Eur J Cancer Care (Engl), 2017. **26**(2).
22. Williams, L.A., et al., *Measuring the symptom burden associated with the treatment of chronic myeloid leukemia.* Blood, 2013. **122**(5): p. 641-7.
23. Zulbaran-Rojas, A., et al., *A prospective analysis of symptom burden for patients with chronic myeloid leukemia in chronic phase treated with frontline second- and third-generation tyrosine kinase inhibitors.* Cancer Med, 2018. **7**(11): p. 5457-5469.
24. Smit, Y., et al., *Patient-reported toxicity symptoms during tyrosine kinase inhibitor treatment in chronic myeloid leukemia: a systematic review and meta-analysis.* Support Care Cancer, 2025. **33**(5): p. 446.
25. Efficace, F., et al., *Patient- versus physician-reporting of symptoms and health status in chronic myeloid leukemia.* Haematologica, 2014. **99**(4): p. 788-93.
26. Smit, Y., et al., *Measuring chronic myeloid leukaemia TKI-related toxic effects in the real world: a systematic review and critical assessment of content validity of patient-reported outcome measures.* Lancet Haematol, 2023.
27. Tong, A., P. Sainsbury, and J. Craig, *Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups.* Int J Qual Health Care, 2007. **19**(6): p. 349-57.
28. von Elm, E., et al., *The Strengthening The Reporting of OBservational studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies.* Lancet, 2007. **370**(9596): p. 1453-7.
29. Baum, F., C. MacDougall, and D. Smith, *Participatory action research.* J Epidemiol Community Health, 2006. **60**(10): p. 854-7.
30. Ector, G.I., et al., *The development of a web-Based, patient-centered intervention for patients with chronic myeloid leukemia (CMyLife): design thinking development approach.* J Med Internet Res, 2020. **22**(5): p. e15895.
31. Verweij, L., et al., *Effectiveness of digital care platform CMyLife for patients with chronic myeloid leukemia: results of a patient-preference trial.* BMC Health Serv Res, 2023. **23**(1): p. 228.
32. Efficace, F., et al., *International development of an EORTC questionnaire for assessing health-related quality of life in chronic myeloid leukemia patients: the EORTC QLQ-CML24.* Qual Life Res, 2014. **23**(3): p. 825-36.

33. Sodergren, S.C., et al., *Developing symptom lists for people with cancer treated with targeted therapies.* Targeted Oncology, 2021. **16**(1): p. 95-107.
34. *EORTC Item Library*. [cited 2022 October 3]; Available from: https://www.eortc.be/itemlibrary/#home/tab/official.
35. Bantug, E.T., et al., *Graphical displays of patient-reported outcomes (PRO) for use in clinical practice: What makes a pro picture worth a thousand words?* Patient Educ Couns, 2016. **99**(4): p. 483-490.
36. Albers, E.A.C., et al., *Visualization formats of patient-reported outcome measures in clinical practice: a systematic review about preferences and interpretation accuracy.* J Patient Rep Outcomes, 2022. **6**(1): p. 18.
37. Snyder, C., et al., *Making a picture worth a thousand numbers: recommendations for graphically displaying patient-reported outcomes data.* Qual Life Res, 2019. **28**(2): p. 345-356.
38. *Educational level [In Dutch]*. May 19 2025]; Available from: https://www.cbs.nl/nl-nl/nieuws/2019/33/verschil-levensverwachting-hoog-en-laagopgeleid-groeit/opleidingsniveau.
39. *Iclusig : EPAR - Product Information*. 2022 [cited 2025 February 6]; Available from: https://www.ema.europa.eu/en/documents/product-information/iclusig-epar-product-information_en.pdf.
40. *Scemblix : EPAR - Product Information*. 2024 [cited 2025 February 6]; Available from: https://www.ema.europa.eu/en/documents/product-information/scemblix-epar-product-information_en.pdf.
41. *Sprycel : EPAR - Product Information*. 2024 [cited 2025 February 6]; Available from: https://www.ema.europa.eu/en/documents/product-information/sprycel-epar-product-information_en.pdf.
42. *Glivec : EPAR - Product Information*. 2024 [cited 2025 February 6]; Available from: https://www.ema.europa.eu/en/documents/product-information/glivec-epar-product-information_en.pdf.
43. *Bosulif : EPAR - Product Information*. 2025; Available from: https://www.ema.europa.eu/en/documents/product-information/bosulif-epar-product-information_en.pdf.
44. *Tasigna : EPAR - Product Information*. 2024 [cited 2025 February 6]; Available from: https://www.ema.europa.eu/en/documents/product-information/tasigna-epar-product-information_en.pdf.
45. Schoenbeck, K.L. and K.E. Flynn, *Health-related quality of life of patients with chronic myeloid leukemia as measured by patient-reported outcomes: current state and future directions.* Curr Hematol Malig Rep, 2021. **16**(6): p. 491-499.
46. Niscola, P., F. Efficace, and E. Abruzzese, *Sexual health in patients with hematological malignancies: a neglected issue.* Support Care Cancer, 2018. **26**(6): p. 1699-1701.
47. de Vries, S.T., et al., *Sex differences in adverse drug reactions reported to the National Pharmacovigilance Centre in the Netherlands: An explorative observational study.* Br J Clin Pharmacol, 2019. **85**(7): p. 1507-1515.
48. Mauvais-Jarvis, F., et al., Sex and gender: modifiers of health, disease, *and medicine.* Lancet, 2020. **396**(10250): p. 565-582.
49. Hantschel, O., U. Rix, and G. Superti-Furga, *Target spectrum of the BCR-ABL inhibitors imatinib, nilotinib and dasatinib.* Leuk Lymphoma, 2008. **49**(4): p. 615-9.
50. Haguet, H., et al., *BCR-ABL Tyrosine Kinase Inhibitors: Which Mechanism(s) May Explain the Risk of Thrombosis?* TH Open, 2018. **2**(1): p. e68-e88.

51. Hillman, S.L., et al., *Evaluation of the value of attribution in the interpretation of adverse event data: a North Central Cancer Treatment Group and American College of Surgeons Oncology Group investigation.* J Clin Oncol, 2010. **28**(18): p. 3002-7.
52. Phillips, K.M., et al., *Quality of life outcomes in patients with chronic myeloid leukemia treated with tyrosine kinase inhibitors: a controlled comparison.* Support Care Cancer, 2013. **21**(4): p. 1097-103.
53. Hyland, K.A., et al., *Cognition in patients treated with targeted therapy for chronic myeloid leukemia: a controlled comparison.* Leuk Lymphoma, 2023. **64**(2): p. 415-423.
54. Dahlen, T., et al., *Adverse outcomes in chronic myeloid leukemia patients treated with tyrosine kinase inhibitors: Follow-up of patients diagnosed 2002-2017 in a complete coverage and nationwide agnostic register study.* Am J Hematol, 2022. **97**(4): p. 421-430.
55. Mokkink, L.B., et al., *COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures.* Qual Life Res, 2018. **27**(5): p. 1171-1179.
56. Djodikromo, M.F., et al., *Patient-guided dose reduction of tyrosine kinase inhibitors in chronic myeloid leukaemia (RODEO study): study protocol for a prospective, multicentre, single-arm trial.* BMC Cancer, 2023. **23**(1): p. 231.
57. Smit, Y., et al., *Benefits and limitations of real-world patient-reported toxicity-symptom monitoring for guidelines and care, as perceived by patients, clinicians and guideline developers [submitted]*. 2025.

Supplementary material

Table 1 COREQ (COnsolidated criteria for REporting Qualitative research) Checklist [1]

Topic	Item no.	Guide questions/description	This study
Domain 1: Research team and reflexivity			
Personal characteristics			
Interviewer/ facilitator	1	Which author/s conducted the interview or focus group?	Yolba Smit, Alexander Currie
Credentials	2	What were the researcher's credentials? E.g. PhD, MD	YS: MD MSc; AC: BM
Occupation	3	What was their occupation at the time of the study?	YS: Policy advisor, PhD student; AC: Medical student
Gender	4	Was the researcher male or female?	YS: Female; AC: Male
Experience and training	5	What experience or training did the researcher have?	YS: Trained as a medical doctor and MSc in epidemiology. AC: Trained as a medical doctor, half-way Master phase. Both: Received recent training in conducting focus groups and individual interviews for qualitative research.
Relationship with participants			
Relationship established	6	Was a relationship established prior to study commencement?	YS: knew two patients, one nurse specialist and 5 haematologists/guideline developers form previous collaborations. One of the guideline developer/ haematologists is the promotor in her PhD trajectory. AC: no prior relationship
Participant knowledge of the interviewer	7	What did the participants know about the researcher? e.g. personal goals, reasons for doing the research	All participants were made aware of the PhD position of YS and of the research questions.
Interviewer characteristics	8	What characteristics were reported about the inter viewer/facilitator? e.g. Bias, assumptions, reasons and interests in the research topic	YS's main interest is applied guideline development. The focus of her PhD is to evaluate how patients' experiences can best be incorporated in to the guideline development process in a systematic way. She assumes that guidelines in general can benefit to some extent from capturing patients experiences, such as side effects, systematically.

Table 1 Continued

Topic	Item no.	Guide questions/description	This study
Domain 2: Study design			**Study protocols: Zenya WebShare - Radboudumc**
Theoretical framework			
Methodological orientation and Theory	9	What methodological orientation was stated to underpin the study? e.g. grounded theory, discourse analysis, ethnography, phenomenology, content analysis	Participatory action research was the methodological orientation that underpinned this study. Participatory action research aims to understand reality, as well as also improve it through change, and in doing so wants to empower the research participants [2].
Participant selection			
Sampling	10	How were participants selected? e.g. purposive, convenience, consecutive, snowball	See Methods "Study population and recruitment"
Method of approach	11	How were participants approached? e.g. face-to-face, telephone, mail, email	See Methods "Study population and recruitment"
Sample size	12	How many participants were in the study?	See Results, Table 1
Non-participation	13	How many people refused to participate or dropped out? Reasons?	See Results
Setting			
Setting of data collection	14	Where was the data collected? e.g. home, clinic, workplace	See Results
Presence of non-participants	15	Was anyone else present besides the participants and researchers?	A facilitator from the CMyLife team was present during all focus groups, and at two focus groups and some interviews with professionals an apprentice was present
Description of sample	16	What are the important characteristics of the sample? e.g. demographic data, date	See Results Table 1
Data collection			
Interview guide	17	Were questions, prompts, guides provided by the authors? Was it pilot tested?	See Methods "Sessions", and Interview guides in Supplement
Repeat interviews	18	Were repeat inter views carried out? If yes, how many?	No

Table 1 Continued

Topic	Item no.	Guide questions/description	This study
Audio/visual recording	19	Did the research use audio or visual recording to collect the data?	See Methods "Sessions"
Field notes	20	Were field notes made during and/or after the interview or focus group?	Field notes were made during and/or after sessions
Duration	21	What was the duration of the inter views or focus group?	See Results
Data saturation	22	Was data saturation discussed?	See Methods "Qualitative data analysis"
Transcripts returned	23	Were transcripts returned to participants for comment and/or correction?	No
Domain 3: analysis and findings			
Data analysis			
Number of data coders	24	How many data coders coded the data?	Two
Description of the coding tree	25	Did authors provide a description of the coding tree?	See Supplementary material 'Detailed results development phase', with tables, and similar for the qualitative validation phase
Derivation of themes	26	Were themes identified in advance or derived from the data?	See Methods "Data analysis"
Software	27	What software, if applicable, was used to manage the data?	See Methods "Data analysis"
Participant checking	28	Did participants provide feedback on the findings?	No
Reporting			
Quotations presented	29	Were participant quotations presented to illustrate the themes/findings? Was each quotation identified? e.g. participant number	See Supplementary material 'Detailed results development phase', with tables, and similar for the qualitative validation phase
Data and findings consistent	30	Was there consistency between the data presented and the findings?	Results section
Clarity of major themes	31	Were major themes clearly presented in the findings?	Table 2, Results section
Clarity of minor themes	32	Is there a description of diverse cases or discussion of minor themes?	Results and Discussion sections, plus detailed description ins Supplement

Table 2 STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) checklist [3] for quantitative validation and pilot data analyses

Topic	Item No	Recommendation	This study
Title and abstract			
	1	(a) Indicate the study's design with a commonly used term in the title or the abstract	Abstract
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	Abstract
Introduction			
Background/ rationale	2	Explain the scientific background and rationale for the investigation being reported	Introduction
Objectives	3	State specific objectives, including any prespecified hypotheses	Introduction
Methods			
Study design	4	Present key elements of study design early in the paper	See Methods. Study protocols: Zenya WebShare - Radboudumc
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	Described in: Quantitative validation, Recruitment and surveys
Participants	6	(a) Cohort study—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	Described in: Quantitative validation, Recruitment and surveys
		Cross-sectional study—Give the eligibility criteria, and the sources and methods of selection of participants	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	Described in: Quantitative validation paragraph
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	Described in: Quantitative validation paragraph
Bias	9	Describe any efforts to address potential sources of bias	None
Study size	10	Explain how the study size was arrived at	The objective was to include as many participants as possible

Table 2 Continued

Topic	Item No	Recommendation	This study
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	Described in: Analyses quantitative validation, and Pilot data paragraphs
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding	Described in: Analyses quantitative validation, and Pilot data paragraphs
		(b) Describe any methods used to examine subgroups and interactions	
		(c) Explain how missing data were addressed	
		(d) Cohort study–If applicable, explain how loss to follow-up was addressed Cross-sectional study–If applicable, describe analytical methods taking account of sampling strategy	
		(e) Describe any sensitivity analyses	
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study–e.g. numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	Results section, Paragraph Quantitative validation, response, feedback
		(b) Give reasons for non-participation at each stage	Unknown
		(c) Consider use of a flow diagram	Figure 1
Descriptive data	14*	(a) Give characteristics of study participants (e.g. demographic, clinical, social) and information on exposures and potential confounders	Table 1
		(b) Indicate number of participants with missing data for each variable of interest	Not reported
		(c) *Cohort study*–Summarise follow-up time (e.g., average and total amount)	6 months from t1 to T4
Outcome data	15*	*Cohort study*–Report numbers of outcome events or summary measures over time	Results section and Supplement
		Cross-sectional study–Report numbers of outcome events or summary measures	Results section and Supplement
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included	Results section, table 4, 5 and Supplement
		(*b*) Report category boundaries when continuous variables were categorised	Methods section

Table 2 Continued

Topic	Item No	Recommendation	This study
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	Not relevant
Other analyses	17	Report other analyses done–e.g. analyses of subgroups and interactions, and sensitivity analyses	No other analyses done
Discussion			
Key results	18	Summarise key results with reference to study objectives	Discussion section
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision	
		Discuss both direction and magnitude of any potential bias	
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	
Generalizability	21	Discuss the generalisability (external validity) of the study results	
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	Declaration of interest paragraph

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies

Methods

Development phase

Study population and recruitment

We advertised for adult chronic-phase CML patients willing to participate in online video or telephone interviews, either individually or in groups, hereafter referred to as sessions, through www.cmycml.nl. We stated we wanted to discuss TKI-related toxicity monitoring. Fifty-nine applicants were purposefully sampled to reflect differences in sex, age, time since diagnosis, education, and hospital type where treatment took place. To address the low number of low-educated applicants, we also recruited through the haematologists involved in this research (NB, JJ, EP).

We recruited haematology nurse specialists (in training) and haematologists working with CML patients though convenience sampling (until saturation) via the CMyLife network. We prioritised CML nurse specialists, assuming they discuss toxicity symptoms more in more detail than haematologists, whose primary focus is disease control. Once the CML nurse specialist pool was exhausted we began inviting haematologists. We invited nine nurse specialists and four haematologists in person, by phone, email, or through their professional society.

Sessions

All sessions were conducted online due to initial COVID restrictions and later for convenience, though telephone or live interviews were offered as alternatives. Initially, we preferred patient focus groups, assuming that interaction would provide valuable insights. However, after three focus groups, we switched to individual interviews, as nurse specialists and haematologists suggested that sexual symptoms were relevant and better discussed in one-on-one sessions.

Two researchers (YS, AC) conducted all sessions using semi-structured interview guides (provided in the Supplementary material), focusing on comprehensibility, comprehensiveness, and relevance as main focal points for content validity [4], and acceptability, workflow and graphic concepts. The core team developed the guides, and the patient version was pilot-tested on a patient representative. For the item-list and workflow, participants received the test instrument 2-5 days prior to the session—patients were asked to complete it, and professionals to review it. Participants were informed that the instrument was intended for routine use within the CMyLife-platform. Informed consent was obtained prior to sessions, which were audio-recorded and transcribed verbatim, anonymizing data. Transcripts were not returned for comments or corrections.

Qualitative data analysis

Two researchers (either YS, JT, AC and/or HS) independently coded text fragments using inductive coding in Atlas.ti, reaching consensus through discussion. Pre-set topics were comprehensibility, comprehensiveness, relevance, workflow and graphic concepts. Codes were categorised into pre-set topics or new topics as needed. Saturation could not be assessed in the development phase due to the focus group format.

Consensus meeting

Major design decisions were made in a consensus meeting with involvement of patient representatives, individual patients, haematologists, researchers and the CMyLife team. The consensus meeting was held right after the development phase, with the goal to consolidate proposals– proposals that originated from participants perspectives in the qualitative analysis - and lay down the principles for future decision making. Participants of the consensus meeting were a patient representative, four patients, three haematologists, and five team members (researchers and members of the CMyLife team). Patients and haematologists had also participated in the development phase.

Qualitative validation phase

Study population and recruitment, sessions and qualitative data analysis were similar to the methodology of the development phase, with the exception of saturation. In the validation phase, saturation was defined as no new topics emerging in the last two sessions after the sixth session, with assessment and discussion after each session. We chose two sessions after the sixth session to start saturation assessment, as it is unlikely saturation is reached within fewer than eight sessions.

Quantitative validation phase

Recruitment and surveys

At **T1**, we administered the EORTC IL234 (provided in the Supplementary material) along with feedback questions on completion time, preferred timing, willingness to use, instrument attitude, comprehensiveness, comprehensibility, and other feedback. At **T2**, patients completed the IL234 and provided open-text feedback. At **T3**, we administered 17 questionnaires (EORTC IL254-270) along with 17 filter questions to assess whether a questionnaire applied. In addition, feedback was requested on completion time, instrument and filter-questions attitude, comprehensiveness and comprehensibility, and other feedback. At **T4**, 24 questionnaires (EORTC IL255, 258-261, 263-266, 281-295) were administered, along with 24 filter questions to assess whether a questionnaire applied, an anchoring question for each item (see responsiveness below), and feedback questions.

Analyses

All analyses were conducted using the R-statistical package 4.4.1, except for the meta-analyses, which were performed in STATA 18.5. Missing data were excluded. Response proportions and feedback were analysed descriptively.

Validity filter-items: at T1, we evaluated the use of two overarching EORTC items, for skin and eye problems, as filter-questions. They were considered valid if ≤5% of patients who responded 'not at all' or 'a little bit' to the overarching item responded 'quite a bit' or 'very much' on any corresponding detailed item – indicating that patients interpreted the overarching item as inclusive of its sub-items. At T3, we used custom filter questions. These were considered preliminarily valid if <10% of patients who reported no symptoms in a given category at T3 had reported 'quite a bit' or 'very much' for any relevant item at T1. At T4, we revised the custom filter questions and administered all detailed items to every patients, regardless of their response to the filter question. Validity was assessed by comparing responses to each filter question with its corresponding items, using the same threshold as at T3. **Redundancy** was defined as a high inter-item Spearman's rank correlation (>0.9) [5], and such items were considered for removal. **Item-level reliability** was assessed using weighted kappa, comparing T1-T2 scores for patients on the same TKI and dosage. We assumed that if symptoms were TKI-related, reliable items would yield consistent responses. A weighted kappa ≥0.40 was considered fair to good and ≥0.75 excellent [5]. **Measurement error** was defined as the percentage positive agreement between T1-T2 scores in patients on the same TKI and dosage. In the absence of a gold standard, **responsiveness** was assessed using self-reported change via item-specific anchor questions ('Is that different from three months ago?'). If the Spearman's rank correlation between self-reported change and mean symptom severity change from T3 to T4 was acceptable (≥0.30 [6]), we compared changes across patients who reported 'improvement', 'stability' or 'deterioration' using the Kruskal-Wallis rank-sum test.

Pilot-data

Scale construction and characteristics: we constructed symptom scales post hoc using T1 data and exploratory parallel factor analysis with oblimin rotation. Analysis was performed if the Kaiser-Meyer-Olkin test of sampling adequacy was ≥0.5 and Bartlett's test of sphericity p<0.05. Items were excluded if they had only low (<0.32) or high (>0.90) inter-item correlations, low prevalence (<10%), or were gender-specific. Scales (factors) were retained and iteratively constructed based on a combination of: eigen value>1, scree plot patterns, interpretability, and factor loadings >0.30. Internal consistency was assessed using Cronbach's alpha, with values >0.70 [5] considered acceptable. **(Comparative) analyses**: missing items in scales were ignored if ≥50% of items were available, as all inter-total correlations exceeded 0.47. Item and scales scores were linearly transformed [7]. We described symptom prevalence and

severity, and assessed differences in mean scores across current treatments with >30 patients using linear models adjusted for age, gender, year of diagnosis, time on current TKI, and TKI dosage. Data from T1 were chosen, as T1 provided the most respondents, except for six items introduced at T4 (Pale/cold fingers/toes; Nail problems; Other nose problems; Back pain; Airway infections; and Other infections). TKI dosage was categorised as below norm or at/above norm. Norm-dosages (milligram/day) were: asciminib 80, bosutinib 400, dasatinib 100, imatinib 400, nilotinib 600, and ponatinib 45. Analyses were exploratory and tested for robustness to multiple comparisons using the Benjamini-Hochberg procedure [8]. We conducted random-effects **meta-analyses** of mean individual symptom severities, comparing TKI subgroups using both our pilot data and systematically identified studies that used EORTC instruments in ≥30 patients treated with dasatinib, imatinib or nilotinib. **Cluster analysis**: To create a simplified indicator for toxicity symptoms we applied K-means clustering. To this end symptom severity scores for all points in time were used. To be able to apply K-means clustering missing values had to be dealt with. Symptoms only applicable for one gender were imputed with a score of 0 for the other gender. For six symptoms the fraction missing was very high (>75%), these were excluded from the cluster analyses. This concerned Nose problems, Back pain, Pale cold skin, Nails, Airway infections and Other infections. For all other symptoms, missing values were imputated using the mice package (v3.17.0), using all measurements as predictors for the missing values. Applying the silhouette procedure, two clusters were found to be optimal (Packages cluster (v2.1.8) and factorextra (v1.0.7)).

Qualitative validation of the graphical overview

To validate the graphical overview, we purposefully sampled patients who participated in the longitudinal quantitative validation, ensuring diversity in age, gender, current treatment, and missing answers. These patients were invited by email. Patients received their own longitudinal data before the interview.

Topics semi-structured interview guides

Each session, and each topic within a session, ended with an open question for other feedback, and whether participants wanted to raise additional issues. Prompts used were: Could you elaborate?; Could you give me an example?; How do you feel about ... ? What needs to be different/left out/added?

Development

Acceptability, comprehensibility, comprehensiveness, relevance

- Feedback on introduction of instrument
- Time needed for completion, opinion on number of items, type of questions
- Comprehensibility (patients only)
- Were there questions you preferred not to answer? (patients only)
- Order of questions
- Missing questions
- Relevance of questions
- Willingness to complete questionnaire (patients only)
- Open text field

Administration, accessibility and privacy

- Opinion on administration in app, feedback on mock-up screens app
- Frequency of administration
- Patient-types to administer to
- Preferences for invites and reminders (patients only)
- Other languages
- Out-of-app administration, low health literacy
- Privacy concerns

Graphical concepts

- Preferences for information in overview, graph types
- Feedback on multiple prototypes (design, scoring (thinking aloud method), legend, sharing)
- Interpretation: check of different prototype mock-up graphs (change over time, in relation to TKI-dosage)

Validation

Acceptability, comprehensibility, comprehensiveness, relevance

- Feedback on introduction of instrument
- Time needed for completion, opinion on number of items, type of questions
- Comprehensibility (patients only)
- Order of questions, missing questions
- Willingness to complete questionnaire (patients only)
- Open text field

Graphical concept symptoms overview (own data, patients only)

- Feedback on concept, design, also per item (legend, graphs, icons, layout)
- Comprehensibility, comprehensiveness, perceived usefulness

Table 3 The 72 items included in the EORTC IL234, with potential filter items in dark blue, items pertaining to potential EORTC filter questions in light blue, and the 25 items added during the development and qualitative validation phases in light grey

Tiredness	Burning eyes	Vomiting
Needing to rest	Discomfort eyes bright light	Diarrhoea
Lacking energy	Blurred vision	Constipation
Lethargy	Red eyes	Abdominal pains/cramps
Weakness	Hearing problems	Problems gas/flatulence
Drowsiness	Nose bleeds	Frequent urination
Trouble sleeping	Dry mouth	Muscle cramps
Headaches	Pain/soreness mouth	Aches/pains muscles/joints
Dizziness	Teeth problems	Muscle weakness
Difficulty remembering	Bleeding gums	Trouble walking pain
Difficulty concentrating	Different taste food/drink	Tingling/numbness hands/feet
Trouble thinking clear	Lack appetite	Difficulty small objects
Easy bruising	Depressed	Swelling body parts
Skin problems	Irritable	Pain
Skin colour change	Tense	Fevers/chills
Itchy skin	Worrying	Excessive sweating
Skin rash	Afraid	Problems tolerating heat/cold
Dry/flaking/cracked skin	Palpitations	Hot flushes
Sore/painful skin	Pain in chest	Sore/enlarged nipples/breasts
Wound healing problems	Shortness of breath	Less interest sex
Hair loss	Coughing	Less sexual enjoyment
Swelling face/eyes	Bloated abdomen	Pain during/after sexual activity
Eye problems	Acid indigestion/heartburn	Difficulty getting/maintaining erection
Watery eyes	Nausea	Dry vagina

Detailed results development phase

Themes that arose from the perspectives of participants with regard to the development of the instrument were: (1) Comprehensibility; (2) Comprehensiveness; (3) (Personal) relevance; (4) Frequency of administration; (5) Order of items; (6) Open text-box; (7) Recall period; and (8) Graphical concepts. There were no issues with the instructions. Five items had issues with **(1) Comprehensibility** (Table 4): two items were retained; one item was added; one alternative item was chosen. The first item was retained because, though patients did not understand a Flemish Dutch word, the comprehension of the item was not affected; for one item there was no alternative in the EORTC Item Library but an additional item was added to be tested in the validation phase; for one item an alternative was chosen; two items were retained as screening items though participants indicated that the meaning was vague to them. Regarding **(2) Comprehensiveness**, 18 topics were raised (Table 5). For 8 topics suitable items were available in the EORTC Item Library, for which 12 items were added. Sexuality was considered a missing topic by professionals, who suggested to add four screening items. None of the patient participants directly acknowledged sexuality as a missing topic. For eight topics no suitable items were available in the EORTC Item Library. One missing topic was weight change, however, professionals indicated that they would prefer to have the actual weight of patients. Lastly, the missing topic of tooth pain was considered covered by the overarching pain item. With regard to **(3) (Personal) relevance**, most patients were enthusiastic and considered the list appropriate, though some patients and professionals mentioned it was a (too) long list (further reported on in the Results, and in detail in Table 6). With regard to the **(4) Frequency of administration**, perspectives differed again (further reported on in the Results, and in detail in Table 7). Some patients would like to fill in symptoms on a weekly base, while others found a once-a-year administration sufficient. Concerning the **(5) Order of items**, this is reported on in the Results, and in detail in Table 8. Another topic that patients came up with was that an **(6) Open text-box**, reported on in the Results, and in detail in Table. Patients mentioned in the chat that the **(7) Recall period** of a week is rather short. With regard to the **(8) Graphical concepts**, this is reported on in the Results.

Consensus meeting

The design decisions made/principles laid down were:

Comprehensibility:

- Choose items from the EORTC Item Library with a more clear phrasing, if comprehensibility or user experience is affected

Comprehensiveness:

- Add items identified as missing, for which a suitable item is available in the EORTC Item Library
- Do not develop new items because of time and budget constraints, instead check the EORTC Item Library yearly
- Add the actual weight of patients in the electronic administration in some way

(Personal) relevance:

- Comprehensiveness is considered more important than brevity by patients and therefore no items will be left out at this stage
- Future evaluation on the prevalence of items and patient participation may alter this decision
- Quantitative evaluation might indicate which items are not discriminative enough and can be left out
- Enable to decline invitation to administration, with reason for decline
- Investigate options for personalisation, e.g., (more) filter questions as several EORTC items are included already (validate these in quantitatively) or categories with filter questions, or only administer items scored on previously (computer adaptive testing)

Frequency of administration:

- Link administration to BCR::ABL1 monitoring
- Administer more frequently at the start of a (new) TKI
- Also administer during treatment-free remission
- Enable on-demand administration on patients' own initiative, and enable reminders set by patients

Order of items:

- Put most prevalent symptoms first
- Keep similar items together

Open text-box:

- Enable open text-box at the end

Graphical representation:

- Line graph is preferred option for individual graphs
- All items in separate line graphs in the overview, per category, is the preferred option for the overview
- Enable graph and open text-box sharing with clinicians

Table 4 Comprehensibility issues raised in the development phase, with illustrative quotes, rationale and decisive action

	Issue	Illustrative quote	Action	Rationale
1	Flemish term for acid indigestion possibly not understood by patients	[indigestion] yes, a lot of people don't know what that is. They know about heartburn, but I think a lot of people think: I have no idea [Pr]	Item retained	Word 'indigestion' not Netherlands Dutch. Comprehensibility item by patients not affected; change if suitable alternative is available
2	Swelling unclear in Dutch, fluid retention preferred	'Have you had swelling in a certain part of the body', I think it is very difficult for people to say, yes, what do we mean by that? [Pr] Uhm, and I thought fluid retention might be a better term [Pr]	Item retained and more specific item on swelling around eyes/in face added	Swelling is very general in Dutch and not necessarily associated with oedema. Fluid retention would be a better phrase. No suitable alternative item available, but specific item on swelling of eyes and face available
3	"Pain on chest' is common in Dutch, not 'in chest'	[...] but I would call that pain on the chest [...] not in the chest [Pr]	Alternative item chosen	Better suitable alternative available
4	'Worrying' had a vague meaning to professionals	"Are you worried?" I don't really know what that means, because about what or how or what? Yes, of course it is a questionnaire and not qualitative, yes, I understand that, but, uhm, yes, I thought that was vague [Pr]	Left in as a screening item	No more specific item available
5	'Hearing problems' too general	Problems with hearing, I thought that was a bit general. I think you have to give people something to think about: "I hear that, or I have that, or I find that annoying" [P]	Left in as a screening item	No more specific item available

Abbreviations: P: patient; Pr: professional

Table 5 Comprehensiveness issues raised in the development phase, with illustrative quotes, rationale and decisive action

	Issue	Illustrative quote	Action	Rationale
1	Dental pain	But perhaps also dental problems that could be involved [...] toothache [P]	None	Considered covered by item pain in general
2	Bleeding tendency	My blood is thinner so I bleed worse [Pa]	Check yearly	No suitable item available
3	Change thirst	[...] whether people would drink more yes or no [...] People also tend to be more thirsty or less thirsty [P]	Check yearly	No suitable item available
4	Facial mask	[...] because with imatinib, for example, I also have that uhm that I sometimes had a very red mask on my face. Uhm, how should I say that, uhm, during exertion or a stressful event, my whole face became really bright and bright red, everything. And that continued for a while until it happened again, until I was alone and calmed down. And then it was over again [P]	Check yearly	No suitable item available
5	Fertility problems	See sexual dysfunction	Check yearly	No suitable item available
6	Itching/painful scalp	I regularly hear from patients about itchy scalp. Now I saw that itching was mentioned, as in total, itchy skin. But I don't think people think about the scalp so quickly [Pr] But yes, my hair is not just hair loss. I think, well, that it could also be that people have pain complaints in their scalp. [...] I also know that when I moved my hair back and forth, it was really annoying [P]	Check yearly	No suitable item available
7	Skin lets loose	Skin lets loose [P, comment in chat of videoconferencing software]	Check yearly	No suitable item available
8	Sun sensitivity	I miss uh the question about sensitivity to sunlight and UV uh with TKIs. For example, I suffer from this very much. [...] I can hardly ... I hardly ever sit in the sun but ... if I do and it's intense then uh, uh I immediately get a rash [P]	Check yearly	No suitable item available in Dutch
9	Tinnitus	Well, I had tinnitus, among other things. [...] I am currently suffering from tinnitus [P]	Check yearly	No suitable item available

Table 5 Continued

	Issue	Illustrative quote	Action	Rationale
10	Weight change	It is weight loss but also gain. That's why I also ... it is not just fluid, but it is certainly also that women often gain a lot of weight [Pr]	None	Professionals prefer actual weight
11	Bloated stomach	I also suffer a lot from a swollen stomach	Added	Suitable item available
		[...] a bloated feeling, it is really full [P]		
12	Cough	Yes, for me I miss coughing as a complaint [P]	Added	Suitable item available
13	Unintentional flatulence	And uh what I also suffer from is uh flatulence. Uh yes, I'm just going to be very honest and that is really, uh, a side effect that also makes your social contacts very difficult because I just can't stop it, that's real, and it comes at a time when you just don't expect it , just like that, out of the blue [P]	Added	Suitable item available
14	Mental capacity	what I also miss is a reduced ability to think - it is slightly different from concentration: at the end of the day it is not possible to do something difficult [P]	Added	Suitable item available
15	Sensitivity cold/heat	I would like to know [...] whether anyone also suffers from cold, cold legs and cold arms? [...] not being able to withstand cold and not being able to withstand heat [P]	2 items added	Suitable item available
		Different temperature (hot-cold) [P, in chat]		
16	Stamina/strength	Could your fitness level be added [...] Hey, how do you experience your condition at the moment, 'can you climb the stairs without a break?' or uh uh I also wrote something like 'can you just carry out your daily activities or do you have to skip there'? [Pr]	Added	Item on muscle weakness added, to be checked in validation phase
17	Tingling hands/feet	Tingling and pricking in toes, hands, feet and fingers and lower legs, really, uh, so bad that it caused me a lot of uh pain [P]	Added	Suitable item available
18	Sexual dysfunction	What I really miss, especially about complaints that they want to discuss, uh, which is of course always a difficult complaint, is sexual dysfunction. Uh, I definitely see erection and libido disorders with the TKIs. Just like fertility problems are an issue and that is not reflected in this list. While that is also a very complicated subject, one that many people have to talk about [Pr]	4 Items added	Suitable items available

Abbreviations: P: patient; Pr: professional

Table 6 Relevance issues raised in the development phase, with illustrative quotes, rationale and decisive action

Issue	Illustrative quote	Action	Rationale
Appropriate but long list	I just think this is a good list. I have seen a list in the past and I really had my doubts whether they were focused on CML. [...]Well, I don't think so with this one. I just think this is a good list [P] In that respect, I think it is very nice and good that such a list exists. Because then you think, hey, I'm not the only one who has had this problem [P] And I think it's a nice list, but I notice that I think it's a lot, a big list [P] I really thought it was a lot, yes, I had it, I also showed it to the doctor yesterday and he also said: this is really way too much [Pr] So I thought it was quite a lot of questions [Pr]	See below	No items stood out as irrelevant. Also, many participants wished to add missing items
Patients value comprehensiveness over brevity	Better too many than too few [P] I wouldn't mind answering 50 questions, right? You can, it's yes or no, right? I mean, you don't have to answer them all in detail. It's a short moment of 'yes, I have it' or, 'I don't have it'. You don't have to explain it. [...] I would prefer completeness [P]	Be as comprehensive as possible	Patients' wish for comprehensiveness outweighs wish for brevity
Personalisation is desired	[...] break down into those questions. So indeed what you do with skin problems, and what you do with eyes for example, yes [Pr] A first question might be uh, do you have any side effects? And then yes or no. Because there are people and they really don't have any problems. And if they answer no, then the rest is not necessary, right? [Pr] And you can indicate this in advance, but you can of course also do it afterwards. That you say, well, that doesn't apply at all, and this one doesn't apply, well then uh you just keep what applies to you [P]	Evaluate possibilities for personalization: Allow to opt out (with reason) Apply (more) filter questions Apply computer adaptive testing to administer items scored on previously	Participants express wish for personalization of list, for better user experience

Abbreviations: P: patient; Pr: professional

Table 7 Frequency of administration issues raised in the development phase, with illustrative quotes, rationale and decisive action

Issue	Illustrative quote	Action	Rationale
Link administration to BCR::ABL1 monitoring	I might do it at the same time as the BCR-ABL check [Pr]	Link administration to BCR::ABL1 monitoring	
	That's minimal actually, right? That is practical and you can relate it to an objective measure. Disease measure [BCR::ABL1] [Pr]		
	I would actually find it quite useful if the patient had completed such a questionnaire standardly, every time, before coming to see me [Pr]		
Administer more frequently at start (new) TKI	Because you have people who can't tolerate anything, and I then put them on a new drug, hey, an unregistered drug, asciminib for example, and then I would also really like to see how that works, how much impact it has [Pr]	Administer more frequently at start (new) TKI	Times around starting, switching and stopping are the most interesting
	I'm still a bit clinging to those 6 months, uhm, where, look, when you're just starting something, yes, you speak to each other more often than every 3 months, right? But for the first year, you could do it every 3 months, and then every six months. [...] because it is precisely in that first year that it is most often you switch and all that, people have side effects [Pr]		
Administer during treatment-free remission	If someone stops, for example, I thought it would be interesting to see that [toxicity symptom monitoring] more often, because people can then get those withdrawal complaints [Pr]	Administer during treatment-free remission	See above
Enable on-demand administration on patients' own initiative	I [would like to] fill it out every month [P]	Enable on-demand administration on patients' own initiative	Some patients wish to enter toxicity symptoms very frequently
	I would fill it out once a month [P]		
	Complete 1 x per month or 1 x every 2 months (not more often) [P]		
	Once a week is fine [P]		

Abbreviations: P: patient; Pr: professional

Table 8 Order of items issues raised in the development phase, with illustrative quotes, rationale and decisive action

Issue	Illustrative quote	Action	Rationale
Keep similar items together	Perhaps because I am currently in training, I miss a logical order, um, because in my view it is completely mixed up. And I think it would be better if you did it in order of the tract anamnesis [Pr]	Reorder and check in validation phase	Some participants felt order of items was not logical (enough)
Start with most prevalent symptoms	I can also imagine that you want to do it in order of likelihood of complaints. Uhm, even then I think it's not quite in the right order I know myself, if I have to answer 48 questions, and the first 20 don't apply to me, then I don't feel like it anymore, then I drop out. Then I'm really not going to finish it. And if I start with the questions that I think are recognizable, recognizable, I have that too, then I am more motivated to finish it. So hence my suggestion to start with uh with the common points [P]	Reorder and check in validation phase	Starting with most prevalent symptoms ensures better user experience
Move from physical to emotional items	Sequence [...] from physical to emotional [P]	Reorder and check in validation phase	The questionnaire will start with the most prevalent symptoms (fatigue) and keep emotional symptoms such as mood symptoms somewhere in the middle to stress their importance. In the consensus meeting it was preferred that sexuality items would be moved to the middle as well, not to leave them as the last stop during consultations. However, EORTC recommended against as the recall period for most sexuality items is four weeks vs. one week for all other items. If inserted at the end the recall period changes only once

Abbreviations: P: patient; Pr: professional

Table 9 Open text-box issue raised in the development phase, with illustrative quotes, rationale and decisive action

Issue	Illustrative quote	Action	Rationale
Provide an open text-box	I missed something when filling it out, but I can't remember that now. So open field would be nice [P] Perhaps you can add your own question and answer [P] I would really like to see [...] why I had a cramp in my thigh on Thursday evening at seven o'clock. And I want to be able to read that [...] on that Thursday evening [...] I made a lot of physical effort, which caused the cramp [...] I must say that [...] what you did that day and the side effect you have might be related. Then you cannot necessarily say that it could be due to the medicines. So then you actually rule out some things [P]	Provide an open text field at the end	Patients indicated they wanted to enter missing symptoms, questions and/or answers, and lifestyle factors that might relate to (the severity of) symptoms

Abbreviations: P: patient; Pr: professional

Table 10 Recall period issue raised in the development phase, with illustrative quotes, rationale and decisive action

Issue	Illustrative quote	Action	Rationale
A week is short to look back	I think last week is so short. Might do for a longer time [P] A week [is] short to look back [P]	None	A week is a generally accepted recall period, also used with the PRO-CTCAE for example

Abbreviations: P: patient; Pr: professional

Detailed results qualitative validation

New issues were mentioned within the topics: (1) Comprehensibility; (2) Comprehensiveness; and (3) Graphical overview. A summary of all items added, and the final set after the qualitative validation phase is shown in Table 3. Twelve items had issues with **(1) Comprehensibility** (Table 11): eleven items were retained; and one item was switched for a better suitable alternative. Of the eleven retained items, six did not have a better suitable alternative available in the EORTC Item Library, though for one item an additional option was available that might be useful and this item was added. Of the other five items, one item was an overarching item which captured both suggestions of the participant; one item was to be moved to a better suited cluster of items; one item included both complicated and simple phrasing in Dutch and was therefore considered understandable; one item with a double negative phrasing was phrased in that sense to keep all answering options in the same direction (a higher score indicates more symptoms); one item was considered a milder version of the suggestion done by the participant. Regarding **(2) Comprehensiveness**, 21 new topics were raised (Table 12). For 11 topics suitable items were available in the EORTC Item Library, for which 11 items were added. For eight topics no suitable items were available in the EORTC Item Library. Two topics were considered interference items, and were not added due to the many items already included. With regard to the **(3) Graphical overview**, patients discussed their own graphical overview: some text was redundant and could be left out; explanation in text and graphical form was required to be able to understand the spark lines; category symbols contributed to easier navigation; inclusion of medication (change) was valuable context; option to highlight symptoms for discussion with professionals would give focus; difference between stable scores and not scoring not visible in spark lines so zooming option is preferred; though colours were perceived as more appealing, they add complexity in design and interpretation; opinions differed on whether to leave out symptoms that only scored 'not at all' so this may be a personalised option.

Table 11 New comprehensibility issues raised in the development phase, with illustrative quotes, rationale and decisive action

	Issue	Illustrative quote	Action	Rationale
1	Short of breath unclear to some patients according to professionals	And short of breath [...] Because sometimes people consider shortness of breath [Nl; kortademigheid] and feeling stuffy [Nl: benauwdheid] to be very different things. I think it also depends a bit on the region, what terminology people give to it. [...] No, because I even notice that sometimes I ask people: are you short of breath? No, but stuffy. So I thought that was striking [Pr] You could also start there with whether there are problems with the lungs. They sometimes find the question of shortness of breath very difficult [Pr]	None	No better suitable alternative available; likely to be understood by most patients as this seems a regional problem
2	Skin discolouration can have different meanings	Skin discolouration, because yeah, that's a bit vague I think, skin discolouration. [...] It can either become very white, but you can also get red skin [P]	None	Overarching question that captures both a change in redness and whiteness
3	Feeling weak not to be positioned with fatigue items	I couldn't deal with that question, because 'weak'... I immediately thought: 'That relates to muscles,' but that has nothing to do with this [P]	To be moved from energy to muscle items in the future	Participant associates feeling weak more with 'muscles' than with 'energy'
4	Lacking energy can be better phrased	I would interpret that as listless, so to speak. Yes, it is more clear to me than a lack of energy, yes [P]	Added	Lethargy item available which might suit some patients better
5	Vague meaning 'urinating frequently'	And I thought frequent urination was a bit of a vague thing, because what is 'often'? I think you should add: 'more than so many times' [Pr]	None	No better suitable alternative available. Comprehension not affected
6	Constipation difficult phrasing	Just like constipation. That's way too difficult. Did you have difficulty passing stool? [P]	None	Both the difficult and more simple phrasing present in Dutch item, thus comprehension not affected

Table 11 Continued

	Issue	Illustrative quote	Action	Rationale
7	Problems (with coughing) considered unclear	Well, then I would rather ask 'have you been coughing, have you coughed excessively, other than with a cold?' [...] Yes, have you had problems with coughing, then I would think [...] what kind of problem could I have? That I can't cough it out? Do you understand? [P]	Alternative item chosen	Better suitable alternative (without problems) available
8	Tolerating (heat or cold) considered vague	How do you get to [this] question? With 'tolerating'? I understand that you can't handle the cold or heat very well... but what do you do if you have problems with tolerating it? That's just: 'I can't stand the cold or the heat', right? [P]	None	No better suitable alternative available; comprehension not affected
9	Double negative phrasing considered complicated	It says: 'are you less interested'? And then: 'not at all'. Then it means you are not less interested at all. Or are you saying that I am not interested in sex at all? That's a bit unclear. Do you understand what I mean? It's actually a double negative then. That makes it difficult [...] because it assumes that you were at least interested in sex [P]	None	The 'double negative' ensures that answering options all point in the same direction
10	Skin rash unclear meaning	What exactly do you mean by rash? [...] I wonder what exactly is that? Are those bumps, or something [...]? [P]	None	No alternative available. Retained as rash (Nl: uitslag) is a very common word
11	(Difficulty remembering) things considered vague	I didn't like that question very much. "Have you had trouble remembering things?" [...] 'things' is also a somewhat difficult word [...] 'Are you forgetful?' perhaps that is a better word [P]	None	Alternative available (forgetfulness), however, trouble remembering is more mild compared to forgetfulness so retained
12	Hot flushes Flemish phrase incomprehensible	And then further along there came 'vapeurs', yes, that is really Flemish then [Pr]	None	No alternative available. The item is still understandable as the Flemish word is an addition between brackets to the Netherlands Dutch phrasing

Abbreviations: Nl: Dutch; P: patient; Pr: professional

Table 12 New comprehensiveness issues raised in the development phase, with illustrative quotes, rationale and decisive action

	Issue	Illustrative quote	Action	Rationale
1	Dry vagina	Then the last question of "do you have difficulty getting/maintaining an erection?". then I think yes, I can imagine when people have sex, uh, that women can also suffer from, for example, a dry vagina, that that might be annoying [P]	Added	Suitable item available
2	Leg pain during exercise	And oh yes, there is one more that I missed, uhm, at least I thought I hadn't seen it, pain in my legs after walking for a while. So claudication complaints [P]	Added	Suitable item available
3	Palpitations	I don't see any palpitations here yet, for example. And that doesn't hurt, but it is a side effect. I didn't really come across that later on either. [...] palpitations can feel alarming or be a side effect that makes the haematologist say: maybe we'll do an ECG next time you have it again? [P]	Added	Suitable item available
4	Muscle stiffness	Only later did I realise that there are side effects that I did not find in your questionnaire, such as stiff muscles, poorer wound healing (for me even all mosquito bites become infected), nosebleeds, menstruation may change [P]	Check yearly	No suitable item available
5	Slow wound healing	See muscle stiffness	Added	Suitable item available
6	Skin infections, esp. acne	I suffered very badly from acne. And I hear that more often. [...] In terms of painful inflammation under the skin. Particularly under the scalp [P]	Check yearly	No suitable item available
7	Nose bleeds	See muscle stiffness	Added	Suitable item available
8	Change menstruation	See muscle stiffness	Check yearly	No suitable item available
9	Fatigue interference	[...] because that lack of energy, or at least, you also have a score list for that. Sometimes a question is also asked in a different way, because can you still do all your activities as you were used to? [Pr]	None	Interference items not to be included due to many items, therefore interference due to pain also left out
10	Petechiae	You ask about bruises, you're not actually asked about, yes, you could say 'have you had a rash?' But I know my doctor at one point said if you get spot bleeds everywhere, then it's time to really sound the alarm [P]	Check yearly	No suitable item available

Table 12 Continued

	Issue	Illustrative quote	Action	Rationale
11	Change hair structure	By hair texture I mean I used to have straight hair and now I have curly hair [P]	Check yearly	Not available in Dutch
12	Change hair colour	We know that hair can also change colour [Pr] My hair turned grey [P]	Check yearly	Not available in Dutch
13	Red eyes		Added	Suitable item available
14	Sensitivity sound	One of the things I really notice is that a lot of noise makes me very tired. So when I've been there on a school day, I'm so devastated afterwards because I've heard a lot of noise. So if I'm in a room having lunch with a whole group, I better decide to get out, otherwise I'll be completely devastated. I can't filter that anymore [P]	Check yearly	No suitable item available
15	Increased appetite	It is a lack of appetite or an increase in appetite, because it also causes an increase [Pr]	Check yearly	No suitable item available
16	Muscular pain after sex	What I personally experience is that due to more muscle tension, I have really bad muscles for two or three days, different than before. Do you understand what I mean? So I experience more discomfort afterwards [P]	Added	Suitable item available
17	Anxiety	And then I actually miss it anxiety. [...] anxiety and depression [...] those are actually often the emotions for which you can also have a diagnosis. This also allows you to do intervention, treatment, and referral, so it is practically relevant to make that distinction [Pr]	Added	Suitable item available
18	Interference muscle cramps		None	Interference items not to be included due to many items
19	Fine motor skills apraxia	I have polyneuropathy as a result of the medication and that is a very serious obstacle, I notice [...] It becomes less easy to button buttons, it becomes more difficult to remove the meat from the packaging, and I also notice that I have become less handy with doing odd jobs, for example [P]	Added	Suitable item available

Table 12 Continued

	Issue	Illustrative quote	Action	Rationale
20	Teeth problems (tooth growth)	Yes, I'm missing a part about teeth [...] but it seems to happen to more people that they have problems with their teeth or molars. Everyone has a cavity now and then and those kinds of things happen from time to time ... but I also notice that my teeth have become weaker [...] for example this tooth comes up and the dentist says: "I really don't understand it at all." [P]	Added	Suitable item available
21	Bleeding gums	Then I also had something like "well, maybe people have problems with their teeth" [...] perhaps also dental problems [...] toothache or maybe it's uh, gum bleeding, uhm, sensitive gums [P]	Added	Suitable item for bleeding gums available

Abbreviations: P: patient; Pr: professional

Detailed results quantitative validation

Comprehensiveness: At T1 and T2, 15 new issues were raised: for eight issues no suitable item was available (in Dutch), for six issues a suitable item was available and added; for one issue a better alternative was available and this item was switched (Table 13). At T3 no missing issues were raised. At T4, one missing issue was raised (dry eyes). This item had been included in the potential filter item from EORTC, which had been eliminated at T4 for lack of validity as a filter question.

