Evidence Generation in the Clinical Development of Medicines for Leukaemia

Submitted in partial fulfilment

of the requirements of the

Degree of Doctorate in Pharmacy

DYLAN SAID

Department of Pharmacy

University of Malta

2019

UNIVERSITY OF MALTA
FACULTY OF MEDICINE AND SURGERY

DECLARATION OF AUTHENTICITY FOR DOCTORAL STUDENTS

Student's I.D. /Code: 338692(M)

Student's Name & Surname: Dylan Said

Course: Doctorate in Pharmacy

Title of Dissertation/Thesis: Evidence Generation in the Clinical Development of Medicines for Leukaemia

I hereby declare that I am the legitimate author of this Dissertation/Thesis and that it is my original work.

No portion of this work has been submitted in support of an application for another degree or qualification of this or any other university or institution of higher education.

I hold the University of Malta harmless against any third party claims with regard to copyright violation, breach of confidentiality, defamation and any other third party right infringement.

□ As a Ph.D. student, as per Regulation 49 of the Doctor of Philosophy Regulations, I accept that my thesis be made publicly available on the University of Malta Institutional Repository.

□ As a Professional Doctoral student, as per Regulation 54 of the Professional Doctorate Regulations, I accept that my dissertation be made publicly available on the University of Malta Institutional Repository.

□ As a Doctor of Sacred Theology student, as per Regulation 17 of the Doctor of Sacred Theology Regulations, I accept that my dissertation be made publicly available on the University of Malta Institutional Repository.

□ As a Doctor of Music student, as per Regulation 24 of the Doctor of Music Regulations, I accept that my dissertation be made publicly available on the University of Malta Institutional Repository.

_____ _____

Signature of Student Date

Dedicated to my dear mother, who unconditionally stood

by my side every step of the way along my academic journey.

Acknowledgements

Sincere gratitude goes to my thesis supervisor Professor Anthony Serracino-Inglott for his encouragement to further my studies and for wholly believing in this research project. I would like to extend my special thanks to the rest of the research team, namely co-supervisor Dr Maresca Attard Pizzuto and advisor Professor John Joseph Borg for their invaluable insights throughout the course of the project. A note of appreciation goes to the members of the Department of Pharmacy at the University of Malta for their sturdy support, particularly Head of Department Professor Lilian M. Azzopardi for her professional guidance during my years as an undergraduate and postgraduate student. I thank you all for always being there to lend me a helping hand, your expertise has been instrumental in reaching this milestone.

A special mention goes to my most important pillar, my family. I am beyond grateful for your words of wisdom and for constantly pushing me to excel and become the best possible version of me.

Abstract

Leukaemia accounts for the highest age-standardised mortality rate among haematological malignancies in Europe. Evidence of efficacy for antineoplastic agents may be valued differently by regulatory and health technology assessment (HTA) bodies in the European Union (EU), impacting decision-making and access to novel medicines.

The aims of the study were to analyse the evolvement of efficacy parameters studied in leukaemia clinical trials (CTs), to explore scientific expert opinions on evidence generated and clinical assessments for antineoplastic therapies and to determine core efficacy outcomes prioritised by EU decision-makers for leukaemia CTs.

Part I data collection on trends in efficacy parameters involved the following: (1) Phase II to Phase IV leukaemia CTs were identified from the EU Clinical Trials Register database throughout an 11-year period (2007-2017), (2) CTs were screened against inclusion criteria, (3) efficacy endpoints were extracted and grouped according to type of measurement, (4) data mining of trends was performed using descriptive and inferential statistics. Part II data collection exploring scientific expert opinions consisted of these steps: (1) the Response Evaluation in Leukaemia (ReVALeu) surveying tool was developed and tested for validity and reliability, (2) the tool was disseminated in an e-Delphi process with two independent panels composed of regulatory and HTA oncology experts, (3) core efficacy outcomes reaching consensus were determined.