Table 13 New comprehensiveness issues raised in the development phase, with remark made, rationale and decisive action

	Remark	Action	Rationale
1	Raynaud's phenomenon	Added	Suitable item available
2	Chronic bladder infection	Check yearly	No suitable item available
3	Easily susceptible to common infectious diseases	Added	Suitable item available
4	Easily susceptible to upper respiratory tract infection	Added	Suitable item available
5	Sour taste	Switched items	More suitable item available
6	Nerve pain	Check yearly	No suitable item available
7	Absence of menstruation	Check yearly	No suitable item available
8	Slow heartbeat	Check yearly	No suitable item available
9	Greasy hair	Check yearly	No suitable item available
10	Rumbling and grunting intestines	Check yearly	Not available in Dutch
11	Receding gums	Check yearly	No suitable item available
12	Back pain	Added	Suitable item available
13	Problems with nails	Added	Suitable item available
14	Coughing with mucus	Check yearly	No suitable item available
15	Sneezing	Added	Suitable item available
16	Dry eyes	Added	Suitable item available

Table 14 Final set of filter questions, translated from Dutch, and status of validation*

Category name	Filter question [$]	Validated	Invalid responses at T4	Adapted filter question, to be validated	Filters for EORTC
Fatigue	… fatigue (6 questions: tired, need to rest, lacked energy, lethargic, weak, drowsy)	Yes	0	Na	IL281
Head and brain	… head or brain (5 questions: headache, dizziness, remembering, concentrating, thinking clearly)	Yes	2 (5%)	Na	IL255
Skin	… skin (6 questions: bruises, discolouration, itch, rash, dry/flaky/cracked, sore/painful)	Yes	2 (6.5%)	Na	IL282
Wound healing	… wound healing (1 question: wound healing)	Yes	0	Na	IL283
Hair and nails	… hair or nails (2 questions: hair loss, nail problems)	Yes	0	Na	IL284
Eyes	… eyes (5 questions: watery, burning, light sensitive, blurry, red)	Yes[£]	1 (3.6%)	… eyes (6 questions: watery, dry, burning, light sensitive, blurry, red)	IL387[£]
Oedema	… oedema (2 questions: swelling eyes/face, swelling body parts)	Yes	0	Na	IL286
Hearing	… hearing (1 question: hearing problems)	Yes	0	Na	IL258
Nose bleeds and nose problems	… nose bleeds or nose problems (2 questions: nose bleeds, nose problems)	No	14 (20.6%)	… nose bleeds or nose problems (2 questions: nose bleeds, nose problems (smell, runny/blocked, sneezing, sore, dry))	IL259
Mouth and teeth	… mouth or teeth (4 questions: dry mouth, pain/sore mouth, teeth problems, bleeding gums)	Yes	1 (2.4%)	Na	IL260
Food and taste	… food or taste (2 questions: taste, appetite)	Yes	3 (5%)	Na	IL261

Table 14 Continued

Category name	Filter question [$]	Validated	Invalid responses at T4	Adapted filter question, to be validated	Filters for EORTC
Mood	… mood (6 questions: depressed, irritable, tense, worried, afraid, difficulty sleeping)	Yes	3 (7.3%)	Na	IL287
Heart	… heart (2 questions: palpitations, pain in chest)	Yes	0	Na	IL263
Breathing	… breathing (2 questions: short of breath, coughing)	Yes	0	Na	IL264
Stomach and intestines	… stomach or intestines (8 questions: bloated, indigestion, nausea, vomiting, diarrhoea, constipation, abdominal pain/ cramps, flatulence)	Yes	1 (3.7%)	Na	IL265
Urination	… urination (1 question: frequent urination)	Yes	1 (2.6%)	Na	IL266
Muscles and joints	… muscles or joints (3 questions: muscle cramps, muscle/ joint ache/pain, muscle weakness)	Yes	0	Na	IL288
Fingers and toes	… fingers or toes (3 questions: pale/cold, tingling/numb, dexterity)	Yes	1 (2.9%)	Na	IL289
Pain	… pain (3 questions: pain walking, back pain, pain in general)	Yes	1 (3.2%)	Na	IL290
Fever and sweating	… fever or sweating (2 questions: fever/chills, excessive sweating)	Yes	3 (5.1%)	Na	IL291
Warmth and cold	… warmth or cold (2 questions: tolerating warmth/cold, hot flushes)	Yes	3 (7.5%)	Na	IL292
Infection	… infections (2 questions: airway infections, other infections)	Yes	1 (1.6%)	Na	IL293
Nipples and breasts	… nipples or breasts (1 question: sore/enlarged nipples/ breasts)	Yes	0	Na	IL294

Table 14 Continued

Category name	Filter question [$]	Validated	Invalid responses at T4	Adapted filter question, to be validated	Filters for EORTC
Sexuality	... sexuality (5 questions: interest, pleasure, pain, problems erection, dry vagina)	No	6 (12.8%)	... sexuality (5 questions: interest in sex, pleasure from sex, pain, problems erection, dry vagina)	IL295

*Validation in Dutch only. Validity defined as: <10% of patients, whom indicated not to have symptoms in a certain category, scored 'quite a bit' or 'very much' on any of the category items. [$]All questions started: 'During the past week: did you have any signs or symptoms of:'. For the Sexuality category this was: 'During the past 4 weeks: did you have any signs or symptoms of:'. Answering options were: 'Yes' and 'No, I have no symptoms or complaints on this topic'. [£] In the final version (EORTC IL387) an item for dry eyes was added. Abbreviations: na: not applicable

Table 15 Reliability (weighted kappa) and measurement error (% positive agreement) of 70 items tested in both T1 and T2, in patients on the same TKI and same dosage

Item	Weighted kappa			% positive agreement		
	Weighted kappa	Lower 95% confidence interval	Upper 95% confidence interval	% positive agreement	Lower 95% confidence interval	Upper 95% confidence interval
Tired	0.69	0.55	0.82	64.1	53.0	73.9
Need to rest	0.62	0.47	0.78	57.1	46.0	67.6
Lacked energy	0.69	0.58	0.80	53.8	42.9	64.5
Lethargic	0.57	0.42	0.73	54.5	43.5	65.2
Weak	0.50	0.31	0.69	54.7	43.4	65.4
Drowsy	0.54	0.40	0.67	48.0	37.1	59.1
Headaches	0.51	0.29	0.74	70.4	59.0	79.8
Dizzy	0.69	0.55	0.82	80.3	69.6	87.9
Difficulty remembering	0.48	0.28	0.68	58.6	46.9	69.4
Difficulty concentrating	0.76	0.65	0.88	66.7	54.9	76.6
Trouble thinking clearly	0.56	0.39	0.72	63.8	52.0	74.1
Bruised easily	0.66	0.47	0.86	72.5	61.0	81.6
Skin colour change	0.51	0.25	0.77	76.9	65.4	85.5
Itchy skin	0.60	0.43	0.76	59.4	47.1	70.5
Skin rash	0.64	0.42	0.85	73.4	61.5	82.7
Dry/flaking/cracked skin	0.47	0.29	0.64	45.3	33.7	57.4
Sore/painful skin	0.42	0.15	0.69	68.8	56.6	78.8
Wounds healing problems	0.45	0.11	0.78	86.4	76.1	92.7
Hair loss	0.72	0.58	0.86	76.9	65.4	85.5

Table 15 Continued

Item	Weighted kappa			% positive agreement		
	Weighted kappa	Lower 95% confidence interval	Upper 95% confidence interval	% positive agreement	Lower 95% confidence interval	Upper 95% confidence interval
Watery eyes	0.64	0.45	0.82	61.5	49.4	72.4
Burning eyes	0.41	0.16	0.66	53.1	41.1	64.8
Discomfort bright light	0.74	0.62	0.87	64.1	51.8	74.7
Blurred vision	0.47	0.23	0.71	51.6	39.6	63.4
Red eyes	0.82	0.69	0.95	88.7	78.5	94.4
Swelling face/eyes	0.76	0.65	0.86	65.2	53.1	75.5
Swelling body parts	0.72	0.56	0.88	75.0	62.8	84.2
Hearing problems	0.76	0.60	0.92	85.7	75.0	92.3
Nose bleeds	0.51	0.10	0.91	85.7	75.0	92.3
Dry mouth	0.80	0.69	0.92	68.8	56.6	78.8
Pain/soreness mouth	0.64	0.47	0.81	71.9	59.9	81.4
Teeth problems	0.69	0.53	0.85	75.0	63.2	84.0
Bleeding gums	0.76	0.56	0.96	84.1	73.2	91.1
Taste change	0.49	0.27	0.71	68.8	56.6	78.8
Appetite loss	0.79	0.71	0.87	73.0	61.0	82.4
Depressed	0.66	0.51	0.80	67.2	55.0	77.4
Irritable	0.51	0.31	0.71	58.7	46.4	70.0
Tense	0.72	0.57	0.86	65.6	53.4	76.1
Worried	0.67	0.50	0.84	68.8	56.6	78.8
Afraid	0.71	0.54	0.88	77.8	66.1	86.3

Table 15 Continued

Item	Weighted kappa			% positive agreement		
	Weighted kappa	Lower 95% confidence interval	Upper 95% confidence interval	% positive agreement	Lower 95% confidence interval	Upper 95% confidence interval
Difficulty sleeping	0.67	0.50	0.83	58.6	46.9	69.4
Palpitations	0.72	0.61	0.82	64.1	51.8	74.7
Pain in chest	0.61	0.35	0.88	81.0	69.6	88.8
Short of breath	0.69	0.55	0.83	65.1	52.8	75.7
Coughed	0.51	0.30	0.72	69.4	57.0	79.4
Feeling bloated	0.59	0.44	0.75	50.8	38.8	62.7
Indigestion	0.74	0.62	0.87	63.5	51.1	74.3
Nausea	0.78	0.65	0.91	71.0	58.7	80.8
Vomited	0.40	-0.09	0.88	85.7	75.0	92.3
Diarrhoea	0.78	0.65	0.90	73.0	61.0	82.4
Obstipation	0.73	0.55	0.90	73.0	61.0	82.4
Abdominal pain/cramps	0.74	0.59	0.88	67.7	55.4	78.0
Flatulence	0.81	0.71	0.90	65.6	53.0	76.3
Frequent urination	0.73	0.61	0.85	57.4	44.9	69.0
Muscle cramps	0.64	0.47	0.82	50.8	38.6	62.9
Aches/pains muscles/joints	0.74	0.64	0.83	50.8	38.6	62.9
Muscle weakness	0.48	0.26	0.70	51.7	39.3	63.8
Tingling/numbness hands/ feet	0.75	0.63	0.86	63.9	51.4	74.8
Problems finger dexterity	0.60	0.43	0.77	72.9	60.4	82.6
Trouble walking pain	0.70	0.53	0.87	75.4	63.3	84.5
Pain	0.80	0.68	0.92	78.3	66.4	86.9

Table 15 Continued

Item	Weighted kappa			% positive agreement		
	Weighted kappa	Lower 95% confidence interval	Upper 95% confidence interval	% positive agreement	Lower 95% confidence interval	Upper 95% confidence interval
Fever/chills	0.62	0.39	0.85	83.3	72.0	90.7
Sweating excessively	0.65	0.46	0.84	65.0	52.4	75.8
Problems tolerating heat/cold	0.68	0.48	0.89	66.7	54.1	77.3
Hot flushes	0.66	0.39	0.94	76.3	64.0	85.3
Sore/enlarged nipples/ breasts	0.54	0.14	0.93	93.3	84.1	97.4
Less interest in sex	0.66	0.49	0.84	67.2	54.4	77.9
Less sexual enjoyment	0.42	0.19	0.64	53.4	40.8	65.7
Pain during/after sex	0.82	0.66	0.98	86.2	75.1	92.8
Difficulty getting/ maintaining erection	0.58	0.30	0.86	66.7	48.8	80.8
Dry vagina	0.89	0.80	0.98	78.6	60.5	89.8

Responsiveness

50 Symptoms/scales were available both in T3-T4 and as anchor item/scales. Five symptoms/scales had only constant data in one of the pairs (either change from T3 to T4 or the anchor item) so correlation could not be evaluated. 25 out of 45 remaining symptoms/scales (56.6%) had an acceptable (≥0.30) association between the anchor and the change score. For these 25 symptoms/scales, groups with stable symptoms, worse, stable and better symptoms were compared (Table 16). The median change across these 25 symptom/scales in improved groups was 29.6%, in stable groups 7.7%, and -18.5% in worsened groups. In 14/25 symptoms/scales the between-group difference was significant. In 21/25 symptoms/scales all mean changes differed in the expected directions, while in four symptoms/scales the improved group improved less than the stable group, though differences were very small (<3%) or only one data point was available for the improved group.

Table 16 Mean change in mean symptom severity scores for patients whom reported improved, stable or worsened symptoms between T3 and T4, for symptoms with an acceptable (>0.30) Spearman's rank correlation between item/scale score changes and anchor scores

	Mean change symptom severity (95%CI)			
Symptom/scale	**Improved**	**Stable**	**Worsened**	**p value***
Fatigue (scale)	18.5 (2.5 to 34.5)	8.5 (4.5 to 12.5)	2.5 (-6.7 to 11.8)	0.08
Cognitive problems(scale)	29.6 (13.7 to 45.5)	6.8 (3.1 to 104)	-16.7 (NA)	0.03
Hair loss	50.0 (-8.7 to 108.7)	-1.0 (-4.0 to 2.1)	-8.3 (-25.8 to 9.1)	<0.01
Headaches	22.2 (-25.6 to 70.0)	4.5 (-1.1 to 10.1)	-66.6 (NA)	0.07
Eye problems (scale)	2.2 (-2.6 to 7.0)	3.3 (0.5 to 6.1)	-5.2 (-12.4 to 2.0)	0.17
Swelling face/eyes	16.7 (-20.8 to 54.1)	2.6 (-1.7 to 7.0)	-20.0 (-47.5 to 7.6)	0.10
Hearing problems	66.6 (NA)	1.7 (-4.1 to 7.4)	-16.7 (-228.2 to 194.9)	0.13
Nose problems	NA	22.2 (15.3 to 29.1)	0 (0 to 0)	0.27
Appetite loss	NA	27.1 (19.2 to 35.0)	-33.3 (NA)	0.08
Mood problems (scale)	38.9 (32.0 to 45.8)	3.7 (-1.8 to 9.2)	-0.7 (-6.5 to 5.0)	<0.01
Pain in chest	0 (NA)	11.1 (7.3 to 14.9)	-16.7 (-36.8 to 3.5)	<0.05
Short of breath	33.3 (-110.1 to 176.8)	10.7 (5.6 to 15.7)	-9.5 (-23.1 to 4.1)	0.08
Feeling bloated	NA	12.3 (7.6 to 17.0)	-19.1 (-37.8 to -0.3)	<0.01
Indigestion	NA	4.8 (-0.6 to 10.2)	-29.7 (-57.2 to 3.8)	<0.01
Vomited	NA	2.3 (-27.5 to 65.6)	-33.3 (-33.3 to -33.3)	<0.001
Diarrhoea	19.0 (-27.5 to 65.6)	9.3 (5.2 to 13.3)	-33.3 (-33.3 to -33.3)	<0.01
Abdominal pain/cramps	33.4 (32.7 to 34.0)	13.7 (7.3 to 20.1)	-18.5 (-33.9 to -3.1)	<0.01
Flatulence	33.4 (NA)	5.4 (-0.5 to 11.3)	-14.3 (-25.6 to -3.0)	0.02
Muscle/joints cramps/pain	5.0 (-3.5 to 13.5)	8.0 (3.4 to 12.5)	-1.1 (-7.5 to 5.3)	0.20
Pale/cold fingers/toes	NA	-0.0 (-4.7 to 4.7)	-47.6 (-65.2 to -30.0)	<0.001

Table 16 Continued

	Mean change symptom severity (95%CI)			
Symptom/scale	**Improved**	**Stable**	**Worsened**	**p value***
Tingling/numbness hands/ feet	NA	7.4 (0.6 to 14.2)	-18.5 (-29.1 to -7.9)	0.01
Back pain	0 (NA)	11.8 (7.2 to 16.4)	-18.5 (-42.7 to 5.6)	0.01
Airway infections	22.2 (-88.2 to 132.6)	10.6 (4.4 to 16.7)	-50.0 (-87.4 to -12.6)	0.03
Other infections	50 (-125.7 to 225.7)	11.1 (2.9 to 19.3)	-66.7 (-183.8 to 50.5)	0.05
Sore/enlarged nipples/ breasts	100 (NA)	4.6 (1.3 to 7.9)	NA	0.02
Median	**29.6**	**7.7**	**-18.5**	

* Kruskal-Wallis rank sum test. Shaded areas indicate a significant ($p<0.05$) differences between groups. NA: not available

Table 17 Prevalence (% with 95%confidence intervals) of symptom scores for seven scales plus 76 single items, in 12 patients on asciminib on T1

Item/scale	Asciminib					
	Not at all	A little	Quite a bit	A lot	Any severity	Moderate-severe $
Fatigue (scale)	8.3 (0.2-38.5)	Na	Na	Na	91.7 (61.5-99.8)	Na
Tired	16.7 (2.1-48.4)	33.3 (9.9-65.1)	25.0 (5.5-57.2)	25.0 (5.5-57.2)	83.3 (51.6-97.9)	50.0 (21.1-78.9)
Needed rest	16.7 (2.1-48.4)	41.7 (15.2-72.3)	41.7 (15.2-72.3)	0.0 (0.0-26.5)	83.3 (51.6-97.9)	41.7 (15.2-72.3)
Lacked energy	9.1 (0.2-41.3)	45.5 (16.7-76.6)	36.4 (10.9-69.2)	9.1 (0.2-41.3)	90.9 (58.7-99.8)	45.5 (16.7-76.6)
Felt lethargic	30.0 (6.7-65.2)	40.0 (12.2-73.8)	20.0 (2.5-55.6)	10.0 (0.3-44.5)	70.0 (34.8-93.3)	30.0 (6.7-65.2)
Felt weak	10.0 (0.3-44.5)	70.0 (34.8-93.3)	10.0 (0.3-44.5)	10.0 (0.3-44.5)	90.0 (55.5-99.7)	20.0 (2.5-55.6)
Sleepy	20.0 (2.5-55.6)	70.0 (34.8-93.3)	0.0 (0.0-30.8)	10.0 (0.3-44.5)	80.0 (44.4-97.5)	10.0 (0.3-44.5)
Headaches	30.0 (6.7-65.2)	60.0 (26.2-87.8)	10.0 (0.3-44.5)	0.0 (0.0-30.8)	70.0 (34.8-93.3)	10.0 (0.3-44.5)
Dizziness	44.4 (13.7-78.8)	22.2 (2.8-60.0)	22.2 (2.8-60.0)	11.1 (0.3-48.2)	55.6 (21.2-86.3)	33.3 (7.5-70.1)
Cognitive problems (scale)	11.1 (0.3-48.2)	Na	Na	Na	88.9 (51.8-99.7)	Na
Difficulty remembering	33.3 (7.5-70.1)	44.4 (13.7-78.8)	22.2 (2.8-60.0)	0.0 (0.0-33.6)	66.7 (29.9-92.5)	22.2 (2.8-60.0)
Difficulty concentrating	22.2 (2.8-60.0)	66.7 (29.9-92.5)	11.1 (0.3-48.2)	0.0 (0.0-33.6)	77.8 (40.0-97.2)	11.1 (0.3-48.2)
Trouble thinking clearly	55.6 (21.2-86.3)	33.3 (7.5-70.1)	11.1 (0.3-48.2)	0.0 (0.0-33.6)	44.4 (13.7-78.8)	11.1 (0.3-48.2)
Bruised easily	55.6 (21.2-86.3)	11.1 (0.3-48.2)	11.1 (0.3-48.2)	22.2 (2.8-60.0)	44.4 (13.7-78.8)	33.3 (7.5-70.1)
Skin problems (scale)	11.1 (0.3-48.2)	Na	Na	Na	91.7 (61.5-99.8)	Na
Skin colour change	77.8 (40.0-97.2)	11.1 (0.3-48.2)	0.0 (0.0-33.6)	11.1 (0.3-48.2)	22.2 (2.8-60.0)	11.1 (0.3-48.2)
Skin itched	22.2 (2.8-60.0)	55.6 (21.2-86.3)	11.1 (0.3-48.2)	11.1 (0.3-48.2)	77.8 (40.0-97.2)	22.2 (2.8-60.0)

Table 17 Continued

Item/scale	Asciminib					
	Not at all	A little	Quite a bit	A lot	Any severity	Moderate-severe $
Skin rash	22.2 (2.8-60.0)	66.7 (29.9-92.5)	0.0 (0.0-33.6)	11.1 (0.3-48.2)	77.8 (40.0-97.2)	11.1 (0.3-48.2)
Skin dry/flaky/ cracked	22.2 (2.8-60.0)	44.4 (13.7-78.8)	22.2 (2.8-60.0)	11.1 (0.3-48.2)	77.8 (40.0-97.2)	33.3 (7.5-70.1)
Skin sore/painful	55.6 (21.2-86.3)	33.3 (7.5-70.1)	11.1 (0.3-48.2)	0.0 (0.0-33.6)	44.4 (13.7-78.8)	11.1 (0.3-48.2)
Wounds healing problems	77.8 (40.0-97.2)	22.2 (2.8-60.0)	0.0 (0.0-33.6)	0.0 (0.0-33.6)	22.2 (2.8-60.0)	0.0 (0.0-33.6)
Hair loss	55.6 (21.2-86.3)	33.3 (7.5-70.1)	0.0 (0.0-33.6)	11.1 (0.3-48.2)	44.4 (13.7-78.8)	11.1 (0.3-48.2)
Nail problems	67.9 (56.4-78.1)	20.5 (12.2-31.2)	9.0 (3.7-17.6)	2.6 (0.3-9.0)	32.1 (21.9-43.6)	11.5 (5.4-20.8)
Eye problems (scale)	22.2 (2.8-60.0)	Na	Na	Na	77.8 (40.0-97.2)	Na
Watery eyes	33.3 (7.5-70.1)	22.2 (2.8-60.0)	44.4 (13.7-78.8)	0.0 (0.0-33.6)	66.7 (29.9-92.5)	44.4 (13.7-78.8)
Burning eyes	33.3 (7.5-70.1)	22.2 (2.8-60.0)	44.4 (13.7-78.8)	0.0 (0.0-33.6)	66.7 (29.9-92.5)	44.4 (13.7-78.8)
Light sensitive eyes	55.6 (21.2-86.3)	0.0 (0.0-33.6)	33.3 (7.5-70.1)	11.1 (0.3-48.2)	44.4 (13.7-78.8)	44.4 (13.7-78.8)
Blurry eyes	33.3 (7.5-70.1)	55.6 (21.2-86.3)	11.1 (0.3-48.2)	0.0 (0.0-33.6)	66.7 (29.9-92.5)	11.1 (0.3-48.2)
Red eyes	75.0 (34.9-96.8)	25.0 (3.2-65.1)	0.0 (0.0-36.9)	0.0 (0.0-36.9)	25.0 (3.2-65.1)	0.0 (0.0-36.9)
Swelling face/eyes	77.8 (40.0-97.2)	11.1 (0.3-48.2)	0.0 (0.0-33.6)	11.1 (0.3-48.2)	22.2 (2.8-60.0)	11.1 (0.3-48.2)
Swelling body parts	87.5 (47.3-99.7)	12.5 (0.3-52.7)	0.0 (0.0-36.9)	0.0 (0.0-36.9)	12.5 (0.3-52.7)	0.0 (0.0-36.9)
Hearing problems	75.0 (34.9-96.8)	0.0 (0.0-36.9)	25.0 (3.2-65.1)	0.0 (0.0-36.9)	25.0 (3.2-65.1)	25.0 (3.2-65.1)
Nose bleeds	50.0 (15.7-84.3)	37.5 (8.5-75.5)	0.0 (0.0-36.9)	12.5 (0.3-52.7)	50.0 (15.7-84.3)	12.5 (0.3-52.7)
Nose problems	40.0 (5.3-85.3)	0.0 (0.0-52.2)	60.0 (14.7-94.7)	0.0 (0.0-52.2)	60.0 (14.7-94.7)	60.0 (14.7-94.7)
Dry mouth	25.0 (3.2-65.1)	50.0 (15.7-84.3)	12.5 (0.3-52.7)	12.5 (0.3-52.7)	75.0 (34.9-96.8)	25.0 (3.2-65.1)

Table 17 Continued

Item/scale	Asciminib					
	Not at all	A little	Quite a bit	A lot	Any severity	Moderate-severe $
Pain/soreness mouth	50.0 (15.7-84.3)	25.0 (3.2-65.1)	25.0 (3.2-65.1)	0.0 (0.0-36.9)	50.0 (15.7-84.3)	25.0 (3.2-65.1)
Teeth problems	62.5 (24.5-91.5)	25.0 (3.2-65.1)	12.5 (0.3-52.7)	0.0 (0.0-36.9)	37.5 (8.5-75.5)	12.5 (0.3-52.7)
Bleeding gums	87.5 (47.3-99.7)	0.0 (0.0-36.9)	12.5 (0.3-52.7)	0.0 (0.0-36.9)	12.5 (0.3-52.7)	12.5 (0.3-52.7)
Taste change	37.5 (8.5-75.5)	50.0 (15.7-84.3)	12.5 (0.3-52.7)	0.0 (0.0-36.9)	62.5 (24.5-91.5)	12.5 (0.3-52.7)
Appetite loss	37.5 (8.5-75.5)	37.5 (8.5-75.5)	12.5 (0.3-52.7)	12.5 (0.3-52.7)	62.5 (24.5-91.5)	25.0 (3.2-65.1)
Mood problems (scale)	10.0 (0.3-44.5)	Na	Na	Na	90.0 (55.5-99.7)	Na
Depressed	50.0 (15.7-84.3)	25.0 (3.2-65.1)	12.5 (0.3-52.7)	12.5 (0.3-52.7)	50.0 (15.7-84.3)	25.0 (3.2-65.1)
Irritable	25.0 (3.2-65.1)	50.0 (15.7-84.3)	12.5 (0.3-52.7)	12.5 (0.3-52.7)	75.0 (34.9-96.8)	25.0 (3.2-65.1)
Tense	25.0 (3.2-65.1)	37.5 (8.5-75.5)	25.0 (3.2-65.1)	12.5 (0.3-52.7)	75.0 (34.9-96.8)	37.5 (8.5-75.5)
Worried	25.0 (3.2-65.1)	25.0 (3.2-65.1)	37.5 (8.5-75.5)	12.5 (0.3-52.7)	75.0 (34.9-96.8)	50.0 (15.7-84.3)
Afraid	50.0 (15.7-84.3)	25.0 (3.2-65.1)	12.5 (0.3-52.7)	12.5 (0.3-52.7)	50.0 (15.7-84.3)	25.0 (3.2-65.1)
Insomnia	30.0 (6.7-65.2)	30.0 (6.7-65.2)	20.0 (2.5-55.6)	20.0 (2.5-55.6)	70.0 (34.8-93.3)	40.0 (12.2-73.8)
Palpitations	37.5 (8.5-75.5)	37.5 (8.5-75.5)	12.5 (0.3-52.7)	12.5 (0.3-52.7)	62.5 (24.5-91.5)	25.0 (3.2-65.1)
Pain in chest	50.0 (15.7-84.3)	37.5 (8.5-75.5)	0.0 (0.0-36.9)	12.5 (0.3-52.7)	50.0 (15.7-84.3)	12.5 (0.3-52.7)
Short of breath	25.0 (3.2-65.1)	25.0 (3.2-65.1)	37.5 (8.5-75.5)	12.5 (0.3-52.7)	75.0 (34.9-96.8)	50.0 (15.7-84.3)
Coughed	25.0 (3.2-65.1)	25.0 (3.2-65.1)	37.5 (8.5-75.5)	12.5 (0.3-52.7)	75.0 (34.9-96.8)	50.0 (15.7-84.3)
Feeling bloated	50.0 (15.7-84.3)	0.0 (0.0-36.9)	12.5 (0.3-52.7)	37.5 (8.5-75.5)	50.0 (15.7-84.3)	50.0 (15.7-84.3)
Indigestion	37.5 (8.5-75.5)	37.5 (8.5-75.5)	12.5 (0.3-52.7)	12.5 (0.3-52.7)	62.5 (24.5-91.5)	25.0 (3.2-65.1)
Nausea	50.0 (15.7-84.3)	12.5 (0.3-52.7)	37.5 (8.5-75.5)	0.0 (0.0-36.9)	50.0 (15.7-84.3)	37.5 (8.5-75.5)

Table 17 Continued

Item/scale	Asciminib					
	Not at all	**A little**	**Quite a bit**	**A lot**	**Any severity**	**Moderate-severe $**
Diarrhoea	75.0 (34.9-96.8)	0.0 (0.0-36.9)	0.0 (0.0-36.9)	25.0 (3.2-65.1)	25.0 (3.2-65.1)	25.0 (3.2-65.1)
Abdominal pain/ cramps	37.5 (8.5-75.5)	12.5 (0.3-52.7)	37.5 (8.5-75.5)	12.5 (0.3-52.7)	62.5 (24.5-91.5)	50.0 (15.7-84.3)
Vomited	87.5 (47.3-99.7)	0.0 (0.0-36.9)	12.5 (0.3-52.7)	0.0 (0.0-36.9)	12.5 (0.3-52.7)	12.5 (0.3-52.7)
Obstipation	37.5 (8.5-75.5)	50.0 (15.7-84.3)	0.0 (0.0-36.9)	12.5 (0.3-52.7)	62.5 (24.5-91.5)	12.5 (0.3-52.7)
Flatulence	37.5 (8.5-75.5)	50.0 (15.7-84.3)	0.0 (0.0-36.9)	12.5 (0.3-52.7)	62.5 (24.5-91.5)	12.5 (0.3-52.7)
Muscle/joint cramps/pain (scale)	0.0 (0.0-36.9)	Na	Na	Na	100.0 (63.1-100.0)	Na
Frequent urination	50.0 (15.7-84.3)	12.5 (0.3-52.7)	12.5 (0.3-52.7)	25.0 (3.2-65.1)	50.0 (15.7-84.3)	37.5 (8.5-75.5)
Muscle cramps	37.5 (8.5-75.5)	50.0 (15.7-84.3)	0.0 (0.0-36.9)	12.5 (0.3-52.7)	62.5 (24.5-91.5)	12.5 (0.3-52.7)
Aches/pains muscles/joints	12.5 (0.3-52.7)	25.0 (3.2-65.1)	50.0 (15.7-84.3)	12.5 (0.3-52.7)	87.5 (47.3-99.7)	62.5 (24.5-91.5)
Muscle weakness	42.9 (9.9-81.6)	42.9 (9.9-81.6)	14.3 (0.4-57.9)	0.0 (0.0-41.0)	57.1 (18.4-90.1)	14.3 (0.4-57.9)
Pale/cold fingers/ toes	80.0 (28.4-99.5)	0.0 (0.0-52.2)	0.0 (0.0-52.2)	20.0 (0.5-71.6)	20.0 (0.5-71.6)	20.0 (0.5-71.6)
Tingling/numbness hands/feet	50.0 (15.7-84.3)	37.5 (8.5-75.5)	12.5 (0.3-52.7)	0.0 (0.0-36.9)	50.0 (15.7-84.3)	12.5 (0.3-52.7)
Problems finger dexterity	50.0 (15.7-84.3)	37.5 (8.5-75.5)	12.5 (0.3-52.7)	0.0 (0.0-36.9)	50.0 (15.7-84.3)	12.5 (0.3-52.7)
Trouble walking pain	37.5 (8.5-75.5)	62.5 (24.5-91.5)	0.0 (0.0-36.9)	0.0 (0.0-36.9)	62.5 (24.5-91.5)	0.0 (0.0-36.9)
Back pain	40.0 (5.3-85.3)	40.0 (5.3-85.3)	0.0 (0.0-52.2)	20.0 (0.5-71.6)	60.0 (14.7-94.7)	20.0 (0.5-71.6)

Table 17 Continued

Item/scale	Asciminib					
	Not at all	A little	Quite a bit	A lot	Any severity	Moderate-severe $
Pain	25.0 (3.2-65.1)	12.5 (0.3-52.7)	50.0 (15.7-84.3)	12.5 (0.3-52.7)	75.0 (34.9-96.8)	62.5 (24.5-91.5)
Fever/chills	50.0 (15.7-84.3)	37.5 (8.5-75.5)	0.0 (0.0-36.9)	12.5 (0.3-52.7)	50.0 (15.7-84.3)	12.5 (0.3-52.7)
Sweating excessively	62.5 (24.5-91.5)	12.5 (0.3-52.7)	12.5 (0.3-52.7)	12.5 (0.3-52.7)	37.5 (8.5-75.5)	25.0 (3.2-65.1)
Problems tolerating heat/cold	62.5 (24.5-91.5)	25.0 (3.2-65.1)	0.0 (0.0-36.9)	12.5 (0.3-52.7)	37.5 (8.5-75.5)	12.5 (0.3-52.7)
Hot flushes	62.5 (24.5-91.5)	25.0 (3.2-65.1)	0.0 (0.0-36.9)	12.5 (0.3-52.7)	37.5 (8.5-75.5)	12.5 (0.3-52.7)
Airway infections	80.0 (28.4-99.5)	20.0 (0.5-71.6)	0.0 (0.0-52.2)	0.0 (0.0-52.2)	20.0 (0.5-71.6)	0.0 (0.0-52.2)
Other infections	100.0 (47.8-100.0)	0.0 (0.0-52.2)	0.0 (0.0-52.2)	0.0 (0.0-52.2)	0.0 (0.0-52.2)	0.0 (0.0-52.2)
Sore/enlarged nipples/breasts	100.0 (63.1-100.0)	0.0 (0.0-36.9)	0.0 (0.0-36.9)	0.0 (0.0-36.9)	0.0 (0.0-36.9)	0.0 (0.0-36.9)
Sexuality problems (scale)	25.0 (3.2-65.1)	Na	Na	Na	75.0 (34.9-96.8)	Na
Less interest sex	25.0 (3.2-65.1)	37.5 (8.5-75.5)	37.5 (8.5-75.5)	0.0 (0.0-36.9)	75.0 (34.9-96.8)	37.5 (8.5-75.5)
Less pleasure sex	37.5 (8.5-75.5)	37.5 (8.5-75.5)	25.0 (3.2-65.1)	0.0 (0.0-36.9)	62.5 (24.5-91.5)	25.0 (3.2-65.1)
Pain during/after sex	75.0 (34.9-96.8)	25.0 (3.2-65.1)	0.0 (0.0-36.9)	0.0 (0.0-36.9)	25.0 (3.2-65.1)	0.0 (0.0-36.9)
Difficult erection	25.0 (0.6-80.6)	50.0 (6.8-93.2)	25.0 (0.6-80.6)	0.0 (0.0-60.2)	75.0 (19.4-99.4)	25.0 (0.6-80.6)
Dry vagina	75.0 (19.4-99.4)	25.0 (0.6-80.6)	0.0 (0.0-60.2)	0.0 (0.0-60.2)	25.0 (0.6-80.6)	0.0 (0.0-60.2)

* At T4 as not available at T1. $ Moderate-severe includes 'Quite a bit' and 'A lot'. Abbreviations: Na: not applicable

Table 18 Prevalence (% with 95%confidence intervals) of symptom scores for seven scales plus 76 single items, in 14 patients on bosutinib on T1

Item/scale	Bosutinib					
	Not at all	A little	Quite a bit	A lot	Any severity	Moderate-severe $
Fatigue (scale)	0.0 (0.0-23.2)	Na	Na	Na	100.0 (76.8-100.0)	Na
Tired	7.1 (0.2-33.9)	50.0 (23.0-77.0)	21.4 (4.7-50.8)	21.4 (4.7-50.8)	92.9 (66.1-99.8)	42.9 (17.7-71.1)
Needed rest	0.0 (0.0-23.2)	71.4 (41.9-91.6)	28.6 (8.4-58.1)	0.0 (0.0-23.2)	100.0 (76.8-100.0)	28.6 (8.4-58.1)
Lacked energy	7.1 (0.2-33.9)	35.7 (12.8-64.9)	42.9 (17.7-71.1)	14.3 (1.8-42.8)	92.9 (66.1-99.8)	57.1 (28.9-82.3)
Felt lethargic	21.4 (4.7-50.8)	50.0 (23.0-77.0)	28.6 (8.4-58.1)	0.0 (0.0-23.2)	78.6 (49.2-95.3)	28.6 (8.4-58.1)
Felt weak	46.2 (19.2-74.9)	15.4 (1.9-45.4)	30.8 (9.1-61.4)	7.7 (0.2-36.0)	53.8 (25.1-80.8)	38.5 (13.9-68.4)
Sleepy	23.1 (5.0-53.8)	53.8 (25.1-80.8)	15.4 (1.9-45.4)	7.7 (0.2-36.0)	76.9 (46.2-95.0)	23.1 (5.0-53.8)
Headaches	72.7 (39.0-94.0)	18.2 (2.3-51.8)	9.1 (0.2-41.3)	0.0 (0.0-28.5)	27.3 (6.0-61.0)	9.1 (0.2-41.3)
Dizziness	63.6 (30.8-89.1)	36.4 (10.9-69.2)	0.0 (0.0-28.5)	0.0 (0.0-28.5)	36.4 (10.9-69.2)	0.0 (0.0-28.5)
Cognitive problems (scale)	18.2 (2.3-51.8)	Na	Na	Na	75.0 (34.9-96.8)	81.8 (48.2-97.7)
Difficulty remembering	27.3 (6.0-61.0)	36.4 (10.9-69.2)	27.3 (6.0-61.0)	9.1 (0.2-41.3)	72.7 (39.0-94.0)	36.4 (10.9-69.2)
Difficulty concentrating	36.4 (10.9-69.2)	27.3 (6.0-61.0)	18.2 (2.3-51.8)	18.2 (2.3-51.8)	63.6 (30.8-89.1)	36.4 (10.9-69.2)
Trouble thinking clearly	27.3 (6.0-61.0)	36.4 (10.9-69.2)	27.3 (6.0-61.0)	9.1 (0.2-41.3)	72.7 (39.0-94.0)	36.4 (10.9-69.2)
Bruised easily	45.5 (16.7-76.6)	18.2 (2.3-51.8)	27.3 (6.0-61.0)	9.1 (0.2-41.3)	54.5 (23.4-83.3)	36.4 (10.9-69.2)
Skin problems (scale)	18.2 (2.3-51.8)	Na	Na	Na	75.0 (34.9-96.8)	81.8 (48.2-97.7)
Skin colour change	45.5 (16.7-76.6)	36.4 (10.9-69.2)	18.2 (2.3-51.8)	0.0 (0.0-28.5)	54.5 (23.4-83.3)	18.2 (2.3-51.8)
Skin itched	45.5 (16.7-76.6)	27.3 (6.0-61.0)	27.3 (6.0-61.0)	0.0 (0.0-28.5)	54.5 (23.4-83.3)	27.3 (6.0-61.0)

Table 18 Continued

Item/scale	Bosutinib					
	Not at all	A little	Quite a bit	A lot	Any severity	Moderate-severe $
Skin rash	54.5 (23.4-83.3)	36.4 (10.9-69.2)	0.0 (0.0-28.5)	9.1 (0.2-41.3)	45.5 (16.7-76.6)	9.1 (0.2-41.3)
Skin dry/flaky/ cracked	18.2 (2.3-51.8)	27.3 (6.0-61.0)	54.5 (23.4-83.3)	0.0 (0.0-28.5)	81.8 (48.2-97.7)	54.5 (23.4-83.3)
Skin sore/painful	54.5 (23.4-83.3)	27.3 (6.0-61.0)	18.2 (2.3-51.8)	0.0 (0.0-28.5)	45.5 (16.7-76.6)	18.2 (2.3-51.8)
Wounds healing problems	63.6 (30.8-89.1)	27.3 (6.0-61.0)	0.0 (0.0-28.5)	9.1 (0.2-41.3)	36.4 (10.9-69.2)	9.1 (0.2-41.3)
Hair loss	27.3 (6.0-61.0)	54.5 (23.4-83.3)	9.1 (0.2-41.3)	9.1 (0.2-41.3)	72.7 (39.0-94.0)	18.2 (2.3-51.8)
Nail problems	67.9 (56.4-78.1)	20.5 (12.2-31.2)	9.0 (3.7-17.6)	2.6 (0.3-9.0)	32.1 (21.9-43.6)	11.5 (5.4-20.8)
Eye problems (scale)	9.1 (0.2-41.3)	Na	Na	Na	75.0 (34.9-96.8)	90.9 (58.7-99.8)
Watery eyes	45.5 (16.7-76.6)	36.4 (10.9-69.2)	18.2 (2.3-51.8)	0.0 (0.0-28.5)	54.5 (23.4-83.3)	18.2 (2.3-51.8)
Burning eyes	36.4 (10.9-69.2)	45.5 (16.7-76.6)	18.2 (2.3-51.8)	0.0 (0.0-28.5)	63.6 (30.8-89.1)	18.2 (2.3-51.8)
Light sensitive eyes	54.5 (23.4-83.3)	27.3 (6.0-61.0)	9.1 (0.2-41.3)	9.1 (0.2-41.3)	45.5 (16.7-76.6)	18.2 (2.3-51.8)
Blurry eyes	45.5 (16.7-76.6)	45.5 (16.7-76.6)	0.0 (0.0-28.5)	9.1 (0.2-41.3)	54.5 (23.4-83.3)	9.1 (0.2-41.3)
Red eyes	72.7 (39.0-94.0)	27.3 (6.0-61.0)	0.0 (0.0-28.5)	0.0 (0.0-28.5)	27.3 (6.0-61.0)	0.0 (0.0-28.5)
Swelling face/eyes	63.6 (30.8-89.1)	27.3 (6.0-61.0)	9.1 (0.2-41.3)	0.0 (0.0-28.5)	36.4 (10.9-69.2)	9.1 (0.2-41.3)
Swelling body parts	45.5 (16.7-76.6)	18.2 (2.3-51.8)	36.4 (10.9-69.2)	0.0 (0.0-28.5)	54.5 (23.4-83.3)	36.4 (10.9-69.2)
Hearing problems	60.0 (26.2-87.8)	20.0 (2.5-55.6)	10.0 (0.3-44.5)	10.0 (0.3-44.5)	40.0 (12.2-73.8)	20.0 (2.5-55.6)
Nose bleeds	81.8 (48.2-97.7)	9.1 (0.2-41.3)	9.1 (0.2-41.3)	0.0 (0.0-28.5)	18.2 (2.3-51.8)	9.1 (0.2-41.3)
Nose problems	40.0 (5.3-85.3)	0.0 (0.0-52.2)	40.0 (5.3-85.3)	20.0 (0.5-71.6)	60.0 (14.7-94.7)	60.0 (14.7-94.7)
Dry mouth	45.5 (16.7-76.6)	36.4 (10.9-69.2)	18.2 (2.3-51.8)	0.0 (0.0-28.5)	54.5 (23.4-83.3)	18.2 (2.3-51.8)

Table 18 Continued

Item/scale	Bosutinib					
	Not at all	**A little**	**Quite a bit**	**A lot**	**Any severity**	**Moderate-severe $**
Pain/soreness mouth	72.7 (39.0-94.0)	18.2 (2.3-51.8)	0.0 (0.0-28.5)	9.1 (0.2-41.3)	27.3 (6.0-61.0)	9.1 (0.2-41.3)
Teeth problems	70.0 (34.8-93.3)	0.0 (0.0-30.8)	30.0 (6.7-65.2)	0.0 (0.0-30.8)	30.0 (6.7-65.2)	30.0 (6.7-65.2)
Bleeding gums	70.0 (34.8-93.3)	20.0 (2.5-55.6)	0.0 (0.0-30.8)	10.0 (0.3-44.5)	30.0 (6.7-65.2)	10.0 (0.3-44.5)
Taste change	63.6 (30.8-89.1)	18.2 (2.3-51.8)	9.1 (0.2-41.3)	9.1 (0.2-41.3)	36.4 (10.9-69.2)	18.2 (2.3-51.8)
Appetite loss	63.6 (30.8-89.1)	27.3 (6.0-61.0)	0.0 (0.0-28.5)	9.1 (0.2-41.3)	36.4 (10.9-69.2)	9.1 (0.2-41.3)
Mood problems (scale)	25.0 (5.5-57.2)	Na	Na	Na	75.0 (34.9-96.8)	75.0 (42.8-94.5)
Depressed	45.5 (16.7-76.6)	45.5 (16.7-76.6)	9.1 (0.2-41.3)	0.0 (0.0-28.5)	54.5 (23.4-83.3)	9.1 (0.2-41.3)
Irritable	45.5 (16.7-76.6)	36.4 (10.9-69.2)	18.2 (2.3-51.8)	0.0 (0.0-28.5)	54.5 (23.4-83.3)	18.2 (2.3-51.8)
Tense	60.0 (26.2-87.8)	10.0 (0.3-44.5)	30.0 (6.7-65.2)	0.0 (0.0-30.8)	40.0 (12.2-73.8)	30.0 (6.7-65.2)
Worried	40.0 (12.2-73.8)	30.0 (6.7-65.2)	30.0 (6.7-65.2)	0.0 (0.0-30.8)	60.0 (26.2-87.8)	30.0 (6.7-65.2)
Afraid	22.2 (2.8-60.0)	66.7 (29.9-92.5)	11.1 (0.3-48.2)	0.0 (0.0-33.6)	77.8 (40.0-97.2)	11.1 (0.3-48.2)
Insomnia	58.3 (27.7-84.8)	16.7 (2.1-48.4)	25.0 (5.5-57.2)	0.0 (0.0-26.5)	41.7 (15.2-72.3)	25.0 (5.5-57.2)
Palpitations	54.5 (23.4-83.3)	36.4 (10.9-69.2)	0.0 (0.0-28.5)	9.1 (0.2-41.3)	45.5 (16.7-76.6)	9.1 (0.2-41.3)
Pain in chest	63.6 (30.8-89.1)	27.3 (6.0-61.0)	9.1 (0.2-41.3)	0.0 (0.0-28.5)	36.4 (10.9-69.2)	9.1 (0.2-41.3)
Short of breath	45.5 (16.7-76.6)	27.3 (6.0-61.0)	18.2 (2.3-51.8)	9.1 (0.2-41.3)	54.5 (23.4-83.3)	27.3 (6.0-61.0)
Coughed	45.5 (16.7-76.6)	27.3 (6.0-61.0)	27.3 (6.0-61.0)	0.0 (0.0-28.5)	54.5 (23.4-83.3)	27.3 (6.0-61.0)
Feeling bloated	36.4 (10.9-69.2)	18.2 (2.3-51.8)	27.3 (6.0-61.0)	18.2 (2.3-51.8)	63.6 (30.8-89.1)	45.5 (16.7-76.6)
Indigestion	45.5 (16.7-76.6)	18.2 (2.3-51.8)	9.1 (0.2-41.3)	27.3 (6.0-61.0)	54.5 (23.4-83.3)	36.4 (10.9-69.2)
Nausea	70.0 (34.8-93.3)	10.0 (0.3-44.5)	10.0 (0.3-44.5)	10.0 (0.3-44.5)	30.0 (6.7-65.2)	20.0 (2.5-55.6)

Table 18 Continued

Item/scale	Bosutinib					
	Not at all	A little	Quite a bit	A lot	Any severity	Moderate-severe $
Diarrhoea	54.5 (23.4-83.3)	9.1 (0.2-41.3)	18.2 (2.3-51.8)	18.2 (2.3-51.8)	45.5 (16.7-76.6)	36.4 (10.9-69.2)
Abdominal pain/ cramps	50.0 (18.7-81.3)	0.0 (0.0-30.8)	20.0 (2.5-55.6)	30.0 (6.7-65.2)	50.0 (18.7-81.3)	50.0 (18.7-81.3)
Vomited	72.7 (39.0-94.0)	18.2 (2.3-51.8)	9.1 (0.2-41.3)	0.0 (0.0-28.5)	27.3 (6.0-61.0)	9.1 (0.2-41.3)
Obstipation	60.0 (26.2-87.8)	30.0 (6.7-65.2)	0.0 (0.0-30.8)	10.0 (0.3-44.5)	40.0 (12.2-73.8)	10.0 (0.3-44.5)
Flatulence	0.0 (0.0-30.8)	30.0 (6.7-65.2)	30.0 (6.7-65.2)	40.0 (12.2-73.8)	100.0 (69.2-100.0)	70.0 (34.8-93.3)
Muscle/joint cramps/pain (scale)	18.2 (2.3-51.8)	Na	Na	Na	81.8 (48.2-97.7)	Na
Frequent urination	30.0 (6.7-65.2)	50.0 (18.7-81.3)	20.0 (2.5-55.6)	0.0 (0.0-30.8)	70.0 (34.8-93.3)	20.0 (2.5-55.6)
Muscle cramps	27.3 (6.0-61.0)	36.4 (10.9-69.2)	9.1 (0.2-41.3)	27.3 (6.0-61.0)	72.7 (39.0-94.0)	36.4 (10.9-69.2)
Aches/pains muscles/joints	30.0 (6.7-65.2)	0.0 (0.0-30.8)	30.0 (6.7-65.2)	40.0 (12.2-73.8)	70.0 (34.8-93.3)	70.0 (34.8-93.3)
Muscle weakness	36.4 (10.9-69.2)	36.4 (10.9-69.2)	18.2 (2.3-51.8)	9.1 (0.2-41.3)	63.6 (30.8-89.1)	27.3 (6.0-61.0)
Pale/cold fingers/ toes	60.0 (14.7-94.7)	0.0 (0.0-52.2)	0.0 (0.0-52.2)	40.0 (5.3-85.3)	40.0 (5.3-85.3)	40.0 (5.3-85.3)
Tingling/numbness hands/feet	50.0 (18.7-81.3)	40.0 (12.2-73.8)	0.0 (0.0-30.8)	10.0 (0.3-44.5)	50.0 (18.7-81.3)	10.0 (0.3-44.5)
Problems finger dexterity	70.0 (34.8-93.3)	30.0 (6.7-65.2)	0.0 (0.0-30.8)	0.0 (0.0-30.8)	30.0 (6.7-65.2)	0.0 (0.0-30.8)
Trouble walking pain	27.3 (6.0-61.0)	18.2 (2.3-51.8)	27.3 (6.0-61.0)	27.3 (6.0-61.0)	72.7 (39.0-94.0)	54.5 (23.4-83.3)
Back pain	60.0 (14.7-94.7)	20.0 (0.5-71.6)	20.0 (0.5-71.6)	0.0 (0.0-52.2)	40.0 (5.3-85.3)	20.0 (0.5-71.6)

Table 18 Continued

Item/scale	Bosutinib					
	Not at all	**A little**	**Quite a bit**	**A lot**	**Any severity**	**Moderate-severe $**
Pain	54.5 (23.4-83.3)	36.4 (10.9-69.2)	9.1 (0.2-41.3)	0.0 (0.0-28.5)	45.5 (16.7-76.6)	9.1 (0.2-41.3)
Fever/chills	72.7 (39.0-94.0)	18.2 (2.3-51.8)	9.1 (0.2-41.3)	0.0 (0.0-28.5)	27.3 (6.0-61.0)	9.1 (0.2-41.3)
Sweating excessively	63.6 (30.8-89.1)	9.1 (0.2-41.3)	27.3 (6.0-61.0)	0.0 (0.0-28.5)	36.4 (10.9-69.2)	27.3 (6.0-61.0)
Problems tolerating heat/cold	40.0 (12.2-73.8)	40.0 (12.2-73.8)	20.0 (2.5-55.6)	0.0 (0.0-30.8)	60.0 (26.2-87.8)	20.0 (2.5-55.6)
Hot flushes	60.0 (26.2-87.8)	30.0 (6.7-65.2)	10.0 (0.3-44.5)	0.0 (0.0-30.8)	40.0 (12.2-73.8)	10.0 (0.3-44.5)
Airway infections	100.0 (47.8-100.0)	0.0 (0.0-52.2)	0.0 (0.0-52.2)	0.0 (0.0-52.2)	0.0 (0.0-52.2)	0.0 (0.0-52.2)
Other infections	80.0 (28.4-99.5)	0.0 (0.0-52.2)	20.0 (0.5-71.6)	0.0 (0.0-52.2)	20.0 (0.5-71.6)	20.0 (0.5-71.6)
Sore/enlarged nipples/breasts	100.0 (71.5-100.0)	0.0 (0.0-28.5)	0.0 (0.0-28.5)	0.0 (0.0-28.5)	0.0 (0.0-28.5)	0.0 (0.0-28.5)
Sexuality problems (scale)	30.0 (6.7-65.2)	Na	Na	Na	70.0 (34.8-93.3)	Na
Less interest sex	30.0 (6.7-65.2)	20.0 (2.5-55.6)	30.0 (6.7-65.2)	20.0 (2.5-55.6)	70.0 (34.8-93.3)	50.0 (18.7-81.3)
Less pleasure sex	40.0 (12.2-73.8)	40.0 (12.2-73.8)	10.0 (0.3-44.5)	10.0 (0.3-44.5)	60.0 (26.2-87.8)	20.0 (2.5-55.6)
Pain during/after sex	70.0 (34.8-93.3)	20.0 (2.5-55.6)	10.0 (0.3-44.5)	0.0 (0.0-30.8)	30.0 (6.7-65.2)	10.0 (0.3-44.5)
Difficult erection	50.0 (1.3-98.7)	0.0 (0.0-84.2)	50.0 (1.3-98.7)	0.0 (0.0-84.2)	50.0 (1.3-98.7)	50.0 (1.3-98.7)
Dry vagina	37.5 (8.5-75.5)	37.5 (8.5-75.5)	25.0 (3.2-65.1)	0.0 (0.0-36.9)	62.5 (24.5-91.5)	25.0 (3.2-65.1)

* At T4 as not available at T1. $ Moderate-severe includes 'Quite a bit' and 'A lot'. Abbreviations: Na: not applicable

Table 19 Prevalence (% with 95%confidence intervals) of symptom scores for seven scales plus 76 single items, in 70 patients on dasatinib on T1

Item/scale	Dasatinib					
	Not at all	A little	Quite a bit	A lot	Any severity	Moderate-severe $
Fatigue (scale)	8.6 (3.2-17.7)	Na	Na	Na	91.4 (82.3-96.8)	Na
Tired	11.4 (5.1-21.3)	28.6 (18.4-40.6)	38.6 (27.2-51.0)	21.4 (12.5-32.9)	88.6 (78.7-94.9)	60.0 (47.6-71.5)
Needed rest	17.1 (9.2-28.0)	38.6 (27.2-51.0)	30.0 (19.6-42.1)	14.3 (7.1-24.7)	82.9 (72.0-90.8)	44.3 (32.4-56.7)
Lacked energy	16.4 (8.5-27.5)	29.9 (19.3-42.3)	32.8 (21.8-45.4)	20.9 (11.9-32.6)	83.6 (72.5-91.5)	53.7 (41.1-66.0)
Felt lethargic	19.7 (10.6-31.8)	49.2 (36.1-62.3)	18.0 (9.4-30.0)	13.1 (5.8-24.2)	80.3 (68.2-89.4)	31.1 (19.9-44.3)
Felt weak	29.5 (18.5-42.6)	41.0 (28.6-54.3)	16.4 (8.2-28.1)	13.1 (5.8-24.2)	70.5 (57.4-81.5)	29.5 (18.5-42.6)
Sleepy	21.3 (11.9-33.7)	42.6 (30.0-55.9)	26.2 (15.8-39.1)	9.8 (3.7-20.2)	78.7 (66.3-88.1)	36.1 (24.2-49.4)
Headaches	50.8 (37.5-64.1)	32.2 (20.6-45.6)	11.9 (4.9-22.9)	5.1 (1.1-14.1)	49.2 (35.9-62.5)	16.9 (8.4-29.0)
Dizziness	63.8 (50.1-76.0)	29.3 (18.1-42.7)	5.2 (1.1-14.4)	1.7 (0.0-9.2)	36.2 (24.0-49.9)	6.9 (1.9-16.7)
Cognitive problems (scale)	22.4 (12.5-35.3)	Na	Na	Na	77.6 (64.7-87.5)	Na
Difficulty remembering	31.0 (19.5-44.5)	44.8 (31.7-58.5)	19.0 (9.9-31.4)	5.2 (1.1-14.4)	69.0 (55.5-80.5)	24.1 (13.9-37.2)
Difficulty concentrating	45.6 (32.4-59.3)	24.6 (14.1-37.8)	22.8 (12.7-35.8)	7.0 (1.9-17.0)	54.4 (40.7-67.6)	29.8 (18.4-43.4)
Trouble thinking clearly	37.9 (25.5-51.6)	43.1 (30.2-56.8)	15.5 (7.3-27.4)	3.4 (0.4-11.9)	62.1 (48.4-74.5)	19.0 (9.9-31.4)
Bruised easily	70.2 (56.6-81.6)	22.8 (12.7-35.8)	5.3 (1.1-14.6)	1.8 (0.0-9.4)	29.8 (18.4-43.4)	7.0 (1.9-17.0)
Skin problems (scale)	14.3 (6.4-26.2)	Na	Na	Na	85.7 (73.8-93.6)	Na
Skin colour change	58.2 (44.1-71.3)	21.8 (11.8-35.0)	20.0 (10.4-33.0)	0.0 (0.0-6.5)	41.8 (28.7-55.9)	20.0 (10.4-33.0)
Skin itched	28.6 (17.3-42.2)	35.7 (23.4-49.6)	32.1 (20.3-46.0)	3.6 (0.4-12.3)	71.4 (57.8-82.7)	35.7 (23.4-49.6)

Table 19 Continued

Item/scale	Dasatinib					
	Not at all	A little	Quite a bit	A lot	Any severity	Moderate-severe $
Skin rash	34.5 (22.2-48.6)	40.0 (27.0-54.1)	18.2 (9.1-30.9)	7.3 (2.0-17.6)	65.5 (51.4-77.8)	25.5 (14.7-39.0)
Skin dry/flaky/ cracked	41.1 (28.1-55.0)	32.1 (20.3-46.0)	23.2 (13.0-36.4)	3.6 (0.4-12.3)	58.9 (45.0-71.9)	26.8 (15.8-40.3)
Skin sore/painful	51.8 (38.0-65.3)	32.1 (20.3-46.0)	14.3 (6.4-26.2)	1.8 (0.0-9.6)	48.2 (34.7-62.0)	16.1 (7.6-28.3)
Wounds healing problems	77.2 (64.2-87.3)	14.0 (6.3-25.8)	7.0 (1.9-17.0)	1.8 (0.0-9.4)	22.8 (12.7-35.8)	8.8 (2.9-19.3)
Hair loss	46.4 (33.0-60.3)	28.6 (17.3-42.2)	23.2 (13.0-36.4)	1.8 (0.0-9.6)	53.6 (39.7-67.0)	25.0 (14.4-38.4)
Nail problems	67.9 (56.4-78.1)	20.5 (12.2-31.2)	9.0 (3.7-17.6)	2.6 (0.3-9.0)	32.1 (21.9-43.6)	11.5 (5.4-20.8)
Eye problems (scale)	18.2 (9.1-30.9)	Na	Na	Na	81.8 (69.1-90.9)	Na
Watery eyes	61.1 (46.9-74.1)	14.8 (6.6-27.1)	18.5 (9.3-31.4)	5.6 (1.2-15.4)	38.9 (25.9-53.1)	24.1 (13.5-37.6)
Burning eyes	47.2 (33.3-61.4)	24.5 (13.8-38.3)	26.4 (15.3-40.3)	1.9 (0.0-10.1)	52.8 (38.6-66.7)	28.3 (16.8-42.3)
Light sensitive eyes	46.3 (32.6-60.4)	25.9 (15.0-39.7)	18.5 (9.3-31.4)	9.3 (3.1-20.3)	53.7 (39.6-67.4)	27.8 (16.5-41.6)
Blurry eyes	44.4 (30.9-58.6)	37.0 (24.3-51.3)	16.7 (7.9-29.3)	1.9 (0.0-9.9)	55.6 (41.4-69.1)	18.5 (9.3-31.4)
Red eyes	83.3 (70.7-92.1)	13.0 (5.4-24.9)	3.7 (0.5-12.7)	0.0 (0.0-6.6)	16.7 (7.9-29.3)	3.7 (0.5-12.7)
Swelling face/eyes	65.5 (51.4-77.8)	16.4 (7.8-28.8)	10.9 (4.1-22.2)	7.3 (2.0-17.6)	34.5 (22.2-48.6)	18.2 (9.1-30.9)
Swelling body parts	57.4 (43.2-70.8)	29.6 (18.0-43.6)	11.1 (4.2-22.6)	1.9 (0.0-9.9)	42.6 (29.2-56.8)	13.0 (5.4-24.9)
Hearing problems	80.4 (67.6-89.8)	16.1 (7.6-28.3)	3.6 (0.4-12.3)	0.0 (0.0-6.4)	19.6 (10.2-32.4)	3.6 (0.4-12.3)
Nose bleeds	91.1 (80.4-97.0)	8.9 (3.0-19.6)	0.0 (0.0-6.4)	0.0 (0.0-6.4)	8.9 (3.0-19.6)	0.0 (0.0-6.4)
Nose problems	50.0 (27.2-72.8)	25.0 (8.7-49.1)	25.0 (8.7-49.1)	0.0 (0.0-16.8)	50.0 (27.2-72.8)	25.0 (8.7-49.1)
Dry mouth	60.7 (46.8-73.5)	26.8 (15.8-40.3)	7.1 (2.0-17.3)	5.4 (1.1-14.9)	39.3 (26.5-53.2)	12.5 (5.2-24.1)

Table 19 Continued

Item/scale	Dasatinib					
	Not at all	A little	Quite a bit	A lot	Any severity	Moderate-severe $
Pain/soreness mouth	76.8 (63.6-87.0)	19.6 (10.2-32.4)	1.8 (0.0-9.6)	1.8 (0.0-9.6)	23.2 (13.0-36.4)	3.6 (0.4-12.3)
Teeth problems	80.0 (67.0-89.6)	14.5 (6.5-26.7)	5.5 (1.1-15.1)	0.0 (0.0-6.5)	20.0 (10.4-33.0)	5.5 (1.1-15.1)
Bleeding gums	85.7 (73.8-93.6)	10.7 (4.0-21.9)	3.6 (0.4-12.3)	0.0 (0.0-6.4)	14.3 (6.4-26.2)	3.6 (0.4-12.3)
Taste change	80.0 (67.0-89.6)	12.7 (5.3-24.5)	5.5 (1.1-15.1)	1.8 (0.0-9.7)	20.0 (10.4-33.0)	7.3 (2.0-17.6)
Appetite loss	75.0 (61.6-85.6)	17.9 (8.9-30.4)	7.1 (2.0-17.3)	0.0 (0.0-6.4)	25.0 (14.4-38.4)	7.1 (2.0-17.3)
Mood problems (scale)	15.3 (7.2-27.0)	Na	Na	Na	84.7 (73.0-92.8)	Na
Depressed	52.7 (38.8-66.3)	25.5 (14.7-39.0)	16.4 (7.8-28.8)	5.5 (1.1-15.1)	47.3 (33.7-61.2)	21.8 (11.8-35.0)
Irritable	49.1 (35.4-62.9)	32.7 (20.7-46.7)	16.4 (7.8-28.8)	1.8 (0.0-9.7)	50.9 (37.1-64.6)	18.2 (9.1-30.9)
Tense	49.1 (35.4-62.9)	38.2 (25.4-52.3)	9.1 (3.0-20.0)	3.6 (0.4-12.5)	50.9 (37.1-64.6)	12.7 (5.3-24.5)
Worried	44.6 (31.3-58.5)	37.5 (24.9-51.5)	12.5 (5.2-24.1)	5.4 (1.1-14.9)	55.4 (41.5-68.7)	17.9 (8.9-30.4)
Afraid	78.2 (65.0-88.2)	9.1 (3.0-20.0)	9.1 (3.0-20.0)	3.6 (0.4-12.5)	21.8 (11.8-35.0)	12.7 (5.3-24.5)
Insomnia	44.1 (31.2-57.6)	32.2 (20.6-45.6)	13.6 (6.0-25.0)	10.2 (3.8-20.8)	55.9 (42.4-68.8)	23.7 (13.6-36.6)
Palpitations	53.6 (39.7-67.0)	26.8 (15.8-40.3)	16.1 (7.6-28.3)	3.6 (0.4-12.3)	46.4 (33.0-60.3)	19.6 (10.2-32.4)
Pain in chest	80.4 (67.6-89.8)	17.9 (8.9-30.4)	1.8 (0.0-9.6)	0.0 (0.0-6.4)	19.6 (10.2-32.4)	1.8 (0.0-9.6)
Short of breath	50.0 (36.3-63.7)	39.3 (26.5-53.2)	8.9 (3.0-19.6)	1.8 (0.0-9.6)	50.0 (36.3-63.7)	10.7 (4.0-21.9)
Coughed	73.2 (59.7-84.2)	17.9 (8.9-30.4)	5.4 (1.1-14.9)	3.6 (0.4-12.3)	26.8 (15.8-40.3)	8.9 (3.0-19.6)
Feeling bloated	37.5 (24.9-51.5)	37.5 (24.9-51.5)	19.6 (10.2-32.4)	5.4 (1.1-14.9)	62.5 (48.5-75.1)	25.0 (14.4-38.4)
Indigestion	63.6 (49.6-76.2)	20.0 (10.4-33.0)	10.9 (4.1-22.2)	5.5 (1.1-15.1)	36.4 (23.8-50.4)	16.4 (7.8-28.8)
Nausea	67.9 (54.0-79.7)	19.6 (10.2-32.4)	8.9 (3.0-19.6)	3.6 (0.4-12.3)	32.1 (20.3-46.0)	12.5 (5.2-24.1)

Table 19 Continued

Item/scale	Dasatinib					
	Not at all	**A little**	**Quite a bit**	**A lot**	**Any severity**	**Moderate-severe $**
Diarrhoea	60.7 (46.8-73.5)	32.1 (20.3-46.0)	3.6 (0.4-12.3)	3.6 (0.4-12.3)	39.3 (26.5-53.2)	7.1 (2.0-17.3)
Abdominal pain/ cramps	49.1 (35.4-62.9)	36.4 (23.8-50.4)	5.5 (1.1-15.1)	9.1 (3.0-20.0)	50.9 (37.1-64.6)	14.5 (6.5-26.7)
Vomited	92.9 (82.7-98.0)	5.4 (1.1-14.9)	1.8 (0.0-9.6)	0.0 (0.0-6.4)	7.1 (2.0-17.3)	1.8 (0.0-9.6)
Obstipation	57.1 (43.2-70.3)	26.8 (15.8-40.3)	7.1 (2.0-17.3)	8.9 (3.0-19.6)	42.9 (29.7-56.8)	16.1 (7.6-28.3)
Flatulence	40.0 (27.0-54.1)	25.5 (14.7-39.0)	25.5 (14.7-39.0)	9.1 (3.0-20.0)	60.0 (45.9-73.0)	34.5 (22.2-48.6)
Muscle/joint cramps/pain (scale)	18.2 (9.1-30.9)	Na	Na	Na	81.8 (69.1-90.9)	Na
Frequent urination	40.0 (27.0-54.1)	25.5 (14.7-39.0)	21.8 (11.8-35.0)	12.7 (5.3-24.5)	60.0 (45.9-73.0)	34.5 (22.2-48.6)
Muscle cramps	48.1 (34.3-62.2)	35.2 (22.7-49.4)	14.8 (6.6-27.1)	1.9 (0.0-9.9)	51.9 (37.8-65.7)	16.7 (7.9-29.3)
Aches/pains muscles/joints	23.6 (13.2-37.0)	36.4 (23.8-50.4)	27.3 (16.1-41.0)	12.7 (5.3-24.5)	76.4 (63.0-86.8)	40.0 (27.0-54.1)
Muscle weakness	60.0 (45.9-73.0)	27.3 (16.1-41.0)	9.1 (3.0-20.0)	3.6 (0.4-12.5)	40.0 (27.0-54.1)	12.7 (5.3-24.5)
Pale/cold fingers/ toes	45.0 (23.1-68.5)	10.0 (1.2-31.7)	25.0 (8.7-49.1)	20.0 (5.7-43.7)	55.0 (31.5-76.9)	45.0 (23.1-68.5)
Tingling/numbness hands/feet	41.8 (28.7-55.9)	25.5 (14.7-39.0)	18.2 (9.1-30.9)	14.5 (6.5-26.7)	58.2 (44.1-71.3)	32.7 (20.7-46.7)
Problems finger dexterity	75.9 (62.4-86.5)	14.8 (6.6-27.1)	7.4 (2.1-17.9)	1.9 (0.0-9.9)	24.1 (13.5-37.6)	9.3 (3.1-20.3)
Trouble walking pain	67.3 (53.3-79.3)	18.2 (9.1-30.9)	10.9 (4.1-22.2)	3.6 (0.4-12.5)	32.7 (20.7-46.7)	14.5 (6.5-26.7)
Back pain	30.0 (11.9-54.3)	45.0 (23.1-68.5)	25.0 (8.7-49.1)	0.0 (0.0-16.8)	70.0 (45.7-88.1)	25.0 (8.7-49.1)

Table 19 Continued

Item/scale	Dasatinib					
	Not at all	A little	Quite a bit	A lot	Any severity	Moderate-severe $
Pain	45.5 (32.0-59.4)	36.4 (23.8-50.4)	14.5 (6.5-26.7)	3.6 (0.4-12.5)	54.5 (40.6-68.0)	18.2 (9.1-30.9)
Fever/chills	72.2 (58.4-83.5)	20.4 (10.6-33.5)	7.4 (2.1-17.9)	0.0 (0.0-6.6)	27.8 (16.5-41.6)	7.4 (2.1-17.9)
Sweating excessively	67.3 (53.3-79.3)	21.8 (11.8-35.0)	10.9 (4.1-22.2)	0.0 (0.0-6.5)	32.7 (20.7-46.7)	10.9 (4.1-22.2)
Problems tolerating heat/cold	50.9 (37.1-64.6)	36.4 (23.8-50.4)	9.1 (3.0-20.0)	3.6 (0.4-12.5)	49.1 (35.4-62.9)	12.7 (5.3-24.5)
Hot flushes	65.5 (51.4-77.8)	21.8 (11.8-35.0)	9.1 (3.0-20.0)	3.6 (0.4-12.5)	34.5 (22.2-48.6)	12.7 (5.3-24.5)
Airway infections	85.0 (62.1-96.8)	10.0 (1.2-31.7)	5.0 (0.1-24.9)	0.0 (0.0-16.8)	15.0 (3.2-37.9)	5.0 (0.1-24.9)
Other infections	85.0 (62.1-96.8)	10.0 (1.2-31.7)	0.0 (0.0-16.8)	5.0 (0.1-24.9)	15.0 (3.2-37.9)	5.0 (0.1-24.9)
Sore/enlarged nipples/breasts	92.6 (82.1-97.9)	7.4 (2.1-17.9)	0.0 (0.0-6.6)	0.0 (0.0-6.6)	7.4 (2.1-17.9)	0.0 (0.0-6.6)
Sexuality problems (scale)	24.5 (13.8-38.3)	Na	Na	Na	75.5 (61.7-86.2)	Na
Less interest sex	32.1 (19.9-46.3)	32.1 (19.9-46.3)	22.6 (12.3-36.2)	13.2 (5.5-25.3)	67.9 (53.7-80.1)	35.8 (23.1-50.2)
Less pleasure sex	39.2 (25.8-53.9)	37.3 (24.1-51.9)	17.6 (8.4-30.9)	5.9 (1.2-16.2)	60.8 (46.1-74.2)	23.5 (12.8-37.5)
Pain during/after sex	76.5 (62.5-87.2)	15.7 (7.0-28.6)	2.0 (0.0-10.4)	5.9 (1.2-16.2)	23.5 (12.8-37.5)	7.8 (2.2-18.9)
Difficult erection	61.9 (38.4-81.9)	28.6 (11.3-52.2)	0.0 (0.0-16.1)	9.5 (1.2-30.4)	38.1 (18.1-61.6)	9.5 (1.2-30.4)
Dry vagina	51.9 (31.9-71.3)	14.8 (4.2-33.7)	22.2 (8.6-42.3)	11.1 (2.4-29.2)	48.1 (28.7-68.1)	33.3 (16.5-54.0)

* At T4 as not available at T1. $ Moderate-severe includes 'Quite a bit' and 'A lot'. Abbreviations: Na: not applicable

Table 20 Prevalence (% with 95%confidence intervals) of scores for seven scales plus 76 single items, in 84 patients on imatinib on T1

Item/scale	Imatinib					
	Not at all	A little	Quite a bit	A lot	Any severity	Moderate-severe $
Fatigue (scale)	10.7 (5.0-19.4)	Na	Na	Na	89.3 (80.6-95.0)	Na
Tired	13.1 (6.7-22.2)	38.1 (27.7-49.3)	32.1 (22.4-43.2)	16.7 (9.4-26.4)	86.9 (77.8-93.3)	48.8 (37.7-60.0)
Needed rest	23.8 (15.2-34.3)	35.7 (25.6-46.9)	34.5 (24.5-45.7)	6.0 (2.0-13.3)	76.2 (65.7-84.8)	40.5 (29.9-51.7)
Lacked energy	20.7 (12.6-31.1)	31.7 (21.9-42.9)	36.6 (26.2-48.0)	11.0 (5.1-19.8)	79.3 (68.9-87.4)	47.6 (36.4-58.9)
Felt lethargic	30.4 (20.5-41.8)	35.4 (25.0-47.0)	30.4 (20.5-41.8)	3.8 (0.8-10.7)	69.6 (58.2-79.5)	34.2 (23.9-45.7)
Felt weak	39.5 (28.4-51.4)	36.8 (26.1-48.7)	23.7 (14.7-34.8)	0.0 (0.0-4.7)	60.5 (48.6-71.6)	23.7 (14.7-34.8)
Sleepy	28.9 (19.1-40.5)	44.7 (33.3-56.6)	22.4 (13.6-33.4)	3.9 (0.8-11.1)	71.1 (59.5-80.9)	26.3 (16.9-37.7)
Headaches	70.3 (58.5-80.3)	21.6 (12.9-32.7)	8.1 (3.0-16.8)	0.0 (0.0-4.9)	29.7 (19.7-41.5)	8.1 (3.0-16.8)
Dizziness	68.9 (57.1-79.2)	28.4 (18.5-40.1)	2.7 (0.3-9.4)	0.0 (0.0-4.9)	31.1 (20.8-42.9)	2.7 (0.3-9.4)
Cognitive problems (scale)	33.8 (23.2-45.7)	Na	Na	Na	66.2 (54.3-76.8)	Na
Difficulty remembering	45.9 (34.3-57.9)	44.6 (33.0-56.6)	8.1 (3.0-16.8)	1.4 (0.0-7.3)	54.1 (42.1-65.7)	9.5 (3.9-18.5)
Difficulty concentrating	50.0 (38.1-61.9)	33.8 (23.2-45.7)	14.9 (7.7-25.0)	1.4 (0.0-7.3)	50.0 (38.1-61.9)	16.2 (8.7-26.6)
Trouble thinking clearly	54.2 (42.0-66.0)	37.5 (26.4-49.7)	6.9 (2.3-15.5)	1.4 (0.0-7.5)	45.8 (34.0-58.0)	8.3 (3.1-17.3)
Bruised easily	51.4 (39.3-63.3)	33.3 (22.7-45.4)	13.9 (6.9-24.1)	1.4 (0.0-7.5)	48.6 (36.7-60.7)	15.3 (7.9-25.7)
Skin problems (scale)	21.1 (12.3-32.4)	Na	Na	Na	78.9 (67.6-87.7)	Na
Skin colour change	61.4 (49.0-72.8)	30.0 (19.6-42.1)	5.7 (1.6-14.0)	2.9 (0.3-9.9)	38.6 (27.2-51.0)	8.6 (3.2-17.7)
Skin itched	46.5 (34.5-58.7)	43.7 (31.9-56.0)	5.6 (1.6-13.8)	4.2 (0.9-11.9)	53.5 (41.3-65.5)	9.9 (4.1-19.3)

Table 20 Continued

Item/scale	Imatinib					
	Not at all	A little	Quite a bit	A lot	Any severity	Moderate-severe $
Skin rash	64.8 (52.5-75.8)	22.5 (13.5-34.0)	11.3 (5.0-21.0)	1.4 (0.0-7.6)	35.2 (24.2-47.5)	12.7 (6.0-22.7)
Skin dry/flaky/ cracked	48.6 (36.4-60.8)	38.6 (27.2-51.0)	11.4 (5.1-21.3)	1.4 (0.0-7.7)	51.4 (39.2-63.6)	12.9 (6.1-23.0)
Skin sore/painful	50.7 (38.6-62.8)	39.4 (28.0-51.7)	9.9 (4.1-19.3)	0.0 (0.0-5.1)	49.3 (37.2-61.4)	9.9 (4.1-19.3)
Wounds healing problems	73.6 (61.9-83.3)	23.6 (14.4-35.1)	1.4 (0.0-7.5)	1.4 (0.0-7.5)	26.4 (16.7-38.1)	2.8 (0.3-9.7)
Hair loss	72.2 (60.4-82.1)	22.2 (13.3-33.6)	4.2 (0.9-11.7)	1.4 (0.0-7.5)	27.8 (17.9-39.6)	5.6 (1.5-13.6)
Nail problems	67.9 (56.4-78.1)	20.5 (12.2-31.2)	9.0 (3.7-17.6)	2.6 (0.3-9.0)	32.1 (21.9-43.6)	11.5 (5.4-20.8)
Eye problems (scale)	11.3 (5.0-21.0)	Na	Na	Na	88.7 (79.0-95.0)	Na
Watery eyes	35.2 (24.2-47.5)	29.6 (19.3-41.6)	28.2 (18.1-40.1)	7.0 (2.3-15.7)	64.8 (52.5-75.8)	35.2 (24.2-47.5)
Burning eyes	35.2 (24.2-47.5)	46.5 (34.5-58.7)	16.9 (9.0-27.7)	1.4 (0.0-7.6)	64.8 (52.5-75.8)	18.3 (10.1-29.3)
Light sensitive eyes	35.7 (24.6-48.1)	45.7 (33.7-58.1)	14.3 (7.1-24.7)	4.3 (0.9-12.0)	64.3 (51.9-75.4)	18.6 (10.3-29.7)
Blurry eyes	40.0 (28.5-52.4)	44.3 (32.4-56.7)	15.7 (8.1-26.4)	0.0 (0.0-5.1)	60.0 (47.6-71.5)	15.7 (8.1-26.4)
Red eyes	60.0 (47.6-71.5)	25.7 (16.0-37.6)	11.4 (5.1-21.3)	2.9 (0.3-9.9)	40.0 (28.5-52.4)	14.3 (7.1-24.7)
Swelling face/eyes	30.6 (20.2-42.5)	30.6 (20.2-42.5)	30.6 (20.2-42.5)	8.3 (3.1-17.3)	69.4 (57.5-79.8)	38.9 (27.6-51.1)
Swelling body parts	52.9 (40.4-65.2)	30.9 (20.2-43.3)	13.2 (6.2-23.6)	2.9 (0.4-10.2)	47.1 (34.8-59.6)	16.2 (8.4-27.1)
Hearing problems	70.8 (58.9-81.0)	19.4 (11.1-30.5)	6.9 (2.3-15.5)	2.8 (0.3-9.7)	29.2 (19.0-41.1)	9.7 (4.0-19.0)
Nose bleeds	90.1 (80.7-95.9)	9.9 (4.1-19.3)	0.0 (0.0-5.1)	0.0 (0.0-5.1)	9.9 (4.1-19.3)	0.0 (0.0-5.1)
Nose problems	54.2 (32.8-74.4)	29.2 (12.6-51.1)	16.7 (4.7-37.4)	0.0 (0.0-14.2)	45.8 (25.6-67.2)	16.7 (4.7-37.4)
Dry mouth	52.9 (40.6-64.9)	25.7 (16.0-37.6)	14.3 (7.1-24.7)	7.1 (2.4-15.9)	47.1 (35.1-59.4)	21.4 (12.5-32.9)

Table 20 Continued

Item/scale	Imatinib					
	Not at all	**A little**	**Quite a bit**	**A lot**	**Any severity**	**Moderate-severe $**
Pain/soreness mouth	76.1 (64.5-85.4)	19.7 (11.2-30.9)	2.8 (0.3-9.8)	1.4 (0.0-7.6)	23.9 (14.6-35.5)	4.2 (0.9-11.9)
Teeth problems	70.4 (58.4-80.7)	22.5 (13.5-34.0)	4.2 (0.9-11.9)	2.8 (0.3-9.8)	29.6 (19.3-41.6)	7.0 (2.3-15.7)
Bleeding gums	73.2 (61.4-83.1)	23.9 (14.6-35.5)	2.8 (0.3-9.8)	0.0 (0.0-5.1)	26.8 (16.9-38.6)	2.8 (0.3-9.8)
Taste change	70.4 (58.4-80.7)	18.3 (10.1-29.3)	5.6 (1.6-13.8)	5.6 (1.6-13.8)	29.6 (19.3-41.6)	11.3 (5.0-21.0)
Appetite loss	67.1 (54.9-77.9)	18.6 (10.3-29.7)	10.0 (4.1-19.5)	4.3 (0.9-12.0)	32.9 (22.1-45.1)	14.3 (7.1-24.7)
Mood problems (scale)	21.3 (12.7-32.3)	Na	Na	Na	78.7 (67.7-87.3)	Na
Depressed	58.6 (46.2-70.2)	24.3 (14.8-36.0)	11.4 (5.1-21.3)	5.7 (1.6-14.0)	41.4 (29.8-53.8)	17.1 (9.2-28.0)
Irritable	55.7 (43.3-67.6)	27.1 (17.2-39.1)	11.4 (5.1-21.3)	5.7 (1.6-14.0)	44.3 (32.4-56.7)	17.1 (9.2-28.0)
Tense	52.9 (40.6-64.9)	28.6 (18.4-40.6)	11.4 (5.1-21.3)	7.1 (2.4-15.9)	47.1 (35.1-59.4)	18.6 (10.3-29.7)
Worried	40.0 (28.5-52.4)	44.3 (32.4-56.7)	12.9 (6.1-23.0)	2.9 (0.3-9.9)	60.0 (47.6-71.5)	15.7 (8.1-26.4)
Afraid	71.4 (59.4-81.6)	18.6 (10.3-29.7)	7.1 (2.4-15.9)	2.9 (0.3-9.9)	28.6 (18.4-40.6)	10.0 (4.1-19.5)
Insomnia	40.5 (29.3-52.6)	28.4 (18.5-40.1)	21.6 (12.9-32.7)	9.5 (3.9-18.5)	59.5 (47.4-70.7)	31.1 (20.8-42.9)
Palpitations	54.3 (41.9-66.3)	31.4 (20.9-43.6)	14.3 (7.1-24.7)	0.0 (0.0-5.1)	45.7 (33.7-58.1)	14.3 (7.1-24.7)
Pain in chest	79.4 (67.9-88.3)	17.6 (9.5-28.8)	1.5 (0.0-7.9)	1.5 (0.0-7.9)	20.6 (11.7-32.1)	2.9 (0.4-10.2)
Short of breath	44.3 (32.4-56.7)	32.9 (22.1-45.1)	20.0 (11.4-31.3)	2.9 (0.3-9.9)	55.7 (43.3-67.6)	22.9 (13.7-34.4)
Coughed	69.1 (56.7-79.8)	23.5 (14.1-35.4)	5.9 (1.6-14.4)	1.5 (0.0-7.9)	30.9 (20.2-43.3)	7.4 (2.4-16.3)
Feeling bloated	54.3 (41.9-66.3)	27.1 (17.2-39.1)	15.7 (8.1-26.4)	2.9 (0.3-9.9)	45.7 (33.7-58.1)	18.6 (10.3-29.7)
Indigestion	58.0 (45.5-69.8)	23.2 (13.9-34.9)	14.5 (7.2-25.0)	4.3 (0.9-12.2)	42.0 (30.2-54.5)	18.8 (10.4-30.1)
Nausea	48.5 (36.2-61.0)	42.6 (30.7-55.2)	8.8 (3.3-18.2)	0.0 (0.0-5.3)	51.5 (39.0-63.8)	8.8 (3.3-18.2)

Table 20 Continued

Item/scale	Imatinib					
	Not at all	A little	Quite a bit	A lot	Any severity	Moderate-severe $
Diarrhoea	47.8 (35.6-60.2)	24.6 (15.1-36.5)	14.5 (7.2-25.0)	13.0 (6.1-23.3)	52.2 (39.8-64.4)	27.5 (17.5-39.6)
Abdominal pain/ cramps	49.3 (37.0-61.6)	36.2 (25.0-48.7)	14.5 (7.2-25.0)	0.0 (0.0-5.2)	50.7 (38.4-63.0)	14.5 (7.2-25.0)
Vomited	78.3 (66.7-87.3)	20.3 (11.6-31.7)	1.4 (0.0-7.8)	0.0 (0.0-5.2)	21.7 (12.7-33.3)	1.4 (0.0-7.8)
Obstipation	78.3 (66.7-87.3)	17.4 (9.3-28.4)	4.3 (0.9-12.2)	0.0 (0.0-5.2)	21.7 (12.7-33.3)	4.3 (0.9-12.2)
Flatulence	27.9 (17.7-40.1)	41.2 (29.4-53.8)	23.5 (14.1-35.4)	7.4 (2.4-16.3)	72.1 (59.9-82.3)	30.9 (20.2-43.3)
Muscle/joint cramps/pain (scale)	10.3 (4.2-20.1)	Na	Na	Na	89.7 (79.9-95.8)	Na
Frequent urination	42.6 (30.7-55.2)	29.4 (19.0-41.7)	23.5 (14.1-35.4)	4.4 (0.9-12.4)	57.4 (44.8-69.3)	27.9 (17.7-40.1)
Muscle cramps	25.0 (15.3-37.0)	35.3 (24.1-47.8)	23.5 (14.1-35.4)	16.2 (8.4-27.1)	75.0 (63.0-84.7)	39.7 (28.0-52.3)
Aches/pains muscles/joints	23.5 (14.1-35.4)	30.9 (20.2-43.3)	33.8 (22.8-46.3)	11.8 (5.2-21.9)	76.5 (64.6-85.9)	45.6 (33.5-58.1)
Muscle weakness	52.2 (39.7-64.6)	28.4 (18.0-40.7)	19.4 (10.8-30.9)	0.0 (0.0-5.4)	47.8 (35.4-60.3)	19.4 (10.8-30.9)
Pale/cold fingers/ toes	54.2 (32.8-74.4)	20.8 (7.1-42.2)	16.7 (4.7-37.4)	8.3 (1.0-27.0)	45.8 (25.6-67.2)	25.0 (9.8-46.7)
Tingling/numbness hands/feet	45.6 (33.5-58.1)	33.8 (22.8-46.3)	16.2 (8.4-27.1)	4.4 (0.9-12.4)	54.4 (41.9-66.5)	20.6 (11.7-32.1)
Problems finger dexterity	67.6 (55.2-78.5)	25.0 (15.3-37.0)	4.4 (0.9-12.4)	2.9 (0.4-10.2)	32.4 (21.5-44.8)	7.4 (2.4-16.3)
Trouble walking pain	60.3 (47.7-72.0)	23.5 (14.1-35.4)	7.4 (2.4-16.3)	8.8 (3.3-18.2)	39.7 (28.0-52.3)	16.2 (8.4-27.1)
Back pain	58.3 (36.6-77.9)	37.5 (18.8-59.4)	0.0 (0.0-14.2)	4.2 (0.1-21.1)	41.7 (22.1-63.4)	4.2 (0.1-21.1)

Table 20 Continued

Item/scale	Imatinib					
	Not at all	**A little**	**Quite a bit**	**A lot**	**Any severity**	**Moderate-severe $**
Pain	47.1 (34.8-59.6)	38.2 (26.7-50.8)	10.3 (4.2-20.1)	4.4 (0.9-12.4)	52.9 (40.4-65.2)	14.7 (7.3-25.4)
Fever/chills	77.9 (66.2-87.1)	19.1 (10.6-30.5)	1.5 (0.0-7.9)	1.5 (0.0-7.9)	22.1 (12.9-33.8)	2.9 (0.4-10.2)
Sweating excessively	61.8 (49.2-73.3)	26.5 (16.5-38.6)	10.3 (4.2-20.1)	1.5 (0.0-7.9)	38.2 (26.7-50.8)	11.8 (5.2-21.9)
Problems tolerating heat/cold	52.9 (40.4-65.2)	33.8 (22.8-46.3)	8.8 (3.3-18.2)	4.4 (0.9-12.4)	47.1 (34.8-59.6)	13.2 (6.2-23.6)
Hot flushes	82.4 (71.2-90.5)	14.7 (7.3-25.4)	1.5 (0.0-7.9)	1.5 (0.0-7.9)	17.6 (9.5-28.8)	2.9 (0.4-10.2)
Airway infections	66.7 (44.7-84.4)	16.7 (4.7-37.4)	4.2 (0.1-21.1)	12.5 (2.7-32.4)	33.3 (15.6-55.3)	16.7 (4.7-37.4)
Other infections	87.5 (67.6-97.3)	4.2 (0.1-21.1)	4.2 (0.1-21.1)	4.2 (0.1-21.1)	12.5 (2.7-32.4)	8.3 (1.0-27.0)
Sore/enlarged nipples/breasts	86.6 (76.0-93.7)	13.4 (6.3-24.0)	0.0 (0.0-5.4)	0.0 (0.0-5.4)	13.4 (6.3-24.0)	0.0 (0.0-5.4)
Sexuality problems (scale)	40.0 (28.0-52.9)	Na	Na	Na	60.0 (47.1-72.0)	Na
Less interest sex	43.1 (30.8-56.0)	24.6 (14.8-36.9)	20.0 (11.1-31.8)	12.3 (5.5-22.8)	56.9 (44.0-69.2)	32.3 (21.2-45.1)
Less pleasure sex	53.1 (40.2-65.7)	17.2 (8.9-28.7)	20.3 (11.3-32.2)	9.4 (3.5-19.3)	46.9 (34.3-59.8)	29.7 (18.9-42.4)
Pain during/after sex	77.8 (65.5-87.3)	15.9 (7.9-27.3)	3.2 (0.4-11.0)	3.2 (0.4-11.0)	22.2 (12.7-34.5)	6.3 (1.8-15.5)
Difficult erection	61.8 (43.6-77.8)	23.5 (10.7-41.2)	11.8 (3.3-27.5)	2.9 (0.1-15.3)	38.2 (22.2-56.4)	14.7 (5.0-31.1)
Dry vagina	55.2 (35.7-73.6)	20.7 (8.0-39.7)	20.7 (8.0-39.7)	3.4 (0.1-17.8)	44.8 (26.4-64.3)	24.1 (10.3-43.5)

* At T4 as not available at T1. $ Moderate-severe includes 'Quite a bit' and 'A lot'. Abbreviations: Na: not applicable

Table 21 Prevalence (% with 95%confidence intervals) of symptom scores for seven scales plus 76 single items, in 39 patients on nilotinib on T1

Item/scale	Nilotinib					
	Not at all	A little	Quite a bit	A lot	Any severity	Moderate-severe $
Fatigue (scale)	5.1 (0.6-17.3)	Na	Na	Na	94.9 (82.7-99.4)	Na
Tired	12.8 (4.3-27.4)	41.0 (25.6-57.9)	38.5 (23.4-55.4)	7.7 (1.6-20.9)	87.2 (72.6-95.7)	46.2 (30.1-62.8)
Needed rest	12.8 (4.3-27.4)	38.5 (23.4-55.4)	43.6 (27.8-60.4)	5.1 (0.6-17.3)	87.2 (72.6-95.7)	48.7 (32.4-65.2)
Lacked energy	17.9 (7.5-33.5)	38.5 (23.4-55.4)	35.9 (21.2-52.8)	7.7 (1.6-20.9)	82.1 (66.5-92.5)	43.6 (27.8-60.4)
Felt lethargic	43.2 (27.1-60.5)	35.1 (20.2-52.5)	21.6 (9.8-38.2)	0.0 (0.0-9.5)	56.8 (39.5-72.9)	21.6 (9.8-38.2)
Felt weak	36.8 (21.8-54.0)	44.7 (28.6-61.7)	18.4 (7.7-34.3)	0.0 (0.0-9.3)	63.2 (46.0-78.2)	18.4 (7.7-34.3)
Sleepy	21.6 (9.8-38.2)	40.5 (24.8-57.9)	37.8 (22.5-55.2)	0.0 (0.0-9.5)	78.4 (61.8-90.2)	37.8 (22.5-55.2)
Headaches	80.6 (64.0-91.8)	16.7 (6.4-32.8)	2.8 (0.1-14.5)	0.0 (0.0-9.7)	19.4 (8.2-36.0)	2.8 (0.1-14.5)
Dizziness	69.4 (51.9-83.7)	25.0 (12.1-42.2)	5.6 (0.7-18.7)	0.0 (0.0-9.7)	30.6 (16.3-48.1)	5.6 (0.7-18.7)
Cognitive problems (scale)	40.0 (23.9-57.9)	Na	Na	Na	60.0 (42.1-76.1)	Na
Difficulty remembering	48.6 (31.4-66.0)	45.7 (28.8-63.4)	5.7 (0.7-19.2)	0.0 (0.0-10.0)	51.4 (34.0-68.6)	5.7 (0.7-19.2)
Difficulty concentrating	57.1 (39.4-73.7)	37.1 (21.5-55.1)	5.7 (0.7-19.2)	0.0 (0.0-10.0)	42.9 (26.3-60.6)	5.7 (0.7-19.2)
Trouble thinking clearly	60.0 (42.1-76.1)	37.1 (21.5-55.1)	2.9 (0.1-14.9)	0.0 (0.0-10.0)	40.0 (23.9-57.9)	2.9 (0.1-14.9)
Bruised easily	71.4 (53.7-85.4)	28.6 (14.6-46.3)	0.0 (0.0-10.0)	0.0 (0.0-10.0)	28.6 (14.6-46.3)	0.0 (0.0-10.0)
Skin problems (scale)	17.1 (6.6-33.6)	Na	Na	Na	82.9 (66.4-93.4)	Na
Skin colour change	77.1 (59.9-89.6)	20.0 (8.4-36.9)	2.9 (0.1-14.9)	0.0 (0.0-10.0)	22.9 (10.4-40.1)	2.9 (0.1-14.9)
Skin itched	37.1 (21.5-55.1)	22.9 (10.4-40.1)	20.0 (8.4-36.9)	20.0 (8.4-36.9)	62.9 (44.9-78.5)	40.0 (23.9-57.9)

Table 21 Continued

Item/scale	Nilotinib					
	Not at all	A little	Quite a bit	A lot	Any severity	Moderate-severe $
Skin rash	50.0 (32.4-67.6)	29.4 (15.1-47.5)	17.6 (6.8-34.5)	2.9 (0.1-15.3)	50.0 (32.4-67.6)	20.6 (8.7-37.9)
Skin dry/flaky/ cracked	34.3 (19.1-52.2)	34.3 (19.1-52.2)	22.9 (10.4-40.1)	8.6 (1.8-23.1)	65.7 (47.8-80.9)	31.4 (16.9-49.3)
Skin sore/painful	74.3 (56.7-87.5)	17.1 (6.6-33.6)	5.7 (0.7-19.2)	2.9 (0.1-14.9)	25.7 (12.5-43.3)	8.6 (1.8-23.1)
Wounds healing problems	97.1 (85.1-99.9)	2.9 (0.1-14.9)	0.0 (0.0-10.0)	0.0 (0.0-10.0)	2.9 (0.1-14.9)	0.0 (0.0-10.0)
Hair loss	50.0 (32.4-67.6)	35.3 (19.7-53.5)	5.9 (0.7-19.7)	8.8 (1.9-23.7)	50.0 (32.4-67.6)	14.7 (5.0-31.1)
Nail problems	67.9 (56.4-78.1)	20.5 (12.2-31.2)	9.0 (3.7-17.6)	2.6 (0.3-9.0)	32.1 (21.9-43.6)	11.5 (5.4-20.8)
Eye problems (scale)	8.8 (1.9-23.7)	Na	Na	Na	91.2 (76.3-98.1)	Na
Watery eyes	45.5 (28.1-63.6)	45.5 (28.1-63.6)	9.1 (1.9-24.3)	0.0 (0.0-10.6)	54.5 (36.4-71.9)	9.1 (1.9-24.3)
Burning eyes	42.4 (25.5-60.8)	45.5 (28.1-63.6)	9.1 (1.9-24.3)	3.0 (0.1-15.8)	57.6 (39.2-74.5)	12.1 (3.4-28.2)
Light sensitive eyes	40.6 (23.7-59.4)	31.2 (16.1-50.0)	25.0 (11.5-43.4)	3.1 (0.1-16.2)	59.4 (40.6-76.3)	28.1 (13.7-46.7)
Blurry eyes	42.4 (25.5-60.8)	45.5 (28.1-63.6)	9.1 (1.9-24.3)	3.0 (0.1-15.8)	57.6 (39.2-74.5)	12.1 (3.4-28.2)
Red eyes	84.8 (68.1-94.9)	9.1 (1.9-24.3)	6.1 (0.7-20.2)	0.0 (0.0-10.6)	15.2 (5.1-31.9)	6.1 (0.7-20.2)
Swelling face/eyes	73.5 (55.6-87.1)	17.6 (6.8-34.5)	8.8 (1.9-23.7)	0.0 (0.0-10.3)	26.5 (12.9-44.4)	8.8 (1.9-23.7)
Swelling body parts	84.4 (67.2-94.7)	9.4 (2.0-25.0)	6.2 (0.8-20.8)	0.0 (0.0-10.9)	15.6 (5.3-32.8)	6.2 (0.8-20.8)
Hearing problems	81.8 (64.5-93.0)	12.1 (3.4-28.2)	3.0 (0.1-15.8)	3.0 (0.1-15.8)	18.2 (7.0-35.5)	6.1 (0.7-20.2)
Nose bleeds	84.8 (68.1-94.9)	12.1 (3.4-28.2)	0.0 (0.0-10.6)	3.0 (0.1-15.8)	15.2 (5.1-31.9)	3.0 (0.1-15.8)
Nose problems	64.3 (35.1-87.2)	14.3 (1.8-42.8)	21.4 (4.7-50.8)	0.0 (0.0-23.2)	35.7 (12.8-64.9)	21.4 (4.7-50.8)
Dry mouth	63.6 (45.1-79.6)	24.2 (11.1-42.3)	6.1 (0.7-20.2)	6.1 (0.7-20.2)	36.4 (20.4-54.9)	12.1 (3.4-28.2)

Table 21 Continued

Item/scale	Nilotinib					
	Not at all	A little	Quite a bit	A lot	Any severity	Moderate-severe $
Pain/soreness mouth	87.9 (71.8-96.6)	3.0 (0.1-15.8)	9.1 (1.9-24.3)	0.0 (0.0-10.6)	12.1 (3.4-28.2)	9.1 (1.9-24.3)
Teeth problems	90.9 (75.7-98.1)	6.1 (0.7-20.2)	3.0 (0.1-15.8)	0.0 (0.0-10.6)	9.1 (1.9-24.3)	3.0 (0.1-15.8)
Bleeding gums	84.4 (67.2-94.7)	15.6 (5.3-32.8)	0.0 (0.0-10.9)	0.0 (0.0-10.9)	15.6 (5.3-32.8)	0.0 (0.0-10.9)
Taste change	81.8 (64.5-93.0)	15.2 (5.1-31.9)	3.0 (0.1-15.8)	0.0 (0.0-10.6)	18.2 (7.0-35.5)	3.0 (0.1-15.8)
Appetite loss	84.8 (68.1-94.9)	9.1 (1.9-24.3)	6.1 (0.7-20.2)	0.0 (0.0-10.6)	15.2 (5.1-31.9)	6.1 (0.7-20.2)
Mood problems (scale)	21.6 (9.8-38.2)	Na	Na	Na	78.4 (61.8-90.2)	Na
Depressed	75.8 (57.7-88.9)	21.2 (9.0-38.9)	3.0 (0.1-15.8)	0.0 (0.0-10.6)	24.2 (11.1-42.3)	3.0 (0.1-15.8)
Irritable	60.6 (42.1-77.1)	33.3 (18.0-51.8)	6.1 (0.7-20.2)	0.0 (0.0-10.6)	39.4 (22.9-57.9)	6.1 (0.7-20.2)
Tense	63.6 (45.1-79.6)	33.3 (18.0-51.8)	3.0 (0.1-15.8)	0.0 (0.0-10.6)	36.4 (20.4-54.9)	3.0 (0.1-15.8)
Worried	48.5 (30.8-66.5)	45.5 (28.1-63.6)	6.1 (0.7-20.2)	0.0 (0.0-10.6)	51.5 (33.5-69.2)	6.1 (0.7-20.2)
Afraid	81.2 (63.6-92.8)	18.8 (7.2-36.4)	0.0 (0.0-10.9)	0.0 (0.0-10.9)	18.8 (7.2-36.4)	0.0 (0.0-10.9)
Insomnia	43.2 (27.1-60.5)	29.7 (15.9-47.0)	24.3 (11.8-41.2)	2.7 (0.1-14.2)	56.8 (39.5-72.9)	27.0 (13.8-44.1)
Palpitations	69.7 (51.3-84.4)	18.2 (7.0-35.5)	9.1 (1.9-24.3)	3.0 (0.1-15.8)	30.3 (15.6-48.7)	12.1 (3.4-28.2)
Pain in chest	87.9 (71.8-96.6)	6.1 (0.7-20.2)	6.1 (0.7-20.2)	0.0 (0.0-10.6)	12.1 (3.4-28.2)	6.1 (0.7-20.2)
Short of breath	66.7 (48.2-82.0)	24.2 (11.1-42.3)	9.1 (1.9-24.3)	0.0 (0.0-10.6)	33.3 (18.0-51.8)	9.1 (1.9-24.3)
Coughed	63.6 (45.1-79.6)	24.2 (11.1-42.3)	12.1 (3.4-28.2)	0.0 (0.0-10.6)	36.4 (20.4-54.9)	12.1 (3.4-28.2)
Feeling bloated	66.7 (48.2-82.0)	24.2 (11.1-42.3)	9.1 (1.9-24.3)	0.0 (0.0-10.6)	33.3 (18.0-51.8)	9.1 (1.9-24.3)
Indigestion	75.8 (57.7-88.9)	15.2 (5.1-31.9)	9.1 (1.9-24.3)	0.0 (0.0-10.6)	24.2 (11.1-42.3)	9.1 (1.9-24.3)
Nausea	75.0 (56.6-88.5)	18.8 (7.2-36.4)	6.2 (0.8-20.8)	0.0 (0.0-10.9)	25.0 (11.5-43.4)	6.2 (0.8-20.8)

Table 21 Continued

Item/scale	Nilotinib					
	Not at all	**A little**	**Quite a bit**	**A lot**	**Any severity**	**Moderate-severe $**
Diarrhoea	84.4 (67.2-94.7)	12.5 (3.5-29.0)	0.0 (0.0-10.9)	3.1 (0.1-16.2)	15.6 (5.3-32.8)	3.1 (0.1-16.2)
Abdominal pain/ cramps	65.6 (46.8-81.4)	21.9 (9.3-40.0)	9.4 (2.0-25.0)	3.1 (0.1-16.2)	34.4 (18.6-53.2)	12.5 (3.5-29.0)
Vomited	93.8 (79.2-99.2)	6.2 (0.8-20.8)	0.0 (0.0-10.9)	0.0 (0.0-10.9)	6.2 (0.8-20.8)	0.0 (0.0-10.9)
Obstipation	68.8 (50.0-83.9)	15.6 (5.3-32.8)	6.2 (0.8-20.8)	9.4 (2.0-25.0)	31.2 (16.1-50.0)	15.6 (5.3-32.8)
Flatulence	40.6 (23.7-59.4)	25.0 (11.5-43.4)	31.2 (16.1-50.0)	3.1 (0.1-16.2)	59.4 (40.6-76.3)	34.4 (18.6-53.2)
Muscle/joint cramps/pain (scale)	25.0 (11.5-43.4)	Na	Na	Na	75.0 (56.6-88.5)	Na
Frequent urination	53.1 (34.7-70.9)	18.8 (7.2-36.4)	25.0 (11.5-43.4)	3.1 (0.1-16.2)	46.9 (29.1-65.3)	28.1 (13.7-46.7)
Muscle cramps	34.4 (18.6-53.2)	43.8 (26.4-62.3)	15.6 (5.3-32.8)	6.2 (0.8-20.8)	65.6 (46.8-81.4)	21.9 (9.3-40.0)
Aches/pains muscles/joints	40.6 (23.7-59.4)	25.0 (11.5-43.4)	25.0 (11.5-43.4)	9.4 (2.0-25.0)	59.4 (40.6-76.3)	34.4 (18.6-53.2)
Muscle weakness	68.8 (50.0-83.9)	28.1 (13.7-46.7)	3.1 (0.1-16.2)	0.0 (0.0-10.9)	31.2 (16.1-50.0)	3.1 (0.1-16.2)
Pale/cold fingers/ toes	78.6 (49.2-95.3)	21.4 (4.7-50.8)	0.0 (0.0-23.2)	0.0 (0.0-23.2)	21.4 (4.7-50.8)	0.0 (0.0-23.2)
Tingling/numbness hands/feet	56.2 (37.7-73.6)	37.5 (21.1-56.3)	6.2 (0.8-20.8)	0.0 (0.0-10.9)	43.8 (26.4-62.3)	6.2 (0.8-20.8)
Problems finger dexterity	81.2 (63.6-92.8)	18.8 (7.2-36.4)	0.0 (0.0-10.9)	0.0 (0.0-10.9)	18.8 (7.2-36.4)	0.0 (0.0-10.9)
Trouble walking pain	54.8 (36.0-72.7)	35.5 (19.2-54.6)	6.5 (0.8-21.4)	3.2 (0.1-16.7)	45.2 (27.3-64.0)	9.7 (2.0-25.8)
Back pain	42.9 (17.7-71.1)	35.7 (12.8-64.9)	7.1 (0.2-33.9)	14.3 (1.8-42.8)	57.1 (28.9-82.3)	21.4 (4.7-50.8)

Table 21 Continued

Item/scale	Nilotinib					
	Not at all	A little	Quite a bit	A lot	Any severity	Moderate-severe $
Pain	62.5 (43.7-78.9)	21.9 (9.3-40.0)	12.5 (3.5-29.0)	3.1 (0.1-16.2)	37.5 (21.1-56.3)	15.6 (5.3-32.8)
Fever/chills	87.5 (71.0-96.5)	12.5 (3.5-29.0)	0.0 (0.0-10.9)	0.0 (0.0-10.9)	12.5 (3.5-29.0)	0.0 (0.0-10.9)
Sweating excessively	53.1 (34.7-70.9)	28.1 (13.7-46.7)	15.6 (5.3-32.8)	3.1 (0.1-16.2)	46.9 (29.1-65.3)	18.8 (7.2-36.4)
Problems tolerating heat/cold	71.9 (53.3-86.3)	15.6 (5.3-32.8)	9.4 (2.0-25.0)	3.1 (0.1-16.2)	28.1 (13.7-46.7)	12.5 (3.5-29.0)
Hot flushes	78.1 (60.0-90.7)	9.4 (2.0-25.0)	6.2 (0.8-20.8)	6.2 (0.8-20.8)	21.9 (9.3-40.0)	12.5 (3.5-29.0)
Airway infections	85.7 (57.2-98.2)	14.3 (1.8-42.8)	0.0 (0.0-23.2)	0.0 (0.0-23.2)	14.3 (1.8-42.8)	0.0 (0.0-23.2)
Other infections	92.9 (66.1-99.8)	7.1 (0.2-33.9)	0.0 (0.0-23.2)	0.0 (0.0-23.2)	7.1 (0.2-33.9)	0.0 (0.0-23.2)
Sore/enlarged nipples/breasts	87.5 (71.0-96.5)	9.4 (2.0-25.0)	3.1 (0.1-16.2)	0.0 (0.0-10.9)	12.5 (3.5-29.0)	3.1 (0.1-16.2)
Sexuality problems (scale)	35.5 (19.2-54.6)	Na	Na	Na	64.5 (45.4-80.8)	Na
Less interest sex	35.5 (19.2-54.6)	32.3 (16.7-51.4)	22.6 (9.6-41.1)	9.7 (2.0-25.8)	64.5 (45.4-80.8)	32.3 (16.7-51.4)
Less pleasure sex	45.2 (27.3-64.0)	25.8 (11.9-44.6)	16.1 (5.5-33.7)	12.9 (3.6-29.8)	54.8 (36.0-72.7)	29.0 (14.2-48.0)
Pain during/after sex	80.0 (61.4-92.3)	10.0 (2.1-26.5)	10.0 (2.1-26.5)	0.0 (0.0-11.6)	20.0 (7.7-38.6)	10.0 (2.1-26.5)
Difficult erection	25.0 (5.5-57.2)	41.7 (15.2-72.3)	16.7 (2.1-48.4)	16.7 (2.1-48.4)	75.0 (42.8-94.5)	33.3 (9.9-65.1)
Dry vagina	35.3 (14.2-61.7)	41.2 (18.4-67.1)	11.8 (1.5-36.4)	11.8 (1.5-36.4)	64.7 (38.3-85.8)	23.5 (6.8-49.9)

* At T4 as not available at T1. $ Moderate-severe includes 'Quite a bit' and 'A lot'. Abbreviations: Na: not applicable

Table 22 Prevalence (% with 95%confidence intervals) of symptom scores for seven scales plus 76 single items, in 8 patients on ponatinib on T1

Item/scale	Ponatinib					
	Not at all	A little	Quite a bit	A lot	Any severity	Moderate-severe $
Fatigue (scale)	12.5 (0.3-52.7)	Na	Na	Na	87.5 (47.3-99.7)	Na
Tired	12.5 (0.3-52.7)	37.5 (8.5-75.5)	25.0 (3.2-65.1)	25.0 (3.2-65.1)	87.5 (47.3-99.7)	50.0 (15.7-84.3)
Needed rest	12.5 (0.3-52.7)	50.0 (15.7-84.3)	37.5 (8.5-75.5)	0.0 (0.0-36.9)	87.5 (47.3-99.7)	37.5 (8.5-75.5)
Lacked energy	14.3 (0.4-57.9)	42.9 (9.9-81.6)	28.6 (3.7-71.0)	14.3 (0.4-57.9)	85.7 (42.1-99.6)	42.9 (9.9-81.6)
Felt lethargic	37.5 (8.5-75.5)	50.0 (15.7-84.3)	12.5 (0.3-52.7)	0.0 (0.0-36.9)	62.5 (24.5-91.5)	12.5 (0.3-52.7)
Felt weak	37.5 (8.5-75.5)	37.5 (8.5-75.5)	25.0 (3.2-65.1)	0.0 (0.0-36.9)	62.5 (24.5-91.5)	25.0 (3.2-65.1)
Sleepy	14.3 (0.4-57.9)	28.6 (3.7-71.0)	57.1 (18.4-90.1)	0.0 (0.0-41.0)	85.7 (42.1-99.6)	57.1 (18.4-90.1)
Headaches	85.7 (42.1-99.6)	14.3 (0.4-57.9)	0.0 (0.0-41.0)	0.0 (0.0-41.0)	14.3 (0.4-57.9)	0.0 (0.0-41.0)
Dizziness	71.4 (29.0-96.3)	28.6 (3.7-71.0)	0.0 (0.0-41.0)	0.0 (0.0-41.0)	28.6 (3.7-71.0)	0.0 (0.0-41.0)
Cognitive problems (scale)	14.3 (0.4-57.9)	Na	Na	Na	85.7 (42.1-99.6)	Na
Difficulty remembering	28.6 (3.7-71.0)	57.1 (18.4-90.1)	14.3 (0.4-57.9)	0.0 (0.0-41.0)	71.4 (29.0-96.3)	14.3 (0.4-57.9)
Difficulty concentrating	14.3 (0.4-57.9)	71.4 (29.0-96.3)	14.3 (0.4-57.9)	0.0 (0.0-41.0)	85.7 (42.1-99.6)	14.3 (0.4-57.9)
Trouble thinking clearly	28.6 (3.7-71.0)	71.4 (29.0-96.3)	0.0 (0.0-41.0)	0.0 (0.0-41.0)	71.4 (29.0-96.3)	0.0 (0.0-41.0)
Bruised easily	85.7 (42.1-99.6)	14.3 (0.4-57.9)	0.0 (0.0-41.0)	0.0 (0.0-41.0)	14.3 (0.4-57.9)	0.0 (0.0-41.0)
Skin problems (scale)	0.0 (0.0-41.0)	Na	Na	Na	100.0 (59.0-100.0)	Na
Skin colour change	57.1 (18.4-90.1)	42.9 (9.9-81.6)	0.0 (0.0-41.0)	0.0 (0.0-41.0)	42.9 (9.9-81.6)	0.0 (0.0-41.0)
Skin itched	14.3 (0.4-57.9)	57.1 (18.4-90.1)	28.6 (3.7-71.0)	0.0 (0.0-41.0)	85.7 (42.1-99.6)	28.6 (3.7-71.0)
Skin rash	57.1 (18.4-90.1)	14.3 (0.4-57.9)	28.6 (3.7-71.0)	0.0 (0.0-41.0)	42.9 (9.9-81.6)	28.6 (3.7-71.0)
Skin dry/flaky/cracked	0.0 (0.0-41.0)	71.4 (29.0-96.3)	28.6 (3.7-71.0)	0.0 (0.0-41.0)	100.0 (59.0-100.0)	28.6 (3.7-71.0)

Table 22 Continued

Item/scale	Ponatinib					
	Not at all	A little	Quite a bit	A lot	Any severity	Moderate-severe $
Skin sore/painful	42.9 (9.9-81.6)	42.9 (9.9-81.6)	14.3 (0.4-57.9)	0.0 (0.0-41.0)	57.1 (18.4-90.1)	14.3 (0.4-57.9)
Wounds healing problems	83.3 (35.9-99.6)	16.7 (0.4-64.1)	0.0 (0.0-45.9)	0.0 (0.0-45.9)	16.7 (0.4-64.1)	0.0 (0.0-45.9)
Hair loss	71.4 (29.0-96.3)	28.6 (3.7-71.0)	0.0 (0.0-41.0)	0.0 (0.0-41.0)	28.6 (3.7-71.0)	0.0 (0.0-41.0)
Nail problems	67.9 (56.4-78.1)	20.5 (12.2-31.2)	9.0 (3.7-17.6)	2.6 (0.3-9.0)	32.1 (21.9-43.6)	11.5 (5.4-20.8)
Eye problems (scale)	14.3 (0.4-57.9)	Na	Na	Na	85.7 (42.1-99.6)	Na
Watery eyes	57.1 (18.4-90.1)	28.6 (3.7-71.0)	14.3 (0.4-57.9)	0.0 (0.0-41.0)	42.9 (9.9-81.6)	14.3 (0.4-57.9)
Burning eyes	28.6 (3.7-71.0)	71.4 (29.0-96.3)	0.0 (0.0-41.0)	0.0 (0.0-41.0)	71.4 (29.0-96.3)	0.0 (0.0-41.0)
Light sensitive eyes	85.7 (42.1-99.6)	14.3 (0.4-57.9)	0.0 (0.0-41.0)	0.0 (0.0-41.0)	14.3 (0.4-57.9)	0.0 (0.0-41.0)
Blurry eyes	28.6 (3.7-71.0)	57.1 (18.4-90.1)	14.3 (0.4-57.9)	0.0 (0.0-41.0)	71.4 (29.0-96.3)	14.3 (0.4-57.9)
Red eyes	85.7 (42.1-99.6)	0.0 (0.0-41.0)	14.3 (0.4-57.9)	0.0 (0.0-41.0)	14.3 (0.4-57.9)	14.3 (0.4-57.9)
Swelling face/eyes	85.7 (42.1-99.6)	14.3 (0.4-57.9)	0.0 (0.0-41.0)	0.0 (0.0-41.0)	14.3 (0.4-57.9)	0.0 (0.0-41.0)
Swelling body parts	83.3 (35.9-99.6)	16.7 (0.4-64.1)	0.0 (0.0-45.9)	0.0 (0.0-45.9)	16.7 (0.4-64.1)	0.0 (0.0-45.9)
Hearing problems	100.0 (59.0-100.0)	0.0 (0.0-41.0)	0.0 (0.0-41.0)	0.0 (0.0-41.0)	0.0 (0.0-41.0)	0.0 (0.0-41.0)
Nose bleeds	100.0 (59.0-100.0)	0.0 (0.0-41.0)	0.0 (0.0-41.0)	0.0 (0.0-41.0)	0.0 (0.0-41.0)	0.0 (0.0-41.0)
Nose problems	0.0 (0.0-97.5)	100.0 (2.5-100.0)	0.0 (0.0-97.5)	0.0 (0.0-97.5)	100.0 (2.5-100.0)	0.0 (0.0-97.5)
Dry mouth	14.3 (0.4-57.9)	71.4 (29.0-96.3)	14.3 (0.4-57.9)	0.0 (0.0-41.0)	85.7 (42.1-99.6)	14.3 (0.4-57.9)
Pain/soreness mouth	85.7 (42.1-99.6)	14.3 (0.4-57.9)	0.0 (0.0-41.0)	0.0 (0.0-41.0)	14.3 (0.4-57.9)	0.0 (0.0-41.0)
Teeth problems	42.9 (9.9-81.6)	57.1 (18.4-90.1)	0.0 (0.0-41.0)	0.0 (0.0-41.0)	57.1 (18.4-90.1)	0.0 (0.0-41.0)
Bleeding gums	57.1 (18.4-90.1)	42.9 (9.9-81.6)	0.0 (0.0-41.0)	0.0 (0.0-41.0)	42.9 (9.9-81.6)	0.0 (0.0-41.0)
Taste change	83.3 (35.9-99.6)	0.0 (0.0-45.9)	16.7 (0.4-64.1)	0.0 (0.0-45.9)	16.7 (0.4-64.1)	16.7 (0.4-64.1)

Table 22 Continued

Item/scale	Ponatinib					
	Not at all	A little	Quite a bit	A lot	Any severity	Moderate-severe $
Appetite loss	66.7 (22.3-95.7)	33.3 (4.3-77.7)	0.0 (0.0-45.9)	0.0 (0.0-45.9)	33.3 (4.3-77.7)	0.0 (0.0-45.9)
Mood problems (scale)	28.6 (3.7-71.0)	Na	Na	Na	71.4 (29.0-96.3)	Na
Depressed	66.7 (22.3-95.7)	33.3 (4.3-77.7)	0.0 (0.0-45.9)	0.0 (0.0-45.9)	33.3 (4.3-77.7)	0.0 (0.0-45.9)
Irritable	50.0 (11.8-88.2)	50.0 (11.8-88.2)	0.0 (0.0-45.9)	0.0 (0.0-45.9)	50.0 (11.8-88.2)	0.0 (0.0-45.9)
Tense	16.7 (0.4-64.1)	83.3 (35.9-99.6)	0.0 (0.0-45.9)	0.0 (0.0-45.9)	83.3 (35.9-99.6)	0.0 (0.0-45.9)
Worried	16.7 (0.4-64.1)	83.3 (35.9-99.6)	0.0 (0.0-45.9)	0.0 (0.0-45.9)	83.3 (35.9-99.6)	0.0 (0.0-45.9)
Afraid	66.7 (22.3-95.7)	33.3 (4.3-77.7)	0.0 (0.0-45.9)	0.0 (0.0-45.9)	33.3 (4.3-77.7)	0.0 (0.0-45.9)
Insomnia	42.9 (9.9-81.6)	57.1 (18.4-90.1)	0.0 (0.0-41.0)	0.0 (0.0-41.0)	57.1 (18.4-90.1)	0.0 (0.0-41.0)
Palpitations	50.0 (11.8-88.2)	50.0 (11.8-88.2)	0.0 (0.0-45.9)	0.0 (0.0-45.9)	50.0 (11.8-88.2)	0.0 (0.0-45.9)
Pain in chest	83.3 (35.9-99.6)	16.7 (0.4-64.1)	0.0 (0.0-45.9)	0.0 (0.0-45.9)	16.7 (0.4-64.1)	0.0 (0.0-45.9)
Short of breath	83.3 (35.9-99.6)	16.7 (0.4-64.1)	0.0 (0.0-45.9)	0.0 (0.0-45.9)	16.7 (0.4-64.1)	0.0 (0.0-45.9)
Coughed	100.0 (54.1-100.0)	0.0 (0.0-45.9)	0.0 (0.0-45.9)	0.0 (0.0-45.9)	0.0 (0.0-45.9)	0.0 (0.0-45.9)
Feeling bloated	33.3 (4.3-77.7)	33.3 (4.3-77.7)	33.3 (4.3-77.7)	0.0 (0.0-45.9)	66.7 (22.3-95.7)	33.3 (4.3-77.7)
Indigestion	16.7 (0.4-64.1)	66.7 (22.3-95.7)	16.7 (0.4-64.1)	0.0 (0.0-45.9)	83.3 (35.9-99.6)	16.7 (0.4-64.1)
Nausea	66.7 (22.3-95.7)	16.7 (0.4-64.1)	16.7 (0.4-64.1)	0.0 (0.0-45.9)	33.3 (4.3-77.7)	16.7 (0.4-64.1)
Diarrhoea	83.3 (35.9-99.6)	16.7 (0.4-64.1)	0.0 (0.0-45.9)	0.0 (0.0-45.9)	16.7 (0.4-64.1)	0.0 (0.0-45.9)
Abdominal pain/ cramps	50.0 (11.8-88.2)	50.0 (11.8-88.2)	0.0 (0.0-45.9)	0.0 (0.0-45.9)	50.0 (11.8-88.2)	0.0 (0.0-45.9)
Vomited	100.0 (47.8-100.0)	0.0 (0.0-52.2)	0.0 (0.0-52.2)	0.0 (0.0-52.2)	0.0 (0.0-52.2)	0.0 (0.0-52.2)
Obstipation	50.0 (11.8-88.2)	50.0 (11.8-88.2)	0.0 (0.0-45.9)	0.0 (0.0-45.9)	50.0 (11.8-88.2)	0.0 (0.0-45.9)

Table 22 Continued

Item/scale	Ponatinib					
	Not at all	A little	Quite a bit	A lot	Any severity	Moderate-severe $
Flatulence	50.0 (11.8-88.2)	33.3 (4.3-77.7)	16.7 (0.4-64.1)	0.0 (0.0-45.9)	50.0 (11.8-88.2)	16.7 (0.4-64.1)
Muscle/joint cramps/ pain (scale)	0.0 (0.0-45.9)	Na	Na	Na	100.0 (54.1-100.0)	Na
Frequent urination	16.7 (0.4-64.1)	50.0 (11.8-88.2)	33.3 (4.3-77.7)	0.0 (0.0-45.9)	83.3 (35.9-99.6)	33.3 (4.3-77.7)
Muscle cramps	50.0 (11.8-88.2)	33.3 (4.3-77.7)	16.7 (0.4-64.1)	0.0 (0.0-45.9)	50.0 (11.8-88.2)	16.7 (0.4-64.1)
Aches/pains muscles/ joints	0.0 (0.0-45.9)	66.7 (22.3-95.7)	33.3 (4.3-77.7)	0.0 (0.0-45.9)	100.0 (54.1-100.0)	33.3 (4.3-77.7)
Muscle weakness	83.3 (35.9-99.6)	16.7 (0.4-64.1)	0.0 (0.0-45.9)	0.0 (0.0-45.9)	16.7 (0.4-64.1)	0.0 (0.0-45.9)
Pale/cold fingers/toes	0.0 (0.0-97.5)	100.0 (2.5-100.0)	0.0 (0.0-97.5)	0.0 (0.0-97.5)	100.0 (2.5-100.0)	0.0 (0.0-97.5)
Tingling/numbness hands/feet	16.7 (0.4-64.1)	66.7 (22.3-95.7)	16.7 (0.4-64.1)	0.0 (0.0-45.9)	83.3 (35.9-99.6)	16.7 (0.4-64.1)
Problems finger dexterity	66.7 (22.3-95.7)	33.3 (4.3-77.7)	0.0 (0.0-45.9)	0.0 (0.0-45.9)	33.3 (4.3-77.7)	0.0 (0.0-45.9)
Trouble walking pain	66.7 (22.3-95.7)	33.3 (4.3-77.7)	0.0 (0.0-45.9)	0.0 (0.0-45.9)	33.3 (4.3-77.7)	0.0 (0.0-45.9)
Back pain	100.0 (2.5-100.0)	0.0 (0.0-97.5)	0.0 (0.0-97.5)	0.0 (0.0-97.5)	0.0 (0.0-97.5)	0.0 (0.0-97.5)
Pain	33.3 (4.3-77.7)	66.7 (22.3-95.7)	0.0 (0.0-45.9)	0.0 (0.0-45.9)	66.7 (22.3-95.7)	0.0 (0.0-45.9)
Fever/chills	66.7 (22.3-95.7)	33.3 (4.3-77.7)	0.0 (0.0-45.9)	0.0 (0.0-45.9)	33.3 (4.3-77.7)	0.0 (0.0-45.9)
Sweating excessively	66.7 (22.3-95.7)	33.3 (4.3-77.7)	0.0 (0.0-45.9)	0.0 (0.0-45.9)	33.3 (4.3-77.7)	0.0 (0.0-45.9)
Problems tolerating heat/cold	66.7 (22.3-95.7)	33.3 (4.3-77.7)	0.0 (0.0-45.9)	0.0 (0.0-45.9)	33.3 (4.3-77.7)	0.0 (0.0-45.9)
Hot flushes	66.7 (22.3-95.7)	16.7 (0.4-64.1)	16.7 (0.4-64.1)	0.0 (0.0-45.9)	33.3 (4.3-77.7)	16.7 (0.4-64.1)
Airway infections	100.0 (2.5-100.0)	0.0 (0.0-97.5)	0.0 (0.0-97.5)	0.0 (0.0-97.5)	0.0 (0.0-97.5)	0.0 (0.0-97.5)

Table 22 Continued

Item/scale	Ponatinib					
	Not at all	**A little**	**Quite a bit**	**A lot**	**Any severity**	**Moderate-severe $**
Other infections	100.0 (2.5-100.0)	0.0 (0.0-97.5)	0.0 (0.0-97.5)	0.0 (0.0-97.5)	0.0 (0.0-97.5)	0.0 (0.0-97.5)
Sore/enlarged nipples/breasts	100.0 (54.1-100.0)	0.0 (0.0-45.9)	0.0 (0.0-45.9)	0.0 (0.0-45.9)	0.0 (0.0-45.9)	0.0 (0.0-45.9)
Sexuality problems (scale)	0.0 (0.0-45.9)	Na	Na	Na	100.0 (54.1-100.0)	Na
Less interest sex	0.0 (0.0-45.9)	66.7 (22.3-95.7)	33.3 (4.3-77.7)	0.0 (0.0-45.9)	100.0 (54.1-100.0)	33.3 (4.3-77.7)
Less pleasure sex	33.3 (4.3-77.7)	33.3 (4.3-77.7)	33.3 (4.3-77.7)	0.0 (0.0-45.9)	66.7 (22.3-95.7)	33.3 (4.3-77.7)
Pain during/after sex	83.3 (35.9-99.6)	0.0 (0.0-45.9)	16.7 (0.4-64.1)	0.0 (0.0-45.9)	16.7 (0.4-64.1)	16.7 (0.4-64.1)
Difficult erection	66.7 (9.4-99.2)	33.3 (0.8-90.6)	0.0 (0.0-70.8)	0.0 (0.0-70.8)	33.3 (0.8-90.6)	0.0 (0.0-70.8)
Dry vagina	0.0 (0.0-97.5)	100.0 (2.5-100.0)	0.0 (0.0-97.5)	0.0 (0.0-97.5)	100.0 (2.5-100.0)	0.0 (0.0-97.5)

* At T4 as not available at T1. $ Moderate-severe includes 'Quite a bit' and 'A lot'. Abbreviations: Na: not applicable

Table 23 Prevalence (% with 95%confidence intervals) of symptom scores for seven scales plus 76 single items, in 40 patients in treatment-free remission on T1

Item/scale	Treatment-free remission					
	Not at all	A little	Quite a bit	A lot	Any severity	Moderate-severe $
Fatigue (scale)	10.0 (2.8-23.7)	Na	Na	Na	90.0 (76.3-97.2)	Na
Tired	12.5 (4.2-26.8)	30.0 (16.6-46.5)	40.0 (24.9-56.7)	17.5 (7.3-32.8)	87.5 (73.2-95.8)	57.5 (40.9-73.0)
Needed rest	20.0 (9.1-35.6)	42.5 (27.0-59.1)	25.0 (12.7-41.2)	12.5 (4.2-26.8)	80.0 (64.4-90.9)	37.5 (22.7-54.2)
Lacked energy	15.4 (5.9-30.5)	30.8 (17.0-47.6)	38.5 (23.4-55.4)	15.4 (5.9-30.5)	84.6 (69.5-94.1)	53.8 (37.2-69.9)
Felt lethargic	34.2 (19.6-51.4)	42.1 (26.3-59.2)	18.4 (7.7-34.3)	5.3 (0.6-17.7)	65.8 (48.6-80.4)	23.7 (11.4-40.2)
Felt weak	47.2 (30.4-64.5)	38.9 (23.1-56.5)	11.1 (3.1-26.1)	2.8 (0.1-14.5)	52.8 (35.5-69.6)	13.9 (4.7-29.5)
Sleepy	25.0 (12.1-42.2)	52.8 (35.5-69.6)	11.1 (3.1-26.1)	11.1 (3.1-26.1)	75.0 (57.8-87.9)	22.2 (10.1-39.2)
Headaches	72.2 (54.8-85.8)	13.9 (4.7-29.5)	11.1 (3.1-26.1)	2.8 (0.1-14.5)	27.8 (14.2-45.2)	13.9 (4.7-29.5)
Dizziness	71.4 (53.7-85.4)	22.9 (10.4-40.1)	2.9 (0.1-14.9)	2.9 (0.1-14.9)	28.6 (14.6-46.3)	5.7 (0.7-19.2)
Cognitive problems (scale)	30.6 (16.3-48.1)	Na	Na	Na	69.4 (51.9-83.7)	Na
Difficulty remembering	45.7 (28.8-63.4)	42.9 (26.3-60.6)	8.6 (1.8-23.1)	2.9 (0.1-14.9)	54.3 (36.6-71.2)	11.4 (3.2-26.7)
Difficulty concentrating	45.7 (28.8-63.4)	37.1 (21.5-55.1)	11.4 (3.2-26.7)	5.7 (0.7-19.2)	54.3 (36.6-71.2)	17.1 (6.6-33.6)
Trouble thinking clearly	52.8 (35.5-69.6)	38.9 (23.1-56.5)	5.6 (0.7-18.7)	2.8 (0.1-14.5)	47.2 (30.4-64.5)	8.3 (1.8-22.5)
Bruised easily	83.3 (67.2-93.6)	13.9 (4.7-29.5)	2.8 (0.1-14.5)	0.0 (0.0-9.7)	16.7 (6.4-32.8)	2.8 (0.1-14.5)
Skin problems (scale)	17.1 (6.6-33.6)	Na	Na	Na	82.9 (66.4-93.4)	Na
Skin colour change	74.3 (56.7-87.5)	22.9 (10.4-40.1)	2.9 (0.1-14.9)	0.0 (0.0-10.0)	25.7 (12.5-43.3)	2.9 (0.1-14.9)
Skin itched	47.1 (29.8-64.9)	29.4 (15.1-47.5)	20.6 (8.7-37.9)	2.9 (0.1-15.3)	52.9 (35.1-70.2)	23.5 (10.7-41.2)
Skin rash	60.0 (42.1-76.1)	25.7 (12.5-43.3)	8.6 (1.8-23.1)	5.7 (0.7-19.2)	40.0 (23.9-57.9)	14.3 (4.8-30.3)
Skin dry/flaky/cracked	37.1 (21.5-55.1)	31.4 (16.9-49.3)	31.4 (16.9-49.3)	0.0 (0.0-10.0)	62.9 (44.9-78.5)	31.4 (16.9-49.3)
Skin sore/painful	68.6 (50.7-83.1)	20.0 (8.4-36.9)	8.6 (1.8-23.1)	2.9 (0.1-14.9)	31.4 (16.9-49.3)	11.4 (3.2-26.7)

Table 23 Continued

Item/scale	Treatment-free remission					
	Not at all	A little	Quite a bit	A lot	Any severity	Moderate-severe $
Wounds healing problems	86.1 (70.5-95.3)	8.3 (1.8-22.5)	2.8 (0.1-14.5)	2.8 (0.1-14.5)	13.9 (4.7-29.5)	5.6 (0.7-18.7)
Hair loss	69.4 (51.9-83.7)	19.4 (8.2-36.0)	11.1 (3.1-26.1)	0.0 (0.0-9.7)	30.6 (16.3-48.1)	11.1 (3.1-26.1)
Nail problems	67.9 (56.4-78.1)	20.5 (12.2-31.2)	9.0 (3.7-17.6)	2.6 (0.3-9.0)	32.1 (21.9-43.6)	11.5 (5.4-20.8)
Eye problems (scale)	34.3 (19.1-52.2)	Na	Na	Na	65.7 (47.8-80.9)	Na
Watery eyes	68.6 (50.7-83.1)	17.1 (6.6-33.6)	14.3 (4.8-30.3)	0.0 (0.0-10.0)	31.4 (16.9-49.3)	14.3 (4.8-30.3)
Burning eyes	58.8 (40.7-75.4)	29.4 (15.1-47.5)	5.9 (0.7-19.7)	5.9 (0.7-19.7)	41.2 (24.6-59.3)	11.8 (3.3-27.5)
Light sensitive eyes	57.1 (39.4-73.7)	14.3 (4.8-30.3)	22.9 (10.4-40.1)	5.7 (0.7-19.2)	42.9 (26.3-60.6)	28.6 (14.6-46.3)
Blurry eyes	61.8 (43.6-77.8)	29.4 (15.1-47.5)	2.9 (0.1-15.3)	5.9 (0.7-19.7)	38.2 (22.2-56.4)	8.8 (1.9-23.7)
Red eyes	91.2 (76.3-98.1)	5.9 (0.7-19.7)	2.9 (0.1-15.3)	0.0 (0.0-10.3)	8.8 (1.9-23.7)	2.9 (0.1-15.3)
Swelling face/eyes	79.4 (62.1-91.3)	8.8 (1.9-23.7)	5.9 (0.7-19.7)	5.9 (0.7-19.7)	20.6 (8.7-37.9)	11.8 (3.3-27.5)
Swelling body parts	48.4 (30.2-66.9)	35.5 (19.2-54.6)	12.9 (3.6-29.8)	3.2 (0.1-16.7)	51.6 (33.1-69.8)	16.1 (5.5-33.7)
Hearing problems	73.5 (55.6-87.1)	8.8 (1.9-23.7)	8.8 (1.9-23.7)	8.8 (1.9-23.7)	26.5 (12.9-44.4)	17.6 (6.8-34.5)
Nose bleeds	88.2 (72.5-96.7)	8.8 (1.9-23.7)	2.9 (0.1-15.3)	0.0 (0.0-10.3)	11.8 (3.3-27.5)	2.9 (0.1-15.3)
Nose problems	25.0 (3.2-65.1)	62.5 (24.5-91.5)	12.5 (0.3-52.7)	0.0 (0.0-36.9)	75.0 (34.9-96.8)	12.5 (0.3-52.7)
Dry mouth	61.8 (43.6-77.8)	23.5 (10.7-41.2)	14.7 (5.0-31.1)	0.0 (0.0-10.3)	38.2 (22.2-56.4)	14.7 (5.0-31.1)
Pain/soreness mouth	73.5 (55.6-87.1)	23.5 (10.7-41.2)	2.9 (0.1-15.3)	0.0 (0.0-10.3)	26.5 (12.9-44.4)	2.9 (0.1-15.3)
Teeth problems	73.5 (55.6-87.1)	20.6 (8.7-37.9)	5.9 (0.7-19.7)	0.0 (0.0-10.3)	26.5 (12.9-44.4)	5.9 (0.7-19.7)
Bleeding gums	84.8 (68.1-94.9)	9.1 (1.9-24.3)	6.1 (0.7-20.2)	0.0 (0.0-10.6)	15.2 (5.1-31.9)	6.1 (0.7-20.2)
Taste change	91.2 (76.3-98.1)	5.9 (0.7-19.7)	2.9 (0.1-15.3)	0.0 (0.0-10.3)	8.8 (1.9-23.7)	2.9 (0.1-15.3)
Appetite loss	90.9 (75.7-98.1)	6.1 (0.7-20.2)	3.0 (0.1-15.8)	0.0 (0.0-10.6)	9.1 (1.9-24.3)	3.0 (0.1-15.8)

Table 23 Continued

Item/scale	Treatment-free remission					
	Not at all	A little	Quite a bit	A lot	Any severity	Moderate-severe $
Mood problems (scale)	16.7 (6.4-32.8)	Na	Na	Na	83.3 (67.2-93.6)	Na
Depressed	51.5 (33.5-69.2)	45.5 (28.1-63.6)	3.0 (0.1-15.8)	0.0 (0.0-10.6)	48.5 (30.8-66.5)	3.0 (0.1-15.8)
Irritable	45.5 (28.1-63.6)	45.5 (28.1-63.6)	9.1 (1.9-24.3)	0.0 (0.0-10.6)	54.5 (36.4-71.9)	9.1 (1.9-24.3)
Tense	50.0 (31.9-68.1)	43.8 (26.4-62.3)	6.2 (0.8-20.8)	0.0 (0.0-10.9)	50.0 (31.9-68.1)	6.2 (0.8-20.8)
Worried	30.3 (15.6-48.7)	60.6 (42.1-77.1)	9.1 (1.9-24.3)	0.0 (0.0-10.6)	69.7 (51.3-84.4)	9.1 (1.9-24.3)
Afraid	75.8 (57.7-88.9)	21.2 (9.0-38.9)	3.0 (0.1-15.8)	0.0 (0.0-10.6)	24.2 (11.1-42.3)	3.0 (0.1-15.8)
Insomnia	51.4 (34.0-68.6)	28.6 (14.6-46.3)	11.4 (3.2-26.7)	8.6 (1.8-23.1)	48.6 (31.4-66.0)	20.0 (8.4-36.9)
Palpitations	57.6 (39.2-74.5)	36.4 (20.4-54.9)	6.1 (0.7-20.2)	0.0 (0.0-10.6)	42.4 (25.5-60.8)	6.1 (0.7-20.2)
Pain in chest	87.5 (71.0-96.5)	12.5 (3.5-29.0)	0.0 (0.0-10.9)	0.0 (0.0-10.9)	12.5 (3.5-29.0)	0.0 (0.0-10.9)
Short of breath	68.8 (50.0-83.9)	15.6 (5.3-32.8)	12.5 (3.5-29.0)	3.1 (0.1-16.2)	31.2 (16.1-50.0)	15.6 (5.3-32.8)
Coughed	71.9 (53.3-86.3)	21.9 (9.3-40.0)	3.1 (0.1-16.2)	3.1 (0.1-16.2)	28.1 (13.7-46.7)	6.2 (0.8-20.8)
Feeling bloated	68.8 (50.0-83.9)	15.6 (5.3-32.8)	9.4 (2.0-25.0)	6.2 (0.8-20.8)	31.2 (16.1-50.0)	15.6 (5.3-32.8)
Indigestion	68.8 (50.0-83.9)	15.6 (5.3-32.8)	12.5 (3.5-29.0)	3.1 (0.1-16.2)	31.2 (16.1-50.0)	15.6 (5.3-32.8)
Nausea	87.1 (70.2-96.4)	12.9 (3.6-29.8)	0.0 (0.0-11.2)	0.0 (0.0-11.2)	12.9 (3.6-29.8)	0.0 (0.0-11.2)
Diarrhoea	75.0 (56.6-88.5)	18.8 (7.2-36.4)	3.1 (0.1-16.2)	3.1 (0.1-16.2)	25.0 (11.5-43.4)	6.2 (0.8-20.8)
Abdominal pain/cramps	68.8 (50.0-83.9)	28.1 (13.7-46.7)	3.1 (0.1-16.2)	0.0 (0.0-10.9)	31.2 (16.1-50.0)	3.1 (0.1-16.2)
Vomited	100.0 (89.1-100.0)	0.0 (0.0-10.9)	0.0 (0.0-10.9)	0.0 (0.0-10.9)	0.0 (0.0-10.9)	0.0 (0.0-10.9)
Obstipation	74.2 (55.4-88.1)	22.6 (9.6-41.1)	0.0 (0.0-11.2)	3.2 (0.1-16.7)	25.8 (11.9-44.6)	3.2 (0.1-16.7)
Flatulence	56.2 (37.7-73.6)	25.0 (11.5-43.4)	15.6 (5.3-32.8)	3.1 (0.1-16.2)	43.8 (26.4-62.3)	18.8 (7.2-36.4)

Table 23 Continued

Item/scale	Treatment-free remission					
	Not at all	A little	Quite a bit	A lot	Any severity	Moderate-severe $
Muscle/joint cramps/ pain (scale)	6.2 (0.8-20.8)	Na	Na	Na	93.8 (79.2-99.2)	Na
Frequent urination	50.0 (31.9-68.1)	31.2 (16.1-50.0)	15.6 (5.3-32.8)	3.1 (0.1-16.2)	50.0 (31.9-68.1)	18.8 (7.2-36.4)
Muscle cramps	25.0 (11.5-43.4)	50.0 (31.9-68.1)	18.8 (7.2-36.4)	6.2 (0.8-20.8)	75.0 (56.6-88.5)	25.0 (11.5-43.4)
Aches/pains muscles/ joints	6.2 (0.8-20.8)	40.6 (23.7-59.4)	40.6 (23.7-59.4)	12.5 (3.5-29.0)	93.8 (79.2-99.2)	53.1 (34.7-70.9)
Muscle weakness	46.9 (29.1-65.3)	43.8 (26.4-62.3)	9.4 (2.0-25.0)	0.0 (0.0-10.9)	53.1 (34.7-70.9)	9.4 (2.0-25.0)
Pale/cold fingers/toes	37.5 (8.5-75.5)	62.5 (24.5-91.5)	0.0 (0.0-36.9)	0.0 (0.0-36.9)	62.5 (24.5-91.5)	0.0 (0.0-36.9)
Tingling/numbness hands/feet	68.8 (50.0-83.9)	25.0 (11.5-43.4)	6.2 (0.8-20.8)	0.0 (0.0-10.9)	31.2 (16.1-50.0)	6.2 (0.8-20.8)
Problems finger dexterity	61.3 (42.2-78.2)	29.0 (14.2-48.0)	6.5 (0.8-21.4)	3.2 (0.1-16.7)	38.7 (21.8-57.8)	9.7 (2.0-25.8)
Trouble walking pain	50.0 (31.9-68.1)	31.2 (16.1-50.0)	18.8 (7.2-36.4)	0.0 (0.0-10.9)	50.0 (31.9-68.1)	18.8 (7.2-36.4)
Back pain	37.5 (8.5-75.5)	25.0 (3.2-65.1)	25.0 (3.2-65.1)	12.5 (0.3-52.7)	62.5 (24.5-91.5)	37.5 (8.5-75.5)
Pain	48.4 (30.2-66.9)	35.5 (19.2-54.6)	16.1 (5.5-33.7)	0.0 (0.0-11.2)	51.6 (33.1-69.8)	16.1 (5.5-33.7)
Fever/chills	93.5 (78.6-99.2)	6.5 (0.8-21.4)	0.0 (0.0-11.2)	0.0 (0.0-11.2)	6.5 (0.8-21.4)	0.0 (0.0-11.2)
Sweating excessively	58.1 (39.1-75.5)	29.0 (14.2-48.0)	12.9 (3.6-29.8)	0.0 (0.0-11.2)	41.9 (24.5-60.9)	12.9 (3.6-29.8)
Problems tolerating heat/cold	61.3 (42.2-78.2)	29.0 (14.2-48.0)	9.7 (2.0-25.8)	0.0 (0.0-11.2)	38.7 (21.8-57.8)	9.7 (2.0-25.8)
Hot flushes	71.0 (52.0-85.8)	25.8 (11.9-44.6)	3.2 (0.1-16.7)	0.0 (0.0-11.2)	29.0 (14.2-48.0)	3.2 (0.1-16.7)
Airway infections	100.0 (59.0-100.0)	0.0 (0.0-41.0)	0.0 (0.0-41.0)	0.0 (0.0-41.0)	0.0 (0.0-41.0)	0.0 (0.0-41.0)
Other infections	87.5 (47.3-99.7)	12.5 (0.3-52.7)	0.0 (0.0-36.9)	0.0 (0.0-36.9)	12.5 (0.3-52.7)	0.0 (0.0-36.9)

Table 23 Continued

Item/scale	Treatment-free remission					
	Not at all	A little	Quite a bit	A lot	Any severity	Moderate-severe $
Sore/enlarged nipples/ breasts	90.3 (74.2-98.0)	9.7 (2.0-25.8)	0.0 (0.0-11.2)	0.0 (0.0-11.2)	9.7 (2.0-25.8)	0.0 (0.0-11.2)
Sexuality problems (scale)	46.7 (28.3-65.7)	Na	Na	Na	53.3 (34.3-71.7)	Na
Less interest sex	53.3 (34.3-71.7)	20.0 (7.7-38.6)	10.0 (2.1-26.5)	16.7 (5.6-34.7)	46.7 (28.3-65.7)	26.7 (12.3-45.9)
Less pleasure sex	60.0 (40.6-77.3)	13.3 (3.8-30.7)	16.7 (5.6-34.7)	10.0 (2.1-26.5)	40.0 (22.7-59.4)	26.7 (12.3-45.9)
Pain during/after sex	73.3 (54.1-87.7)	20.0 (7.7-38.6)	0.0 (0.0-11.6)	6.7 (0.8-22.1)	26.7 (12.3-45.9)	6.7 (0.8-22.1)
Difficult erection	71.4 (41.9-91.6)	21.4 (4.7-50.8)	7.1 (0.2-33.9)	0.0 (0.0-23.2)	28.6 (8.4-58.1)	7.1 (0.2-33.9)
Dry vagina	40.0 (16.3-67.7)	26.7 (7.8-55.1)	20.0 (4.3-48.1)	13.3 (1.7-40.5)	60.0 (32.3-83.7)	33.3 (11.8-61.6)

* At T4 as not available at T1. $ Moderate-severe includes 'Quite a bit' and 'A lot'. Abbreviations: Na: not applicable

Table 24 Prevalence (% with 95%confidence intervals) of symptom scores for seven scales plus 76 single items, in 227 patients on any TKI on T1

Item/scale	Any TKI					
	Not at all	A little	Quite a bit	A lot	Any severity	Moderate-severe $
Fatigue (scale)	8.4 (5.1-12.8)	Na	Na	Na	91.6 (87.2-94.9)	Na
Tired	12.3 (8.4-17.3)	36.1 (29.9-42.7)	33.9 (27.8-40.5)	17.6 (12.9-23.2)	87.7 (82.7-91.6)	51.5 (44.8-58.2)
Needed rest	17.6 (12.9-23.2)	40.1 (33.7-46.8)	34.8 (28.6-41.4)	7.5 (4.4-11.7)	82.4 (76.8-87.1)	42.3 (35.8-49.0)
Lacked energy	17.3 (12.5-22.9)	33.6 (27.4-40.3)	35.5 (29.1-42.2)	13.6 (9.4-18.9)	82.7 (77.1-87.5)	49.1 (42.3-55.9)
Felt lethargic	29.2 (23.1-35.9)	41.1 (34.4-48.1)	23.9 (18.3-30.3)	5.7 (3.0-9.8)	70.8 (64.1-76.9)	29.7 (23.6-36.4)
Felt weak	35.0 (28.5-41.9)	39.8 (33.1-46.8)	20.4 (15.1-26.5)	4.9 (2.4-8.7)	65.0 (58.1-71.5)	25.2 (19.5-31.7)
Sleepy	24.0 (18.3-30.5)	44.6 (37.7-51.7)	26.0 (20.1-32.6)	5.4 (2.7-9.4)	76.0 (69.5-81.7)	31.4 (25.1-38.2)
Headaches	65.0 (57.9-71.6)	25.4 (19.5-32.1)	8.1 (4.7-12.9)	1.5 (0.3-4.4)	35.0 (28.4-42.1)	9.6 (5.9-14.7)
Dizziness	66.2 (59.0-72.8)	28.2 (22.0-35.1)	4.6 (2.1-8.6)	1.0 (0.1-3.7)	33.8 (27.2-41.0)	5.6 (2.8-9.9)
Cognitive problems (scale)	28.9 (22.6-35.8)	Na	Na	Na	71.1 (64.2-77.4)	Na
Difficulty remembering	39.7 (32.8-46.9)	44.8 (37.7-52.1)	12.9 (8.5-18.4)	2.6 (0.8-5.9)	60.3 (53.1-67.2)	15.5 (10.7-21.3)
Difficulty concentrating	46.6 (39.4-53.9)	34.2 (27.5-41.4)	15.5 (10.7-21.4)	3.6 (1.5-7.3)	53.4 (46.1-60.6)	19.2 (13.9-25.4)
Trouble thinking clearly	47.9 (40.7-55.2)	40.1 (33.1-47.4)	9.9 (6.1-15.0)	2.1 (0.6-5.2)	52.1 (44.8-59.3)	12.0 (7.7-17.4)
Bruised easily	61.8 (54.5-68.7)	26.7 (20.6-33.6)	8.9 (5.3-13.9)	2.6 (0.9-6.0)	38.2 (31.3-45.5)	11.5 (7.4-16.9)
Skin problems (scale)	16.9 (11.9-23.1)	Na	Na	Na	83.1 (76.9-88.1)	Na
Skin colour change	63.1 (55.8-70.0)	25.7 (19.6-32.6)	9.6 (5.8-14.8)	1.6 (0.3-4.6)	36.9 (30.0-44.2)	11.2 (7.1-16.7)
Skin itched	37.0 (30.1-44.3)	37.6 (30.6-44.9)	18.5 (13.3-24.8)	6.9 (3.7-11.5)	63.0 (55.7-69.9)	25.4 (19.4-32.2)
Skin rash	50.3 (42.9-57.6)	31.6 (25.0-38.7)	13.9 (9.3-19.7)	4.3 (1.9-8.3)	49.7 (42.4-57.1)	18.2 (12.9-24.5)
Skin dry/flaky/cracked	38.8 (31.8-46.2)	36.7 (29.8-44.0)	20.7 (15.2-27.2)	3.7 (1.5-7.5)	61.2 (53.8-68.2)	24.5 (18.5-31.3)
Skin sore/painful	55.6 (48.2-62.8)	32.3 (25.7-39.4)	11.1 (7.0-16.5)	1.1 (0.1-3.8)	44.4 (37.2-51.8)	12.2 (7.9-17.7)

Table 24 Continued

Item/scale	Any TKI					
	Not at all	A little	Quite a bit	A lot	Any severity	Moderate-severe $
Wounds healing problems	78.9 (72.5-84.5)	16.8 (11.8-22.9)	2.6 (0.9-6.0)	1.6 (0.3-4.5)	21.1 (15.5-27.5)	4.2 (1.8-8.1)
Hair loss	57.1 (49.8-64.3)	29.1 (22.7-36.1)	10.1 (6.2-15.3)	3.7 (1.5-7.5)	42.9 (35.7-50.2)	13.8 (9.2-19.5)
Nail problems	67.9 (56.4-78.1)	20.5 (12.2-31.2)	9.0 (3.7-17.6)	2.6 (0.3-9.0)	32.1 (21.9-43.6)	11.5 (5.4-20.8)
Eye problems (scale)	13.4 (8.8-19.1)	Na	Na	Na	86.6 (80.9-91.2)	Na
Watery eyes	45.9 (38.6-53.4)	28.1 (21.8-35.2)	21.6 (15.9-28.3)	4.3 (1.9-8.3)	54.1 (46.6-61.4)	25.9 (19.8-32.9)
Burning eyes	39.7 (32.6-47.1)	39.7 (32.6-47.1)	19.0 (13.6-25.4)	1.6 (0.3-4.7)	60.3 (52.9-67.4)	20.7 (15.0-27.2)
Light sensitive eyes	43.7 (36.4-51.2)	32.8 (26.0-40.1)	17.5 (12.3-23.8)	6.0 (3.0-10.5)	56.3 (48.8-63.6)	23.5 (17.6-30.3)
Blurry eyes	41.3 (34.1-48.8)	43.5 (36.2-51.0)	13.6 (9.0-19.4)	1.6 (0.3-4.7)	58.7 (51.2-65.9)	15.2 (10.4-21.2)
Red eyes	73.8 (66.8-80.0)	18.0 (12.8-24.4)	7.1 (3.8-11.8)	1.1 (0.1-3.9)	26.2 (20.0-33.2)	8.2 (4.7-13.2)
Swelling face/eyes	54.8 (47.4-62.0)	22.3 (16.6-29.0)	17.0 (11.9-23.2)	5.9 (3.0-10.2)	45.2 (38.0-52.6)	22.9 (17.1-29.5)
Swelling body parts	62.0 (54.5-69.1)	24.6 (18.5-31.6)	11.7 (7.4-17.4)	1.7 (0.3-4.8)	38.0 (30.9-45.5)	13.4 (8.8-19.3)
Hearing problems	76.3 (69.6-82.3)	15.6 (10.7-21.6)	5.9 (3.0-10.3)	2.2 (0.6-5.4)	23.7 (17.7-30.4)	8.1 (4.6-13.0)
Nose bleeds	87.6 (82.0-92.0)	10.8 (6.7-16.1)	0.5 (0.0-3.0)	1.1 (0.1-3.8)	12.4 (8.0-18.0)	1.6 (0.3-4.6)
Nose problems	52.2 (39.8-64.4)	21.7 (12.7-33.3)	24.6 (15.1-36.5)	1.4 (0.0-7.8)	47.8 (35.6-60.2)	26.1 (16.3-38.1)
Dry mouth	54.1 (46.6-61.4)	29.2 (22.8-36.3)	10.8 (6.7-16.2)	5.9 (3.0-10.4)	45.9 (38.6-53.4)	16.8 (11.7-22.9)
Pain/soreness mouth	77.4 (70.7-83.2)	16.7 (11.6-22.8)	4.3 (1.9-8.3)	1.6 (0.3-4.6)	22.6 (16.8-29.3)	5.9 (3.0-10.3)
Teeth problems	75.5 (68.7-81.6)	17.4 (12.2-23.7)	6.0 (3.0-10.4)	1.1 (0.1-3.9)	24.5 (18.4-31.3)	7.1 (3.8-11.8)
Bleeding gums	78.8 (72.2-84.5)	17.9 (12.7-24.3)	2.7 (0.9-6.2)	0.5 (0.0-3.0)	21.2 (15.5-27.8)	3.3 (1.2-7.0)
Taste change	73.9 (66.9-80.1)	16.8 (11.7-23.1)	6.0 (3.0-10.4)	3.3 (1.2-7.0)	26.1 (19.9-33.1)	9.2 (5.5-14.4)
Appetite loss	71.2 (64.1-77.6)	18.5 (13.1-24.9)	7.6 (4.2-12.4)	2.7 (0.9-6.2)	28.8 (22.4-35.9)	10.3 (6.3-15.7)
Mood problems (scale)	19.5 (14.2-25.7)	Na	Na	Na	80.5 (74.3-85.8)	Na

Table 24 Continued

Item/scale	Any TKI					
	Not at all	A little	Quite a bit	A lot	Any severity	Moderate-severe $
Depressed	59.0 (51.5-66.2)	25.7 (19.5-32.6)	10.9 (6.8-16.4)	4.4 (1.9-8.4)	41.0 (33.8-48.5)	15.3 (10.4-21.3)
Irritable	52.5 (45.0-59.9)	32.2 (25.5-39.5)	12.0 (7.7-17.6)	3.3 (1.2-7.0)	47.5 (40.1-55.0)	15.3 (10.4-21.3)
Tense	51.6 (44.1-59.1)	33.5 (26.7-40.9)	10.4 (6.4-15.8)	4.4 (1.9-8.5)	48.4 (40.9-55.9)	14.8 (10.0-20.8)
Worried	41.5 (34.3-49.0)	42.1 (34.8-49.6)	13.1 (8.6-18.9)	3.3 (1.2-7.0)	58.5 (51.0-65.7)	16.4 (11.3-22.6)
Afraid	71.7 (64.5-78.1)	18.9 (13.5-25.4)	6.7 (3.5-11.4)	2.8 (0.9-6.4)	28.3 (21.9-35.5)	9.4 (5.6-14.7)
Insomnia	42.7 (35.7-49.9)	30.2 (23.9-37.0)	19.1 (13.9-25.3)	8.0 (4.7-12.7)	57.3 (50.1-64.3)	27.1 (21.1-33.9)
Palpitations	56.0 (48.5-63.3)	28.8 (22.4-35.9)	12.5 (8.1-18.2)	2.7 (0.9-6.2)	44.0 (36.7-51.5)	15.2 (10.4-21.2)
Pain in chest	79.1 (72.5-84.8)	17.0 (11.9-23.3)	2.7 (0.9-6.3)	1.1 (0.1-3.9)	20.9 (15.2-27.5)	3.8 (1.6-7.8)
Short of breath	50.5 (43.1-58.0)	32.1 (25.4-39.3)	14.7 (9.9-20.6)	2.7 (0.9-6.2)	49.5 (42.0-56.9)	17.4 (12.2-23.7)
Coughed	67.0 (59.7-73.8)	21.4 (15.7-28.1)	9.3 (5.5-14.5)	2.2 (0.6-5.5)	33.0 (26.2-40.3)	11.5 (7.3-17.1)
Feeling bloated	49.5 (42.0-56.9)	28.3 (21.9-35.4)	16.8 (11.7-23.1)	5.4 (2.6-9.8)	50.5 (43.1-58.0)	22.3 (16.5-29.0)
Indigestion	59.9 (52.4-67.1)	22.5 (16.7-29.3)	12.1 (7.7-17.7)	5.5 (2.7-9.9)	40.1 (32.9-47.6)	17.6 (12.3-23.9)
Nausea	61.1 (53.6-68.3)	27.2 (20.9-34.3)	10.0 (6.0-15.3)	1.7 (0.3-4.8)	38.9 (31.7-46.4)	11.7 (7.4-17.3)
Diarrhoea	61.0 (53.5-68.1)	22.5 (16.7-29.3)	7.7 (4.3-12.6)	8.8 (5.1-13.9)	39.0 (31.9-46.5)	16.5 (11.4-22.7)
Abdominal pain/cramps	51.7 (44.1-59.2)	31.1 (24.4-38.4)	11.7 (7.4-17.3)	5.6 (2.7-10.0)	48.3 (40.8-55.9)	17.2 (12.0-23.5)
Vomited	86.2 (80.3-90.9)	11.6 (7.3-17.2)	2.2 (0.6-5.6)	0.0 (0.0-2.0)	13.8 (9.1-19.7)	2.2 (0.6-5.6)
Obstipation	66.3 (58.9-73.1)	23.2 (17.3-30.0)	5.0 (2.3-9.2)	5.5 (2.7-9.9)	33.7 (26.9-41.1)	10.5 (6.4-15.9)
Flatulence	33.5 (26.7-40.9)	33.0 (26.1-40.4)	24.6 (18.5-31.6)	8.9 (5.2-14.1)	66.5 (59.1-73.3)	33.5 (26.7-40.9)
Muscle/joint cramps/pain (scale)	15.0 (10.1-21.1)	Na	Na	Na	85.0 (78.9-89.9)	Na
Frequent urination	42.5 (35.1-50.1)	27.4 (21.0-34.5)	22.9 (17.0-29.8)	7.3 (3.9-12.1)	57.5 (49.9-64.9)	30.2 (23.5-37.5)

Table 24 Continued

Item/scale	Any TKI					
	Not at all	A little	Quite a bit	A lot	Any severity	Moderate-severe $
Muscle cramps	35.2 (28.2-42.7)	37.4 (30.3-45.0)	17.3 (12.1-23.7)	10.1 (6.1-15.4)	64.8 (57.3-71.8)	27.4 (21.0-34.5)
Aches/pains muscles/ joints	25.7 (19.5-32.8)	30.7 (24.1-38.0)	30.7 (24.1-38.0)	12.8 (8.3-18.7)	74.3 (67.2-80.5)	43.6 (36.2-51.2)
Muscle weakness	57.3 (49.7-64.7)	28.7 (22.1-35.9)	12.4 (7.9-18.1)	1.7 (0.3-4.8)	42.7 (35.3-50.3)	14.0 (9.3-20.0)
Pale/cold fingers/toes	58.0 (45.5-69.8)	15.9 (8.2-26.7)	13.0 (6.1-23.3)	13.0 (6.1-23.3)	42.0 (30.2-54.5)	26.1 (16.3-38.1)
Tingling/numbness hands/feet	45.8 (38.4-53.4)	33.5 (26.7-40.9)	14.0 (9.2-19.9)	6.7 (3.5-11.4)	54.2 (46.6-61.6)	20.7 (15.0-27.3)
Problems finger dexterity	71.9 (64.7-78.4)	21.9 (16.1-28.7)	4.5 (2.0-8.7)	1.7 (0.3-4.8)	28.1 (21.6-35.3)	6.2 (3.1-10.8)
Trouble walking pain	58.7 (51.1-66.0)	25.7 (19.5-32.8)	8.9 (5.2-14.1)	6.7 (3.5-11.4)	41.3 (34.0-48.9)	15.6 (10.7-21.8)
Back pain	46.4 (34.3-58.8)	37.7 (26.3-50.2)	10.1 (4.2-19.8)	5.8 (1.6-14.2)	53.6 (41.2-65.7)	15.9 (8.2-26.7)
Pain	48.3 (40.8-55.9)	34.4 (27.5-41.9)	13.3 (8.7-19.2)	3.9 (1.6-7.8)	51.7 (44.1-59.2)	17.2 (12.0-23.5)
Fever/chills	76.0 (69.0-82.0)	19.6 (14.0-26.1)	3.4 (1.2-7.2)	1.1 (0.1-4.0)	24.0 (18.0-31.0)	4.5 (1.9-8.6)
Sweating excessively	62.2 (54.7-69.3)	23.9 (17.9-30.8)	12.2 (7.8-17.9)	1.7 (0.3-4.8)	37.8 (30.7-45.3)	13.9 (9.2-19.8)
Problems tolerating heat/ cold	55.9 (48.3-63.3)	31.3 (24.6-38.6)	8.9 (5.2-14.1)	3.9 (1.6-7.9)	44.1 (36.7-51.7)	12.8 (8.3-18.7)
Hot flushes	73.7 (66.7-80.0)	17.3 (12.1-23.7)	5.6 (2.7-10.0)	3.4 (1.2-7.2)	26.3 (20.0-33.3)	8.9 (5.2-14.1)
Airway infections	79.7 (68.3-88.4)	13.0 (6.1-23.3)	2.9 (0.4-10.1)	4.3 (0.9-12.2)	20.3 (11.6-31.7)	7.2 (2.4-16.1)
Other infections	88.4 (78.4-94.9)	5.8 (1.6-14.2)	2.9 (0.4-10.1)	2.9 (0.4-10.1)	11.6 (5.1-21.6)	5.8 (1.6-14.2)
Sore/enlarged nipples/ breasts	90.4 (85.1-94.3)	9.0 (5.2-14.2)	0.6 (0.0-3.1)	0.0 (0.0-2.1)	9.6 (5.7-14.9)	0.6 (0.0-3.1)
Sexuality problems (scale)	31.8 (24.9-39.3)	Na	Na	Na	68.2 (60.7-75.1)	Na

Table 24 Continued

Item/scale	Any TKI					
	Not at all	A little	Quite a bit	A lot	Any severity	Moderate-severe $
Less interest sex	35.3 (28.2-42.9)	30.1 (23.3-37.5)	23.1 (17.1-30.1)	11.6 (7.2-17.3)	64.7 (57.1-71.8)	34.7 (27.6-42.3)
Less pleasure sex	45.3 (37.7-53.1)	27.6 (21.1-35.0)	18.8 (13.2-25.5)	8.2 (4.6-13.4)	54.7 (46.9-62.3)	27.1 (20.5-34.4)
Pain during/after sex	77.4 (70.3-83.5)	14.9 (9.9-21.2)	4.8 (2.1-9.2)	3.0 (1.0-6.8)	22.6 (16.5-29.7)	7.7 (4.2-12.9)
Difficult erection	53.9 (42.1-65.5)	28.9 (19.1-40.5)	10.5 (4.7-19.7)	6.6 (2.2-14.7)	46.1 (34.5-57.9)	17.1 (9.4-27.5)
Dry vagina	48.8 (37.9-59.9)	25.6 (16.8-36.1)	18.6 (11.0-28.4)	7.0 (2.6-14.6)	51.2 (40.1-62.1)	25.6 (16.8-36.1)

* At T4 as not available at T1. $ Moderate-severe includes 'Quite a bit' and 'A lot'. Abbreviations: Na: not applicable

Table 25 Mean and median symptoms scores for 7 scales plus 43 single items, in 12 patients on asciminib, and 14 patients on bosutinib at T1

Item	Asciminib						Bosutinib					
	Mean	Lower 95%CI	Higher 95%CI	Median	IQR	Skewness	Mean	Lower 95%CI	Higher 95%CI	Median	IQR	Skewness
Fatigue (scale)	40.2	23.8	56.7	38.9	26.4	0.5	42.8	29.0	56.6	41.6	40.2	0.4
Headaches	26.6	11.6	41.7	33.3	25.0	0.1	12.1	0.0	27.2	0.0	16.7	1.5
Dizzy	33.3	4.7	61.9	33.3	66.6	0.6	12.1	0.8	23.4	0.0	33.3	0.6
Cognitive problems (scale)	25.9	9.9	41.9	22.2	22.2	0.7	39.4	18.7	60.0	33.3	55.5	0.1
Bruised easily	33.3	0.0	67.2	0.0	66.6	0.7	33.3	8.8	57.8	33.3	66.6	0.5
Skin problems (scale)	28.9	11.1	46.6	26.6	20.0	0.7	27.9	12.1	43.6	26.6	33.3	0.3
Wounds healing problems	7.4	0.0	18.7	0.0	0.0	1.3	18.2	0.0	39.1	0.0	33.3	1.8
Hair loss	22.2	0.0	47.8	0.0	33.3	1.5	33.3	13.3	53.3	33.3	16.7	0.9
Nail problems*	16.7	0.0	45.9	0.0	25.0	1.1	16.7	0.0	47.2	16.7	33.3	0.0
Eye problems (scale)	28.5	12.1	44.8	33.3	33.3	-0.3	21.8	10.3	33.3	13.3	20.0	0.6
Swelling face/eyes	14.8	0.0	40.8	0.0	0.0	2.1	15.1	0.0	30.5	0.0	33.3	1.1
Swelling body parts	4.2	0.0	14.0	0.0	0.0	2.3	30.3	9.2	51.4	33.3	66.6	0.2
Hearing problems	16.7	0.0	42.4	0.0	16.7	1.2	23.3	0.0	48.6	0.0	33.3	1.2
Nose bleeds	25.0	0.0	53.8	16.7	33.3	1.3	9.1	0.0	23.5	0.0	0.0	2.1
Other nose problems*	40.0	0.0	85.3	66.6	66.6	-0.4	46.6	0.0	100.0	66.6	66.6	-0.1
Dry mouth	37.5	9.9	65.1	33.3	16.7	0.7	24.2	6.6	41.8	33.3	33.3	0.5
Pain/soreness mouth	25.0	0.3	49.7	16.7	41.6	0.5	15.1	0.0	36.1	0.0	16.7	2.1
Teeth problems	16.7	0.0	37.7	0.0	33.3	1.1	20.0	0.0	43.0	0.0	50.0	0.9
Bleeding gums	8.3	0.0	28.0	0.0	0.0	2.3	16.7	0.0	39.8	0.0	25.0	1.9
Taste change	25.0	5.3	44.7	33.3	33.3	0.3	21.2	0.0	44.2	0.0	33.3	1.4

Table 25 Continued

Item	Asciminib						Bosutinib					
	Mean	Lower 95%CI	Higher 95%CI	Median	IQR	Skewness	Mean	Lower 95%CI	Higher 95%CI	Median	IQR	Skewness
Appetite loss	33.3	3.5	63.1	33.3	41.6	0.8	18.2	0.0	39.1	0.0	33.3	1.8
Mood problems (scale)	37.2	16.2	58.2	33.3	33.3	0.8	22.8	9.2	36.5	16.7	40.2	0.5
Palpitations	33.3	3.5	63.1	33.3	41.6	0.8	21.2	0.5	41.9	0.0	33.3	1.6
Pain in chest	25.0	0.0	53.8	16.7	33.3	1.3	15.1	0.0	30.5	0.0	33.3	1.1
Short of breath	45.8	16.3	75.3	50.0	41.6	0.0	30.3	6.9	53.7	33.3	50.0	0.7
Coughed	45.8	16.3	75.3	50.0	41.6	0.0	27.2	7.7	46.8	33.3	50.0	0.4
Feeling bloated	45.8	3.9	87.8	33.3	100.0	0.1	42.4	15.7	69.1	33.3	66.6	0.2
Indigestion	33.3	3.5	63.1	33.3	41.6	0.8	39.4	9.6	69.1	33.3	83.3	0.5
Nausea	29.1	1.5	56.7	16.7	66.6	0.3	20.0	0.0	45.6	0.0	25.0	1.4
Vomited	8.3	0.0	28.0	0.0	0.0	2.3	12.1	0.0	27.2	0.0	16.7	1.5
Diarrhoea	25.0	0.0	63.7	0.0	25.0	1.2	33.3	5.0	61.6	0.0	66.6	0.6
Obstipation	29.2	1.5	56.8	33.3	33.3	1.2	20.0	0.0	43.0	0.0	33.3	1.7
Abdominal pain/cramps	41.6	9.2	74.1	50.0	66.6	0.1	43.3	9.5	77.1	33.3	91.7	0.2
Flatulence	29.2	1.5	56.8	33.3	33.3	1.2	70.0	49.1	90.9	66.6	58.4	-0.2
Frequent urination	37.5	0.0	75.3	16.7	75.0	0.5	30.0	12.4	47.5	33.3	25.0	0.1
Muscle/joint cramps/pain (scale)	36.2	17.4	55.0	36.6	32.5	0.2	41.3	20.8	61.8	46.6	56.7	-0.2
Pail/cold fingers/toes*	20.0	0.0	75.5	0.0	0.0	1.5	40.0	0.0	100.0	0.0	100.0	0.4
Tingling/numbness hands/ feet	20.8	0.1	41.5	16.7	33.3	0.7	23.3	0.7	45.9	16.7	33.3	1.5
Problems finger dexterity	20.8	0.1	41.5	16.7	33.3	0.7	10.0	0.0	21.5	0.0	25.0	0.9

Table 25 Continued

Item	Asciminib						Bosutinib					
	Mean	Lower 95%CI	Higher 95%CI	Median	IQR	Skewness	Mean	Lower 95%CI	Higher 95%CI	Median	IQR	Skewness
Back pain*	33.3	0.0	84.0	33.3	33.3	0.9	20.0	0.0	57.0	0.0	33.3	0.8
Fever/chills	25.0	0.0	53.8	16.7	33.3	1.3	12.1	0.0	27.2	0.0	16.7	1.5
Sweating excessively	25.0	0.0	57.4	0.0	41.6	1.1	21.2	0.5	41.9	0.0	50.0	0.8
Problems tolerating heat/ cold	20.8	0.0	50.4	0.0	33.3	1.6	26.6	7.8	45.4	33.3	33.3	0.3
Hot flushes	20.8	0.0	50.4	0.0	33.3	1.6	16.7	0.0	33.5	0.0	33.3	1.0
Airway infections*	6.7	0.0	25.2	0.0	0.0	1.5	0.0	NA	NA	0.0	0.0	NA
Other infections*	0.0	NA	NA	0.0	0.0	NA	13.3	0.0	50.3	0.0	0.0	1.5
Sore/enlarged nipples/ breasts	0.0	NA	NA	0.0	0.0	NA	0.0	NA	NA	0.0	0.0	NA
Sexuality problems (scale)	25.0	8.7	41.2	27.8	36.1	-0.2	30.0	8.8	51.2	27.8	47.2	0.7
Difficulty getting/ maintaining erection	33.3	0.0	76.6	33.3	16.7	0.0	33.3	0.0	100.0	33.3	33.3	0.0
Dry vagina	8.3	0.0	34.8	0.0	8.3	1.2	29.1	5.9	52.4	33.3	41.6	0.2

* At T4 as not available at T1. Abbreviations: CI: confidence interval; IQR: inter quartile range. Shaded cells indicate a skewness of ≥2.0. A skewness of <1.0 is seen as excellent, and <2.0 as acceptable [9].

Table 26 Mean and median symptoms scores for 7 scales plus 43 single items, in 70 patients on dasatinib, and 84 patients on imatinib at T1

Item	Dasatinib						Imatinib					
	Mean	Lower 95%CI	Higher 95%CI	Median	IQR	Skewness	Mean	Lower 95%CI	Higher 95%CI	Median	IQR	Skewness
Fatigue (scale)	46.3	39.8	52.9	47.2	31.9	0.1	39.0	33.7	44.3	38.9	44.4	0.0
Headaches	23.7	16.1	31.3	0.0	33.3	1.1	12.6	7.7	17.5	0.0	33.3	1.4
Dizzy	14.9	9.0	20.9	0.0	33.3	1.5	11.3	7.2	15.3	0.0	33.3	1.2
Cognitive problems (scale)	30.5	23.6	37.5	22.2	33.3	0.7	21.1	16.3	25.9	11.1	33.3	0.6
Bruised easily	12.9	6.9	18.8	0.0	33.3	1.8	21.7	15.7	27.8	0.0	33.3	0.9
Skin problems (scale)	28.5	23.4	33.5	26.6	20.0	0.4	19.5	15.1	24.0	13.3	20.0	1.2
Wounds healing problems	11.1	5.0	17.2	0.0	0.0	2.1	10.2	5.7	14.7	0.0	33.3	2.2
Hair loss	26.8	19.1	34.4	33.3	41.6	0.6	11.6	6.6	16.5	0.0	33.3	1.9
Nail problems*	8.8	1.5	16.0	0.0	16.7	1.1	25.0	11.1	38.9	0.0	66.6	0.8
Eye problems (scale)	22.2	16.9	27.5	13.3	26.7	0.6	27.3	22.9	31.8	26.6	26.6	0.5
Swelling face/eyes	20.0	11.4	28.6	0.0	33.3	1.4	38.9	31.3	46.4	33.3	66.6	0.2
Swelling body parts	19.1	12.1	26.1	0.0	33.3	1.1	22.0	15.4	28.7	0.0	33.3	1.0
Hearing problems	7.7	3.2	12.2	0.0	0.0	2.1	13.9	8.0	19.7	0.0	33.3	1.8
Nose bleeds	3.0	0.4	5.5	0.0	0.0	2.9	3.3	0.9	5.6	0.0	0.0	2.7
Other nose problems*	25.0	11.7	38.2	16.7	41.6	0.5	20.8	10.0	31.6	0.0	33.3	0.7
Dry mouth	19.0	11.5	26.6	0.0	33.3	1.5	25.2	17.6	32.8	0.0	33.3	1.0
Pain/soreness mouth	9.5	4.2	14.8	0.0	0.0	2.5	9.9	5.2	14.5	0.0	0.0	2.3
Teeth problems	8.5	3.5	13.4	0.0	0.0	2.1	13.1	7.6	18.7	0.0	33.3	2.0
Bleeding gums	5.9	1.7	10.1	0.0	0.0	2.7	9.8	5.8	13.9	0.0	33.3	1.5
Taste change	9.7	3.8	15.6	0.0	0.0	2.4	15.5	8.8	22.1	0.0	33.3	1.8

Table 26 Continued

Item	Dasatinib						Imatinib					
	Mean	Lower 95%CI	Higher 95%CI	Median	IQR	Skewness	Mean	Lower 95%CI	Higher 95%CI	Median	IQR	Skewness
Appetite loss	10.7	5.3	16.1	0.0	8.3	1.7	17.1	10.4	23.9	0.0	33.3	1.5
Mood problems (scale)	24.2	18.3	30.1	16.7	27.8	1.2	24.3	18.6	30.1	16.7	33.3	1.2
Palpitations	23.2	15.4	31.0	0.0	33.3	1.0	20.0	14.2	25.8	0.0	33.3	0.8
Pain in chest	7.1	3.1	11.2	0.0	0.0	2.0	8.3	3.8	12.8	0.0	0.0	2.7
Short of breath	20.8	14.3	27.3	16.7	33.3	1.0	27.1	20.3	33.9	33.3	33.3	0.6
Coughed	13.1	6.3	19.8	0.0	33.3	2.0	13.2	7.8	18.6	0.0	33.3	1.7
Feeling bloated	30.9	23.0	38.9	33.3	41.6	0.6	22.4	15.6	29.1	0.0	33.3	1.0
Indigestion	19.4	11.3	27.5	0.0	33.3	1.4	21.7	14.6	28.8	0.0	33.3	1.1
Nausea	16.1	8.8	23.3	0.0	33.3	1.6	20.1	14.8	25.3	33.3	33.3	0.6
Vomited	3.0	0.0	6.1	0.0	0.0	4.1	7.7	4.1	11.4	0.0	0.0	1.7
Diarrhoea	16.7	10.1	23.2	0.0	33.3	1.6	30.9	22.3	39.5	33.3	66.6	0.8
Obstipation	22.6	14.1	31.1	0.0	33.3	1.3	8.7	4.4	12.9	0.0	0.0	1.9
Abdominal pain/cramps	24.8	16.5	33.2	33.3	33.3	1.2	21.7	15.9	27.5	33.3	33.3	0.6
Flatulence	34.5	25.4	43.7	33.3	66.6	0.5	36.7	29.5	44.0	33.3	66.6	0.4
Frequent urination	35.7	26.1	45.4	33.3	66.6	0.5	29.9	22.5	37.3	33.3	66.6	0.6
Muscle/joint cramps/pain (scale)	25.5	19.7	31.3	25.0	33.3	0.9	31.3	25.7	36.9	26.6	34.1	0.5
Pail/cold fingers/toes*	40.0	20.6	59.3	33.3	66.6	0.3	26.4	12.0	40.7	0.0	41.6	0.9
Tingling/numbness hands/ feet	35.1	25.3	45.0	33.3	66.6	0.6	26.4	19.4	33.5	33.3	33.3	0.8
Problems finger dexterity	11.7	5.3	18.1	0.0	0.0	2.0	14.2	8.4	20.0	0.0	33.3	1.8

Table 26 Continued

Item	Dasatinib						Imatinib					
	Mean	Lower 95%CI	Higher 95%CI	Median	IQR	Skewness	Mean	Lower 95%CI	Higher 95%CI	Median	IQR	Skewness
Back pain*	31.6	19.8	43.5	33.3	41.6	0.1	16.7	6.5	26.8	0.0	33.3	1.8
Fever/chills	11.7	6.1	17.3	0.0	33.3	1.5	8.8	4.3	13.4	0.0	0.0	2.5
Sweating excessively	14.5	8.3	20.7	0.0	33.3	1.3	17.1	11.2	23.1	0.0	33.3	1.3
Problems tolerating heat/ cold	21.8	14.6	29.0	0.0	33.3	1.1	21.6	14.9	28.2	0.0	33.3	1.2
Hot flushes	17.0	9.6	24.3	0.0	33.3	1.5	7.3	3.0	11.7	0.0	0.0	3.0
Airway infections*	6.7	0.0	14.8	0.0	0.0	2.5	20.8	6.0	35.7	0.0	33.3	1.5
Other infections*	8.3	0.0	19.5	0.0	0.0	3.1	8.3	0.0	18.7	0.0	0.0	2.9
Sore/enlarged nipples/ breasts	2.5	0.1	4.9	0.0	0.0	3.3	4.5	1.7	7.3	0.0	0.0	2.1
Sexuality problems (scale)	29.1	21.3	36.9	22.2	33.3	1.2	25.3	18.5	32.1	22.2	44.4	0.9
Difficulty getting/ maintaining erection	19.0	5.0	33.1	0.0	33.3	1.7	18.6	9.0	28.2	0.0	33.3	1.3
Dry vagina	30.8	16.3	45.4	0.0	66.6	0.7	24.1	12.4	35.8	0.0	33.3	0.9

* At T4 as not available at T1. Abbreviations: CI: confidence interval; IQR: inter quartile range. Shaded cells indicate a skewness of ≥2.0. A skewness of <1.0 is seen as excellent, and <2.0 as acceptable [9].

Table 27 Mean and median symptoms scores for 7 scales plus 43 single items, in 39 patients on nilotinib, and 8 patients on ponatinib at T1

Item	Nilotinib						Ponatinib					
	Mean	Lower 95%CI	Higher 95%CI	Median	IQR	Skewness	Mean	Lower 95%CI	Higher 95%CI	Median	IQR	Skewness
Fatigue (scale)	38.5	31.2	45.7	38.9	33.3	0.0	40.5	20.2	60.8	40.0	31.9	-0.2
Headaches	7.4	1.9	12.9	0.0	0.0	2.1	4.8	0.0	16.4	0.0	0.0	2.0
Dizzy	12.0	5.3	18.7	0.0	33.3	1.4	9.5	0.0	24.5	0.0	16.7	0.9
Cognitive problems (scale)	16.5	10.4	22.6	11.1	33.3	0.9	28.5	11.9	45.2	33.3	11.1	-0.3
Bruised easily	9.5	4.3	14.8	0.0	33.3	0.9	4.8	0.0	16.4	0.0	0.0	2.0
Skin problems (scale)	24.2	16.8	31.7	20.0	36.7	0.6	28.5	13.6	43.5	26.6	26.6	-0.1
Wounds healing problems	1.0	0.0	2.9	0.0	0.0	5.7	5.6	0.0	19.8	0.0	0.0	1.8
Hair loss	24.5	13.7	35.3	16.7	33.3	1.2	9.5	0.0	24.5	0.0	16.7	0.9
Nail problems*	5.1	0.0	12.7	0.0	0.0	1.9	0.0	NA	NA	0.0	0.0	NA
Eye problems (scale)	21.3	15.4	27.2	20.0	24.6	1.4	17.1	4.4	29.9	20.0	16.7	0.4
Swelling face/eyes	11.8	4.2	19.3	0.0	25.0	1.6	4.8	0.0	16.4	0.0	0.0	2.0
Swelling body parts	7.3	0.6	13.9	0.0	0.0	2.4	5.6	0.0	19.8	0.0	0.0	1.8
Hearing problems	9.1	1.1	17.1	0.0	0.0	2.7	0.0	NA	NA	0.0	0.0	NA
Nose bleeds	7.1	0.0	14.2	0.0	0.0	3.4	0.0	NA	NA	0.0	0.0	NA
Other nose problems*	19.0	2.7	35.4	0.0	33.3	0.9	33.3	NA	NA	33.3	0.0	NA
Dry mouth	18.2	7.9	28.4	0.0	33.3	1.6	33.3	15.5	51.1	33.3	0.0	0.0
Pain/soreness mouth	7.1	0.0	14.1	0.0	0.0	2.6	4.8	0.0	16.4	0.0	0.0	2.0
Teeth problems	4.0	0.0	8.9	0.0	0.0	3.5	19.0	2.6	35.5	33.3	33.3	-0.3
Bleeding gums	5.2	0.8	9.6	0.0	0.0	1.9	14.3	0.0	30.7	0.0	33.3	0.3
Taste change	7.1	1.3	12.8	0.0	0.0	2.2	11.1	0.0	39.6	0.0	0.0	1.8

Table 27 Continued

Item	Nilotinib						Ponatinib					
	Mean	Lower 95%CI	Higher 95%CI	Median	IQR	Skewness	Mean	Lower 95%CI	Higher 95%CI	Median	IQR	Skewness
Appetite loss	7.1	0.6	13.5	0.0	0.0	2.5	11.1	0.0	29.1	0.0	25.0	0.7
Mood problems (scale)	16.4	11.1	21.7	11.1	22.2	1.0	16.7	5.6	27.7	22.2	16.7	-0.6
Palpitations	15.1	5.8	24.5	0.0	33.3	1.7	16.7	0.0	35.8	16.7	33.3	0.0
Pain in chest	6.1	0.0	12.3	0.0	0.0	2.8	5.6	0.0	19.8	0.0	0.0	1.8
Short of breath	14.1	6.3	22.0	0.0	33.3	1.3	5.6	0.0	19.8	0.0	0.0	1.8
Coughed	16.1	7.7	24.6	0.0	33.3	1.1	0.0	NA	NA	0.0	0.0	NA
Feeling bloated	14.1	6.3	22.0	0.0	33.3	1.3	33.3	2.0	64.6	33.3	50.0	0.0
Indigestion	11.1	3.5	18.7	0.0	0.0	1.7	33.3	11.2	55.4	33.3	0.0	0.0
Nausea	10.4	3.3	17.5	0.0	8.3	1.7	16.7	0.0	45.9	0.0	25.0	1.1
Vomited	2.1	0.0	5.0	0.0	0.0	3.6	0.0	NA	NA	0.0	0.0	NA
Diarrhoea	7.3	0.0	14.6	0.0	0.0	3.4	5.6	0.0	19.8	0.0	0.0	1.8
Obstipation	18.7	6.9	30.5	0.0	33.3	1.6	16.7	0.0	35.8	16.7	33.3	0.0
Abdominal pain/cramps	16.7	7.0	26.3	0.0	33.3	1.5	16.7	0.0	35.8	16.7	33.3	0.0
Flatulence	32.3	21.1	43.5	33.3	66.6	0.3	22.2	0.0	50.7	16.7	33.3	0.6
Frequent urination	26.0	14.7	37.3	0.0	66.6	0.7	38.9	12.5	65.2	33.3	25.0	-0.2
Muscle/joint cramps/pain (scale)	23.1	15.0	31.1	20.0	28.3	0.9	21.1	10.8	31.4	20.0	5.0	1.3
Pail/cold fingers/toes*	7.1	0.0	15.3	0.0	0.0	1.4	33.3	NA	NA	33.3	0.0	NA
Tingling/numbness hands/ feet	16.7	9.2	24.1	0.0	33.3	0.8	33.3	11.2	55.4	33.3	0.0	0.0
Problems finger dexterity	6.2	1.5	11.0	0.0	0.0	1.6	11.1	0.0	29.1	0.0	25.0	0.7

Table 27 Continued

Item	Nilotinib						Ponatinib					
	Mean	Lower 95%CI	Higher 95%CI	Median	IQR	Skewness	Mean	Lower 95%CI	Higher 95%CI	Median	IQR	Skewness
Back pain*	30.9	10.3	51.6	33.3	33.3	0.9	0.0	NA	NA	0.0	0.0	NA
Fever/chills	4.2	0.1	8.2	0.0	0.0	2.3	11.1	0.0	29.1	0.0	25.0	0.7
Sweating excessively	22.9	12.6	33.2	0.0	33.3	1.0	11.1	0.0	29.1	0.0	25.0	0.7
Problems tolerating heat/ cold	14.6	5.0	24.2	0.0	33.3	1.7	11.1	0.0	29.1	0.0	25.0	0.7
Hot flushes	13.5	3.0	24.0	0.0	0.0	2.1	16.7	0.0	45.9	0.0	25.0	1.1
Airway infections*	4.8	0.0	11.7	0.0	0.0	2.0	0.0	NA	NA	0.0	0.0	NA
Other infections*	2.4	0.0	7.5	0.0	0.0	3.3	0.0	NA	NA	0.0	0.0	NA
Sore/enlarged nipples/ breasts	5.2	0.0	10.6	0.0	0.0	2.9	0.0	NA	NA	0.0	0.0	NA
Sexuality problems (scale)	26.5	16.7	36.4	22.2	44.4	0.6	29.6	6.7	52.5	22.2	25.0	0.9
Difficulty getting/ maintaining erection	41.6	19.3	64.0	33.3	41.6	0.5	11.1	0.0	58.9	0.0	16.7	0.7
Dry vagina	33.3	16.2	50.4	33.3	33.3	0.8	33.3	NA	NA	33.3	0.0	NA

* At T4 as not available at T1. Abbreviations: CI: confidence interval; IQR: inter quartile range. Shaded cells indicate a skewness of ≥2.0. A skewness of <1.0 is seen as excellent, and <2.0 as acceptable [9].

Table 28 Mean and median symptoms scores for 7 scales plus 43 single items, in 40 patients in treatment-free remission at T1, and 227 patients on any TKI at T1

Item	Treatment-free remission						Any TKI					
	Mean	Lower 95%CI	Higher 95%CI	Median	IQR	Skewness	Mean	Lower 95%CI	Higher 95%CI	Median	IQR	Skewness
Fatigue (scale)	40.5	32.3	48.8	41.6	34.0	0.2	41.5	38.2	44.8	38.9	38.9	0.1
Headaches	14.8	5.7	23.9	0.0	33.3	1.7	15.4	12.1	18.7	0.0	33.3	1.5
Dizzy	12.4	4.5	20.3	0.0	33.3	2.1	13.5	10.5	16.5	0.0	33.3	1.5
Cognitive problems (scale)	22.5	14.7	30.4	22.2	33.3	1.3	24.6	21.3	27.9	22.2	33.3	0.8
Bruised easily	6.5	1.2	11.7	0.0	0.0	2.4	17.4	13.8	21.1	0.0	33.3	1.4
Skin problems (scale)	20.6	14.5	26.7	13.3	23.3	0.8	24.3	21.4	27.2	20.0	26.6	0.7
Wounds healing problems	7.4	0.2	14.6	0.0	0.0	3.1	8.9	6.1	11.7	0.0	0.0	2.5
Hair loss	13.9	6.1	21.7	0.0	33.3	1.4	20.1	16.2	24.0	0.0	33.3	1.3
Nail problems*	20.0	0.0	43.0	0.0	33.3	1.7	14.7	8.7	20.7	0.0	33.3	1.5
Eye problems (scale)	16.2	10.1	22.3	13.3	26.6	1.0	24.1	21.4	26.8	20.0	30.0	0.6
Swelling face/eyes	12.7	2.8	22.7	0.0	0.0	2.2	24.6	20.1	29.1	0.0	33.3	1.0
Swelling body parts	23.6	13.6	33.7	33.3	33.3	0.9	17.7	13.9	21.4	0.0	33.3	1.2
Hearing problems	17.6	6.1	29.2	0.0	25.0	1.6	11.3	8.0	14.6	0.0	0.0	2.1
Nose bleeds	4.9	0.0	10.0	0.0	0.0	3.0	5.0	2.8	7.2	0.0	0.0	3.8
Other nose problems*	29.1	11.3	47.0	33.3	8.3	0.1	25.1	18.0	32.1	0.0	66.6	0.6
Dry mouth	17.6	8.9	26.3	0.0	33.3	1.0	22.9	18.6	27.2	0.0	33.3	1.2
Pain/soreness mouth	9.8	3.7	15.9	0.0	25.0	1.5	10.0	7.0	13.1	0.0	0.0	2.3
Teeth problems	10.8	3.9	17.6	0.0	25.0	1.6	10.9	7.8	13.9	0.0	0.0	2.0
Bleeding gums	7.1	0.6	13.5	0.0	0.0	2.5	8.3	5.8	10.9	0.0	0.0	2.2
Taste change	3.9	0.0	8.7	0.0	0.0	3.6	12.9	9.2	16.5	0.0	33.3	2.0

Table 28 Continued

Item	Treatment-free remission						Any TKI					
	Mean	Lower 95%CI	Higher 95%CI	Median	IQR	Skewness	Mean	Lower 95%CI	Higher 95%CI	Median	IQR	Skewness
Appetite loss	4.0	0.0	8.9	0.0	0.0	3.5	13.9	10.3	17.6	0.0	33.3	1.8
Mood problems (scale)	22.8	15.0	30.7	22.2	19.4	2.0	23.1	19.9	26.3	16.7	27.8	1.2
Palpitations	16.1	8.8	23.4	0.0	33.3	0.9	20.6	16.7	24.5	0.0	33.3	1.1
Pain in chest	4.2	0.1	8.2	0.0	0.0	2.3	8.6	5.9	11.3	0.0	0.0	2.5
Short of breath	16.7	6.5	26.8	0.0	33.3	1.5	23.2	19.2	27.1	0.0	33.3	0.9
Coughed	12.5	4.0	21.0	0.0	33.3	2.1	15.6	11.9	19.2	0.0	33.3	1.5
Feeling bloated	17.7	6.7	28.7	0.0	33.3	1.6	26.1	21.6	30.5	33.3	33.3	0.9
Indigestion	16.7	6.5	26.8	0.0	33.3	1.5	21.0	16.7	25.4	0.0	33.3	1.2
Nausea	4.3	0.1	8.5	0.0	0.0	2.2	17.4	13.8	21.0	0.0	33.3	1.3
Vomited	0.0	NA	NA	0.0	0.0	NA	5.3	3.3	7.4	0.0	0.0	2.7
Diarrhoea	11.5	3.0	19.9	0.0	8.3	2.3	21.4	16.8	26.1	0.0	33.3	1.4
Obstipation	10.7	2.8	18.7	0.0	16.7	2.5	16.6	12.5	20.6	0.0	33.3	1.8
Abdominal pain/cramps	11.4	4.9	18.0	0.0	33.3	1.3	23.7	19.4	28.0	0.0	33.3	1.1
Flatulence	21.9	11.5	32.2	0.0	33.3	1.0	36.3	31.5	41.0	33.3	66.6	0.4
Frequent urination	23.9	13.7	34.2	16.7	33.3	0.9	31.6	26.9	36.4	33.3	66.6	0.6
Muscle/joint cramps/pain (scale)	30.9	24.0	37.8	33.3	33.3	0.0	28.6	25.2	31.9	25.8	28.7	0.7
Pail/cold fingers/toes*	20.8	6.4	35.2	33.3	33.3	-0.5	27.0	18.2	35.9	0.0	66.6	1.0
Tingling/numbness hands/ feet	12.5	5.2	19.8	0.0	33.3	1.4	27.2	22.7	31.7	33.3	33.3	0.9
Problems finger dexterity	17.2	7.8	26.6	0.0	33.3	1.5	12.0	8.8	15.2	0.0	33.3	2.0

Table 28 Continued

Item	Treatment-free remission						Any TKI					
	Mean	Lower 95%CI	Higher 95%CI	Median	IQR	Skewness	Mean	Lower 95%CI	Higher 95%CI	Median	IQR	Skewness
Back pain*	37.5	6.1	68.8	33.3	66.6	0.4	25.1	18.2	32.0	33.3	33.3	1.0
Fever/chills	2.1	0.0	5.2	0.0	0.0	3.5	9.9	7.0	12.7	0.0	0.0	2.2
Sweating excessively	18.3	9.4	27.1	0.0	33.3	0.9	17.8	14.0	21.5	0.0	33.3	1.2
Problems tolerating heat/ cold	16.1	7.8	24.4	0.0	33.3	1.0	20.3	16.3	24.3	0.0	33.3	1.3
Hot flushes	10.7	4.1	17.3	0.0	33.3	1.4	12.8	9.2	16.5	0.0	33.3	2.0
Airway infections*	0.0	NA	NA	0.0	0.0	NA	10.6	4.7	16.5	0.0	0.0	2.5
Other infections*	4.2	0.0	14.0	0.0	0.0	2.3	6.8	1.7	11.8	0.0	0.0	3.3
Sore/enlarged nipples/ breasts	3.2	0.0	6.9	0.0	0.0	2.7	3.4	1.8	4.9	0.0	0.0	3.1
Sexuality problems (scale)	23.0	11.5	34.4	11.1	41.6	1.3	27.1	23.0	31.1	22.2	44.4	0.9
Difficulty getting/ maintaining erection	11.9	0.0	24.1	0.0	25.0	1.5	23.2	16.3	30.2	0.0	33.3	1.2
Dry vagina	35.5	15.2	55.8	33.3	66.6	0.5	27.9	21.0	34.8	33.3	58.3	0.8

* At T4 as not available at T1. Abbreviations: CI: confidence interval; IQR: inter quartile range. Shaded cells indicate a skewness of ≥2.0. A skewness of <1.0 is seen as excellent, and <2.0 as acceptable [9].

Our pilot data in the context of published literature

Meta-analysis of severity of symptoms

In a systematic search in PubMed and Embase, up to February 2025 [10] we identified eight other studies that used EORTC-instruments in CML patients, with mean severity data that we could compare to our data (Table 29). We excluded one Indian study from this comparison, as it's mean symptom severities were only a fraction of those reported in other studies and therefore lacked credibility [11]. For example, they reported a mean diarrhoea severity of 1.4, compared to a range of 17.9-21.7 in other studies. Similarly, for dyspnoea, they reported a mean severity of 1.3, compared to a range of 22.3-27.1 in other studies. Five further studies had data that could not be compared to other data, because either: means could not be extracted from the article [12-14]; it was unclear for which symptoms means/medians were reported [15]; or linear transformation had not been performed [16].

Symptom severity could be compared for eight symptoms/scales, for dasatinib, imatinib and nilotinib as current treatments. If only the standard deviation was given, we calculated the 95% confidence interval with the webtool www.openepi.com, with the population size set at 10,000, based on the z-test for sample sizes ≥30, and based on the t-test for sample sizes <30. For the 24 symptom-treatment combinations, there was no significant difference (defined as: overlapping 95%CIs of the unadjusted estimates) between our data and data from the literature in 28 combinations. In six symptom-treatment combinations there was a significant difference between our data and data from the literature, with differences between unadjusted estimates ranging from 6.5 to 21.3%. Notably, we could not compare our data to data collected with a validated questionnaire for most of the items in our instrument.

In a random effects meta-analysis, there was no statistical difference between TKI treatments for fatigue, pain and insomnia (Figure 1, Figure 3, and Figure 5), whereas:

- Patients on imatinib reported more severe nausea and/or vomiting compared to patients on dasatinib or nilotinib (unadjusted mean severity difference 5.4% and 7.9%, Figure 2)
- Patients on dasatinib and imatinib reported more severe dyspnoea than patients on nilotinib (unadjusted mean severity difference 8.8% and 11.4%, Figure 4)

- Patients on imatinib reported more severe appetite loss than patients on nilotinib (unadjusted mean severity difference 10.1%, Figure 6)
- Patients on dasatinib reported more severe constipation than patients on imatinib (unadjusted mean severity difference 12.6%, Figure 7)
- Patients on imatinib reported more severe diarrhoea than patients on nilotinib (unadjusted mean severity difference 17.1%, Figure 8)

Table 29 Comparison of mean symptom severity, measured with EORTC-items in patient groups ≥30 symptoms

Symptom, treatment	Study	Recruitment	Number of included patients	Mean (95%CI))
Fatigue scale EORTC *				
D	Efficace [17]	Hospital-based	94	29.1 (24.4-33.8)
	Foulon [18]	Web-based	66	60.3 (53.0-67.6)
	This study	Web-based	70	47.3 (40.5-54.2)
I	Ali [19]	Hospital-based	84	30.3 (26.7-33.9)
	Efficace [17]	Hospital-based	94	34.9 (29.7-40.1)
	Foulon [18]	Web-based	146	49.2 (44.9-53.5)
	Rashid [20]	Hospital-based	101	#35.8 (31.4-40.2)
	This study	Web-based	84	40.0 (34.4-45.6)
N	Ali [19]	Hospital-based	38	33.7 (27.6-39.8)
	Efficace [21]	Hospital based	124	$19.1 (15.3-22.8)
	Foulon [18]	Web-based	84	48.5 (42.9-54.1)
	Sacha [22]	Hospital-based	177	§ 26.6 (23.6-29.6)
	Nguyen [23]	Hospital-based	121	28.3 (24.3-32.3)
	This study	Web-based	39	40.4 (33.0-47.8)
D, I, N, O	Efficace [24]	Unclear	779	28.0 (27.2-28.8)
D, I, N	This study	Web-based	193	42.7 (39.0-46.5)
B, D, I, N, P, O, TFR	Foulon [18]	Web-based	377	50.0 (47.8-52.2)
B, D, I, N, P, TFR	This study	Web-based	255	42.5 (39.2-45.7)
Nausea/vomiting scale EORTC				
D	Efficace [17]	Hospital-based	94	6.1 (4.0-8.2)
	This study	Web-based	70	9.5 (5.4-13.7)

Table 29 Continued

Symptom, treatment	Study	Recruitment	Number of included patients	Mean (95%CI))
I	Ali [19]	Hospital-based	84	11.8 (9.3-14.3)
	Efficace [17]	Hospital-based	94	10.6 (7.2-14.0)
	Rashid [20]	Hospital-based	101	#13.4 (11.8-15.0)
	This study	Web-based	84	14.0 (10.4-17.6)
N	Ali [19]	Hospital-based	38	10.1 (5.7-14.5)
	Efficace [21]	Hospital based	124	$2.7 (1.3-4.1)
	Nguyen [23]	Hospital-based	121	7.2 (na)
	Sacha [22]	Hospital-based	177	§2.2 (1.3-3.1)
	This study	Web-based	39	6.2 (2.0-10.5)
D, I, N, O	Efficace [24]	Unclear	779	6.2 (5.8-6.6)
D, I, N	This study	Web-based	193	10.8 (8.5-13.2)
B, D, I, N, P, O, TFR	Foulon [18]	Web-based	377	13.7 (12.1-15.3)
B, D, I, N, P, TFR	This study	Web-based	255	9.7 (7.6-11.8)
Pain				
D	Efficace [17]	Hospital-based	94	14.7 (10.3-19.1)
	This study	Web-based	70	25.4 (17.9-33.0)
I	Ali [19]	Hospital-based	84	18.3 (13.6-23.0)
	Efficace [17]	Hospital-based	94	22.0 (17.4-26.6)
	Rashid [20]	Hospital-based	101	#28.7 (24.8-32.6)
	This study	Web-based	84	24.0 (17.3-30.7)
N	Ali [19]	Hospital-based	38	32.0 (23.2-40.8)
	Nguyen [23]	Hospital-based	121	21.1 (16.7-25.5)

Table 29 Continued

Symptom, treatment	Study	Recruitment	Number of included patients	Mean (95%CI))
	Sacha [22]	Hospital-based	177	§ 16.0 (13.0-19.0)
	This study	Web-based	39	18.7 (8.6-28.8)
Dyspnoea				
D	Efficace [17]	Hospital-based	94	21.4 (16.4-26.4)
	This study	Web-based	70	20.8 (14.3-27.3)
I	Ali [19]	Hospital-based	84	18.6 (13.6-23.6)
	Efficace [17]	Hospital-based	94	22.3 (17.0-27.6)
	Rashid [20]	Hospital-based	101	# 26.7 (24.3-29.1)
	This study	Web-based	84	27.1 (20.3-33.9)
N	Ali [19]	Hospital-based	38	15.9 (8.9-22.9)
	Nguyen [23]	Hospital-based	121	14.3 (na)
	Sacha [22]	Hospital-based	177	§ 11.3 (8.5-14.1)
	This study	Web-based	39	14.1 (6.3-22.0)
D, I, N, O	Efficace [24]	Unclear	779	17.0 (16.2-17.8)
D, I, N	This study	Web-based	193	22.2 (18.1-26.3)
B, D, I, N, P, O, TFR	Foulon [18]	Web-based	377	30.1 (27.7-32.8)
B, D, I, N, P, TFR	This study	Web-based	255	21.3 (17.6-25.0)
Insomnia				
D	Efficace [17]	Hospital-based	94	20.4 (15.0-25.8)
	This study	Web-based	70	29.9 (21.3-38.6)
I	Ali [19]	Hospital-based	84	20.8 (15.9-25.7)
	Efficace [17]	Hospital-based	94	27.0 (20.9-33.1)
	Rashid [20]	Hospital-based	101	# 15.7 (14.3-17.1)

Table 29 Continued

Symptom, treatment	Study	Recruitment	Number of included patients	Mean (95%CI))
	This study	Web-based	84	33.3 (25.5-41.1)
N	Ali [19]	Hospital-based	38	17.3 (10.3-24.3)
	Nguyen [23]	Hospital-based	121	22.9 (18.0-27.8)
	Sacha [22]	Hospital-based	177	§ 21.1 (17.9-24.3)
	This study	Web-based	39	28.8 (19.0-38.7)
Appetite loss				
D	Efficace [17]	Hospital-based	94	11.8 (7.0-16.6)
	This study	Web-based	70	10.7 (5.3-16.1)
I	Ali [19]	Hospital-based	84	15.2 (11.0-19.4)
	Efficace [17]	Hospital-based	94	12.8 (8.4-17.2)
	Rashid [20]	Hospital-based	101	# 23.4 (20.6-26.2)
	This study	Web-based	84	17.1 (0.4-23.9)
N	Ali [19]	Hospital-based	38	13.0 (6.8-19.2)
	Efficace [21]	Hospital based	124	$4.1 (1.4-6.8)
	Nguyen [23]	Hospital-based	121	17.1 (na)
	Sacha [22]	Hospital-based	177	§ 7.5 (5.2-9.8)
	This study	Web-based	39	7.1 (0.6-13.5)
D, I, N, O	Efficace [24]	Unclear	779	10.0 (9.2-10.8)
D, I, N	This study	Web-based	193	12.8 (9.0-16.5)
B, D, I, N, P, O, TFR	Foulon [18]	Web-based	377	14.7 (12.8-16.6)
B, D, I, N, P, TFR	This study	Web-based	255	11.6 (8.5-14.8)
Constipation				
D	Efficace [17]	Hospital-based	94	19.5 (14.7-24.3)

Table 29 Continued

Symptom, treatment	Study	Recruitment	Number of included patients	Mean (95%CI))
	This study	Web-based	70	22.6 (14.1-31.1)
I	Ali [19]	Hospital-based	84	2.2 (0.4-4.0)
	Efficace [17]	Hospital-based	94	10.9 (6.9-14.9)
	Rashid [20]	Hospital-based	101	#9.3 (8.5-10.1)
	This study	Web-based	84	8.7 (4.4-12.9)
N	Ali [19]	Hospital-based	38	26.0 (13.7-38.1)
	Nguyen [23]	Hospital-based	121	11.9 (na)
	Sacha [22]	Hospital-based	177	§10.5 (7.6-13.4)
	This study	Web-based	39	18.7 (6.9-30.3)
D, I, N, O	Efficace [24]	Unclear	779	9.9 (9.3-10.5)
D, I, N	This study	Web-based	193	15.7 (11.4-20.0)
B, D, I, N, P, O, TFR	Foulon [18]	Web-based	377	16.7 (14.5-18.9)
B, D, I, N, P, TFR	This study	Web-based	255	15.2 (11.5-18.8)
Diarrhoea				
D	Efficace [17]	Hospital-based	94	8.8 (4.8-12.8)
	This study	Web-based	70	16.7 (10.1-23.2)
I	Ali [19]	Hospital-based	84	24.8 (17.8-31.8)
	Efficace [17]	Hospital-based	94	17.9 (13.4-22.4)
	Rashid [20]	Hospital-based	101	#18.8 (16.2-21.4)
	This study	Web-based	84	30.9 (22.3-39.5)
N	Ali [19]	Hospital-based	38	8.7 (3.0-14.4)
	Nguyen [23]	Hospital-based	121	5.5 (na)
	Sacha [22]	Hospital-based	177	§1.4 (0.5-2.3)

Table 29 Continued

Symptom, treatment	Study	Recruitment	Number of included patients	Mean (95%CI))
	This study	Web-based	39	7.3 (0.0-14.6)
D, I, N, O	Efficace [24]	Unclear	779	12.0 (11.3-12.7)
D, I, N	This study	Web-based	193	21.0 (16.2-25.8)
B, D, I, N, P, O, TFR	Foulon [18]	Web-based	377	24.8 (22.3-27.3)
B, D, I, N, P, TFR	This study	Web-based	255	19.7 (15.6-23.9)

Grey shading indicates non-overlapping 95%CIs between studies. Red shaded areas indicate non-overlapping 95%CIs between studies, and a difference between unadjusted estimates of 15% or greater; * The EORTC QLQ-C30's Fatigue scale consists of three items (needed rest, felt weak, tired), whilst our Fatigue (scale) consists of six items (tired, need to rest, lacked energy, lethargic, weak, drowsy). For this comparison the EORTC's three-item Fatigue scale is used. $: after 24 months of treatment; #: after 12 months of treatment; §: at the sixth visit. Abbreviations: B: bosutinib, CI: confidence interval; D: dasatinib, I: imatinib, N: nilotinib, na: not available; O: other; P: ponatinib; SD: standard deviation; TFR: treatment-free remission

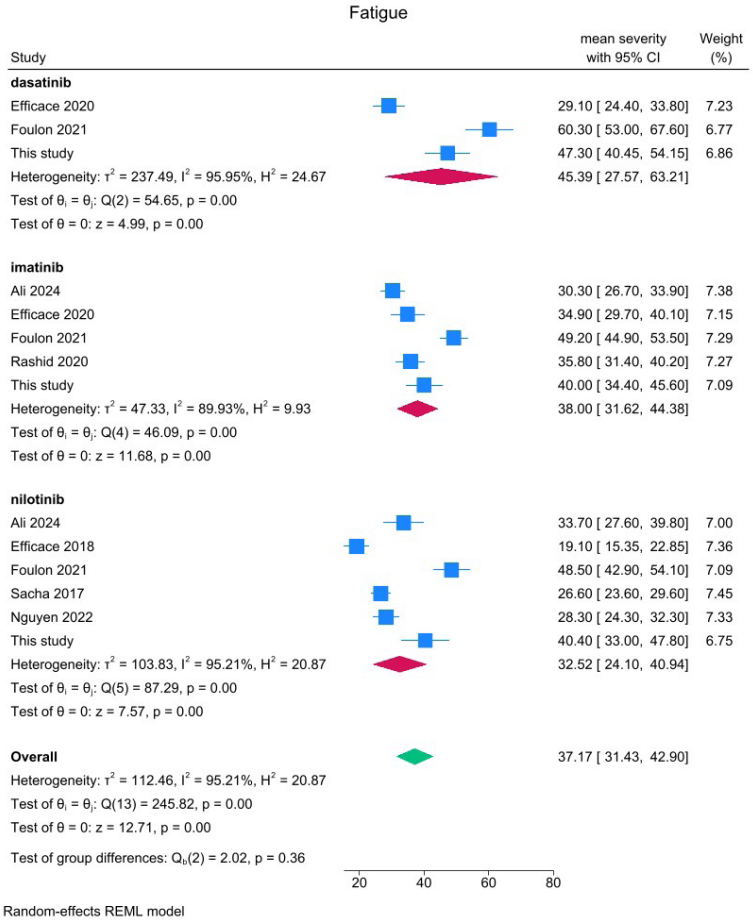


Figure 1 Meta-analysis of mean fatigue severity measured by EORTC instruments in ≥30 patients per treatment group

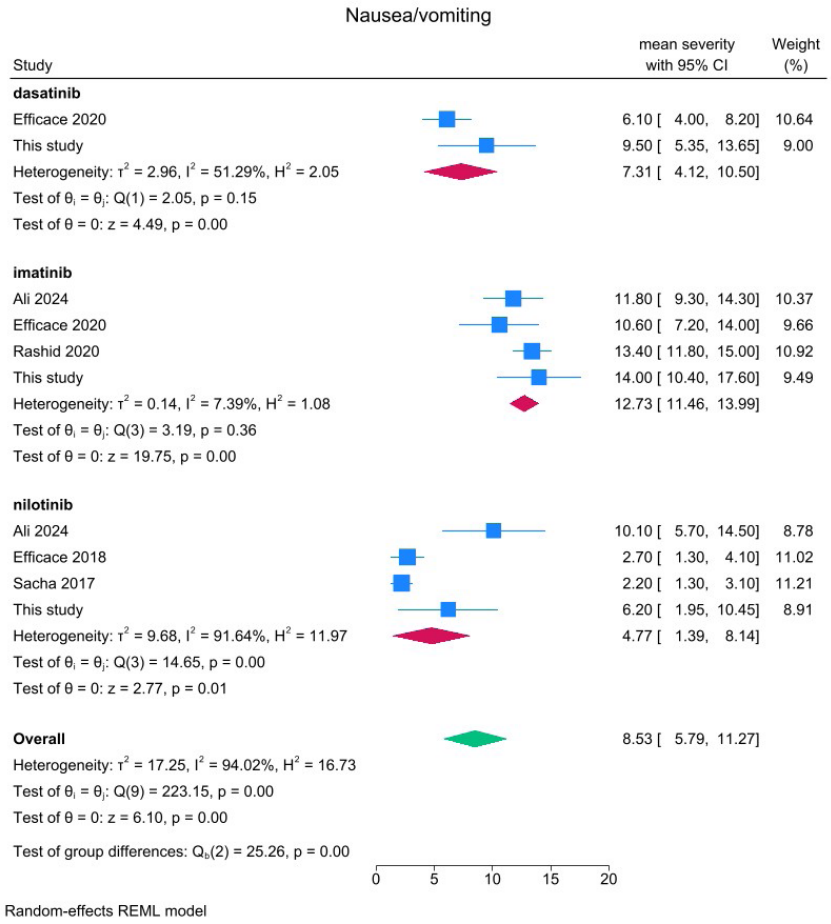


Figure 2 Meta-analysis of mean nausea/vomiting severity measured by EORTC instruments in ≥30 patients per treatment group

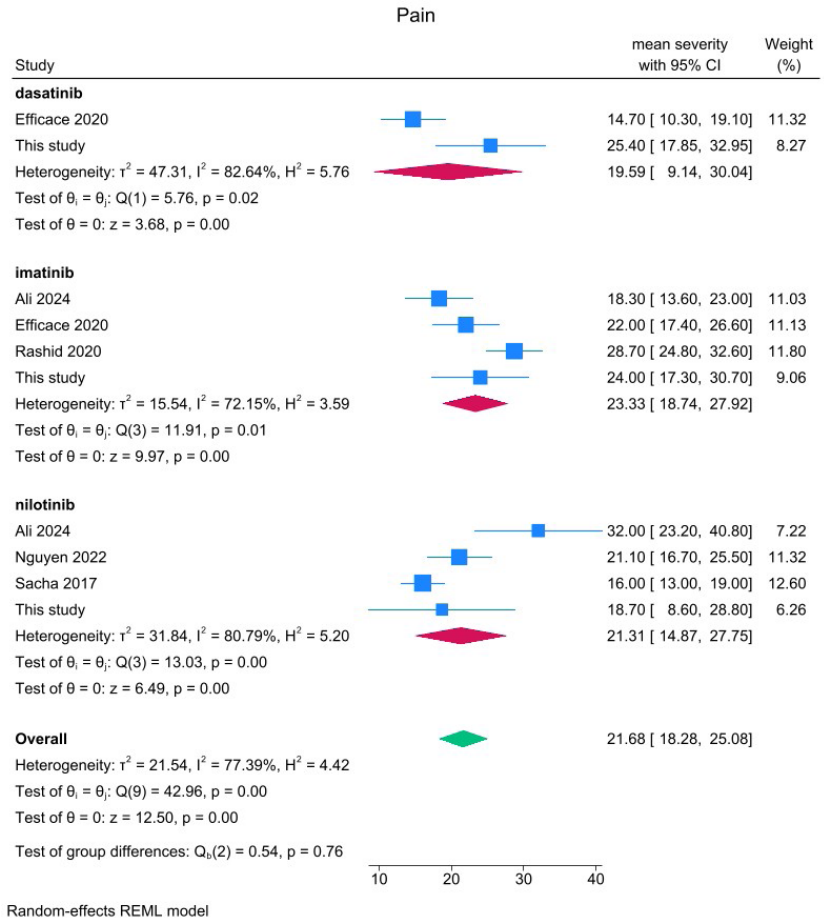


Figure 3 Meta-analysis of mean pain severity measured by EORTC instruments in ≥30 patients per treatment group

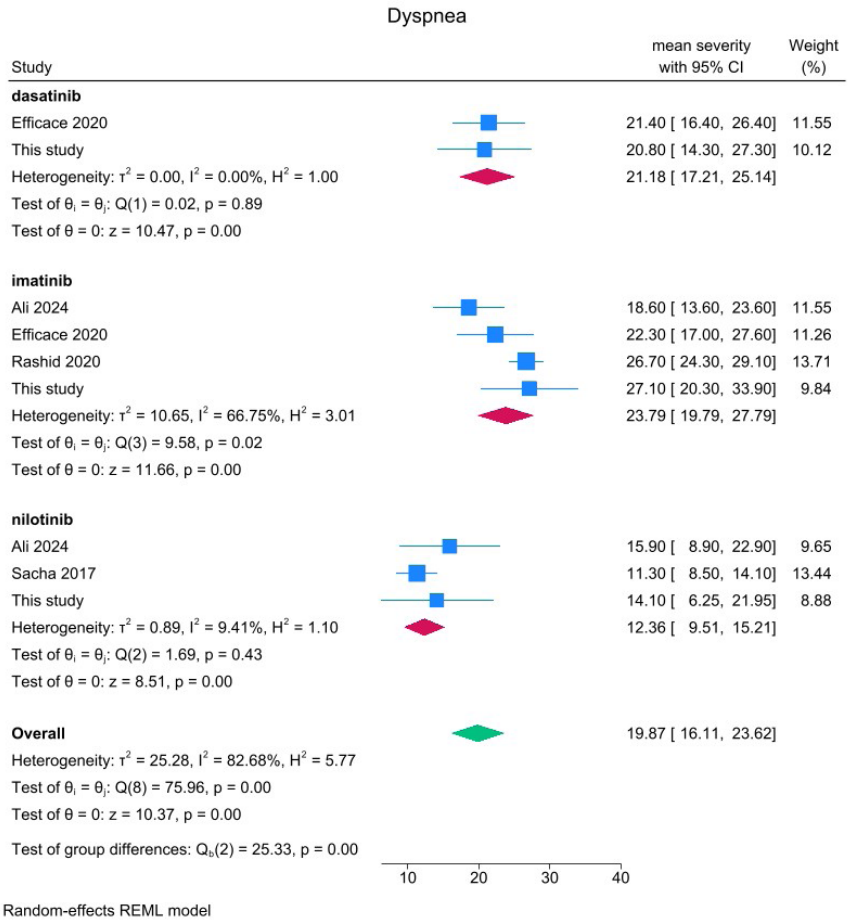


Figure 4 Meta-analysis of mean dyspnoea severity measured by EORTC instruments in ≥30 patients per treatment group

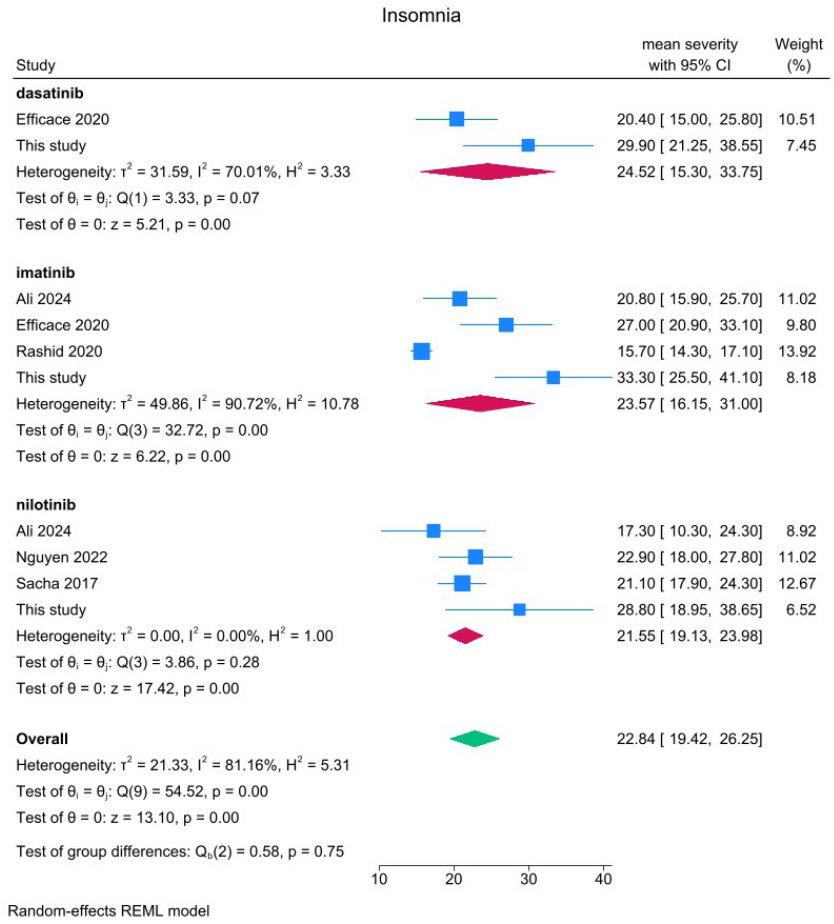


Figure 5 Meta-analysis of mean insomnia severity measured by EORTC instruments in ≥30 patients per treatment group

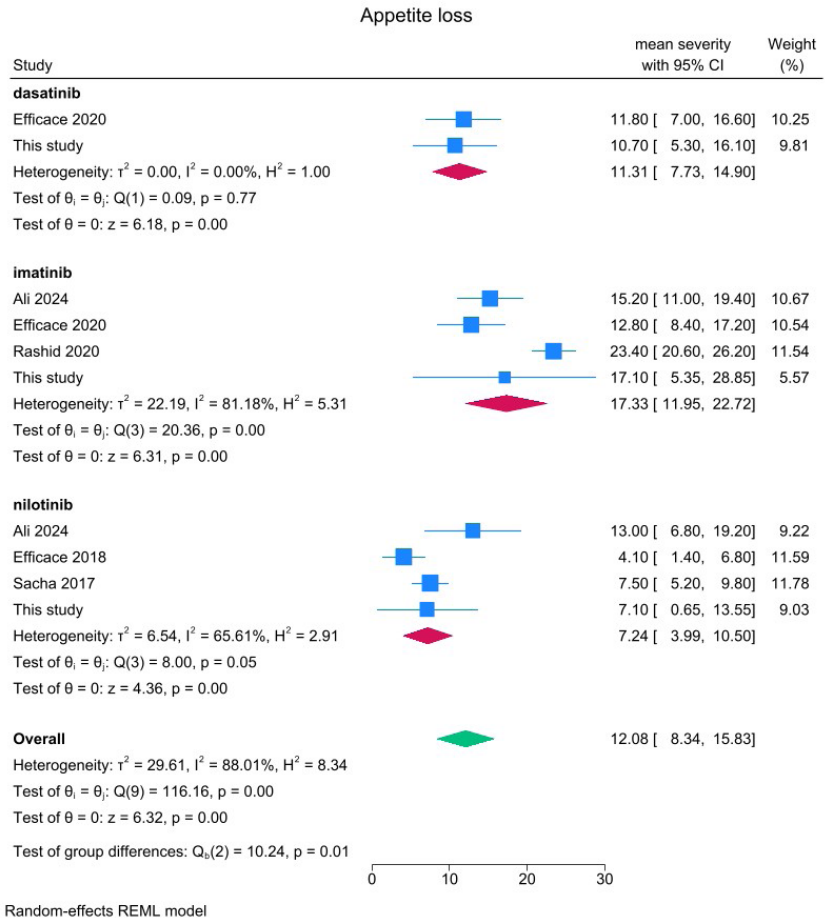


Figure 6 Meta-analysis of mean appetite loss severity measured by EORTC instruments in ≥30 patients per treatment group

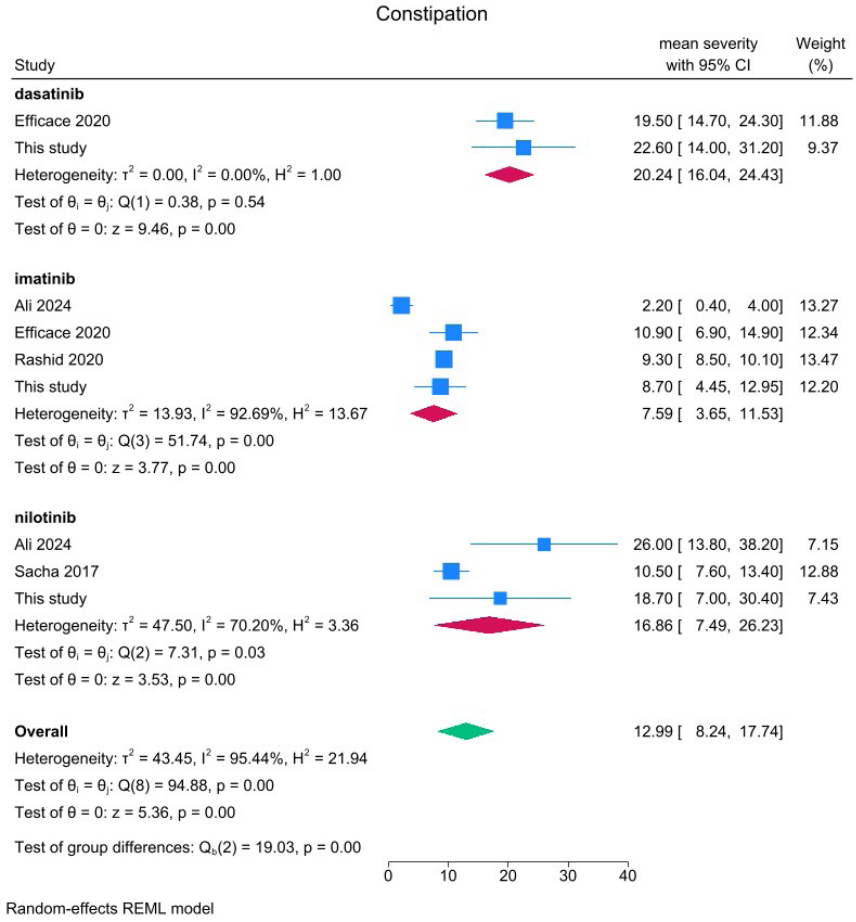


Figure 7 Meta-analysis of mean constipation severity measured by EORTC instruments in ≥30 patients per treatment group

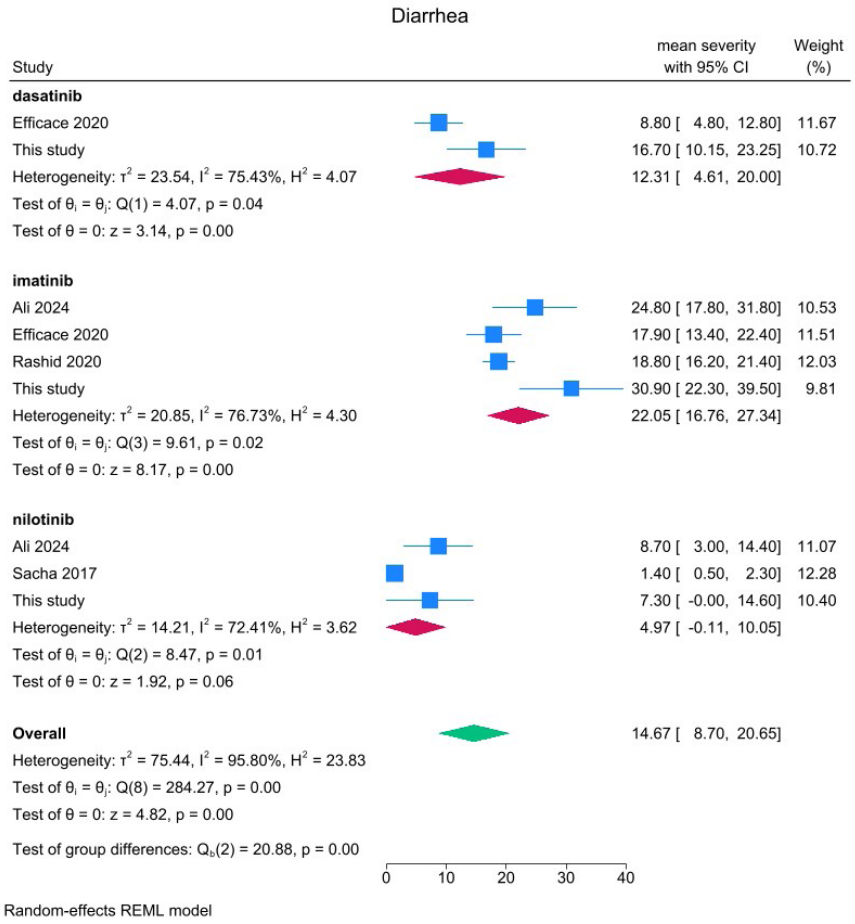


Figure 8 Meta-analysis of mean diarrhoea severity measured by EORTC instruments in ≥30 patients per treatment group

TKI-treatments compared in other studies

In a systematic search in PubMed and Embase, from inception up to February 2025, we identified ten studies that compared symptom prevalence or mean symptom severity between current treatments, measured with validated patient-reported instruments in at least thirty patients (Table 30). Seven studies made comparisons between similar current treatments as we did, of which only one study by Efficace et al. adjusted for covariates [17]. This one study found that patients on dasatinib experienced less nausea/vomiting, pain and diarrhoea compared to imatinib, while patients on imatinib experienced less constipation. Comparing results from our study with this study, our results are similar with regard to diarrhoea and constipation. However, for nausea/vomiting and pain we found no difference between dasatinib and imatinib; while for headaches, skin problems, and hair loss we did find significant differences favouring imatinib. Possibly, differences in matching/covariates adjusted for, and/or population differences, are the cause of these differences. Efficace et al. matched on co-morbidity and previous CML treatment for example, but did not adjust for dosage, as we did. Other differences we identified (Cognitive problems (scale), Swelling face/eyes, Feeling bloated, and Hot flushes) were not evaluated by Efficace et al.

Table 30 Significant differences in symptoms reported by patients, measured with validated instruments, in studies that included over 30 patients and made comparisons between TKI treatments

Article	Study type, setting	Current treatment	Questionnaires	Significant differences between current treatments	Adjusted for
Abulafia 2020 [12]	Cross-sectional, nationwide, recruitment via e-mail patients' organization	Dasatinib (n=45) Imatinib (n=70) Nilotinib (n=24)	EORTC QLQ-CML24	Eight individual symptoms evaluated, three with significant differences in **mean symptom severity**: • Less fatigue in imatinib (p=0.02) vs. dasatinib and nilotinib • Less pain in imatinib vs. nilotinib (p=0.01) • Less worry, stress, depression or nervousness in imatinib vs. nilotinib (p=0.01)	Not adjusted
Ali 2024 [19]	Single centre, observational	Bosutinib (n=39) Imatinib (n=84) Nilotinib (n=38)	EORTC QLQ-CML24	Eight individual symptoms evaluated, three with significant differences in **mean symptom severity** (all p<0.05 after Bonferroni): • Less severe pain in imatinib vs. nilotinib • Less severe constipation in imatinib versus bosutinib and nilotinib • More severe diarrhoea in bosutinib vs. imatinib and nilotinib, and in imatinib vs. nilotinib	Not adjusted

Table 30 Continued

Article	Study type, setting	Current treatment	Questionnaires	Significant differences between current treatments	Adjusted for
Bostan 2020 [15]	Cross-sectional, university medical centre	Dasatinib (n=30) Original imatinib (n=17) Generic imatinib (n=44) Nilotinib (n=30)	EORTC QLQ-CML24, MDASI-CML	28 Individual symptoms evaluated, three with significant differences in their **prevalence**: • Less swelling of ankles, legs or around the eyes in dasatinib or nilotinib vs. imatinib (p=0.02) • Less rash/skin changes in imatinib (p=0.02) vs. dasatinib and nilotinib • Less appetite loss in original imatinib (p=0.03) vs. vs. dasatinib and nilotinib The difference in **median severity** was evaluated for 8 individual symptoms, with no statistical differences between treatment groups	Multiple comparisons, not adjusted
Cho 2025 [13]	Retrospective cohort, single centre	Dasatinib (n=22) Imatinib (n=53) Nilotinib (n=67)	EORTC QLQ-C30	Six individual symptoms for differences in **mean symptom severity**, of which 2 with significant differences: • Nausea/vomiting was less severe in imatinib vs. nilotinib • Diarrhoea was more severe in imatinib and dasatinib vs. nilotinib	Not adjusted
Cortes 2019 [25]	RCT	Bosutinib (n=268) Imatinib (n=268)	EQ-5D	Two symptoms evaluated (pain/discomfort and anxiety/depression), neither with significant differences in **repeated measures mixed-effect models** or **prevalence** at 12 months	Not adjusted (RCT)

Table 30 Continued

Article	Study type, setting	Current treatment	Questionnaires	Significant differences between current treatments	Adjusted for
Efficace 2020 [17]	Propensity matched multicentre case-control study	Dasatinib (n=94) Imatinib (n=94)	EORTC QLQ-CML24	Eight individual symptoms evaluated, four with significant **adjusted mean severity** differences: • Less nausea/vomiting (p=0.02), pain, diarrhoea (p<0.01) in dasatinib • Less constipation in imatinib (p=0.03)	Matching on age, sex, living arrangement, comorbidity, ECOG-performance status, Sokal risk, and previous CML treatment. Adjustment for education and months from treatment start
Réa 2023 [26]	RCT	Asciminib (n=157) Bosutinib (n=76)	EQ-5D, MDASI-CML	20 Individual symptoms evaluated, seven with significant differences in average change **of mean symptom severity** from baseline: Favoured asciminib (as 95%CI did not include 0) for feeling drowsy, lack of appetite, dry mouth, nausea, vomiting, diarrhoea, and pain (p-values not given)	Not adjusted (RCT)

Table 30 Continued

Article	Study type, setting	Current treatment	Questionnaires	Significant differences between current treatments	Adjusted for
Uyanik [14]	Cohort, single centre	Imatinib (n=20) Generic imatinib (n=22) 2nd-generation TKIs (n=10)	EORTC QLQ-C30, HADS	Eight individual symptoms evaluated (including anxiety and depression) in **mean severity**, one with a significant difference: • Patients on generic imatinib reported more severe diarrhoea than patients on imatinib (p=0.012)	Not adjusted
Williams 2013 [27]	Cross-sectional, single centre	Bafetinib (n=1) Bosutinib (n=6) Dasatinib (n=34) Imatinib (n=71) Nilotinib (n=22) Ponatinib (n=5) Rebastinib (n=3) Other (n=3) No treatment (n=7)	MDASI-CML	Twenty individual symptoms tested, three with significant differences in **mean symptom severity**: • Less severe diarrhoea (p<0.01), muscle soreness and cramping (p=0.02), swelling (p=0.04) in dasatinib and nilotinib vs. imatinib	Not adjusted

Table 30 Continued

Article	Study type, setting	Current treatment	Questionnaires	Significant differences between current treatments	Adjusted for
Zulbaran-Rojas 2018 [28]	Cohort study	Dasatinib (n=104) Nilotinib (n=82) Ponatinib (n=33)	MDASI-CML	Twenty individual symptoms tested, 11 with significant differences **in mean symptom severity**: • Dasatinib lower disturbed sleep, diarrhoea (p<0.01) and oedema (p=0.02) than nilotinib • Dasatinib higher pain (p=0.03) than nilotinib • Dasatinib lower disturbed sleep vs. ponatinib (p=0.02) • Nilotinib lower malaise (p=0.01), oedema, pain (p<0.01), shortness of breath (p=0.01) than ponatinib • Dasatinib and nilotinib lower skin rash, muscle cramps, dry mouth (p<0.01) and distress (p<0.01 for dasatinib and p=0.03 for nilotinib) vs. ponatinib	Not adjusted

Shaded cells indicate comparisons including between similar current treatments as in our study

Supplementary references

1. Tong, A., P. Sainsbury, and J. Craig, *Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups.* Int J Qual Health Care, 2007. **19**(6): p. 349-57.
2. Baum, F., C. MacDougall, and D. Smith, *Participatory action research.* J Epidemiol Community Health, 2006. **60**(10): p. 854-7.
3. von Elm, E., et al., *The Strengthening The Reporting of OBservational studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies.* Lancet, 2007. **370**(9596): p. 1453-7.
4. Terwee, C.B., et al., *COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study.* Qual Life Res, 2018. **27**(5): p. 1159-1170.
5. de Vet, H.C., et al., *Measurement in medicine: a practical guide*. Practical Guides to Biostatistics and Epidemiology. 2011, Cambridge: Cambridge University Press.
6. Revicki, D., et al., *Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes.* J Clin Epidemiol, 2008. **61**(2): p. 102-9.
7. Fayers, P.M., et al., *The EORTC QLQ-C30 Scoring Manual (3rd Edition)*, E.Q.o.L. Group, Editor. 2001, European Organisation for Research and Treatment of Cancer: Brussels.
8. Benjamin, Y. and Y. Hochberg, *Controlling the false discovery rate: a practical and powerful approach to multiple testing.* J R Soc Ser B (Methodol), 1995. **57**: p. 289-300.
9. Hair, J., et al., *A primer on partial least squares structural equation modeling (PLS-SEM) 3rd ed.* 2022, Thousand Oaks, CA: Sage.
10. Smit, Y., et al., *Patient-reported toxicity symptoms during tyrosine kinase inhibitor treatment in chronic myeloid leukemia: a systematic review and meta-analysis.* Support Care Cancer, 2025. **33**(5): p. 446.
11. Jain, P., et al., *Comparative study for the efficacy, safety and quality of life in patients of chronic myeloid leukemia treated with Imatinib or Hydroxyurea.* J Res Pharm Pract, 2013. **2**(4): p. 156-61.
12. Shacham Abulafia, A., et al., *Health-related quality of life in patients with chronic myeloid leukemia treated with first- versus second-generation tyrosine kinase inhibitors.* J Clin Med, 2020. **9**(11).
13. Cho, H.J., et al., *Real-world experience in the management of chronic myeloid leukemia patients focused on tyrosine kinase inhibitors intolerance and health-related quality of life.* Acta Haematol, 2025: p. 1-14.
14. Uyanik, M.S., et al., *The impact of new generic formulations of imatinib mesylate on general quality of life in chronic phase chronic myeloid leukaemia patients.* Eur J Cancer Care (Engl), 2017. **26**(2).
15. Bostan, H., et al., *Quality of life and symptom burden with first- and second-generation tyrosine kinase inhibitors in patients with chronic-phase chronic myeloid leukemia.* Clin Lymphoma Myeloma Leuk, 2020. **20**(12): p. 836-842.
16. Efficace, F., et al., *International development of an EORTC questionnaire for assessing health-related quality of life in chronic myeloid leukemia patients: the EORTC QLQ-CML24.* Qual Life Res, 2014. **23**(3): p. 825-36.

17. Efficace, F., et al., *Health-related quality of life of newly diagnosed chronic myeloid leukemia patients treated with first-line dasatinib versus imatinib therapy.* Leukemia, 2020. **34**(2): p. 488-498.
18. Foulon, S., et al., *Health state utility and quality of life measures in patients with chronic myeloid leukemia in France.* Qual Life Res, 2021. **30**(7): p. 2021-2032.
19. Ali, H., Z. Mohamed, and D. Abdullah, *Health-related quality of life in chronic myeloid leukaemia patients receiving long-term therapy with different tyrosine kinase inhibitors in Kurdistan region.* Georgian Med News, 2024(354): p. 173-180.
20. Rashid, A. and Z. Albasry, *Assessing the quality of life in imatinib-treated chronic myeloid leukemia patients: a cross sectional descriptive study.* Int J Pharmaceut Res, 2020. **12**(2): p. 2784-91.
21. Efficace, F., et al., *Health-related quality of life in patients with chronic myeloid leukemia receiving first-line therapy with nilotinib.* Cancer, 2018. **124**(10): p. 2228-2237.
22. Sacha, T., et al., *Quality of life and adherence to therapy in patients with chronic myeloid leukemia treated with nilotinib as a second-line therapy: a multicenter prospective observational study.* Clin Lymphoma Myeloma Leuk, 2017. **17**(5): p. 283-295.
23. Nguyen, C.T.T., et al., *Quality of life among chronic myeloid leukemia patients in the second-line treatment with nilotinib and influential factors.* Qual Life Res, 2022. **31**(3): p. 733-743.
24. Efficace, F., et al., *Validation and reference values of the EORTC QLQ-CML24 questionnaire to assess health-related quality of life in patients with chronic myeloid leukemia.* Leuk Lymphoma, 2021. **62**(3): p. 669-678.
25. Cortes, J.E., et al., *Patient-reported outcomes in the phase 3 BFORE trial of bosutinib versus imatinib for newly diagnosed chronic phase chronic myeloid leukemia.* J Cancer Res Clin Oncol, 2019. **145**(6): p. 1589-1599.
26. Rea, D., et al., *Health-related quality of life of patients with resistant/intolerant chronic phase chronic myeloid leukemia treated with asciminib or bosutinib in the phase 3 ASCEMBL trial.* Leukemia, 2023. **37**(5): p. 1060-1067.
27. Williams, L.A., et al., *Measuring the symptom burden associated with the treatment of chronic myeloid leukemia.* Blood, 2013. **122**(5): p. 641-7.
28. Zulbaran-Rojas, A., et al., *A prospective analysis of symptom burden for patients with chronic myeloid leukemia in chronic phase treated with frontline second- and third-generation tyrosine kinase inhibitors*

Chapter 5

Benefits and limitations of real-world patient-reported toxicity symptom monitoring for guidelines and care, as perceived by patients, clinicians and guideline developers

Yolba Smit, Lynn Verweij, Alexander Currie, Jeroen Janssen, Eduardus Posthuma, Andre Dekker, Rosella Hermens, Nicole Blijlevens

Cancer Medicine. 2025, 14(8);e70880.

Abstract

Background

Toxicity monitoring should be modernised to include real-world patient-reported data. However, little is known on how stakeholders view the incorporation of real-world patient-reported toxicity symptoms into guidelines. This gap hinders the development of a sustained learning health care environment and limits the incorporation of this data into daily care.

Methods

This qualitative study, reported according to COREQ, involved interviews with 29 plus ten chronic myeloid leukaemia (CML) patients and 18 CML clinicians, including eight haematologists/guideline developers. The interviews were audio-recorded, transcribed, and independently coded in Atlas.ti. A framework, adapted from systematically sourced literature, was used for coding. Codes were assessed as either beneficial or limiting. An expert panel of all CML guideline developers completed and prioritised the identified knowledge gaps through a RAND-modified Delphi procedure.

Results

Thirty-one benefits and limitations of systematically monitoring patient-reported toxicity symptoms in the real-world were identified. Compared to an existing framework, novel benefits centred around the use of aggregated data: participants viewed real-world patient-reported toxicity symptoms as a way to systematically include patients' toxicity symptoms in the guidelines; personalise guideline advise; and fill knowledge gaps. The expert panel agreed on 14 knowledge gaps in chronic myeloid leukaemia care that could be addressed through such data. Novel limitations focused on the suitability, acceptance, and applicability of toxicity symptom monitoring in routine clinical practice. Participants felt that this monitoring does not establish a causal link between medication and symptoms, and it has no added value over open conversation.

Conclusions

The benefits and limitations of adopting patient-reported real-world toxicity symptom monitoring need to be leveraged and addressed to ensure maximum value and uptake. Guideline developers viewed aggregated data as beneficial. The identified knowledge gaps provide concrete points of action for CML guideline development.

Introduction

The routine monitoring of patient-reported symptoms, with feedback to health care professionals, has improved both quality of life and survival in solid oncology patients [1-4]. This evidence firmly establishes patient-reported outcome measures: monitoring across the care continuum is now recommended by ESMO to detect and manage treatment toxicities early, as well as relapses, and tailor care [5]. In haematooncology, however, evidence for symptom monitoring is still limited [6]. Nevertheless, there is a need to modernise toxicity assessment by incorporating patient-reported outcomes into real-world care in this field as well [7, 8].

In addition to improving the quality of care, patient-reported toxicity symptom assessment can be systematically aggregated and evaluated as a quality metric [5]; contribute to the scientific debate; and inform the development of evidence-based guidelines. At present, haematooncology guidelines rely almost exclusively on clinical trials. However, these trials rarely incorporate patient-reported outcomes [7], and their selective inclusion criteria limit real-world applicability.

Patient-reported outcomes are particularly relevant in chronic conditions like chronic myeloid leukaemia (CML), where most patients require lifelong tyrosine kinase inhibitors (TKIs). TKI use is associated with substantial low-grade toxicity [9], which is often underestimated by clinicians [10]. Approximately 30% of patients require a TKI switch due to intolerance [11]. If CML guidelines focus too heavily on survival and disease control, they risk losing relevance. For most patients, disease control is adequate, making toxicity management the primary concern.

So, why are CML patient-reported toxicity symptoms still not collected at an aggregated level and used to guide CML care? As a first step, CML guideline developers need to assess whether monitoring CML patient-reported toxicity symptoms is useful for the evolving CML guideline, and if so, in what way?

At the same time the question remains: how should real-world TKI-related toxicity symptoms in CML patients be measured? Patient-reported symptoms do not necessarily imply that the TKI used caused the symptoms (toxicity), much like symptoms or adverse events reported by professionals do not imply causality. The establishment of a patient-reported Common Terminology

Criteria for Adverse Events (PRO-CTCAE) by the National Cancer Institute shows that adverse events and tolerability reported by patients are now considered valuable sources of information on treatment toxicity [12].

In CML, we know that patients report worse symptoms such as depression, dyspnoea, fatigue, pain, and composite symptom-burden scores (e.g., nausea, diarrhoea, itching, skin changes, and swelling of arms or legs) compared to controls or the general population [13, 14]. Since the CML itself is well-controlled in the vast majority of patients, the most plausible explanation at the group level is the TKI-treatment. However, no patient-reported instrument currently has sufficient content validity to measure TKI-related toxicity symptoms [15].

The development of a new instrument needs to consider content validity, feedback on results, and a sustainable workflow [16]. However, these factors alone do not guarantee successful uptake [4]. The views of intended users can create barriers and facilitators for uptake, and these need to be addressed from the outset. While extensive knowledge is available on the perspectives of end-users of patient-reported outcomes in general, there is little focus on toxicity symptom reporting by patients in particular [17-41]. Perspectives from the field of haematooncology are limited, do not focus on toxicity symptom reporting, and reveal specific limitations [23, 42-45].

To prepare care and guideline development for real-world toxicity symptom monitoring, we sought the views of patients, clinicians, and clinician-guideline developers. What do they see as the benefits and limitations of real-world toxicity symptom monitoring, and how do they view the use of its aggregated data in individual care and guideline development? To pave the way for the actual use of such data in CML guideline development, we also identified knowledge gaps in CML care that could be addressed through systematic patient-reported toxicity symptom monitoring.

Material and methods

Design and setting

This qualitative study interviewed CML patients and professionals, reported in accordance with the COREQ checklist [46] (Table 1 in the Supplementary material), and included a Delphi procedure. We chose a qualitative design to capture the full spectrum of views on the uptake of real-world TKI-related

toxicity symptom monitoring and to address knowledge gaps. Ethical approval was waivered by the institutional Medical Ethical Committee, as the study was not subject to the Medical Research Involving Human Subjects Act.

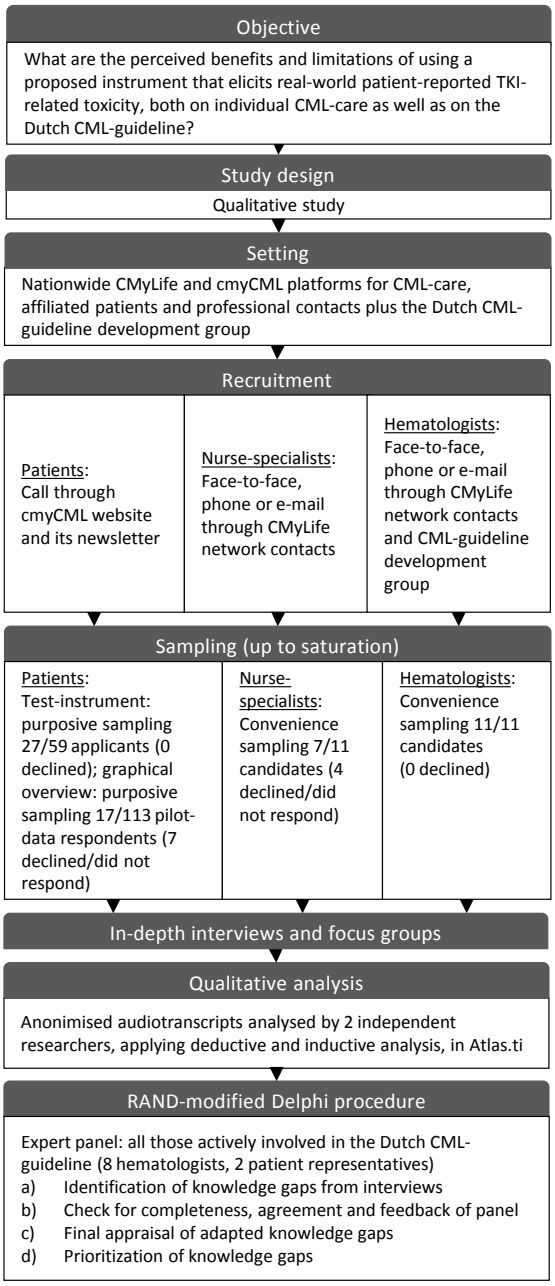


Figure 1 Flowchart of the study

The qualitative interviews were conducted as part of the development and validation phases of an instrument designed to capture real-world patient-reported TKI-related toxicity symptoms, hereafter referred to as the test instrument, described in Table 2 in the Supplementary material. A graphical overview of patients' scores was included as part of the instrument.

Dutch CML patients are treated in eight academic hospitals and 68 general hospitals [47]. In order to give them the tools to manage and monitor their own CML, interpret BCR::ABL1 results, and act on them, the nationwide digital care platform CMyLife was established. CMyLife includes a patient-portal (www.cmycml.nl) and a guideline application [48, 49]. Currently, patients on the platform have access to their own electronic personal health environment, which bridges individual hospitals' electronic patients files to the nationwide CMyLife guideline application. The test instrument is intended to function within the CMyLife digital care platform.

The Delphi procedure was conducted within the national HOVON (Dutch-Belgian Cooperative Trial Group for Haematology-Oncology) CML-MPN working group. This group is the guideline development group for the Dutch CML guideline [50]. Figure 1 summarises the study flow.

Interviews

Recruitment

Recruitment and sampling took place during two phases: (1) the qualitative development and validation phases; and (2) the subsequent quantitative validation phase of the test instrument.

In the first phase, 27 out of 59 patients who responded to a call on the CMyLife platform were purposefully sampled (up to data saturation) to reflect differences in sex, age, time since diagnosis, education, and hospital type. Haematology nurse specialists (in training) were recruited through our network via convenience sampling, and additional haematologists were recruited until saturation was reached.

In the second phase, a further ten patients were recruited for interviews based on the graphical overview constructed from their own data. These ten patients were sampled from a group of 113 patients who responded to another invitational call on the CMyLife platform and participated in pilot data

collection with the test instrument for six months (quantitative validation phase). The sampling was purposefully done to reflect differences in sex, age, time since diagnosis, current treatment and completion of pilot data.

Lastly, all Dutch haematologists actively involved in CML guideline development within the national HOVON CML-MPN working group were recruited for the Delphi procedure.

Sessions

Two researchers (YS, AC) conducted focus groups and individual interviews, hereafter referred to as sessions, using semi-structured interview guides centred on perceived (dis)utility (Supplementary material). Two to five days prior to the sessions, the test instrument and/or graphical overview were sent to participants, asking them to either fill it in (patients) or read it through (professionals and patients for their own graphical overview). Informed consent was obtained before the sessions began, and all sessions were audio-recorded, transcribed, and anonymised.

Analysis

We systematically reviewed the literature and extracted (sub)themes from a recent comprehensive systematic review [17] completing these themes with codes from other systematic reviews plus more recent primary studies [18-41]. We used this framework as a code tree for deductive coding (Table 3 in the Supplementary material). When no suitable codes were available, we applied new codes through inductive coding. We then categorised codes according to pre-existing (sub)themes, or placed them in new categories when existing themes did not apply. If necessary, themes were rearranged to better reflect our data. We labelled subthemes as a benefit, limitation, or mixed. Codes, (the arrangement of) themes, and benefit/limitation labelling were discussed between two researchers, who initially coded independently, until consensus was reached.

RAND-modified Delphi procedure

The expert panel consisted of ten members: eight haematologists, and two patient representatives from patient organization Hematon. This group included all haematologists actively involved in the Dutch CML guideline, as part of the HOVON CML-MPN working group. The procedure consisted of four rounds (described in detail in the Supplementary material): (1) identification of knowledge gaps from the interviews; (2) verification of completeness of,

and agreement with, the identified knowledge gaps, followed by adjustments if needed; (3) appraisal of the set; and (4) rating of priorities.

Results

Interviews

We interviewed 27 patients, plus ten additional patients, seven nurse specialists (one in training) and 11 haematologists, eight of whom were involved in CML guideline development. Participants' characteristics are described in Table 1. Compared to the existing evidence framework [17-41], four new benefits, two new limitations and three new mixed benefit/limitation themes emerged (Table 2). An overview of these new subthemes with illustrative quotations is provided in Table 3 (other subthemes can be found in the Supplementary material). Here, we describe the new subthemes, as well as those that emerged with an additional focus compared to the existing evidence. Table 3 in the Supplementary material provides an overview of the differences between the existing evidence and this study.

Table 1 Characteristics of participants in sessions

Patients (n=27) † Test instrument, including prototype graphical overview	
Female (%)	16 (59%)
Mean age (range)	59 years (33 to 77 years)
Mean time since diagnosis (range)	6 years (0 to 16 years)
Treatment hospital	University medical centre: 9 (30%) General hospital: 17 (70%)
Education level	Low educated: 1 (4%)
	Middle educated: 3 (11%)
	High educated: 19 (70%)
	Unknown: 4 (15%)
Current treatment	Not available
Patients (n=10) † Graphical overview own pilot data	
Female (%)	6 (60%)
Mean age (range)	60 years (47 to 84 years)
Mean time since diagnosis (range)	7 years (3 to 10 years)
Treatment hospital	Not available
Education level	Not available
Current treatment	Asciminib (1), bosutinib (1), dasatinib (3), imatinib (3), nilotinib (1), and ponatinib (1)
Nurse specialists (in training) (n=7) Test instrument, including prototype graphical overview	
Female (%)	7 (100%)
Hospital	University medical centre: 3 (43%)
	General hospital: 4 (57%)
Haematologists (n=11) (8 of whom guideline developers) Test instrument, including prototype graphical overview	
Female (%)	6 (55%)
Hospital	University medical centre: 6 (54.5%)
	General hospital: 5 (45.5%)
Involved in Dutch CML guideline development	8 (72%)

†These groups may have overlapped to some extent

Table 2 Framework of perceived benefits and limitations of real-world patient-reported TKI-related toxicity symptom monitoring, arranged according to five themes that emerged after qualitative analysis of sessions with CML patients, nurse specialists and haematologists/guideline developers †

1. Active patient involvement and partnership
• Enables greater awareness and reflection (Benefit)
• *Objectifies subjective experience (Benefit)*
• Encourages patient involvement (Benefit)
• Facilitates goal setting and shared decision making (Benefit)
• Influences honesty (Limitation)
• Permits discussion of sensitive topics (Benefit)
• Enables *self-prevention* and self-care (Benefit)
• Leads to worse symptom experience (Limitation)
• Fulfils desire to help others (Benefit)
2. Focus of consultation
• Helpful as a screening tool (Benefit)
• Prioritises patients' needs (Benefit)
• Provides reassurance that clinicians care (Benefit)
• Provides one piece of the picture (Limitation)
• Structures consultations and improves efficiency (Benefit)
• Provides redundant information (Limitation)
• *Suits paramedical consultation (Mixed)*
3. Quality of individualpatient care
• Helps determine *and monitor* side effects of treatment (Benefit)
• Prompts appropriate, standardised action *and acceptance* (Benefit)
• Assists in learning (Benefit)
• Can inaccurately estimate the problem/not specific enough to be clinically meaningful (Limitation)
• *Does not establish a causal link (Limitation)*
4. Suitability and acceptability
• Suitability for all patients*, clinicians, and workflows (Limitation)*
• Confronts too much with disease (Limitation)
• *Usefulness depends on (change in) symptom severity (Mixed)*
• *Has no added value over open questions or conversation (Limitation)*
5. Improving care and clinical guidelines with aggregated data
• *Real-world data has pros and cons (Mixed)*
• Benchmarks clinical trials and medication costs (Benefit)
• Puts individual experience in context (Benefit)
• *Systematically includes patients' experiences in guidelines (Benefit)*
• *Personalises guideline advise (Benefit)*
• *Fills in knowledge gaps (Benefit)*

† Newly identified themes, subthemes, and foci compared to previous studies[17-41] *in italics*

(1) Active patient involvement and partnership

Within this theme, the new subtheme, 'Objectifies subjective experience', emerged from patients who expressed that the graphical overview can clarify and strengthen their perspective, enhancing their confidence in discussing their symptoms with professionals and peers. Additionally, an further focus emerged: patients mentioned the use of self-prevention techniques to prevent symptoms, such as increasing fluid intake before exercise. Based on this, we renamed the subtheme to: 'Enables self-prevention and self-care'.

(2) Focus of consultation

Within this theme, an additional focus, 'Suits paramedical consultation', emerged, which combined two previous subthemes ('Shifts away from the main medical problem' and 'Raises unrealistic expectations for care') were merged. Patients acknowledged that quality of life issues are not typically the core responsibility of medical doctors, and haematologists admitted that they do not always discuss toxicity symptoms with patients, especially if no medical action is required. Haematologists viewed toxicity symptoms as more within the realm of nurse specialists, and patients shared positive experiences with nurse specialist consultations on toxicity symptoms. When a nurse specialist is available, shifting responsibilities can benefit the focus of consultation of haematologists. However, if no nurse specialist is available, this shift in focus may be a limitation for haematologists' consultations, as it diverts attention from the main medical issue.

(3) Quality of individual patient care

Within this theme an additional focus emerged. Both patients and professionals found that the test instrument facilitated the acceptance of symptoms that cannot be resolved or are unlikely to improve. Understanding that certain symptoms were related to TKI toxicity rather than the CML itself helped patients move forward. As a result, we renamed the subtheme 'Prompts appropriate, standardised action *and acceptance*' to better reflect the role of acceptance in this process.

Additionally, the subtheme 'Does not establish a causal link' emerged. Patients expressed uncertainty about whether their symptoms were caused by CML medication or other factors, such as the CML itself, aging, or other diseases or medications. Some patients self-censored symptoms they attributed to other causes.

(4) Suitability and acceptance

Patients mentioned that not all clinicians were digitally skilled enough to work with the test instrument. Professionals mentioned that stand-alone applications have disadvantages, as they require separate installations and do not integrate with the hospital workflow. We therefore renamed the subtheme 'Suitability for all patients' to include clinicians and workflow considerations.

Moreover, two new subthemes were identified, First, some patients felt the test instrument did not apply to them because they either did not experience symptoms or attributed their symptoms to other causes, such as aging. The instrument and graphical overview were considered most useful for severe symptoms or changes in symptoms, for example at the start of a (new) therapy. This led to the subtheme 'Usefulness depends on (change in) symptom severity' (mixed benefit and limitations).

Additionally, some patients and professionals felt that open-ended questions or conversation were more effective and that the test instrument had no added value. This resulted in a new subtheme: 'Has no added value over open questions or conversation' (limitation).

(5) Improving care and clinical guidelines with aggregated data

Within this theme, an additional focus and three new subthemes emerged. As an additional focus, haematologists/guideline developers stated that real-world patient-reported data could provide diverse and long-term insights, including information on drug-drug interactions, data that can only be gathered from real-world settings. However, they also noted that real-world data are collected in a less structured manner compared to clinical trials, making it more difficult to draw reliable conclusions. Despite this limitation, they acknowledged that large datasets could still reveal meaningful patterns, leading to the subtheme 'Real-world data has pros and cons' (mixed benefit/limitation).

Another subtheme emerged from the recognition that patient-reported data systematically provide insight into patients' experiences, which was considered valuable in itself. This led to the new benefit: 'Systematically includes patients' experiences in guidelines'. Building on this, haematologists/guideline developers suggested that the test instrument has the potential to 'Personalises guideline advise' and 'Fills in knowledge gaps' (new benefits). Identified knowledge gaps with the potential to be addressed by the test instrument were incorporated into the Delphi procedure.

Though patients were not specifically interviewed about the use of aggregated data, they expressed a strong interest in learning from it. They wanted to know how to prevent and/or mitigate symptoms - an important perspective which was not addressed by professionals. As a result, this knowledge gap was also included in the Delphi procedure.

Table 3 Novel themes and subthemes† with an assessment of benefit or limitation, with illustrative quotes

Themes and subthemes	Benefit or limitation	Illustrative quotes
Active patient involvement and partnership		
Objectifies subjective experience	Benefit	It [graphical overview] does somewhat reflect how you feel, how you experience the side effects. Because you can explain it, but then he can also see it in the charts how things are (P)
		You can see it. It clearly says there are symptoms. In any case, it's not a figment of the imagination. I find that difficult sometimes (P)
Enables *self-prevention* and (self-)care	Benefit	And so, I also know that with this medicine, at least for me actually, I might need to increase my fluid intake a lot compared to before 2004, so I have fewer muscle cramps when exercising (P)
		But I just think, I keep finding it interesting for myself that I can keep control where I can, also uhm the prevention of side effects (P)
Focus of consultation		
Suits paramedical consultation	Mixed	Yes, the doctor is indeed mainly concerned with, uhm, is the disease well suppressed, uhm, and a little bit about quality of life, but yes, not very extensively, because that is not the core business [P]
		So that is the nurse specialist, and uh who asks the questions, looks at where the scores are high and sees whether she can come up with something herself to alleviate that symptoms or uh or consult with the pharmacy or consult with us, what can we do about that (H/GD)
		But they [toxicity symptoms] are sometimes not important enough to discuss, for various reasons. Because you don't need it at the moment, or because you as a doctor can't do anything about it, and because as a doctor you are trained to screen for side effects that require action
Quality of *individual patient* care		
Helps determine *and monitor* side effects of treatment	Benefit	And uhm yes, it also seems pretty important to me to keep track of, and especially for new patients, so they know that, yes, gosh, in the beginning it can be very intense, it can become less or there can be symptoms that only occur after a few months (P)
		So, I can imagine that it's nice to see this [test] instrument every three, four months of that patient. Then I can compare them and then I think, well, this has remained stable with this patient. But if he suddenly says he cannot walk three meters anymore, then I have a reason to pick that out, you know (NS)

Table 3 Continued

Themes and subthemes	Benefit or limitation	Illustrative quotes
Helps determine *and monitor* side effects of treatment	Benefit	Because you have patients that couldn't take anything, and then I put them on this new medication, you know, an unregistered medication, asciminib for example, and then I want to know how that works and what the impact is (H/GD) For example, there is a side effect and that may be a reason to switch treatment, and then you can see, look, just like that in a score, look, indeed, that switch did work or not (H/GD)
Prompts appropriate, standardised action *and acceptance*	Benefit	Not all problems can be resolved of course, but if you know where things come from, then you can interpret, uhm, more easily, uhm, for yourself, I mean, oh, okay, you know, I don't have to worry that I feel this, I don't have to worry that I have that (P) With a lot of side effects is not about the doctor's action at all, but to be heard, to get advice on how to deal with them in life, to have, uhm, named them, to get confirmation that they are non-lethal (H/GD) If the patient thinks: "Okay, it's not from the disease, okay, I don't have to do anything about it, well, then I'll see how I deal with it" (H/GD)
Does not establish a causal link	Limitation	Like, yeah, where does it come from. And that is very difficult. Yes. And then again, sometimes I think, eh, I have a lot of night sweats, and then you think: is it the menopause or not. But me, yes, that has never bothered me that much, but yes, I am just soaking wet at night [P] The problem I have is that sometimes I no longer know where the complaints come from. Is it a side effect of the drug, is it from the CML? I also have a hernia … [P] Complaints or side effects also occur with other diseases, I did not really know what to complete […] I try to distinguish what is due to CML and what is due to another disease [P]
Suitability and acceptability		
Suitability for all patients, *clinicians, and workflows*	Limitation	[…] because not everybody has a computer, so yes, then you'll need old fashioned paper, or an app (NS) We have also started with stand-alone applications, uhm, in which you, uhm, set out these kinds of questionnaires outside your own electronic health record environment and visualise them. This has enormous disadvantages, because people must download separate apps for this, we often cannot see it, it is not part of your workflow if you have your electronic patient file open in your outpatient clinic (H/GD) If you want to implement it nationally with four different electronic health record systems and 80 different ways of implementation, I'm curious if that will work (H/GD)

Table 3 Continued

Themes and subthemes	Benefit or limitation	Illustrative quotes
Usefulness depends on (change in) symptom severity	Mixed	Well, I don't have any side effects, so yes, the side effects I have then I think: that has to do with my age, I'm getting older. Uh, I don't know what to keep track of either [...] no, [I] do not [want to] keep track (P)
		I feel that this questionnaire does not apply to me now. It did how it was, but not how it is now (P)
		Yes and well, that's my personal opinion, that it especially helps those people with a lot of complaints (H/GD)
		Well, especially when starting with those pills. So, the first three months I would be more focused on that list (H/GD)
Has no added value over open questions or conversation	Limitation	An open question, that that would work better (P)
		Well, why should a patient need a lead to go to a doctor with his complaints? Because I think, if you have a complaint or a symptom or a side effect of a medication or whatever [...]. You also always have your outpatient visits or contact with your doctors [...] and then you can just discuss that. I don't really need a proof for that, let's put it that way (P)
		So, I find that difficult. What does such a questionnaire do, which you cannot simply pick up in a conversation? (H/GD)
Improving care *and clinical guidelines* with aggregated data		
Real-world data has pros and cons	Mixed	Uhm, that's the Achilles' heel of real-world research, right, that, uhm, that data are collected in a less structured manner. That makes it really very difficult to draw reliable conclusions, other that general patterns (H/GD)
		[...] and there are so many snapshots that due to the large number of measurements you make, overall common denominators can still be distilled (H/GD)
Real-world data has pros and cons	Mixed	[...] gives information in patients with an enormous diversity of variables that have not become visible in normal guideline development ... uhm ... gives information about multi-drug use, because it is not only the TKI, but it is also the combination with treatment that can give or reduce or worsen symptoms, so that interaction (H/GD)
		So that means that uhm the standard guideline contains data of short duration, observations from studies mainly, occasionally from observational studies. Collecting patient-reported outcomes for guideline use provides information about real-world (H/GD)
Systematically includes patients' experiences	Benefit	It is just the value of experience [...] the fact that we uhm familiarise uhm ourselves with what patients experience [...] I call it patients' knowledge that needs to be brought in to the uhm, public domain (H/GD)

Table 3 Continued

Themes and subthemes	Benefit or limitation	Illustrative quotes
Personalises guideline advise	Benefit	[…] and that is personalizing, right? Of treatment choices based on the guideline. So that personalization, uhm, that's often based on side effects or risk profiles, but here we are talking side effects. So, if we know better what bothers patients, yes, that, that would, well, yes, you could make different therapy choices so to speak (H/GD)
		And for guidelines, look, whatever we're developing now is decision aids. So, if you choose a certain strategy treatment or something, when, yes, what trade-offs do you make? Yes, it's important for that too. […] Whether this medicine affects sex quality or temperature sense, just to mention 2 categories. […] The guideline does not go that deep. The guideline is still very much focused on effectiveness. […] That requires a considerable cultural change (H)
Fills in knowledge gaps	Benefit	[…] the solutions for those side effects. Because that is another thing, […] because now it is more about signalling, recognizing complaints, actually. One [care professional] says: go to a physiotherapist, the other gives you a pot of magnesium, a jar of calcium, well, what can be done about it? And what options are there? (P)
		Well, I think that there are unanswered questions. Things were we, that we don't know well yet. […] for example, I don't know, say, I have a patient on imatinib, and then I switch to uhm a second generation TKI […], what are the chances that the symptoms disappear? (H/GD)
		Can you expect something when you lower the dosage? (H/GD)
		I can hardly imagine that this would really have an impact on the clinical guideline. […] there is also quite a bit of literature about side effects in general, in real-world registries […] the patterns are clear to us, but the extent is really very difficult to evaluate reliably [enough to decide] that I have to take that or that medicine (H/GD)

† Novel findings compared to previous studies [17-41] *in italics*. Identifiers in brackets after quotes: H/GD: haematologist/guideline developer; NS: nurse specialist; P: patient

RAND-modified Delphi procedure

In the first round, seven out of ten expert panel members responded. They gave feedback on the completeness of nine knowledge gaps identified from the interviews. Their input led to the addition of five new knowledge gaps.

In the second round, seven respondents approved of the revised set. In the third round, the 14 knowledge gaps were prioritised based on their relevance to the Dutch CML guideline (Table 4). The knowledge gap 'What is the burden of single toxicity symptoms, and over what time period do toxicity symptoms appear and/or disappear?' was ranked as the highest priority by five out of ten expert panel members involved in the prioritization process.

Table 4 Knowledge gaps in CML care that can be solved (in part) by real-world patient-reported toxicity symptom monitoring, as identified and prioritised by the Dutch CML guideline working group and patient representatives

Order of priority	Knowledge gap
1	Per TKI: what is the burden of single toxicity symptoms, and over what time period do toxicity symptoms appear and/or disappear?
2 †	How many patients switch from which TKI, due to intolerability, and for what symptoms?
2 †	To what extent do patients and clinicians recognise TKI-related toxicity symptoms, and can standardised instruments help recognise TKI-related toxicity?
4	Which toxicity symptoms change in severity after a change in TKI dosage, or remain equal or newly appear, and by how much?
5	Which toxicity symptoms change in severity after switching from another TKI, or remain equal or newly appear, and by how much?
6	How can toxicity symptoms be prevented and/or alleviated?
7	Which toxicity symptoms change in severity after stopping TKI treatment, or remain equal or newly appear, by how much, and over what time period?
8 †	To what extent does TKI-related toxicity influence medication adherence by CML patients?
8 †	To what extent is the burden of toxicity symptoms affected by patient- and (co) medication factors?
10 †	Which toxicity symptoms change after restarting TKI therapy, or newly appear, and by how much?
10 †	Per TKI: how does disease response (BCR::ABL1 control) weigh against tolerability?
12	How much does food in combination with TKI intake influence toxicity symptoms, as experienced by patients?
13	What is the relationship between TKI plasma concentration and experienced toxicity symptoms? How useful are TKI plasma concentration measurements in guiding TKI dose reduction?
14	What is the effect of growth factors because of haematological toxicity on toxicity symptoms?

† Ex aequo

Discussion

This study explored stakeholders' perspectives on capturing real-world patient-reported TKI-related toxicity symptoms using a pilot instrument. From these insights, we developed a comprehensive framework outlining the perceived benefits and limitations for both care and guideline development. While many viewpoints aligned with existing literature on patient-reported outcomes in general, we identified nine new benefits and/or limitations, as along with four extensions to previously recognised ones. Here, we discuss the implications for clinical care and guideline improvement in CML and beyond.

Three newly identified limitations related to 'Suitability and acceptability'. Future efforts should address these limitations, for example, by incorporating a single overarching question to filter out patients who do not experience symptoms or prefer not to participate. While such approaches require validation, they could help prevent overestimation at the group level. Notably, concerns were raised about the digital skills of both patients and professionals, as our approach relies on a digital care platform. This platform offers the advantage of directly involving patients while bypassing siloed and unstandardised provider data. However, its limitation lies in its purely digital workflow, which operates outside the hospital's usual digital environment. To avoid inclusion bias, it is essential to support less digitally skilled patients in participating, possibly through non-digital alternatives. Additionally, ensuring the ongoing engagement of professionals requires integrating the platform within their hospital's digital workflow.

To maximise the value of TKI-related toxicity symptom monitoring, collaboration with both patients and professionals during instrument development is essential. Identified benefits should be leveraged through strategies that inform, motivate, and support users. These efforts can also help persuade non-users of the benefits they may not yet recognise.

Two extensions to the framework of existing studies [17-41] centred on further patient-empowerment. First, patients indicated that discussing toxicity symptoms helped them accept symptoms that could not be alleviated. While professionals expressed frustration over their inability to provide a solution, patients reported that simply understanding their symptoms helps them move forward.

Second, patients wanted to identify symptom patterns to better prevent or manage their symptoms. Some were already tracking their symptoms, reflecting characteristics of so-called "e-patients" (equipped, enabled, empowered, and engaged patients) or even "expert-patients" (those seeking deeper understanding and participation in science). This highlights a desire for self-management that goes beyond what current health care practice typically offers [51, 52].

It is important to note that, at the individual level, reported symptoms do not imply causality-a limitation acknowledged by participants. However, consistently collected and standardised real-world patient-reported toxicity data are considered unique and valuable for toxicity characterisation in haematooncology [7].

Most importantly, we identified three new benefits and one new mixed benefit/limitation under the theme 'Improving care and clinical guidelines with aggregated data'. While most research focuses on the clinical, individual use of patient-reported outcomes, we envision a more comprehensive approach that maximised the use of aggregated real-world patient-reported toxicity data [53]. This has the potential to enhance the 'evidence ecosystem' by providing insight into daily practice and producing evidence directly from patients.

The evidence ecosystem functions as a feedback loop in which evidence is synthesised, guidance is created, disseminated, implemented and evaluated - ultimately informing the production of new evidence (Figure 2) [54, 55]. This cycle can increase value and reduce waste within a true learning health care environment [8]. This is especially important in fields like CML care, where the focus has shifted from survival to preventing and managing chronic toxicity [56].

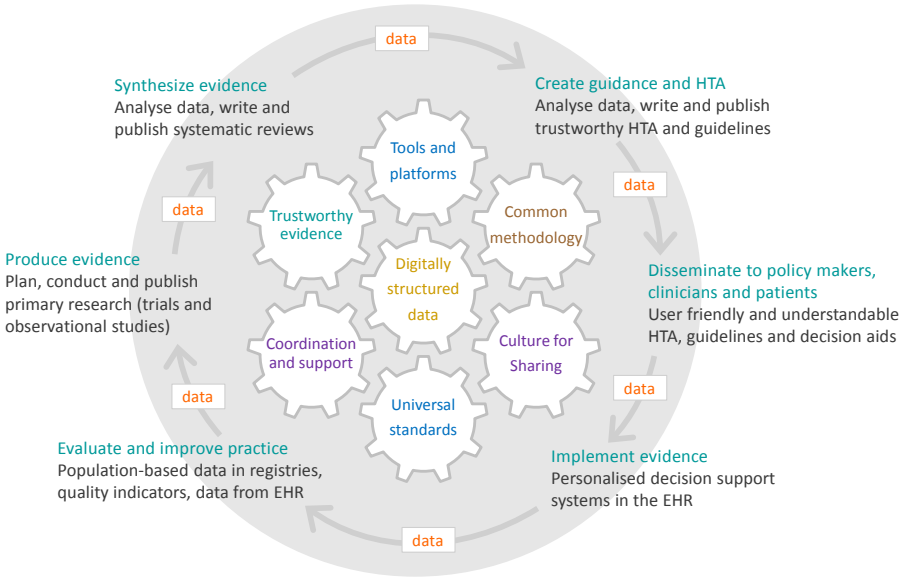


Figure 2 The concept of the 'Digital and Trustworthy Evidence Ecosystem' visualised with the production of evidence upstream and its impact downstream (reproduced with permission[57]). Abbreviations: HER: electronic health records; HTA: health technology assessment

Although the benefits of real-world evidence in CML care have been acknowledged [58], real-world CML patient-reported toxicity has not yet been systematically evaluated. Such data provide insights into treatment satisfaction and adherence [59], and the impact of toxicities on quality of life [58, 60].

Additionally, it may help address the limited stakeholder involvement and applicability of current guidelines [61]. Professionals noted that the test instrument could personalise guideline recommendations by identifying which symptoms most impact patients, and by informing the development of decision aids based on this knowledge.

The fact that CML guideline developers view real-world patient-reported toxicity symptoms as beneficial is an important first step toward realising their potential. However, the extent to which these benefits will be realised depends on addressing challenges such as the unstructured collection of real-world data. Transparency will likely be key, as it enables end-users to evaluate the quality and applicability of evidence resulting from real-world data [62, 63].

To promote transparency and reproducibility, a harmonised template for evaluating real-world evidence on treatment effects has been developed, with pilot testing underway [64]. Still, unresolved issues remain, including the frequency of data collection; integration with other databases; data audit; data ownership; electronic data capture; and the impact of the clinical pathway on data collection [65]. One potential solution is a 'fit for use' assessment, which evaluates the relevance and reliability of real-world data [63, 66]. This approach can be applied to identified knowledge gaps to help guide the collection of relevant and reliable data.

Regarding the identified novel benefit of benchmarking clinical trials and medication costs, a related concept - quality-adjusted time without symptoms or toxicities (Q-TWIST) – has been described in oncology [67]. However, it remains a relatively a coarse instrument, as it cannot capture real observed health-status differences between different TKI medication regimens.

A broader application of patient-reported toxicity data is its potential use in benchmarking medication costs, which has led to a new field of research. Intervention pharmaco-economics uses dose-optimization trials to reduce both toxicity and costs, a practice which has now been applied to CML care as well [68-71]. Standardised real-world patient-reported toxicity symptom monitoring could contribute to this field by identifying interventions and/or patient groups most likely to benefit from reduced toxicity.

Strengths

Key strengths of this study include the application of a qualitative methodology in an under-explored field, allowing us to capture both positive and negative perspectives with nuance. This serves as a first step towards integrating patient-reported toxicity symptoms into guideline development. We adhered to the COREQ guidelines, conducting interviews with semi-structured interview guides; and had two researchers independently analyse the data. The analysis was based on a robust framework derived from 25 systematically sourced studies.

Another major strength is the participation of the HOVON CML-MPN working group, which has the potential to bridge care, knowledge gaps, research, and guideline development within the evidence ecosystem. This national working group not only initiates and coordinates clinical trials but also develops guidelines, facilitating the integration of evidence into guidelines, while linking knowledge gaps to new research. The Delphi procedure enabled the HOVON CML-MPN working group to elaborate on the knowledge gaps identified during the qualitative interviews. As a result, this study translates conceptual research findings into concrete actions for CML guideline development.

Limitations

A limitation of our study is the underrepresentation of lower-educated patients. This may have biased our results toward more active and involved patients, as those recruited through online cancer communities tend to seek more active participation compared to those recruited through population-based methods [72]. However, several patients did indicate they would not fill out the test instrument, so we did capture the perspectives of less enthusiastic patients as well.

Additionally, three subthemes from previous studies were not identified in our research. First, the subtheme 'Fuels privacy concerns' [28] may not have emerged because our study took place in a research setting with an informed consent procedure in place. Second, 'Ensures holistic care' may have been less relevant in our setting, as our test instrument did not cover certain aspects of quality of life (e.g., social, spiritual factors) that were included in previous studies [17]. Third, 'Inhibits interaction and rapport' may not have been identified because we did not study the test instrument in active use. Though our new limiting subtheme 'Has no added value over open questions or conversation' shares some common ground, it is narrower in scope.

Conclusions

In conclusion, we identified 31 benefits and/or limitations that should be leveraged and/or addressed through targeted implementation strategies to maximise the value and uptake of systematic toxicity symptom reporting by patients. Most importantly, this study is the first to explore the views of guideline developers on real-world toxicity symptom reporting by patients. It also identified 14 CML-related knowledge gaps that could be addressed using such data, translating conceptual research findings into concrete actions for CML guideline development.

Further research and consensus are needed to develop coordinated methods and sustainable governance for the robust inclusion of real-world patient-reported toxicity data in professional guidelines. Such data will complement clinical trial findings, enhancing their applicability to real-world practice.

References

1. Gibbons, C., et al., *Routine provision of feedback from patient-reported outcome measurements to healthcare providers and patients in clinical practice.* Cochrane Database Syst Rev, 2021. **10**: p. CD011589.
2. Basch, E., et al., *Overall survival results of a trial assessing patient-reported outcomes for symptom monitoring during routine cancer treatment.* JAMA, 2017. **318**(2): p. 197-198.
3. Basch, E., et al., *Effect of electronic symptom monitoring on patient-reported outcomes among patients with metastatic cancer: a randomized clinical trial.* JAMA, 2022. **327**(24): p. 2413-2422.
4. Lai-Kwon, J., et al., *The feasibility, acceptability, and effectiveness of electronic patient-reported outcome symptom monitoring for immune checkpoint inhibitor toxicities: a systematic review.* JCO Clin Cancer Inform, 2023. **7**: p. e2200185.
5. Di Maio, M., et al., *The role of patient-reported outcome measures in the continuum of cancer clinical care: ESMO Clinical Practice Guideline.* Ann Oncol, 2022. **33**(9): p. 878-892.
6. Ruland, C.M., et al., *Effects of a computer-supported interactive tailored patient assessment tool on patient care, symptom distress, and patients' need for symptom management support: a randomized clinical trial.* J Am Med Inform Assoc, 2010. **17**(4): p. 403-10.
7. Thanarajasingam, G., et al., *Beyond maximum grade: modernising the assessment and reporting of adverse events in haematological malignancies.* Lancet Haematol, 2018. **5**(11): p. e563-e598.
8. Thanarajasingam, G., et al., *Reaching beyond maximum grade: progress and future directions for modernising the assessment and reporting of adverse events in haematological malignancies.* Lancet Haematol, 2022. **9**(5): p. e374-e384.
9. Scheuter, P., et al., *Patient-reported toxicity symptoms during tyrosinekinase inhibitor treatment in chronic myeloid leukemia: a systematic review and meta-analysis [submitted].* 2024.
10. Efficace, F., et al., *Patient- versus physician-reporting of symptoms and health status in chronic myeloid leukemia.* Haematologica, 2014. **99**(4): p. 788-93.
11. Hochhaus, A., et al., *European LeukemiaNet 2020 recommendations for treating chronic myeloid leukemia.* Leukemia, 2020. **34**(4): p. 966-984.
12. Basch, E., et al., *Development of the National Cancer Institute's patient-reported outcomes version of the common terminology criteria for adverse events (PRO-CTCAE).* J Natl Cancer Inst, 2014. **106**(9).
13. Foulon, S., et al., *Health state utility and quality of life measures in patients with chronic myeloid leukemia in France.* Qual Life Res, 2021. **30**(7): p. 2021-2032.
14. Phillips, K.M., et al., *Quality of life outcomes in patients with chronic myeloid leukemia treated with tyrosine kinase inhibitors: a controlled comparison.* Support Care Cancer, 2013. **21**(4): p. 1097-103.
15. Smit, Y., et al., *Measuring chronic myeloid leukaemia TKI-related toxic effects in the real world: a systematic review and critical assessment of content validity of patient-reported outcome measures.* Lancet Haematol, 2023.
16. Foster, A., et al., *The facilitators and barriers to implementing patient reported outcome measures in organisations delivering health related services: a systematic review of reviews.* J Patient Rep Outcomes, 2018. **2**: p. 46.

17. Campbell, R., et al., *Perceived benefits and limitations of using patient-reported outcome measures in clinical practice with individual patients: a systematic review of qualitative studies.* Qual Life Res, 2021.
18. Boyce, M.B., J.P. Browne, and J. Greenhalgh, *The experiences of professionals with using information from patient-reported outcome measures to improve the quality of healthcare: a systematic review of qualitative research.* BMJ Qual Saf, 2014. **23**(6): p. 508-18.
19. Carfora, L., et al., *Patients' experiences and perspectives of patient-reported outcome measures in clinical care: A systematic review and qualitative meta-synthesis.* PLoS One, 2022. **17**(4): p. e0267030.
20. Nic Giolla Easpaig, B., et al., *What are the attitudes of health professionals regarding patient reported outcome measures (PROMs) in oncology practice? A mixed-method synthesis of the qualitative evidence.* BMC Health Serv Res, 2020. **20**(1): p. 102.
21. Glenwright, B.G., et al., *Facilitators and barriers to implementing electronic patient-reported outcome and experience measures in a health care setting: a systematic review.* J Patient Rep Outcomes, 2023. **7**(1): p. 13.
22. Greenhalgh, J., et al., *Functionality and feedback: a realist synthesis of the collation, interpretation and utilisation of patient-reported outcome measures data to improve patient care*. 2017: Southampton (UK).
23. Bennink, C., et al., *Improving outcome-driven care in multiple myeloma using patient-reported outcomes: a qualitative evaluation study.* Patient, 2023. **16**(3): p. 255-264.
24. Lai-Kwon, J., et al., *Using implementation science frameworks to guide the use of electronic patient-reported outcome symptom monitoring in routine cancer care.* JCO Oncol Pract, 2024.
25. Lombi, L., S. Alfieri, and C. Brunelli, *'Why should I fill out this questionnaire?' A qualitative study of cancer patients' perspectives on the integration of e-PROMs in routine clinical care.* Eur J Oncol Nurs, 2023. **63**: p. 102283.
26. Sandhu, S., et al., *Implementation of electronic patient-reported outcomes in routine cancer care at an academic center: identifying opportunities and challenges.* JCO Oncol Pract, 2020. **16**(11): p. e1255-e1263.
27. Tolstrup, L.K., et al., *Patient-reported outcomes during immunotherapy for metastatic melanoma: mixed methods study of patients' and clinicians' experiences.* J Med Internet Res, 2020. **22**(4): p. e14896.
28. Tran, C., et al., *Utilizing digital health to collect electronic patient-reported outcomes in prostate cancer: single-arm pilot trial.* J Med Internet Res, 2020. **22**(3): p. e12689.
29. Smith, T.G., et al., *Oncology team perspectives on distress screening: a multisite study of a well-established use of patient-reported outcomes for clinical assessment.* Support Care Cancer, 2022. **30**(2): p. 1261-1271.
30. Brooks, J.V., et al., *The role of a symptom assessment tool in shaping patient-physician communication in palliative care.* J Pain Symptom Manage, 2020. **59**(1): p. 30-38.
31. Garcia Farina, E., et al., *Barriers to electronic patient-reported outcome measurement among patients with cancer and limited english proficiency.* JAMA Netw Open, 2022. **5**(7): p. e2223898.
32. Govindaraj, R., et al., *Assessing patient-reported outcomes in routine cancer clinical care using electronic administration and telehealth technologies: realist synthesis of potential mechanisms for improving health outcomes.* J Med Internet Res, 2023. **25**: p. e48483.

33. Kennedy, F., et al., *'We do need to keep some human touch'-Patient and clinician experiences of ovarian cancer follow-up and the potential for an electronic patient-reported outcome pathway: A qualitative interview study.* Eur J Cancer Care (Engl), 2022. **31**(2): p. e13557.
34. Lapen, K., et al., *Development and pilot implementation of a remote monitoring system for acute toxicity using electronic patient-reported outcomes for patients undergoing radiation therapy for breast cancer.* Int J Radiat Oncol Biol Phys, 2021. **111**(4): p. 979-991.
35. McCready, T.M., et al., *A remote symptom monitoring tool as part of ambulatory cancer surgery recovery: a qualitative analysis of patient experience.* JCO Oncol Pract, 2023. **19**(8): p. 595-601.
36. Pompili, C., et al., *Patients' views of routine quality of life assessment following a diagnosis of early-stage non-small cell lung cancer.* Interact Cardiovasc Thorac Surg, 2020. **31**(3): p. 324-330.
37. Samuel, C.A., et al., *Racial differences in user experiences and perceived value of electronic symptom monitoring in a cohort of black and white bladder and prostate cancer patients.* Qual Life Res, 2021. **30**(11): p. 3213-3227.
38. Shunmugasundaram, C., et al., *Development and implementation barriers of a new patient-reported measure: The Radiation therapy-related Inconvenience Questionnaire (RIQ).* J Med Imaging Radiat Oncol, 2023. **67**(7): p. 777-788.
39. Skåre, T.S., et al., *Barriers and facilitators when implementing electronic patient-reported outcome measures at a municipal cancer care unit: a qualitative study.* Cancer Nurs, 2023. **46**(4): p. E268-e275.
40. Snyder, C., et al., *A PRO-cision medicine intervention to personalize cancer care using patient-reported outcomes: intervention development and feasibility-testing.* Qual Life Res, 2022. **31**(8): p. 2341-2355.
41. Sung, J.H., et al., *Acceptability and accuracy of patient-reported outcome measures (PROMs) for surveillance of breathlessness in routine lung cancer care: A mixed-method study.* Lung Cancer, 2020. **147**: p. 1-11.
42. Thestrup Hansen, S., et al., *"I am sure that they use my PROM data for something important." A qualitative study about patients' experiences from a hematologic outpatient clinic.* Cancer Nurs, 2020. **43**(5): p. E273-E282.
43. Thestrup Hansen, S., et al., *User experiences on implementation of patient reported outcome measures (PROMs)in a Haematological outpatient clinic.* J Patient Rep Outcomes, 2020. **4**(1): p. 87.
44. Thestrup Hansen, S., et al., *Nurses' experiences when introducing patient-reported outcome measures in an outpatient clinic: an interpretive description study.* Cancer Nurs, 2021. **44**(2): p. E108-E120.
45. Thestrup Hansen, S., et al., *Haematologists' experiences implementing patient reported outcome measures (PROMs) in an outpatient clinic: a qualitative study for applied practice.* J Patient Rep Outcomes, 2019. **3**(1): p. 74.
46. Tong, A., P. Sainsbury, and J. Craig, *Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups.* Int J Qual Health Care, 2007. **19**(6): p. 349-57.
47. Geelen, I.G.P., et al., *Impact of hospital experience on the quality of tyrosine kinase inhibitor response monitoring and consequence for chronic myeloid leukemia patient survival.* Haematologica, 2017. **102**(12): p. e486-e489.

48. Ector, G.I., et al., *The development of a web-Based, patient-centered intervention for patients with chronic myeloid leukemia (CMyLife): design thinking development approach.* J Med Internet Res, 2020. **22**(5): p. e15895.
49. Verweij, L., et al., *Effectiveness of digital care platform CMyLife for patients with chronic myeloid leukemia: results of a patient-preference trial.* BMC Health Serv Res, 2023. **23**(1): p. 228.
50. Janssen, J.J.W.M., et al., *Guideline chronic myeloid leukemia [In Dutch]*. 2023, HOVON.
51. Anampa-Guzman, A., et al., *The rise of the expert patient in cancer: from backseat passenger to co-navigator.* JCO Oncol Pract, 2022. **18**(8): p. 578-583.
52. Bodenheimer, T., et al., *Patient self-management of chronic disease in primary care.* JAMA, 2002. **288**(19): p. 2469-75.
53. Calvert, M., et al., *Maximising the impact of patient reported outcome assessment for patients and society.* BMJ, 2019. **364**: p. k5267.
54. Vandvik, P.O. and L. Brandt, *Future of evidence ecosystem series: evidence ecosystems and learning health systems: why bother?* J Clin Epidemiol, 2020. **123**: p. 166-170.
55. Cartabellotta, A. and J.K. Tilson, *The ecosystem of evidence cannot thrive without efficiency of knowledge generation, synthesis, and translation.* J Clin Epidemiol, 2019. **110**: p. 90-95.
56. Steegmann, J.L., et al., *European LeukemiaNet recommendations for the management and avoidance of adverse events of treatment in chronic myeloid leukaemia.* Leukemia, 2016. **30**(8): p. 1648-71.
57. *MAGIC Evidence Ecosystem Foundation. The Digital and trustworthy evidence ecosystem*.
58. Webster, J. and B.D. Smith, *The case for real-world evidence in the future of clinical research on chronic myeloid leukemia.* Clin Ther, 2019. **41**(2): p. 336-349.
59. Efficace, F., et al., *Patient-reported symptom monitoring and adherence to therapy in patients with newly diagnosed chronic myeloid leukemia.* Cancer, 2023.
60. Cella, D., C.J. Nowinski, and O. Frankfurt, *The impact of symptom burden on patient quality of life in chronic myeloid leukemia.* Oncology, 2014. **87**(3): p. 133-47.
61. Moleman, M., et al., *Evidence-basing for quality improvement; bringing clinical practice guidelines closer to their promise of improving care practices.* J Eval Clin Pract, 2022.
62. Orsini, L.S., et al., *Improving transparency to build trust in real-world secondary data studies for hypothesis testing-why, what, and how: recommendations and a road map from the real-world evidence transparency initiative.* Value Health, 2020. **23**(9): p. 1128-1136.
63. Lerro, C.C., et al., *The bar is high: evaluating fit-for-use oncology real-world data for regulatory decision making.* JCO Clin Cancer Inform, 2024. **8**: p. e2300261.
64. Wang, S.V., et al., *Harmonized protocol template to enhance reproducibility of hypothesis evaluating real-world evidence studies on treatment effects: a good practices report of a joint ISPE/ISPOR task force.* Value Health, 2022. **25**(10): p. 1663-1672.
65. Maruszczyk, K., et al., *Systematic review of guidance for the collection and use of patient-reported outcomes in real-world evidence generation to support regulation, reimbursement and health policy.* J Patient Rep Outcomes, 2022. **6**(1): p. 57.
66. Castellanos, E.H., B.K. Wittmershaus, and S. Chandwani, *Raising the bar for real-world data in oncology: approaches to quality across multiple dimensions.* JCO Clin Cancer Inform, 2024. **8**: p. e2300046.
67. Solem, C.T., et al., *Systematic review and benchmarking of Quality-Adjusted Time Without Symptoms or Toxicity (Q-TWiST) in oncology.* Expert Rev Pharmacoecon Outcomes Res, 2018. **18**(3): p. 245-253.

68. Tannock, I.F., et al., *Patient-centered, self-funding dose optimization trials as a route to reduce toxicity, lower cost, and improve access to cancer therapy.* Ann Oncol, 2023.
69. Djodikromo, M.F., et al., *Patient-guided dose reduction of tyrosine kinase inhibitors in chronic myeloid leukaemia (RODEO study): study protocol for a prospective, multicentre, single-arm trial.* BMC Cancer, 2023. **23**(1): p. 231.
70. van Ommen-Nijhof, A., et al., *A revolving research fund to study efficient use of expensive drugs: big wheels keep on turning.* Ann Oncol, 2021. **32**(10): p. 1212-1215.
71. Kantarjian, H., et al., *The influence of drug prices, new availability of inexpensive generic imatinib, new approvals, and post-marketing research on the treatment of chronic myeloid leukaemia in the USA.* Lancet Haematol, 2022. **9**(11): p. e854-e861.
72. van Eenbergen, M.C., et al., *Comparing survivors of cancer in population-based samples with those in online cancer communities: cross-sectional questionnaire study.* JMIR Cancer, 2022. **8**(1): p. e19379.

Supplementary material

Table 1 COREQ (COnsolidated criteria for REporting Qualitative research) checklist [1]

Topic	Item no.	Guide questions/description	This study
Domain 1: Research team and reflexivity			
Personal characteristics			
Interviewer/facilitator	1	Which author/s conducted the interview or focus group?	Yolba Smit, Alexander Currie
Credentials	2	What were the researcher's credentials? E.g., PhD, MD	YS: MD MSc. AC: MSc
Occupation	3	What was their occupation at the time of the study?	YS: Policy advisor, PhD student. AC: Medical student
Gender	4	Was the researcher male or female?	YS: Female, AC: male
Experience and training	5	What experience or training did the researcher have?	YS: trained as a medical doctor and MSc in epidemiology. AC: trained as a medical doctor (halfway through Master phase). Both received recent training in conducting focus groups and individual interviews for qualitative research.
Relationship with participants			
Relationship established	6	Was a relationship established prior to study commencement?	YS knew two patients; one nurse specialist and five haematologists/ guideline developers form previous collaborations. One of the haematologists/guideline developer is the promotor in her PhD trajectory.AC: none.
Participant knowledge of the interviewer	7	What did the participants know about the researcher? e.g., personal goals, reasons for doing the research	All participants were made aware of the PhD position of YS and of the research questions.
Interviewer characteristics	8	What characteristics were reported about the inter viewer/facilitator? e.g., Bias, assumptions, reasons, and interests in the research topic	YS's main interest is applied guideline development. The focus of her PhD is to evaluate how patients' experiences can best be incorporated into the guideline development process in a systematic way. She assumes that guidelines in general can benefit to some extent from capturing patients' experiences, such as side effects, systematically.

Table 1 Continued

Topic	Item no.	Guide questions/description	This study
Domain 2: Study design			
Theoretical framework			
Methodological orientation and Theory	9	What methodological orientation was stated to underpin the study? e.g., grounded theory, discourse analysis, ethnography, phenomenology, content analysis	Participatory action research was the methodological orientation that underpinned this study, because it aims not only to understand reality but also improve it through change, and in doing so wants to empower the research participants [2].
Participant selection			
Sampling	10	How were participants selected? e.g., purposive, convenience, consecutive, snowball	See Methods "Study population and recruitment". In addition, we purposefully selected as many nurse specialists as possible because we assumed they discuss toxicity symptoms more in depth with patients, compared to haematologists whose primary focus is disease control. When we exhausted the pool of haematology nurse specialist, we proceeded by recruiting haematologists working in general hospitals. The number of nurse specialists working with CML patients in the Netherlands is limited (exact figure unknown).
Method of approach	11	How were participants approached? e.g., face-to-face, telephone, mail, email	See Methods "Study population and recruitment". In addition, We advertised for chronic phase CML patients willing to participate in either online video- or telephone-, group- or individual interviews through the website www.cmycml.nl and its newsletter. The call stated that we wanted to discuss the monitoring of toxicity symptoms related to TKI-treatment. We contacted all professionals face-to-face, by telephone, email, or through their professional society.
Sample size	12	How many participants were in the study?	See Table 1 and Figure 1 Flow chart of the study.

Table 1 Continued

Topic	Item no.	Guide questions/description	This study
Non-participation	13	How many people refused to participate or dropped out? Reasons?	See Figure 1 Flow chart of the study. Three nurse specialists declined participation (2 due to time pressure, one due to parental leave) and one did not respond. For the graphical overview interviews, six patients did not respond and one declined.
Setting			
Setting of data collection	14	Where was the data collected? e.g., home, clinic, workplace	For the test instrument sessions, patients were interviewed in three online focus group sessions of seven to nine participants, and three individual interviews of which one was by telephone and two online (one patient did not want to participate in a group, one patient preferred a telephone interview, and one patient for unclear reasons). For the graphical overview sessions, patients were interviewed in individual sessions (one face-to-face upon the patients request in a walk-in house near the patient's home, and nine online video meetings). Nurse specialists were interviewed in three individual online sessions and in two online sessions with two nurse specialists present. Haematologists were interviewed in seven individual online sessions and two groups of two participants each. Sessions with professionals were planned as individual or group sessions according to convenience. Participants were made aware that the test instrument was considered for routine administration within the cmyCML platform. At the start we preferred focus groups because we assumed that the interaction between participants would add valuable information. If focus groups were not practically feasible or if participants preferred individual interviews, that was possible as well.
Presence of non-participants	15	Was anyone else present besides the participants and researchers?	A facilitator from the CMyLife team was present during all focus groups, and at two focus groups and some interviews with professionals an apprentice was present.
Description of sample	16	What are the important characteristics of the sample? e.g., demographic data, date	See Table 1 Results.

Table 1 Continued

Topic	Item no.	Guide questions/description	This study
Data collection			
Interview guide	17	Were questions, prompts, guides provided by the authors? Was it pilot tested?	Semi-structured interview-guides were developed by the project group, for which the patient version was pilot tested on a patient representative.
Repeat interviews	18	Were repeat interviews carried out? If yes, how many?	No
Audio/visual recording	19	Did the research use audio or visual recording to collect the data?	See Methods "Interviews".
Field notes	20	Were field notes made during and/or after the interview or focus group?	Field notes were made during and/or after sessions.
Duration	21	What was the duration of the inter views or focus group?	Sessions took from 20 minutes (individual telephone interviews) to two hours (online focus groups).
Data saturation	22	Was data saturation discussed?	Saturation was defined as no new topics arising in the last two sessions, and was assessed and discussed after each session.
Transcripts returned	23	Were transcripts returned to participants for comment and/or correction?	Transcripts were not returned to participants for comments and/or corrections. The exception being that identified knowledge gaps entered the RAND-modified Delphi procedure, in which eight out of ten experts had been interviewed themselves.
Domain 3: analysis and findings			
Data analysis			
Number of data coders	24	How many data coders coded the data?	Two.
Description of the coding tree	25	Did authors provide a description of the coding tree?	Table 2 Results section.
Derivation of themes	26	Were themes identified in advance or derived from the data?	See Methods "Data analysis".

Table 1 Continued

Topic	Item no.	Guide questions/description	This study
Software	27	What software, if applicable, was used to manage the data?	See Methods "Data analysis".
Participant checking	28	Did participants provide feedback on the findings?	No, except for feedback on the knowledge gaps identified, through the RAND-modified Delphi procedure.
Reporting			
Quotations presented	29	Were participant quotations presented to illustrate the themes/findings? Was each quotation identified? e.g., participant number	Table 2 in Results and Table 2 in Supplementary material.
Data and findings consistent	30	Was there consistency between the data presented and the findings?	Results section.
Clarity of major themes	31	Were major themes clearly presented in the findings?	Table 2, Results section.
Clarity of minor themes	32	Is there a description of diverse cases or discussion of minor themes?	Results and Discussion sections.

Test instrument

We constructed a 48-item instrument out of all symptom items taken from the EORTC QLQ-CML-24 plus all CML symptoms from the EORTC Symptom Set [1-3]. During development and validation phases (not reported on here) the instrument grew to 61 items, the EORTC IL196.

Table 2 The 61 items included in the test instrument (EORTC IL196) at the end of the study

Skin problems	Feeling weak	Unintentional gas/flatulence
Skin colour change	Lacking energy	Pain
Itchy skin	Feeling drowsy	Pain interference
Skin rash	Trouble sleeping	Pain in chest
Dry/flaking/cracked skin	Headaches	Muscle cramps
Sore/painful skin	Dizziness	Aches/pains muscles/joints
Easy bruising	Difficulty remembering	Muscle weakness
Hair loss	Difficulty concentrating	Swelling body parts
Eye problems	Trouble thinking clear	Tingling/numbness hands/feet
Watery eyes	Feeling depressed	Shortness of breath
Burning eyes	Feeling irritable	Coughing problems
Discomfort eyes bright light	Feeling tense	Fevers/chills
Blurred vision	Worrying	Excessive sweating
Swelling face/eyes	Frequent urination	Problems tolerating heat/cold
Hearing problems	Bloated feeling abdomen	Hot flushes
Dry mouth	Acid indigestion/heartburn	Sore/enlarged nipples/breasts
Pain/soreness mouth	Feeling nauseated	Less interest sex
Different taste food/drink	Vomiting	Less sexual enjoyment
Lack appetite	Diarrhoea	Difficulty getting/maintaining erection
Tiredness	Feeling constipated	
Needing to rest	Abdominal pains/cramps	

Semi-structured interview guide patient focus groups or individual interviews on test instrument

- Prompts:
  - Can you elaborate?
  - Can you give me an example?
  - How do you feel about [example given by other patients]?
- What do you want to know about your side effects?
- Would you use this questionnaire to keep score of your side effects? Why?
- What insights would the answers to the questionnaire give you, or to what insights might repeated answers lead?
- To which outcomes or actions would these insights lead?
- Would you like to compare your answers to your own previous answers, or to other CML patients, or otherwise?
- *Semi-structured interview guide patient focus groups or individual interviews on graphical overview*What is your opinion on the graphical overview?
- What were your first thoughts?
- What do you think of the design?
  - In general, legend, graphs, icons
  - What can be improved and how?
  - What could be left out?
  - What could be added?
  - Experiencing symptoms?
  - Change in anxiety, stress?
  - Anything else?
- The symptoms that you personally scored 'not at all' have been left out. What is your opinion on that?
- Some professionals stated they would like patients to indicate which symptoms they want to discuss during consultation. What do you think?
- How does the graphical overview influence your wellbeing?
- Would you like to discuss your overview with others? If yes, with whom and why? If not, why not?

Semi-structured interview guide nurse practitioner focus groups or individual interviews

- Prompts:
  - Can you elaborate?
  - Can you give me an example?
  - How do you feel about [example given in another workshop]?
- The answers to the questions in the questionnaire might be useful during consultations. What is your opinion on that?
- What would you do with the information from the questionnaire?
- Would you discuss the information during consultation? Why yes/no?
- Would you also like the information from the questionnaire if patients switched TKI, or changed TKI dosage, or stopped their TKI?
- How can we reach patients with low health-literacy when we want to use this questionnaire?

Semi-structured interview guide haematologist & guideline developers focus groups or individual interviews

In addition to above, questions relevant to the CML guideline were:

- As a guideline developer/haematologist, what is your opinion on the collection of symptoms of side effects of CML patients?
- Will the information benefit the guideline? Why yes/no?
- Which evidence gaps might this information help solve?
- Are there advantages/disadvantages to using this instrument, when developing the guideline?
- Are there requirements for the collection of symptom data for guideline development?
- Would information on the symptoms of patients that stopped their TKI be useful for the guideline?

Table 3 Framework of previous studies and this study, with new (sub)themes and their assessment in italics

Previous studies[4-28]	**This study**
Active patient involvement and partnership	**Idem**
Enables greater awareness and reflection (Benefit)	Idem
Encourages patient involvement (Benefit)	Idem
Facilitates goal setting and shared decision making (Benefit)	Idem
Influences honesty (Mixed)	Idem
Permits discussion of sensitive topics (Benefit)	Idem
	Enables *self-prevention* and self-care (Benefit) [£]
Leads to worse symptom experience (Limitation)	Idem
Fulfils desire to help others (Benefit)	Idem
Fuels privacy concerns (Limitation)	Not identified *Objectifies subjective experience (Benefit)*
Focus of consultation	**Idem**
Helpful as a screening tool (Benefit)	Idem
Prioritises patients' needs (Benefit)	Idem
Provides one piece of the picture (Limitation)	Idem
Structures consultations and improves efficiency (Benefit)	Idem
Shifts away from the main medical problem (Limitation) [$]	*Suits paramedical consultation (Mixed)* [$]
Raises unrealistic expectations for care (Limitation) [$]	Merged with 'Suits paramedical consultation' Provides redundant information[§] Provides reassurance that clinicians care[¥]
Quality of care	**Quality of individual patient care**
Assists diagnosis and enables tailored (self-) care (Benefit) [£]	Moved to 'Active patient involvement and partnership'. Care part covered by 'Helps determine and monitor side effects' and 'Prompts appropriate, standardised action and acceptance'
Ensures holistic care (Benefit)	Not identified
Can inaccurately estimate the problem (Limitation)	… /not specific enough to be clinically meaningful [#]
Prompts appropriate, standardised action (Benefit)	… *and acceptance*
Is an opportunity for the education of patients and professionals (Benefit)	Assists in learning (Benefit)

Table 3 Continued

Previous studies[4-28]	This study
	Helps determine *and monitor* side effects of treatment *
	Does not establish a causal link (Limitation)
Standardised monitoring of patient outcomes over time	**Moved to 'Quality of individual patient care'**
Helps determine effectiveness and side effects of treatment (Benefit)	Merged, moved to 'Quality of individual patient care' *
Useful for monitoring changes and tracking progress (Benefit)	Merged, moved to 'Quality of individual patient care' *
Patient-clinician relationship	**Moved or not identified**
Provides reassurance that clinicians care (Benefit)	Moved to 'Focus of consultation' ¥
Inhibits interaction and rapport (Limitation)	Not identified
Lack of valuable information	**Merged/moved to other themes**
PRO data is not specific enough to be clinically meaningful (Limitation)	Merged with 'Can inaccurately estimate the problem' #
Provides redundant information (Limitation)	Moved to 'Focus of consultation'§
Suitability for all patients	**Suitability and acceptability**
Suitability for all patients (Limitation)	*..., clinicians and workflows*
Confronts too much with disease (Limitation)	Confronts too much with disease (Limitation)
	Usefulness depends on (change in) symptom severity (Mixed)
	Has no added value over open questions or conversation (Limitation)
Improving care with aggregated data	**Improving care and clinical guidelines with aggregated data**
Data interpretation is challenging (Limitation)	*Real-world data has pros and cons (Mixed)*
Benchmarks hospital performance and clinical trials which will eventually lead to lower costs (Benefit)	Benchmarks clinical trials and medication costs (Benefit)
Puts individual experience in context (Benefit)	Puts individual experiences in context (Benefit)
	Systematically includes patients' experiences in guidelines (Benefit) Personalises guideline advise (Benefit)
	Fills in knowledge gaps (Benefit)

Main themes and subthemes were taken from Campbell et al.[4] and completed with (sub) themes from other relevant studies. 'Idem' indicates that we used the exact same phrasing for the (sub)theme as previous studies did. Special characters (£, $, *, #, §, ¥) indicate merged with, or moved to other (sub)themes

Methods RAND-modified Delphi procedure

We used a RAND-modified Delphi method to reach consensus on, and prioritise knowledge gaps on TKI toxicity, that may be addressed (in part) by aggregated data from the test instrument [29]. In step (1) all knowledge gaps mentioned during the interviews were identified. For step (2) all panel members received a questionnaire by email, asking after completeness of, and agreement with, the formulated knowledge gaps (see the questionnaires below). All panel members had the opportunity to add new knowledge gaps and to formulate feedback. Answers were analysed by adding all new knowledge gaps and by reformulating knowledge gaps, based on the feedback received. For step (3) the set was sent to panel members for final appraisal, including the feedback received and how feedback was processed. In step (4) panel members were asked to prioritise knowledge gaps in the final set, in order of relevance for the Dutch CML guideline.

Questions administered in the RAND-modified Delphi procedure step (b)

1. Which knowledge gaps are you missing in this list of nine knowledge gaps, that you consider relevant for the CML guideline and that could be (partly) answered with patient-reported TKI-related toxicity?
2. Which of the nine knowledge gap(s) can be better formulated, and in what way?
3. Do you have general feedback regarding these knowledge gaps?

Questions administered in the RAND-modified Delphi procedure step (c)

1. Is the feedback properly processed?
2. Do you have any feedback on the knowledge gaps as they are currently formulated?

Questions administered in the RAND-modified Delphi procedure step (d)

Please indicate below what priority you think each relevant knowledge gap has

- Enter a number between 1 (highest priority) and 14 (lowest priority) behind each knowledge gap
- Use each number from 1 to 14 only once

Table 4 Themes and subthemes previously identified*, with an assessment of benefit or limitation, with illustrative quotes from this study

Themes and subthemes	Benefit or limitation	Illustrative quotes
Active patient involvement and partnership		
Enables greater awareness and reflection	Benefit	So, seeing a list like this helps to be able to identify what symptoms are, or could indeed possibly be, CML-related, before I sweep them under the carpet or put them in a drawer and don't want to know anything more about them (P)
		Maybe there are things that don't, yes, that don't come to my mind or that I don't think about, that I might have to tick (P)
		Yes, that's possible, look, I, I keep track of my own side effects and then I try to relate those to things I have done or that I do (P)
		So, I have tried for myself to uhm get in a sort of biological rhythm in taking the medication [...] to see what fits me best. And that's why I've kept notes of all my side effects, to come to my ideal situation (P)
Encourages patient involvement	Benefit	I think it is also a good summary for them [patients] to prepare, for the outpatient clinic, or for when they speak to us [nurse specialists] or a specialist (NS)
		I can also imagine that it is nice for patients, so, okay, the scores are high so I will talk about that in the doctor's office (H/GD)
		Well, I do think that if, for example, some items keep popping up that really trouble you, then you can bring it to your haematologist and say: "Well, listen, this or that bothers me for such a long time, can we do anything about it?" (P)
Facilitates goal setting and shared decision making	Benefit	Well look, the challenge is more: the extent to which the information provided by the patients PROs can influence decision-making in the doctor's office. A good response [to treatment] with as few side effects as possible, that is the trick. And the normal guidelines are focused on the best possible response. And side effects are not discounted there. While you can also opt for a slightly less optimal response, but with less misery and a better quality of life (H/GD)
Influences honesty	Limitation	If you have a patient who says yes to everything, then I don't do anything with it. [...] You can never have everything. [...] you have patients who supposedly suffer from everything. And you must get that from a conversation and not from a questionnaire. There are people who [...] do not dare to say that it is going well, because otherwise they don't feel heard. [...] I always find that the complicated part of these kinds of questionnaires. That you have no subjective feeling about whether this is realistic (H/GD)
		I used to really work a lot [with questionnaires] and then I would ask my open question and something different always emerged. [...] So that I thought: "oh yes but why do you still fill this in if you give me completely different answers?" So, I would keep asking my open question. [...] So, for me open questions give me way more than that list for those people, but I don't know why that is (NS)

Table 4 Continued

Themes and subthemes	Benefit or limitation	Illustrative quotes
Permits discussion of sensitive topics	Benefit	But, sometimes uhm, I think gosh, I must defend myself by saying, uhm, guys, the treatment is there and working great, but uhm, can we talk side effects because uhm they do influence quality of live negatively to a great extent. And uhm, I often feel guilty talking about this with people, or I think oh, I must defend myself (P)
		What I really do miss, particularly about symptoms they [patients] want to discuss, is sexual dysfunction. [..] While that is a really complex topic for a lot of people to discuss (NS)
Leads to worse symptom experience	Limitation	I actually started feeling everything (NS)
		But they [the questions] are negatively formulated [...] they are not positively formulated in the context of positive health (H/GD)
Fulfils desire to help others	Benefit	I wouldn't fill it in for me, but for others. [...] to, to, just like this, help them to gain insight [...] I would like to help, uhm, others (P)
Focus of consultation		
Helpful as a screening tool	Benefit	Look, I see this as a screening instrument, right? [...] So, you check, are there things that have not been reported spontaneously that are relevant, and uhm, if a flag is raised, so, somebody experiences a symptom than you go into that, checking if there is a relationship at all and continue questioning the impact (H/GD)
		The medical doctor must see it when it isn't stated that action is still needed [...] so I think more like prior notice (H/GD)
Prioritises patients' needs	Benefit	Well, I think it is especially useful at the time when you, uhm discuss with your treating physician, that they get a better insight in whether you are doing fine in the long term (P)
		For the physician, the main issue is whether the disease is under control, uhm, and also a little bit about quality of life, but no, not a lot, because that is not core business, so, at that moment I think (P)
		If a doctor doesn't want to hear about it, that whining, that the patient can say: "yes, but look, it's very important to me because this is a complaint that comes back every time" (H/GD)
		So, the patient [gets] the possibility to discuss it [their symptoms] (H/GD)
Provides reassurance that clinicians care	Benefit	It is certainly a starting point to talk about with a health care provider. If I tick a lot, the [specialist] might say: okay, it's worse than I thought. And you also come to talk about quality of life. So, you can discuss: okay, are there things we can do? [...] So, you can address that better that way. Then she [nurse specialist] also has a little insight into how I experience it (P)
Provides one piece of the picture	Limitation	But that's the problem with questionnaires, you don't ask everything, and you might miss things (H/GD)

Table 4 Continued

Themes and subthemes	Benefit or limitation	Illustrative quotes
Structures consultations and improves efficiency	Benefit	Well, I think it is especially useful when you talk to your treating physician, that he gains a bit more insight in whether you are doing well in the long term. It's often only 10 minutes you get with him [...] before you know it, you're outside again (P)
		And in our hospital, I see that in the anaesthesia department, they do that very well, they send all questionnaires to patients preoperatively and that works great. That improves the consultation throughput, right? So, in the end, uhm, it provides a benefit for health care (NS)
		And [...] shorten contact time. Or even, in my ideal situation, skip it, you know. What I want is on demand consultations. So, if you fill out the [test] instrument, and I see it and its's stable, and you have no further questions, then we can skip it the next three to four months. Also, just because, as you [other participant] just said, not focus on the disease too much (NS)
Provides redundant information	Limitation	As a researcher I want to know as much as possible, as a medical doctor I do not want to know everything. Because I can't do anything about it, I can't ... then I think, well, all those symptoms ... (H/GD)
		It's a bit more my personal feeling. Like: "gee, what are we going to do with this data? What does it lead to? Does it really have added value?." On the other hand, if you have a decent conversation with a patient, a lot of things are already discussed (H/GD)
		But I don't really need to know everything that bothers patients just a little bit. [...] and this is not such a politically correct comment, but I have fifteen minutes for them, and I see them 3 times a year or so, and I want to know the highlights [...] then I really just want to know the highlights and know what really bothers them (H/GD)
		Well look, [...] I usually do outpatients, patients who are involved in all kinds of things and have to deal with all kinds of things, and this provides so much information that I think: well, I ask 6 questions: are you tired, do you have muscle problems, do you have problems with your skin, do you have problems with your eyes, how are things going in your private life and are there any changes, can I help you with anything? ? Then I also get the information I need. Then I think this is just too much stimulus, too much information (H/GD)
Quality of *individual patient* care		
Assists in learning	Benefit	Yes, then I'm looking for how it is for others. I try to learn from that, I find it interesting [...] let's say 100 people keep track of all [their] side effects, then a certain profile emerges, well, I find that interesting. Because that can help me, but maybe others too (P)
		Indeed, to start recognizing patterns [P]
		What I would find nice myself is if there are people who, when you get the data, yes, I always have such and such a side effect, but this works great. And that I didn't know that yet [H/GD]

Table 4 Continued

Themes and subthemes	Benefit or limitation	Illustrative quotes
Can inaccurately estimate the problem/ not specific enough to be clinically meaningful	Limitation	Because were one person fills out 'a little', the other person states 'quite a bit', which maybe has more to do with personality than with the severity of the itch (NS) So, because, if I'm a squeaker, I mind everything and if I'm a tough boy, I don't mind at all (H/GD) Are you short of breath? […] "Well, yesterday it was muggy. […] So, if the humidity is over 80%, I feel a tightness in my chest, but I am not short of breath" (H/GD) Al lot of patients indicate they are tired all the time, but the persons I see have aged 10 years in the meantime, do you understand? So, uhm, age moves up. So, age is crucial and the interpretation of symptoms […] [is] different for women than for men (H/GD) Do you ever feel short of breath? Yeah well, as I walked up the stairs this morning, I thought, oh god, I have to work again, oh, I'm tired. Do you understand? (H/GD)
Suitability and acceptability		
Confronts too much with disease	Limitation	And not that it doesn't matter, or that I can't think about it [i.e., symptoms of toxicity], but yes, I prefer not to think about it (P) That they [patients] say: "yes, I kind of stopped because I was confronted too much with my symptoms. I started to feel everything" So that's the downside. So that is also the freedom that patients should have, of whether or not to use [the questionnaire] (H/GD) The complicated thing is that you start thinking about everything. That's my experience too. When a doctor asks me something, I think, 'Oh, do I have it or don't I?' And then I say no, for example, but then I think, 'But do I really never have that?' While the question would never have occurred to me if it hadn't been an issue. […] burdened with questions about what is possible. That's like reading the package leaflet, you know? Then you are waiting to get all those things, while that may not apply to you at all (P)
Improving care and clinical guidelines with aggregated data		
Benchmarks clinical trials and medication costs	Benefit	You know that perspective can be brought forward by saying, yes, but in practice you see that this is the trouble, you know, the burden which this medication causes (H/GD) And especially because this is a disease that [people] carry with them for the rest of their lives and need to take medication for the rest of their lives, so are condemned the rest of their lives, I think, we shouldn't just look at maximum disease response, so normal survival […] so that perspective needs to be taken into account by saying, yes, well, this is the burden in real-world from this medication. And that, I think, needs to be assessed scientifically, how bad that burden is, and then relate that to uhm, uhm, response, and then weigh (H/GD) […] has a relationship with financial toxicity, in that uhm, we pay too much for the medication […] because something is promised with a certain response […] but that ignores toxicity and impact on [quality] of life, and that is not discounted in the price we must pay (H/GD)

Table 4 Continued

Themes and subthemes	Benefit or limitation	Illustrative quotes
Puts individual experience in context	Benefit	In the discussion with the patient, that's when the benchmark comes into play [H/GD]
		Do you know what they have for CLL [chronic lymphoid leukaemia]? Uhm uh, they have bars for, well, your own complaints, and uh there is a bar next to it, and that is your peer group, namely people with CLL, and there is a bar next to that for the, well, normal population. That is also very reassuring [...] because then you have a lot of people who say yes: I am tired, but it's actually not too bad for me because those other people suffer much more [H/GD]

* Main themes and subthemes were taken from Campbell et al. [1] and completed with (sub)themes from other relevant studies [2-25]. Identifiers in brackets after quotes: H/GD: haematologist/guideline developer; NS: nurse specialist; P: patient

Supplementary references

1. Tong, A., P. Sainsbury, and J. Craig, *Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups.* Int J Qual Health Care, 2007. **19**(6): p. 349-57.
2. Baum, F., C. MacDougall, and D. Smith, *Participatory action research.* J Epidemiol Community Health, 2006. **60**(10): p. 854-7.
3. Efficace, F., et al., *International development of an EORTC questionnaire for assessing health-related quality of life in chronic myeloid leukemia patients: the EORTC QLQ-CML24.* Qual Life Res, 2014. **23**(3): p. 825-36.
4. Efficace, F., et al., *Validation and reference values of the EORTC QLQ-CML24 questionnaire to assess health-related quality of life in patients with chronic myeloid leukemia.* Leuk Lymphoma, 2021. **62**(3): p. 669-678.
5. Sodergren, S.C., et al., *Developing symptom lists for people with cancer treated with targeted therapies.* Targeted Oncology, 2021. **16**(1): p. 95-107.
6. Campbell, R., et al., *Perceived benefits and limitations of using patient-reported outcome measures in clinical practice with individual patients: a systematic review of qualitative studies.* Qual Life Res, 2021.
7. Boyce, M.B., J.P. Browne, and J. Greenhalgh, *The experiences of professionals with using information from patient-reported outcome measures to improve the quality of healthcare: a systematic review of qualitative research.* BMJ Qual Saf, 2014. **23**(6): p. 508-18.
8. Carfora, L., et al., *Patients' experiences and perspectives of patient-reported outcome measures in clinical care: A systematic review and qualitative meta-synthesis.* PLoS One, 2022. **17**(4): p. e0267030.
9. Glenwright, B.G., et al., *Facilitators and barriers to implementing electronic patient-reported outcome and experience measures in a health care setting: a systematic review.* J Patient Rep Outcomes, 2023. **7**(1): p. 13.
10. Greenhalgh, J., et al., *Functionality and feedback: a realist synthesis of the collation, interpretation and utilisation of patient-reported outcome measures data to improve patient care*. 2017: Southampton (UK).
11. Nic Giolla Easpaig, B., et al., *What are the attitudes of health professionals regarding patient reported outcome measures (PROMs) in oncology practice? A mixed-method synthesis of the qualitative evidence.* BMC Health Serv Res, 2020. **20**(1): p. 102.
12. Bennink, C., et al., *Improving outcome-driven care in multiple myeloma using patient-reported outcomes: a qualitative evaluation study.* Patient, 2023. **16**(3): p. 255-264.
13. Lai-Kwon, J., et al., *Using implementation science frameworks to guide the use of electronic patient-reported outcome symptom monitoring in routine cancer care.* JCO Oncol Pract, 2024.
14. Lombi, L., S. Alfieri, and C. Brunelli, *'Why should I fill out this questionnaire?' A qualitative study of cancer patients' perspectives on the integration of e-PROMs in routine clinical care.* Eur J Oncol Nurs, 2023. **63**: p. 102283.
15. Sandhu, S., et al., *Implementation of electronic patient-reported outcomes in routine cancer care at an academic center: identifying opportunities and challenges.* JCO Oncol Pract, 2020. **16**(11): p. e1255-e1263.

16. Tolstrup, L.K., et al., *Patient-reported outcomes during immunotherapy for metastatic melanoma: mixed methods study of patients' and clinicians' experiences.* J Med Internet Res, 2020. **22**(4): p. e14896.
17. Tran, C., et al., *Utilizing digital health to collect electronic patient-reported outcomes in prostate cancer: single-arm pilot trial.* J Med Internet Res, 2020. **22**(3): p. e12689.
18. Smith, T.G., et al., *Oncology team perspectives on distress screening: a multisite study of a well-established use of patient-reported outcomes for clinical assessment.* Support Care Cancer, 2022. **30**(2): p. 1261-1271.
19. Brooks, J.V., et al., *The role of a symptom assessment tool in shaping patient-physician communication in palliative care.* J Pain Symptom Manage, 2020. **59**(1): p. 30-38.
20. Garcia Farina, E., et al., *Barriers to electronic patient-reported outcome measurement among patients with cancer and limited english proficiency.* JAMA Netw Open, 2022. **5**(7): p. e2223898.
21. Govindaraj, R., et al., *Assessing patient-reported outcomes in routine cancer clinical care using electronic administration and telehealth technologies: realist synthesis of potential mechanisms for improving health outcomes.* J Med Internet Res, 2023. **25**: p. e48483.
22. Kennedy, F., et al., *'We do need to keep some human touch'-Patient and clinician experiences of ovarian cancer follow-up and the potential for an electronic patient-reported outcome pathway: A qualitative interview study.* Eur J Cancer Care (Engl), 2022. **31**(2): p. e13557.
23. Lapen, K., et al., *Development and pilot implementation of a remote monitoring system for acute toxicity using electronic patient-reported outcomes for patients undergoing radiation therapy for breast cancer.* Int J Radiat Oncol Biol Phys, 2021. **111**(4): p. 979-991.
24. McCready, T.M., et al., *A remote symptom monitoring tool as part of ambulatory cancer surgery recovery: a qualitative analysis of patient experience.* JCO Oncol Pract, 2023. **19**(8): p. 595-601.
25. Pompili, C., et al., *Patients' views of routine quality of life assessment following a diagnosis of early-stage non-small cell lung cancer.* Interact Cardiovasc Thorac Surg, 2020. **31**(3): p. 324-330.
26. Samuel, C.A., et al., *Racial differences in user experiences and perceived value of electronic symptom monitoring in a cohort of black and white bladder and prostate cancer patients.* Qual Life Res, 2021. **30**(11): p. 3213-3227.
27. Shunmugasundaram, C., et al., *Development and implementation barriers of a new patient-reported measure: The Radiation therapy-related Inconvenience Questionnaire (RIQ).* J Med Imaging Radiat Oncol, 2023. **67**(7): p. 777-788.
28. Skåre, T.S., et al., *Barriers and facilitators when implementing electronic patient-reported outcome measures at a municipal cancer care unit: a qualitative study.* Cancer Nurs, 2023. **46**(4): p. E268-e275.
29. Snyder, C., et al., *A PRO-cision medicine intervention to personalize cancer care using patient-reported outcomes: intervention development and feasibility-testing.* Qual Life Res, 2022. **31**(8): p. 2341-2355.
30. Sung, J.H., et al., *Acceptability and accuracy of patient-reported outcome measures (PROMs) for surveillance of breathlessness in routine lung cancer care: A mixed-method study.* Lung Cancer, 2020. **147**: p. 1-11.
31. Dalkey, N. and O. Helmer, *An experimental application of the delphi method to the use of experts.* Management Science, 1963. **9**(3): p. 458-67.

Chapter 6

General discussion

This thesis focused on making real-world patient-reported toxicity symptoms accessible for CML care and the CML guidelines. Its aim was to develop and validate a new instrument, along with its workflow, to capture real-world patient-reported TKI toxicity symptoms in CML, test its application, and prepare for implementation.

In this final chapter, we first summarise our main findings before reflecting on them from a broader perspective: How to choose an instrument to capture patient-reported outcome measures in CML, in the absence of a gold standard? And what have we learned about patients' comparative experiences of toxicity, and what is the way to go forward? Finally, we discuss the implications for guidelines and future research.

Summary of main findings

- In our systematic review and meta-analysis of 11 studies, including 2,987 CML patients, the prevalence of 13 out of 47 symptoms differed significantly between dasatinib, imatinib, and nilotinib. However, robust evidence on the prevalence and severity of patient-reported toxicity symptoms - based on direct or adjusted comparisons between TKI types - is still lacking (**Chapter 2**).
- None of the six instruments to measure patient-reported TKI-related toxicity in CML patients, identified in a systematic review, had sufficient content validity based on our assessment (**Chapter 3**).
- At least 278 CML patients and 18 CML professionals participated in the development of a comprehensive new EORTC instrument and its workflow. Pilot data showed significant severity differences between dasatinib, imatinib, and nilotinib, in a third of symptoms and scales, with differences up to 30% (**Chapter 4**). These data can be used in shared decision-making.
- The perspectives of patients and health care professionals on the benefits and limitations of real-world patient-reported toxicity monitoring for both clinical guidelines and patient care, were explored in **Chapter 5**. Compared to an existing framework, novel benefits centred on the use of aggregated data. Through a Delphi procedure, the Dutch CML guideline development group identified fourteen key knowledge gaps in CML care that could be addressed through systematic patient-reported toxicity monitoring.

Our research shows that the evaluation of patient-reported symptoms in CML has, up to now, provided scattered, mostly unadjusted, often inconsistent, and sometimes conflicting evidence, and mainly on the three eldest, and thereby most-used, TKIs. This obviates a call for the widespread use of

patient-reported outcome measures in CML research, in a standardised manner. Standardised, because the use of different instruments and different evaluation methodologies makes it difficult to summarise and interpret existing evidence.

Choosing an instrument in the absence of a gold standard

Unfortunately, there is no consensus on which instrument is the gold standard to measure patient-reported outcome measures in CML. With our newly developed instrument, there are now six patient-reported outcome instruments that have been validated to some extent: the new EORTC IL386 [1]; the EORTC QLQ-CML24 [2, 3]; the EORTC Symptom Set [4]; the FACT-LEU [5]; the HM-PRO [6-9]; and the MDASI-CML [10]. The last three are not sufficiently specific for TKI-treated CML, and, in our opinion, should not be used. The three EORTC instruments are all feasible and have the benefit of a 0-100 score, which facilitates clinical interpretation because points can be interpreted as percentage-points. The three instruments differ in four characteristics, which provide options depending on the intended use: (1) Quality of life versus symptoms, (2) Symptom comprehensiveness, (3) Design for clinical care versus research, and (4) Measurement properties beyond content validity.

If **quality of life** assessment is required, the EORTC QLQ-CML24 is the preferred instrument, which particularly applies to research purposes. It is less suited for practical clinical use, as the evaluation and interpretation of scales requires computation, skills and experience. In addition, abstract concepts such as 'physical functioning' or 'symptom burden' lack application in clinical care. In cases where both quality of life and symptom assessment are needed, the new EORTC IL407 can be added to the EORTC QLQ-CML24. Combined, the QLQ-CML24 and the IL407 include all symptoms from the EORTC IL386.

The new EORTC IL386 is the most **comprehensive** instrument, offering greater detail on fatigue and mood problems while also including items on for example wound healing—a symptom that the other two EORTC instruments do not cover. Most importantly, the new EORTC IL386 is the only CML-specific instrument that includes items on sexuality. Sexuality has not been assessed using validated instruments in CML patients, so we do not know how to interpret the high prevalence and severity of sexuality problems we found. Two studies have used generic measures for sexuality problems. The first, a Chinese study, evaluated 'decrease in sexual desire', which, while prevalent, was not among the most frequently reported symptoms [11]. The second study, conducted

in Israel, assessed 'sexual function impairment', which was reported by 40% of patients [12]. This scarcity of data, combined with the high prevalence of sexuality-related symptoms in our study, highlights the need for further investigation into sexual health in CML. Sexuality is an understudied area, not only in CML but also within haematology more broadly [13, 14].

Even though comprehensive, the IL386 is still not as comprehensive as patients would like. Notably, menstrual changes are not included, as suitable items were lacking in the EORTC Item Library. Menstrual disorders were only reported in a single study that used an unvalidated generic Chinese instrument. In that study, half of the women under 50 years of age reported experiencing either amenorrhea, hypomenorrhea, or hypermenorrhoea [11]. This certainly indicates that menstrual problems should be evaluated further.

Both the EORTC QLQ-CML24 and the EORTC Symptom Set were designed primarily for research, whereas the EORTC IL386 was **designed for both research and clinical care** involving patients not only in the development of the instrument itself, but also in its workflow. This led to adaptations to patients' preferences, such as the grouping of items in categories, with generic filter questions that lead to skipping or entering categories, and a prototype graphical overview. The EORTC QLQ-C30 (the most used EORTC Quality of Life instrument), is part of the QLQ-CML24, and has computer adaptive testing available, which allows the questionnaire to be tailored to the individual [15]. However, the primary goal of this approach is to increase precision and enable the use of smaller sample sizes in research, without sacrificing statistical power. For each symptom in the QLQ-C30, additional items have been added to the computer adaptive testing item bank, which can be triggered based on patients' responses. For example, for the item diarrhoea, one item in the QLQ-C30, 13 additional items can be triggered [16], providing a more detailed overview of the diarrhoea dimension. This contrasts with our generic filter questions, which apply to broad categories to enhance patient's experience. Whether these options will improve the user- experience, and, ultimately, increase response rates remains to be seen.

Beyond content validity, both the IL386 [1] and the QLQ-CML24 [2, 3] have undergone validation to some extent, which the Symptom Set has not. Content validity is considered the most important measurement property [17] because items must be relevant, comprehensive and comprehensible, and all three instruments have undergone content validation. However, other

measurement properties are also important when determining the quality of an instrument: construct validity, criterion validity, reliability, responsiveness and interpretability (Figure 1), though not all properties apply to individually scored items [18, 19].

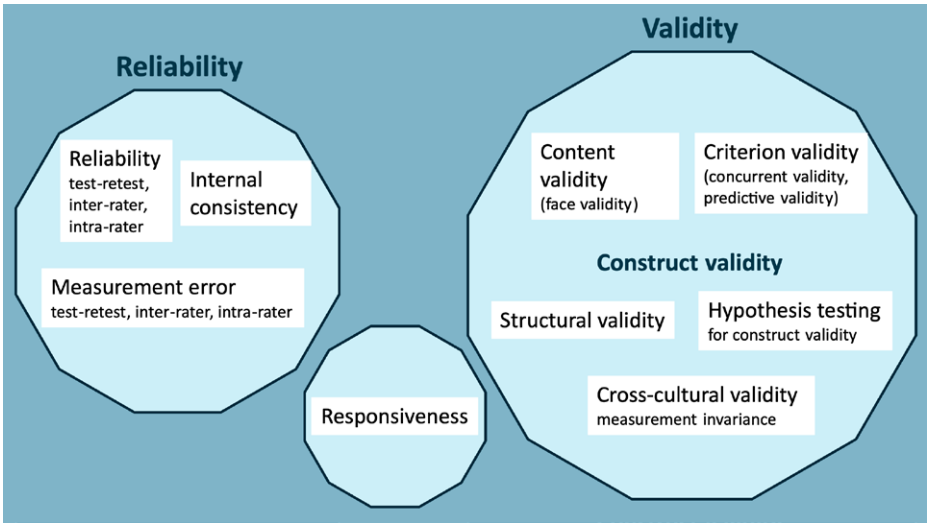


Figure 1 The nine measurement properties of outcome measurement instruments according to COSMIN, adapted from [18]

A description of the measurement properties of the three EORTC instruments is given in Table 1, which shows that validation of individual items has only taken place for the new IL386, whereas the validation of the QLQ-CML24 has focused solely on its scales. It seems plausible that the validation of certain measurement properties—such as reliability, measurement error, hypothesis testing, cross-cultural validity, criterion validity, and responsiveness—in one instrument could be extrapolated to the same (combination of) individual items in the other instrument. If this extrapolation is possible, validation of individual items for the QLQ-CML24 and Symptom Set can be deduced from the validation of individual items of the IL386 and seem satisfactory for most measurement properties evaluated. Validation of criterion validity – the degree to which the scores of an instrument are an adequate reflection of a gold standard [18], has not taken place yet.

Table 1 Measurement properties of three EORTC instruments, feasible for the monitoring of real-world toxicity symptoms in TKI-treated CML patients [$]

Measurement property	Instrument		
	EORTC IL386 [1]	**EORTC QLQ-CML24 [2, 3] ***	**EORTC Symptom Set [4]**
Reliability			
Internal consistency	Not applicable to individually scored items (formative model). For the 7 symptom scales[&], internal consistency (Cronbach's alpha) ranged from 0.73 to 0.93	Not applicable to individually scored items (formative model). Internal consistency (Cronbach's alpha) was 0.83 for symptom burden scale	Not applicable as all symptoms are scored individually (formative model)
Reliability	Weighted kappa calculated in patients on same TKI in same dosage 2 weeks apart: fair to excellent reliability for all 70 items evaluated	Intraclass correlation coefficient calculated in patients 3 to 7 days apart: 0.83 for symptom burden scale. Not evaluated for individual items	Not evaluated
Measurement error	Correct classification evaluated across 70 items, ranging from 45.3 to 93.3% of patients, on same TKI in same dosage 2 weeks apart	Not evaluated	Not evaluated
Validity			
Content validity	Not assessed	Assessed as Inconsistent [20]	Assessed as Inconsistent [20]
Construct validity			
• *Structural validity*	• Not applicable to a formative model, in which the items together form the construct, but are not all related. For the 7 different symptom scales an exploratory factor analysis was performed	• Confirmatory factor analysis performed for the symptom burden scale	• Not applicable to a formative model, in which the items together form the construct but are not all related.

Table 1 Continued

Measurement property	Instrument		
	EORTC IL386 [1]	**EORTC QLQ-CML24 [2, 3] ***	**EORTC Symptom Set [4]**
• *Hypotheses testing*	• A third of symptoms/ scales differed statistically significantly, and up to 30% in severity, between dasatinib, imatinib and/or nilotinib, adjusted for covariates, in directions that can be explained through pathophysiology	• The symptom burden scale differed statistically significantly between patients with higher vs. lower Karnofsky performance status; comorbidities vs. no comorbidities; ECOG performance status high vs. low; and imatinib vs. 2nd generation TKI, all in the expected direction • The symptom burden scale (QLQ-CML24) was correlated (>0.50) with the functioning, fatigue, and pain scales of the QLQ-C30, in the expected directions	• Not evaluated
• *Cross-cultural validity*	• Not evaluated. The instrument was developed in Dutch patients, though each item was sourced from the EORTC Item Library which contains items developed in culturally differing populations	• Not evaluated. Participants from 10 countries participated in development	• Not evaluated. Participants from 16 countries participated in development
Criterion validity	Not evaluated	Not evaluated	Not evaluated
Responsiveness			
Responsiveness	Indication of responsiveness in half of 50 symptoms/ scales, when compared to anchor questions	No significant difference in symptom scale for patients who achieved complete cytogenetic response. Patients who reported a clinically meaningful deterioration or improvement in the EORTC QLQ-C30 global health status/QL scale, had statistically significant mean changes in the symptom burden scale in the expected directions	Not evaluated

[§] Content validity is the only measurement property which has been assessed according to the COSMIN methodology. Otherwise, descriptive only; * Including the accompanying EORTC QLQ-C30; [&] Fatigue, Mood problems, Muscle/joint pain/cramps, Skin-, Eye-, Cognitive- and Sexuality problems

Besides criterion validation, validation of individual items through hypothesis testing appears to be the most pressing need. In our pilot data we did find that the identified differences in symptom severity between various TKIs were all in directions that can be explained through pathophysiological mechanisms. However, we did not formulate a priori hypotheses which we subsequently tested. A formal evaluation of measurement properties using the COSMIN methodology, beyond content validity, could aid in instrument selection and help identify the most critical areas for further validation. By nature, instruments designed to monitor toxicity symptoms must remain flexible, allowing for the inclusion of new symptoms and the adaptation of personalised features, such as filter questions. This flexibility should not hinder their use; instead, further validation and development should progress alongside their clinical, and also research, application. Compared to more generic and static instruments, such as those measuring quality of life, symptom instruments will need to be both more adaptable and more specific.

Standardised comparisons across TKIs

The key question is which TKI causes less severe symptoms of a particular type in an individual patient, enabling patients to make informed decisions when faced with intolerability of that symptom. In our meta-analysis (**Chapter 2**) and pilot data (**Chapter 4**), approximately one-third of symptoms differed among the three most commonly used treatments. However, the specific symptoms that showed significant differences varied between these analyses. Furthermore, even when significant differences were found for similar symptoms, the results were often inconsistent. For instance, our meta-analysis indicated that moderate to severe abdominal pain was significantly more prevalent in patients on imatinib compared to nilotinib. In contrast, our pilot data suggested that abdominal pain was significantly more severe in patients on dasatinib, but not imatinib, compared to nilotinib.

The differences between the two studies are likely due to variations in outcome measures (prevalence vs. severity), the adjustment for covariates in our pilot data (which was not done in our meta-analysis), and potentially low patient numbers in the pilot study. Nevertheless, we believe that mean severity is more relevant than prevalence for individual patients and that adjusting for covariates in mean severity scores provides a more reliable approach to evaluating differences between treatments. Furthermore, we consider the findings of our pilot data to be consistent with the underlying pathophysiological mechanisms.

Comparisons with existing literature are limited, as only one study has adjusted for covariates, focusing solely on dasatinib versus imatinib across eight symptoms [21]. To provide context for our pilot data, we conducted a meta-analysis of mean unadjusted severity scores, combining our study with eight systematically sourced studies that utilised EORTC instruments [1]. However, this analysis was only feasible for the eight symptoms available in those studies. Currently, our pilot data offer the most extensive assessment of symptom severity across most symptoms and TKIs. Unfortunately, we were unable to provide reliable estimates for newer TKIs, as our sample sizes were relatively small.

To fully inform patients, prospectively collected data on how treatment and dose adjustments affect specific symptoms are essential. Such data are gradually becoming available–for example, the ASCEMBL trial compared switching to asciminib or bosutinib in resistant/intolerant CML patients previously treated with ≥2 TKIs [22]. Similarly, the RODEO trial will help determine whether dose reductions influence symptom severity [23]. Nonetheless, these isolated research efforts are unlikely to meet all the informational needs required for shared decision-making when choosing or switching TKIs. Only 6% of leukaemia trials conducted between 2004 and 2026 included some form of patient-reported outcome measures [24]. Even when such measures are incorporated, they are often reported separately, and with significant delays, as seen in the BFORE trial [25], or limited to sub-studies only, as in the ENEST-trial [26]. In other haematological areas this same problem has been identified. For example, a systematic review on lymphoma research found that patient-reported outcome data were under-collected, under-reported, and often positively framed despite a lack of improvement [27].

Therefore, trial data needs to be complemented by large scale and long-term prospective data collection in routine care. Hereby it is essential to use validated instruments, report on covariates in a transparent and standardised manner, and include patients on the newest TKIs as well as those in treatment-free remission. At a minimum, covariates should include age, sex, and dosage; and preferably, they should also encompass comorbidities, comedications, time of diagnosis, BCR::ABL1 levels; previous CML treatments, and time on the current CML treatment. The future use of the EORTC IL386 within the CMyLife digital care platform will generate such data over the coming years. At present, the lack of large-scale, adjusted data, using content-validated instruments,

with transparent reporting across TKIs, is hampering knowledge development. An individual patient data meta-analysis of between-TKI differences, adjusted for covariates, could provide valuable insights, though it would require significant effort and collaboration among multiple research groups.

Broadening the impact, and future directions

The Dutch CML evidence ecosystem

Our research demonstrated that CML guideline developers view patient-reported toxicity monitoring as beneficial, as it systematically incorporates patients' experiences into the guidelines, personalises the advice provided, and helps fill existing knowledge gaps. The identified and prioritised knowledge gaps offer concrete action points for future research using such data, ultimately enhancing the CML guidelines. Additionally, this data is essential for phase 4 pharmacovigilance studies involving longer follow-up periods and the detailed reporting of individual symptoms, which help to accurately assess the true tolerability of the treatment [28]. To translate the findings from patients' experiences to what to do in an individual context, we need causal outcome prediction modelling. Which patients are affected by which symptoms and adverse effects?

Ideally, real-world patient-reported toxicity symptoms will complement real-world CML outcome information, such as the National Cancer Registry data, the PHAROS-CML registry and Hemobase, which are used to monitor CML care quality indicators in the Netherlands [29-31]. At present, efforts are underway to link CML outcome information with pharmaceutical supply data. Comorbidities, for which medications are prescribed, can be derived from pharmaceutical supply data. Integrating information from these various sources could lead to the creation of a nationwide *trias* dataset which would include (1) CML treatment and outcome data, (2) comorbidity information, and (3) patient-reported symptoms. To the best of our knowledge, such a dataset would be unique, as nationwide research databases are typically based on national health insurance data and either do not contain or cannot be linked to patient-reported symptoms [32]. Moreover, only through the integration of context, exemplified by multiple data levels as shown in Textbox 1, can we address some of the complex challenges we face in modern medicine [33]. 'The least toxic, best tolerated CML treatment pathway for me/my patient' is one such complex problem.

To increase knowledge for complex problems we need a more diverse research approach that supplements the statistical thinking (deterministic causal effects) of clinical RCTs. We must incorporate complexity thinking into ecological/population-based studies or pragmatic trials, where algorithms can identify outcome patterns, and individuals can be clustered according to their characteristics, with these clusters correlating to specific outcome patterns [33]. We made a first attempt to implement this approach with our pilot data, but to fully unlock the potential of such methods, much larger, integrated data sets are needed. The envisioned *trias* dataset, described above, with CML treatment and outcome data, comorbidities, and patient-reported symptoms, will be a crucial next step. Of note, patient-reported outcomes can be viewed either as an outcome in themselves or in combination with other outcomes, or as a marker/predictor of other outcomes [34, 35], underpinning the interrelatedness of factors within a network of markers and outcomes.

Textbox 1 Context derived from multiple data levels is needed to solve complex medical problems such as 'What is the least toxic, best tolerated CML treatment pathway for me/my patient?'

Data levels relevant to CML:

- Demographics (age, gender)
- Socioeconomics (income, education,)
- Lifestyle markers (smoking, exercise)
- Psychological markers (personality traits, anxiety, depression, adherence)
- Care markers (organisation and professional characteristics)
- Comorbidities and comedication
- Biological markers (body mass index, pharmacokinetics)
- Proteomics (pathways involved in pharmacokinetics and resistance)
- Genetic markers (mutations, treatment resistance)

In relation to CML outcomes*:

- Survival
- Disease control (BCR::ABL1)
- Attainment of treatment-free remission
- Attainment of personalised treatment goals
- Quality of life
- Adverse events
- Patient-reported toxicities

* Depending on the research question, outcomes may be considered prognostic markers or vice versa

Information from patients and patient-professional interactions must be integrated into the CML evidence ecosystem. Such an information feedback loop, depicted in Figure 2, should ideally contain a CML dataset that can be used for both specific scientific research questions and broader real-world evidence evaluation. The use of this CML dataset will lead to scientific

publications, which will then be incorporated into the CML guideline. In turn, the CML guideline will support decision aids, from which treatment choices will be made. This rigorous process in also a slow process. It would therefore be worthwhile to explore how data iterations from each individual patient could inform practice. This would be worthwhile especially in emerging therapies where toxicities are underexplored.

Data on treatment choices, patient-reported toxicity symptoms, and consent for data sharing can be collected within the CMyLife digital care platform and flow to the personal health environment of patients. Care data, collected in electronic patient files of care organisations, will also, at least in the future, be integrated into patients' personal health environments. Data from these personal health environments can then be used to feed into the CML dataset. Possibly, CumuluZ could serve as a platform for data exchange across Dutch health care organisations. The European Health Data Space regulation, with secure and trustworthy reuse of health data for research as one of its cornerstones, will hopefully help realise and sustain these future developments. Alternatively, federated learning techniques, such as the personal health train, could question separate datasets and return meaningful answers.

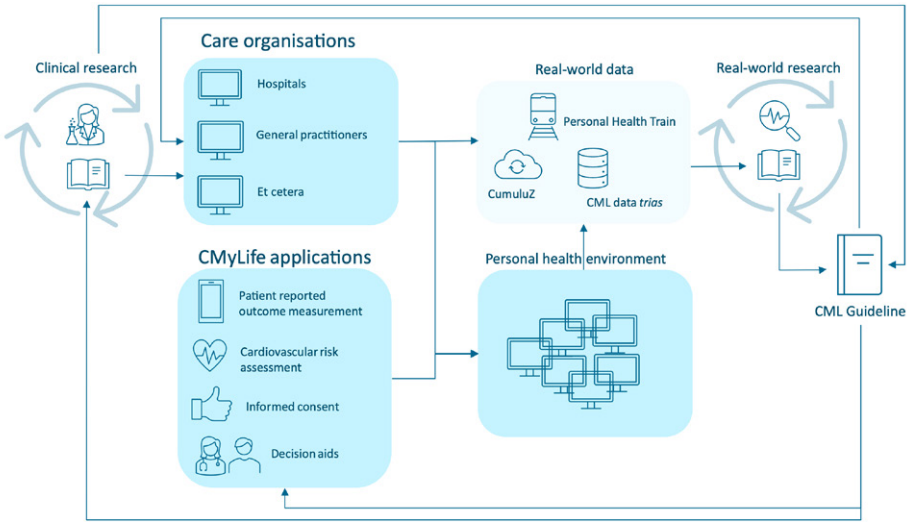


Figure 2 Information feedback loop within the CML evidence ecosystem

Beyond CML

The work presented in this thesis has served as a both a pilot project and a catalyst for the integration of real-world patient-reported symptoms into the Dutch haemato-oncology evidence ecosystem. The lessons learned are now being used as a framework for the development of real-world symptom monitoring in other blood cancers. Specifically, for chronic lymphoid leukaemia (CLL) and myeloproliferative neoplasms (MPN), CMyLife recently developed and implemented applications for the digital monitoring of patients' symptoms. The development of instruments (using the EORTC Item Library), design (including filter questions) and workflow (such as graphical feedback to patients) draws heavily on the lessons learned from this thesis. At the time of writing (June 12, 2025), these applications have been available for six weeks. In that time, 79 CLL and 99 MPN patients have entered their symptoms 295 times, with the number of participants growing. This nationwide effort aligns with the leading vision of comprehensive adverse event evaluation, that accurately characterises the effects of emerging cancer therapies, supporting patients and clinicians in shared decision-making, and supplying high-quality data to regulators responsible for approving safe and effective treatments [24, 28, 36-39].

Real-world evidence is more than just a buzz-word, though we concur that so-called real-world evidence often deviates from the overall patient population [40, 41]. The challenge will be to include all patients in real-world data sets. So, to what extent will the symptom monitoring data, when collected, represent the general haemato-oncologic population? The CMyLife platform has the advantage of directly involving patients and bypassing siloed and unstandardised provider data. However, its disadvantage lies in being a purely digital workflow, outside of the hospital's usual digital environment. We know that patients active on the CMyLife digital care platform tend to be younger and better educated [1, 42, 43], which may introduce bias. We have therefore directed recent efforts at low health/digital literacy patients, to make the platform more inclusive [44]. Enabling less digitally skilled patients to participate in the platform - possibly even through non-digital methods - is important to avoid inclusion bias.

Symptom monitoring may, by default, attract patients who experience more severe symptoms. Ideally, the instrument and its workflow will provide enough value to the clinic to encourage the participation of patients who experience fewer or no symptoms, though this needs to be evaluated. Especially in patients with few or no symptoms, consultations could become more efficient if

professionals can quickly see, in a single overview, that experienced symptoms need not be discussed as per patient preference. The ongoing involvement of health care professionals requires the integration of CMyLife services within the digital workflow of care organisations. This is especially important for smaller (non-university) centres, where separate workflows (in separate systems or for separate disease entities) may not be attractive, as only a small number of patients are served per disease entity.

The unstructured collection and quality of real-world data remains a challenge. Transparency may be key, as it enables end-users to evaluate the quality and applicability of evidence derived from real-world data [45, 46]. Unresolved issues include the frequency of data collection; integration with other databases; data audit; data ownership; electronic data capture; and the impact of the clinical pathway on data collection [47]. One approach to address these issues is a 'fit for use' assessment, which evaluates the relevance and reliability of real-world data [46, 48]. This approach can be applied to identified knowledge gaps, guiding the collection of relevant and reliable data. Ultimately, and most importantly, we believe that these challenges should not hinder the widespread use of real-world patient-reported toxicity symptoms, both in daily clinical care and in research.

Concluding remarks

In this thesis, the development and validation of an instrument is described, along with its personalised workflow, to monitor CML patient-reported TKI toxicity symptoms in the real world. Furthermore, the knowledge gaps in CML care that can be addressed by using such data are highlighted. These are the tools needed to improve current CML guidelines, shifting the focus beyond disease control and the attainment of treatment-free remission, towards the tolerability of treatment. The lessons learned are currently being applied in the development, validation, and implementation of real-world symptom monitoring in other haemato-oncology diseases.

A firm step-up in the collection and analysis of patient-reported toxicity symptoms in the real world is needed to realise the potential of incorporating these data into guidelines and shared decision-making. Key prerequisites include content-validated instruments, standardised reporting (such as EORTC-instruments, mean severity, and covariates), covering all available treatment regimens - especially newer TKIs and treatment-free remission - longitudinal data, and integrated analyses with other levels of data.

References

1. Smit, Y., et al., *Development, validation, and pilot testing of a newly developed patient-reported instrument for monitoring real-world TKI-related toxicity in CML care (submitted)*. 2025.
2. Efficace, F., et al., *International development of an EORTC questionnaire for assessing health-related quality of life in chronic myeloid leukemia patients: the EORTC QLQ-CML24.* Qual Life Res, 2014. **23**(3): p. 825-36.
3. Efficace, F., et al., *Validation and reference values of the EORTC QLQ-CML24 questionnaire to assess health-related quality of life in patients with chronic myeloid leukemia.* Leuk Lymphoma, 2021. **62**(3): p. 669-678.
4. Sodergren, S.C., et al., *Developing symptom lists for people with cancer treated with targeted therapies.* Targeted Oncology, 2021. **16**(1): p. 95-107.
5. Cella, D., et al., *Measuring health-related quality of life in leukemia: the Functional Assessment of Cancer Therapy--Leukemia (FACT-Leu) questionnaire.* Value Health, 2012. **15**(8): p. 1051-8.
6. Goswami, P., et al., *Development of a novel hematological malignancy specific patient-reported outcome measure (HM-PRO): content validity.* Frontiers in Pharmacology, 2020. **11 (article 209)**.
7. Goswami, P., et al., *Paper and electronic versions of HM-PRO, a novel patient-reported outcome measure for hematology: an equivalence study.* J Comp Eff Res, 2019. **8**(7): p. 523-533.
8. Goswami, P., et al., *Reliability of a novel hematological malignancy specific patient-reported outcome measure: HM-PRO.* Front Pharmacol, 2020. **11**: p. 571066.
9. Goswami, P., et al., *Hematological malignancy specific patient-reported outcome measure (HM-PRO): construct validity study.* Frontiers in Pharmacology, 2020. **11 (aricle 1308)**.
10. Williams, L.A., et al., *Measuring the symptom burden associated with the treatment of chronic myeloid leukemia.* Blood, 2013. **122**(5): p. 641-7.
11. Yu, L., et al., *Variables associated with patient-reported symptoms in persons with chronic phase chronic myeloid leukemia receiving tyrosine kinase inhibitor therapy.* Medicine (Baltimore), 2019. **98**(48): p. e18079.
12. Shacham Abulafia, A., et al., *Health-related quality of life in patients with chronic myeloid leukemia treated with first- versus second-generation tyrosine kinase inhibitors.* J Clin Med, 2020. **9**(11).
13. Schoenbeck, K.L. and K.E. Flynn, *Health-related quality of life of patients with chronic myeloid leukemia as measured by patient-reported outcomes: current state and future directions.* Curr Hematol Malig Rep, 2021. **16**(6): p. 491-499.
14. Niscola, P., F. Efficace, and E. Abruzzese, *Sexual health in patients with hematological malignancies: a neglected issue.* Support Care Cancer, 2018. **26**(6): p. 1699-1701.
15. Petersen, M.A., et al., *International validation of the EORTC CAT Core: a new adaptive instrument for measuring core quality of life domains in cancer.* Qual Life Res, 2020. **29**(5): p. 1405-1417.
16. Petersen, M.A., et al., *The EORTC CAT Core-The computer adaptive version of the EORTC QLQ-C30 questionnaire.* Eur J Cancer, 2018. **100**: p. 8-16.

17. Terwee, C.B., et al., *COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study.* Qual Life Res, 2018. **27**(5): p. 1159-1170.
18. Mokkink, L.B., et al., *The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes.* J Clin Epidemiol, 2010. **63**(7): p. 737-45.
19. Mokkink, L.B., E. Elsman, and C.B. Terwee, *Conducting a systematic review of patient-reported outcome measures. Manual version 2.0*. 2024, Amsterdam Public health research institute, Amsterdam UMC: Amsterdam.
20. Smit, Y., et al., *Measuring chronic myeloid leukaemia TKI-related toxic effects in the real world: a systematic review and critical assessment of content validity of patient-reported outcome measures.* Lancet Haematol, 2023.
21. Efficace, F., et al., *Health-related quality of life of newly diagnosed chronic myeloid leukemia patients treated with first-line dasatinib versus imatinib therapy.* Leukemia, 2020. **34**(2): p. 488-498.
22. Rea, D., et al., *Health-related quality of life of patients with resistant/intolerant chronic phase chronic myeloid leukemia treated with asciminib or bosutinib in the phase 3 ASCEMBL trial.* Leukemia, 2023. **37**(5): p. 1060-1067.
23. Djodikromo, M.F., et al., *Patient-guided dose reduction of tyrosine kinase inhibitors in chronic myeloid leukaemia (RODEO study): study protocol for a prospective, multicentre, single-arm trial.* BMC Cancer, 2023. **23**(1): p. 231.
24. Thanarajasingam, G., et al., *Beyond maximum grade: modernising the assessment and reporting of adverse events in haematological malignancies.* Lancet Haematol, 2018. **5**(11): p. e563-e598.
25. Cortes, J.E., et al., *Patient-reported outcomes in the phase 3 BFORE trial of bosutinib versus imatinib for newly diagnosed chronic phase chronic myeloid leukemia.* J Cancer Res Clin Oncol, 2019. **145**(6): p. 1589-1599.
26. Yu, L., et al., *Achieving optimal response at 12 months is associated with a better health-related quality of life in patients with chronic myeloid leukemia: a prospective, longitudinal, single center study.* BMC Cancer, 2018. **18**(1): p. 782.
27. Milrod, C.J., et al., *Quality-of-life endpoints collection, reporting, and framing in randomised trials of indolent lymphomas: a systematic review.* Lancet Haematol, 2025. **12**(4): p. e312-e317.
28. Brockelmann, P.J., et al., *Beyond maximum grade: tolerability of immunotherapies, cellular therapies, and targeted agents in haematological malignancies.* Lancet Haematol, 2025. **12**(6): p. e470-e481.
29. Ector, G., et al., *Adherence to quality indicators in chronic myeloid leukemia care: results from a population-based study in The Netherlands.* Leuk Lymphoma, 2023. **64**(2): p. 424-432.
30. Geelen, I.G.P., et al., *Impact of hospital experience on the quality of tyrosine kinase inhibitor response monitoring and consequence for chronic myeloid leukemia patient survival.* Haematologica, 2017. **102**(12): p. e486-e489.
31. Geelen, I.G.P., et al., *Treatment outcome in a population-based, 'real-world' cohort of patients with chronic myeloid leukemia.* Haematologica, 2017. **102**(11): p. 1842-1849.
32. Hsing, A.W. and J.P. Ioannidis, *Nationwide population science: lessons from the taiwan national health insurance research database.* JAMA Intern Med, 2015. **175**(9): p. 1527-9.

33. Sturmberg, J.P. and M. Mercuri, *Every problem is embedded in a greater whole.* J Eval Clin Pract, 2025. **31**(1): p. e14139.
34. Paudel, R., et al., *Use of Patient-Reported Outcomes in Risk Prediction Model Development to Support Cancer Care Delivery: A Scoping Review.* JCO Clin Cancer Inform, 2024. **8**: p. e2400145.
35. Efficace, F., et al., *Patient-reported outcomes as independent prognostic factors for survival in oncology: systematic review and meta-analysis.* Value Health, 2021. **24**(2): p. 250-267.
36. Thanarajasingam, G., et al., *Reaching beyond maximum grade: progress and future directions for modernising the assessment and reporting of adverse events in haematological malignancies.* Lancet Haematol, 2022. **9**(5): p. e374-e384.
37. Thanarajasingam, G., et al., *Beyond maximum grade: introduction to The Lancet Haematology Adverse Events Reporting Series.* Lancet Haematol, 2025. **12**(6): p. e403-e406.
38. Major, A., A.C. Dueck, and G. Thanarajasingam, *Beyond maximum grade: advancing the measurement and analysis of adverse events in malignant haematology trials in the modern era.* Lancet Haematol, 2025. **12**(6): p. e451-e462.
39. Bhatnagar, V., et al., *Beyond maximum grade: using patient-generated data to inform tolerability of treatments for haematological malignancies.* Lancet Haematol, 2025. **12**(6): p. e463-e469.
40. Ioannidis, J.P.A., *We certainly need real life research-but what is real life?* J Eval Clin Pract, 2025. **31**(1): p. e70027.
41. Berger, M.L., et al., *When Will Real-World Data Fulfill Its Promise to Provide Timely Insights in Oncology?* JCO Clin Cancer Inform, 2024. **8**: p. e2400039.
42. Verweij, L., et al., *Effectiveness of digital care platform CMyLife for patients with chronic myeloid leukemia: results of a patient-preference trial.* BMC Health Serv Res, 2023. **23**(1): p. 228.
43. Smit, Y., et al., *Benefits and limitations of real-world patient-reported toxicity-symptom monitoring for guidelines and care, as perceived by patients, clinicians and guideline developers [submitted].* 2025.
44. van Donkelaar, I., *What do chronic leukemia patients with low health literacy need to understand from disease-related information: a qualitative study*, in *IQHealthcare*. 2023, Radboud university medical center.
45. Orsini, L.S., et al., *Improving transparency to build trust in real-world secondary data studies for hypothesis testing-why, what, and how: recommendations and a road map from the real-world evidence transparency initiative.* Value Health, 2020. **23**(9): p. 1128-1136.
46. Lerro, C.C., et al., *The bar is high: evaluating fit-for-use oncology real-world data for regulatory decision making.* JCO Clin Cancer Inform, 2024. **8**: p. e2300261.
47. Maruszczyk, K., et al., *Systematic review of guidance for the collection and use of patient-reported outcomes in real-world evidence generation to support regulation, reimbursement and health policy.* J Patient Rep Outcomes, 2022. **6**(1): p. 57.
48. Castellanos, E.H., B.K. Wittmershaus, and S. Chandwani, *Raising the bar for real-world data in oncology: approaches to quality across multiple dimensions.* JCO Clin Cancer Inform, 2024. **8**: p. e2300046.

Appendices

Summary

At present, clinical practice guidelines in haemato-oncology primarily focus on efficacy as demonstrated by randomised controlled trials. However, daily practice is far less standardised and considerably more diverse. Both patients and health care professionals often do not adhere to guideline recommendations, as these frequently fail to address the real-world dilemmas they encounter.

In chronic myeloid leukaemia (CML), unwanted variations in care are common. The real-world setting is not as controlled as clinical trials and is inherently more complex. This is illustrated by the fact that in 60% to 80% of cases, timely monitoring of the BCR::ABL1 biomarker is not adhered to, and by the fact that one in four to one in five patients switch to a different type of tyrosine kinase inhibitor (TKI) treatment due to intolerability.

In this thesis, we explored how real-world patient-reported toxicity symptoms can be made accessible for use in both CML care and guideline development.

Overview of patient-reported toxicity symptoms

Given that a substantial number of CML patients switch treatments due to intolerability, it is essential to understand which types of TKIs are associated with which specific toxicity symptoms in an individual patient. In **Chapter 2,** a systematic search and selection was conducted to collect the best available evidence on patient-reported toxicity symptoms associated with each TKI.

PubMed and Embase databases were searched from inception up to February 2025 for studies reporting the prevalence of individual symptoms in CML patients, stratified by TKI type. We identified 11 studies, including 2,987 CML patients and reporting on the prevalence of 47 different symptoms of any severity. No data were available for asciminib and ponatinib, and only very limited data were found for bosutinib.

The prevalence of low-grade patient-reported symptoms was high among patients treated with dasatinib, imatinib, and nilotinib. Fatigue was the most frequently reported symptom across all three treatments, with 64 to 71% of patients experiencing fatigue of any severity.

In indirect, unadjusted comparisons, the prevalence of 13 out of 47 symptoms differed significantly between dasatinib, imatinib, and nilotinib. However, robust evidence on the prevalence and severity of patient-reported toxicity symptoms - based on direct or adjusted comparisons between TKI types - is still lacking.

The content validity of existing instruments

In **Chapter 3,** the content validity of existing instruments to measure patient-reported TKI-related toxicity in CML patients was critically evaluated. We focused on content validity exclusively, as it forms the foundation of any reliable and valid patient-reported outcome measure.

First systematic search of the PubMed and Embase databases up to October 2022 was performed. Eligible studies included those that utilised or developed instruments for CML patients and that assessed and reported individual patient-reported symptoms by TKI type. Content validity was assessed according to the Consensus-Based Standards for the Selection of Health Measurement Instruments (COSMIN).

Six instruments were identified, but none had sufficient content validity to reliably monitor patient-reported TKI-related toxicity, based on our assessment. Five instruments - the EORTC QLQ-CML24, the EORTC Symptom Set, the FACT-LEU, the HM-PRO and the MDASI-CML - were rated as having inconsistent content validity due to one or more limitations: lack of evaluation by professionals post-development, minimal involvement of chronic myeloid leukaemia patients, or omission of key symptoms. The remaining instrument, a generic Chinese questionnaire, was rated as having insufficient content validity. These judgements were supported by evidence of moderate to very low quality.

Of note, the two EORTC instruments were the only ones that did not omit key toxicity symptoms, such as muscle cramps. However, their relevance was rated as inconsistent: the EORTC QLQ-CML24 includes questions on health-related quality of life, while the EORTC Symptom Set contains items derived from solid cancer treatments.

A new instrument to capture TKI-related toxicity symptoms in CML patients

In **Chapter 4,** a new instrument to capture TKI-related toxicity symptoms in CML patients was developed and validated. Moreover, pilot-data were collected and explored. The development process began with all CML-specific items from the two previously mentioned EORTC instruments.

A total of at least 278 CML patients and 18 CML professionals (including haematology nurse specialists and haematologists) participated in development and validation, as well as in pilot data collection. This collaborative effort resulted in a comprehensive 77-item EORTC instrument. For 23 issues, no corresponding EORTC item was available; these will need to be added in future versions.

An alternative format of the instrument was also developed, comprising 24 EORTC questionnaires combined with 24 generic filter questions. This format was preferred by 87% of patients. The instrument demonstrated fair to excellent psychometric properties in terms of reliability. Robust scales could be constructed for seven item groups, yielding 50 individual scales or symptoms that are suitable for research purposes. Preliminary evidence also indicated responsiveness in half of these 50 symptoms/scales.

Pilot-data analyses

Sexuality, although highly prevalent and severely experienced in our pilot data, was not included in existing instruments—highlighting the importance of a comprehensive approach. In our analysis, one-third of the symptoms/scales showed up to a 30% difference in severity between dasatinib, imatinib, and/or nilotinib, after adjusting for covariates. These findings support more informed shared decision-making when switching TKIs due to intolerability.

Notably, for most individual symptoms, no other studies provided data suitable for comparing mean severity across different TKIs. Due to the limited number of participants receiving asciminib, bosutinib, and ponatinib, we were unable to compare symptom severity for these treatments. Additionally, a higher proportion of women were found in the higher severity symptom cluster compared to the lower severity cluster.

Perceived benefits, limitations and knowledge gaps

In **Chapter 5,** the perceptions of patients and health care professionals on the benefits and limitations of real-world patient-reported toxicity monitoring, for

both clinical guidelines and patient care, were explored. This was conducted through focus groups and in-depth interviews with 27 patients, seven nurse specialists and 11 haematologists and guideline developers. Thirty-one benefits and limitations were identified.

Compared to an existing framework, novel benefits centred on the use of aggregated data. Participants viewed real-world patient-reported toxicity monitoring valuable because it systematically incorporates patients' experiences into guidelines, personalises guideline recommendations, and addresses knowledge gaps.

Novel limitations related to the suitability, acceptance, and applicability of toxicity symptom monitoring in routine clinical practice. These benefits and limitations provide valuable direction for targeted implementation strategies to enhance the uptake and impact of toxicity symptom monitoring.

Through a Delphi procedure, the Dutch CML guideline development group identified fourteen key knowledge gaps in CML care that could be addressed through systematic patient-reported toxicity monitoring. The top three priority knowledge gaps were:

- What is the burden of individual toxicity symptoms?
- How many patients switch from which TKI due to intolerability, and for which symptoms?
- How do toxicity symptoms change in severity after TKI dosage, adjustment, and by how much?

These knowledge gaps translate directly into actionable points for future CML guideline development.

Discussion

In **Chapter 6,** a reflection on the main findings of our research is presented in the context of recent literature. In addition, their implications for daily clinical practice, as well as directions for future research, are discussed. The lessons learned through this thesis are already being applied to the development, validation, and implementation of real-world symptom monitoring in other haemato-oncology diseases.

Samenvatting

Op dit moment richten richtlijnen in de hemato-oncologie vooral over hoe goed behandelingen werken, zoals blijkt uit speciale onderzoeken, de gerandomiseerde gecontroleerde studies. Maar in de dagelijkse zorg gaat het vaak anders dan in deze speciale onderzoeken. Er is namelijk veel meer variatie in patiënten en in hun ziekten, dan dat onderzocht wordt. Patiënten en zorgverleners houden zich vaak niet precies aan de richtlijnen, omdat die niet altijd passen bij de situaties die ze in het echt tegenkomen.

Bij chronische myeloïde leukemie (CML) komen vaak verschillen in de zorg voor die niet gewenst zijn. De zorg in de praktijk is minder strak geregeld dan in speciale onderzoeken en is meestal ingewikkelder. Dit blijkt bijvoorbeeld uit het feit dat bij 60% tot 80% van de patiënten de BCR::ABL1-biomarker niet op tijd wordt gecontroleerd. Ook verandert één op de vier tot één op de vijf patiënten van type tyrosinekinaseremmer (TKI) omdat ze last hebben van symptomen van bijwerkingen.

In dit proefschrift is onderzocht hoe meldingen van symptomen van bijwerkingen door patiënten beter gebruikt kunnen worden in de zorg voor CML en bij het maken van richtlijnen.

Overzicht van symptomen van bijwerkingen die patiënten zelf melden

Omdat veel CML-patiënten van behandeling veranderen door symptomen van bijwerkingen, is het belangrijk om te weten welke TKI's welke symptomen van bijwerkingen geven. In **Hoofdstuk 2** is daarom een systematisch literatuuronderzoek gedaan naar de best beschikbare informatie over welke symptomen van bijwerkingen patiënten melden bij elk type TKI.

Er is gezocht in de databases PubMed en Embase, van het begin tot februari 2025, naar onderzoeken die beschrijven hoe vaak verschillende symptomen voorkomen bij CML-patiënten, per type TKI. Er zijn elf onderzoeken gevonden met samen 2.987 CML-patiënten, waarin 47 verschillende symptomen werden genoemd. Er waren geen gegevens over de TKI's asciminib en ponatinib, en maar weinig gegevens over bosutinib.

Milde bijwerkingen kwamen vaak voor bij patiënten die dasatinib, imatinib of nilotinib gebruikten. Vermoeidheid was het meest voorkomende symptoom

bij al deze drie behandelingen. 64 tot 71% van de patiënten zei hier last van te hebben.

In indirecte, niet-gecorrigeerde vergelijkingen verschilden 13 van de 47 symptomen significant in prevalentie tussen dasatinib, imatinib en nilotinib. Sterk bewijs over de prevalentie en ernst van symptomen van bijwerkingen – gebaseerd op directe of gecorrigeerde vergelijkingen tussen TKI's – ontbreekt echter nog steeds.

Toen dasatinib, imatinib en nilotinib werden vergeleken, bleken 13 van de 47 symptomen duidelijk anders te zijn. Er is echter nog geen heel sterk bewijs over hoe vaak en hoe erg deze bijwerkingen verschillen. Dat komt vooral omdat deze TKIs op een indirecte manier met elkaar zijn vergeleken. Ook kunnen de berekeningen nog niet met alle factoren rekening houden die deze verschillen kunnen veroorzaken.

De inhoud van bestaande meetinstrumenten

In **Hoofdstuk 3** is gekeken hoe goed bestaande meetinstrumenten zijn. Deze instrumenten meten symptomen van TKI-behandelingen die CML-patiënten zelf melden. We keken hier vooral naar de inhoud van de vragenlijsten: begrijpen patiënten de vragen? Is de vragenlijst volledig? En zijn de vragen belangrijk genoeg? Dit heet inhoudsvaliditeit. De inhoud van de vragenlijsten is namelijk het belangrijkste voor een betrouwbare en goede vragenlijst.

Als eerste werd een systematisch literatuuronderzoek gedaan in de databanken PubMed en Embase, tot oktober 2022. Onderzoeken die meetinstrumenten voor CML-patiënten gebruikten of maakten, en waarin de door patiënten gemelde symptomen per type TKI werden beschreven, kwamen in aanmerking. De inhoudsvaliditeit werd beoordeeld volgens de COSMIN-criteria (Consensus-Based Standards for the Selection of Health Measurement Instruments).

Van de zes gevonden instrumenten had geen enkele een goede inhoudsvaliditeit om symptomen van bijwerkingen door TKI's betrouwbaar te meten. Vijf instrumenten – de EORTC QLQ-CML24, de EORTC Symptom Set, de FACT-LEU, de HM-PRO en de MDASI-CML – kregen de beoordeling 'inconsistent' omdat ze één of meer problemen hadden: ze werden niet gecontroleerd door professionals na ontwikkeling, er waren weinig CML-patiënten betrokken, of er

misten belangrijke symptomen. Het zesde instrument, een Chinese vragenlijst, kreeg de beoordeling 'onvoldoende'. Deze oordelen werden ondersteund door bewijs van matige tot zeer lage kwaliteit.

Opvallend is dat de twee EORTC-instrumenten als enige geen belangrijke symptomen misten, zoals spierkrampen. Toch kregen ze de beoordeling 'inconsistent'. Dit komt omdat de EORTC QLQ-CML24 ook vragen over kwaliteit van leven bevat, en de EORTC Symptom Set ook is gemaakt voor klachten bij andere kankersoorten dan CML.

Een nieuw instrument voor symptomen van bijwerkingen van TKI's bij CML-patiënten

In **Hoofdstuk 4** is een nieuw meetinstrument gemaakt en getest om symptomen van TKI-behandelingen te meten, zoals CML-patiënten die zelf melden. De ontwikkeling begon met alle CML-specifieke vragen uit de twee eerder genoemde EORTC-instrumenten. Het nieuwe meetinstrument is ook getest op inhoudsvaliditeit en meeteigenschappen. Daarnaast zijn er proefgegevens verzameld en bekeken.

In totaal deden minstens 278 CML-patiënten en 18 CML-zorgverleners mee, zoals verpleegkundig specialisten hematologie en hematologen. Zij werkten mee aan de ontwikkeling, het testen en het verzamelen van proefgegevens. Dit resulteerde in een uitgebreid EORTC-instrument met 77 vragen. Voor 23 onderwerpen was nog geen goede vraag beschikbaar; deze moeten in toekomstige versies worden toegevoegd.

Er werd ook een andere versie van het meetinstrument gemaakt. Deze bestond uit 24 EORTC-vragenlijsten met 24 algemene filtervragen. 87% van de patiënten vond deze versie het beste. Het instrument bleek betrouwbaar te zijn. Voor zeven groepen van vragen konden goede meetschalen worden gemaakt. Zo ontstonden 50 schalen of losse symptomen die gebruikt kunnen worden voor onderzoek. Van deze 50 was er voor de helft al een eerste bewijs dat ze goed reageren op veranderingen (responsiviteit).

Analyse van de proefgegevens

Seksualiteit kwam vaak voor in de proefgegevens en werd als ernstig ervaren, maar staat niet in de bestaande meetinstrumenten (**Hoofstuk 4**). Dit laat zien hoe belangrijk het is om een compleet meetinstrument te hebben. In onze analyse bleek dat bij een derde van de symptomen of meetschalen het verschil

in ernst tussen dasatinib, imatinib of nilotinib wel 30% kon zijn. Ook nadat we rekening hielden met andere factoren. Deze resultaten kunnen helpen bij het samen kiezen van de beste TKI als patiënten van behandeling wisselen vanwege bijwerkingen.

Opvallend is dat voor de meeste losse symptomen geen andere studies geschikt waren om de gemiddelde ernst tussen verschillende TKI's te vergelijken. Omdat er maar weinig mensen waren die asciminib, bosutinib of ponatinib gebruikten, kon er geen vergelijking worden gemaakt voor de ernst van symptomen voor deze middelen. Ook bleek dat vrouwen vaker in de groep met ernstigere symptomen zaten dan in de groep met mildere symptomen.

Voordelen, beperkingen en ontbrekende kennis

In **Hoofdstuk 5** is onderzocht wat patiënten en zorgverleners vinden van het monitoren van door patiënten gemelde symptomen van bijwerkingen in de dagelijkse zorg. Dit gaat om het gebruik van deze informatie voor richtlijnen en voor de zorg zelf. Dit hebben we gedaan met gesprekken in groepen en met diepte-interviews met 27 patiënten, zeven verpleegkundig specialisten en 11 hematologen en mensen die richtlijnen maken. In totaal werden 31 voordelen en beperkingen genoemd.

In vergelijking met een eerdere overzichten van voordelen en beperkingen, kwamen er nieuwe voordelen naar voren over het gebruik van samengevoegde data. De deelnemers vonden het belangrijk dat deze manier van monitoring ervoor zorgt dat ervaringen van patiënten beter meegenomen worden in richtlijnen. Dit maakt de richtlijn persoonlijker en helpt om ontbrekende kennis – kennishiaten – op te vullen.

Nieuwe beperkingen gingen vooral over hoe geschikt, geaccepteerd en toepasbaar het is om bijwerkingen in de dagelijkse praktijk te volgen. Deze inzichten geven ideeën voor gerichte plannen om het gebruik en het effect van deze monitoring beter te maken.

Met een Delphi-procedure gaf de Nederlandse CML-richtlijnwerkgroep veertien belangrijke kennishiaten aan in de zorg voor CML. Deze kunnen opgelost worden door systematisch de symptomen van bijwerkingen te monitoren die patiënten zelf melden. De drie belangrijkste kennishiaten waren:

- Hoe ernstig zijn de verschillende symptomen van bijwerkingen en hoeveel last hebben patiënten ervan?
- Hoeveel patiënten wisselen van TKI door bijwerkingen, en welke symptomen spelen daarbij een rol?
- Hoe veranderen de symptomen van bijwerkingen in ernst als de TKI-dosering wordt aangepast, en hoeveel veranderen ze?

Deze kennishiaten zorgen direct voor duidelijke punten waar aan gewerkt kan worden bij het maken van nieuwe CML-richtlijnen.

Discussie

In Hoofdstuk 6 kijken we naar de belangrijkste uitkomsten van ons onderzoek en vergelijken die met recente literatuur. We bespreken wat deze resultaten betekenen voor de dagelijkse zorg en welke mogelijkheden er zijn voor toekomstig onderzoek. De lessen uit dit proefschrift worden nu al gebruikt bij het maken, testen en invoeren van symptoommonitoring voor andere ziekten binnen de hemato-oncologie.

Research data management

Ethics and privacy

This thesis is based on the results of existing data from published papers (Chapters 3 and 4) and research involving human participants (Chapters 5 and 6), conducted in accordance with relevant national and international legislation and regulations, guidelines, codes of conduct and Radboud university medical centre policies.

A statement that the study was not subject to the Dutch Medical Research Involving Human Subjects Act (WMO), was obtained from the recognised Medical Ethics Review Committee 'METC Oost-Nederland' (file numbers 2021-13053, 2022-16118, and 2023-16311).

Participation in these studies was voluntary. All participating patients and professionals gave oral or written informed consent to collect and process their data for this research project. All data were anonymised, no identifying information is present, and can be shared for reuse.

Data collection and storage

Publicly available data were used for the papers as presented in Chapters 3 and 4. Original data was collected during focus group discussions and individual interviews (Chapters 5 and 6). The transcripts of the focus groups and interviews were transcribed verbatim, fully anonymised and analysed using Atlas.ti software on the department server and are only accessible by project members working at the Radboudumc. Original audio tapes were deleted. Also, notes taken during these sessions and anonymised results from the online surveys are stored digitally.

Pilot data for chapter 5 was obtained by using Castor EDC for secured online questionnaires, and was thereafter temporarily stored and analysed in workspaces in the Azure DRE (ID-dws-1968-CMG) (DRE Portal (mydre.org)). This data is now stored on the department server. These secure storage options safeguard the availability, integrity and confidentiality of the data.

Availability of data

All studies are or will be published open access, except for Chapter 4 for which this was not feasible according to the journal's policies for reviews. The datasets from Chapters 5 and 6 are published in Data Sharing Collections

(DSC's) in the Radboud Data Repository (DOI 10.34973/9288-ds63), with restricted access. Requests for access will be checked by the Committee of health Care Renewal and Innovation of the Dutch Society for Haematology. Data were made reusable by adding sufficient documentation (research protocol, codebook and a readme file), and by using preferred and sustainable data formats.

All collected data will be stored for 15 years after termination of the study.

PhD portfolio of Yolba Smit

Department: **Haematology**
PhD period: **01/04/2020 – 16/06/2025**
PhD Supervisors: **Prof. Dr. N.M.A. Blijlevens, Prof. Dr. R.P.M.G. Hermens, Prof. Dr. Ir. A.L.A.J. Dekker**

Training activities	Hours
Courses	
• Qualitative research methods in healthcare (2021)	28.00
• RIHS - Introduction course for PhD candidates (2021)	15.00
• RU - Qualitative Research Methods and Analysis (2022)	84.00
• Radboudumc - Scientific integrity (2023)	20.00
• Radboudumc - eBROK course (2023)	42.00
• Implementation research (2023)	12.00
• Analysis longitudinal and multilevel data using R (2023)	96.00
Seminars	
• Workshop and pitch Brilliant business models (2021)	4.00
• Meet the expert: communicating your science: vlogging & blogging for beginners (2021)	1.50
• Meet the expert: Data visualisation; including reading: Storytelling with data (2021)	11.50
• Meet the Expert; Managing Your Interns; Blessing or curse (2022)	2.50
• Boost your writing skills (2022)	1.00
• Industry involvement in science (2022)	2.00
Conferences	
• AQUA Guideline working group conference (2021)	4.00
• Dutch Hematology Congress 2022 (2022)	6.00
• Modernising haematological malignancy adverse event assessment and reporting (2022)	1.00
• Projectleidersbijeenkomst Complexe interventies (2023)	4.00
Other	
• Reviewer ZonMw Implementation Science Practitioner Fellowship applications (2022)	2.00
• Proposal Junior Researcher 2022 (2022)	80.00
• Input and feedback on various grant proposals for CMyLife (2024, 2025)	40.00
• Writing data management plan overarching CMyGuideline project (2024)	20.00
• Daily supervision 2 student assistants WP4 CMyGuideline project (2024)	20.00
• Supervision non-academic personnel and BMS consultancy student WP1 CMyGuideline project (2024)	64.00

Teaching activities	
Lecturing	
• Oral presentation at AQUA workgroup (2020)	5.50
• Presentation skills MIN18 Hemato-Oncology: From concept to cure (2021)	4.00
• Writing skills MIN18 Hemato-Oncology: From concept to cure (2021)	12.00
• Guest Lecture "Communicatie in professionele contexten" (2022)	6.00
• Masterclass How to connect with your audience (2022)	16.00
• Presentation skills MIN18 Hemato-Oncology: From concept to cure (2022)	4.00
• Writing skills MIN18 Hemato-Oncology: From concept to cure (2022)	12.00
• Presentation skills MIN18 Hemato-oncology: from concept to cure (2023)	4.00
• Writing skills MIN18 Hemato-oncology: from concept to cure (2023)	12.00
• Presentation skills MIN18 Hemato-oncology: from concept to cure (2024)	4.00
• Writing skills MIN18 Hemato-oncology: from concept to cure (2024)	12.00
Supervision of internships / other	
• Daily supervision BMS literature thesis (2022)	6.00
• Daily supervision Medical scientific internship (2022)	20.00
• Daily supervision Medical scientific internship (2022)	20.00
• Daily supervision Master BMS internship (2022)	34.00
• Daily supervision Master Internship Management Policy Analysis (2022)	34.00
• Daily supervision Master Internship Communication, health and life sciences (Wageningen) (2022)	24.00
• Internship supervision Communication students (2023)	20.00
• Daily supervision 1st year Master BMS (2023)	34.00
• Daily supervision 2nd year Master BMS (2023)	34.00
• Daily supervision 2nd year Master BMS (2023)	34.00
• Daily supervision BMS literature thesis (2023)	6.00
• Supervision review CLL-treatment-related patient reported toxicity instruments (2024)	15.00
• Daily supervision Medical Scientific internship (2024)	20.00
Total	**953.00**

Dankwoord

Mijn dank is groot. Mijn promotietraject is in stoffelijke vorm gegoten, de afronding nadert. Dit is het resultaat van veel hard werk door velen: mensen die mij zijn voor gegaan en de weg hebben gewezen, en mensen die naast mij hebben gestaan. Patiënten, collega's, vrienden, familie. Dank aan jullie allen voor de hulp, steun en afleiding.

Allereerst wil ik mijn promotieteam bedanken. Nicole, je bent een groot inspirator voor me. Je liet me van de zee dromen zodat ik een boot ging bouwen. Flexibel als je bent gaf je me de kans om te promoveren. Je denken is scherp als een mes en een paar stappen vooruit, je begeleiding is betrouwbaar en gedegen. Dat komt samen in scherpe en consistente feedback, soms een paar regels of woorden, maar altijd geleverd indien nodig of gevraagd. Als ik wat te ver weg droomde haalde je me weer terug, en toch heb ik altijd heel veel ruimte ervaren om mijn nieuwsgierigheid te mogen volgen. Rosella, je bent de partner in crime van Nicole in veel projecten en ondertussen haar evenknie geworden. Je oratie en het feest daarna lieten zien hoe mensen jou zien en dat is ook zoals ik je heb mogen ervaren: je hebt een groot hart en je houdt van sociale contacten. Je was net zo consistent, betrouwbaar en flexibel als Nicole tijdens mijn promotietraject. Je hield de lijntjes kort en was een oneindige vraagbaak voor alle grote en kleine vragen. Andre, je was nieuw de nieuwe noot in de samenwerking. Je hebt me geïnspireerd met je visie om een wereldwijde infrastructuur voor toegankelijke gezondheid databases te willen bouwen. Ik hoop in de toekomst nog een keer samen te werken en meer van je te leren.

Graag wil ik de manuscriptcommissie, prof. Aura Timen, prof. Thijs Merkx en dr. Saskia Klein bedanken voor het kritisch lezen en beoordelen van mijn proefschrift, en voor hun flexibiliteit in het kiezen van een datum voor mijn verdediging. Dank ook aan alle leden van de promotiecommissie om zitting te nemen in de corona.

Dan wil ik graag alle patiënten bedanken en dan denk ik in de eerste plaats aan Jenny Borst. Jenny, je werkt als vrijwilliger voor Hematon en dat doe je met hart en ziel. Ik kon je alles vragen, niets is je te veel, en je draagt onze gezamenlijk projecten uit als een waar boegbeeld. En natuurlijk wil ik ook alle individuele patiënten van harte bedanken. Jullie hebben je tijd gegeven om met mij in gesprek te gaan, me dingen te leren, en mijn enquêtes in te vullen.

Het duurt altijd te lang voordat de resultaten er zijn maar op deze resultaten kunnen jullie heel erg trots zijn. Jullie ervaringen en verhalen draag ik met me mee en hebben dit werk vorm gegeven.

Naast alle patiënten en mijn promotieteam was dit proefschrift niet mogelijk geweest zonder de vele (aanstormende) wetenschappers die hebben geholpen financiering te verkrijgen, medische, statistische of farmacologische expertise hebben gedeeld, of hebben samengewerkt op het terrein van patient reported outcomes, kwalitatief onderzoek, of data science: Jeroen Janssen, Ward Posthuma, Charlotte Bekker, Pien Scheuter, Myrthe Lange, Iris Walraven, Sanne Metsemakers, Lynn Verweij, Jacolien Timmerman, Aleander Currie, Hatem Soltan, Inger Abma, Rianne Fijten, Chang Sun, en Ewald Bronkhorst.

Ook alle mensen die hebben meegewerkt aan CMyLife: heel hartelijk dank! Nicole Blijlevens stond aan de wieg van CMyLife, en ergens vlak daarna kwam Ankie Sneek: Ankie, we hebben ontzettend gelachen en – vooral jij – heel hard gewerkt. Je doorzettingsvermogen en relativeringsvermogen zijn aanstekelijk. Uiteraard wil ik ook alle collega's van 'het CMyLife-team' bedanken. In al die jaren zijn velen betrokken geweest, maar in het bijzonder bedank ik: Anton Tenzer, Corien Kromkamp, Pien Schaeffer, Esri Wener, Sabine ter Wolbeek en Lynn Verweij. Anton, je verdient sowieso mijn eeuwige dank omdat je me van een incapabele ICT-partner hebt afgeholpen. Daarnaast heb ik veel van je geleerd op het gebied van projectmatig werken en communiceren. Je communicatie in dia's en praatplaatjes kwam me eerst wat zonderling over, maar blijken in de praktijk erg handig om ideeën over te brengen en uit te wisselen. Corien, Pien, Esri, Sabine en Lynn: voor dit onderzoek hebben heel jullie veel mede georganiseerd en uitgevoerd, meegedacht over de opzet en de benadering van de patiënten, en geholpen met communicatie van de resultaten. Jullie leidden alles in goede banen met een oneindig geduld en onbetaalbaar goed humeur.

Dan kom ik bij mijn studenten: Myrthe, Pien, Nina, Eveline, Hanna, Alexander, Hatem, Jacolien, Ilse, Daphne, Djuna, Sophie, Christian, en Daan. Jullie zijn de belangrijkste reden dat ik wilde promoveren, omdat ik zonder dr. voor mijn naam niet zelfstandig studenten mag begeleiden. Jullie hebben stuk voor stuk hard meegewerkt aan de verschillende onderdelen van mijn onderzoek of het bredere project. Soms met lichte tegenzin – *'helaas is een wetenschappelijke stage verplicht'* – al viel het achteraf wel mee kreeg ik dan te horen; maar meestal met heel veel enthousiasme en elan. Jullie waren onmisbaar als

onafhankelijk tweede onderzoeker bij literatuur reviews en kwalitatief onderzoek, en sommigen hebben delen van het onderzoek nagenoeg zelfstandig verricht en daarmee een auteurschap verdiend. Jullie hebben me veel vreugde en een weinig kopzorgen gegeven, en jullie spontane vragen lieten me altijd dieper nadenken. Ik ben trots op jullie allemaal en wens jullie alle goeds in jullie verdere carrière.

Mijn lieve en geestige kamergenoten en collegae promovendi zijn een stimulans geweest om überhaupt aan promoveren te gaan denken: Charlotte de Mooij, Joline Saes, Geneviève Ector en Peter Karel. Na al die gezellige jaren tussen jullie in kon ik toch niet helemaal inschatten hoe intensief het promoveren daadwerkelijk zou zijn en heb ik alleen maar meer respect voor jullie gekregen. Lynn, Sanne, Melissa, Eva M, Eva H, Ayla, Nienke, Björn, Arnoud, Bauke, Asna, Daan, Roxanne, en CoPilot: dankzij jullie ben ik helemaal op de hoogte van – in willekeurige volgorde – de sneltoets voor de Euro, waar Ospel ligt, de sokken-controverse tussen Gen Z en Gen Alpha, hoe je vrienden maakt in de trein, wat echt goede vlaai is, hoe je het beste kan bonden met collega's, en dat alle kleine en grote problemen in het leven uiteindelijk weer goed komen. Heel veel dank mijn geklaag aan te horen, samen te lunchen, mij nieuwe dingen te leren (command L!) en me uit de brand te helpen. Tel daar de (kerst)borrels, etentjes, pub quizzen (2e plaats!) bij op en dan zijn jullie toch echt wel de leuke kant van onderzoek doen.

Naast mijn kamergenoten zijn er ook veel andere collega's die me mentaal en concreet gesteund hebben, zoals het management, het secretariaat, de stafleden, en de verdere collega's. Ik noem hier enkelen die er altijd met hun goede humeur (Jacqueline, Lindsay) en warme belangstelling (Jackie, Corine, Olga, Bas, Jesse, Emile, Ruben, Carin en Stijn) tijdens mijn PhD traject voor mij waren.

Dan rest mij nog mijn twee paranimfen te bedanken voor hun steun en vriendschap naar aanloop van, en op mijn verdediging op 13 maart: Liesbeth (jarig op 12 maart) en Linfee (jarig op 14 maart). Liesbeth, ik ken je al meer dan 40 jaar als een intelligente, grappige en sociale vrouw. Ik waardeer ook je rust, nuchterheid en boekenkast ten zeerste. Hoe zit het nou met die *dad* slippers? Linfee, onze vriendschap en jouw liefde voor ceremonies en het rituele maken je tot mijn paranimf. Ik hou van je gevoelige, recht door zee, en ondernemende karakter.

En dan kom ik bij meest nabije familieleden: pappa en mamma, Robbert en Floris. Jullie zijn er altijd voor mij en hebben heel erg veel geduld met mij. Ik kan met jullie lachen, wandelen, vakantie vieren (kamperen!), uitpuffen, lekker eten (hotpot?), bergen beklimmen, goede gesprekken voeren (de comfort temperatuur van slaapzakken), thuis komen (wat eten we?). Jullie zijn mijn oppepper en klaagmuur, steun en toeverlaat in barre tijden. We lopen samen een groot deel van onze levensweg en een heel klein deel daarvan komt nu ten einde met dit boekje.

About the author

Yolba Smit graduated as a medical doctor from Vrije Universiteit Amsterdam in 1999. She gained practical experience as a general practitioner in both Chile and the Netherlands while simultaneously pursuing a master's degree in epidemiology at the London School of Hygiene and Tropical Medicine. During her studies, she completed research internships in Eritrea and Bhutan. After earning her degree in epidemiology, Yolba balanced her career between working as a researcher and practicing as a general medical doctor.

In 2008, Yolba Smit transitioned to working as a freelance independent researcher, focusing exclusively on evidence-based medicine and guideline development. Over the years, she contributed to more than 10 guideline development projects for the Netherlands Comprehensive Cancer Organisation (IKNL) and the Belgian Healthcare Knowledge Centre (KCE). Additionally, she worked on over 10 research projects for the National Health Care Institute (ZiNL), supporting its role in advising the Minister on the health insurance package. She also developed and taught courses in evidence-based medicine at the European Centre for Disease Prevention and Control (ECDC) and the European Society for Paediatric Gastroenterology, Hepatology, and Nutrition (ESPGHAN). Furthermore, she co-initiated and chaired the Guideline Assessment Committee for the Professional Nursing Association (V&VN).

In 2016, seeking a deeper intellectual challenge and the collaborative environment of team science, Yolba Smit joined the Department of Haematology at Radboud university medical center as a medical writer. Her primary focus was on securing grants for research and innovation. To date, she has helped secure over 3.3 million euros in funding, both as a co-applicant, main applicant, and project lead. One of the major projects she contributed to was *CMyGuideline*, which also provided the opportunity for her to pursue PhD research. Yolba continues to work wholeheartedly at the Department of Haematology, where she combines her roles as a researcher and proposal writer.

9 789465 151106

Radboudumc university medical center

IN DEI NOMINE FELICITER

Radboud University