Thirty-six unique efficacy endpoints were identified from the final dataset of CTs (N=431) and grouped into clusters of survival (n=5), time-to-event (n=6), response rates and biomarkers (n=16) and other (n=9). Complete response rate was the most studied

primary endpoint (CTs: 19%, n=81), with progression-free survival (PFS) and minimal residual disease (MRD) registering the highest frequency change pre- and post-2012 (PFS: 8%, p=0.01; MRD: 8%, p=0.003). Thirty-six panellists were recruited in the e-Delphi; 24 regulatory representatives from the European Medicines Agency (EMA) and 12 experts from HTA bodies in 9 EU countries. Opinions on the quality of pre- and post-authorisation evidence generated for antineoplastic agents were statistically different (pre-authorisation: p=0.01, post-authorisation: p=0.04). Six efficacy endpoints achieving consensus were common to both regulatory and HTA groups of decision-makers and identified as the core outcomes.

Biomarker-based endpoints are emerging as primary efficacy measures in leukaemia CTs. Decision-makers perceive the quality of evidence generated for antineoplastic therapy differently. The identification of core efficacy outcomes may potentially optimise CT data packages for regulatory and reimbursement approvals.

Keywords

Clinical trials – Core outcomes – Evidence – Health technology assessment – Leukaemia – Regulatory

Table of Contents

List of Figures

List of Tables

List of Appendices

Glossary

Term	Definition
Accelerated assessment	The reduction of assessment timeframe by the European Medicines Agency (EMA) Committee for Medicinal Products for Human Use (CHMP) from 210 days to 150 days for medicinal products which are of major interest from the point of view of public health and in particular from the viewpoint of therapeutic innovation. **Citation:** European Commission. Regulation No. 726/2004 of the European Parliament and of the Council of 31 March 2004 laying down community procedures for the authorisation and supervision of medicinal products for human and veterinary use and establishing a European Medicines Agency [Internet]. Official Journal of the European Union 2004; L136:1-33 [cited 2019 May 08]. Available from: https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-1/reg_2004_726/reg_2004_726_en.pdf
Adaptive pathways	The adaptive pathways concept is an approach to medicines approval that aims to improve patients' access to medicines in cases of high unmet medical need. To achieve this goal, several approaches are envisaged: identifying small populations with severe disease where a medicine's benefit-risk balance could be favourable; making more use of real world data where appropriate to support clinical trial data; and involving health technology assessment (HTA) bodies early in development to increase the chance that medicines will be recommended for payment and ultimately covered by national healthcare systems. **Citation**: European Medicines Agency. Adaptive Pathways Workshop [Internet]. December 2016 [cited 2019 May 08]. Available from: https://www.ema.europa.eu/en/documents/report/adaptive-pathways-workshop-report-meeting-stakeholders-8-december-2016_en.pdf
Added therapeutic value	The incremental "therapeutic value" brought by a new drug or intervention compared with the best available treatment options already on the market. The therapeutic value can be defined in terms of patient-relevant endpoints and relevant levels of effectiveness, efficacy and safety. **Citation**: European Parliament: Environment, Public Health and Food Safety Committee (ENVI). Towards a harmonised EU assessment of added therapeutic value of medicines [Internet]. June 2015 [cited 2018 Dec 10]. Available from: http://www.europarl.europa.eu/RegData/etudes/STUD/2015/542219/IPOL_STU(2015)542219_EN.pdf
Centrally authorised medicinal product	A medicine with a single marketing authorisation issued by the European Commission and valid across the European Union. **Citation**: European Medicines Agency. Authorisation of Medicines [Internet]. 2019 [cited 2019 May 08]. Available from: https://www.ema.europa.eu/en/about-us/what-we-do/authorisation-medicines

Clinical endpoint	An aspect of a patient's clinical or health status that is measured to assess the benefit or harm of a treatment. A clinical endpoint describes a valid measure of clinical benefit due to treatment: the impact of treatment on how a patient feels, functions and survives. **Citation:** European Network for Health Technology Assessment. Endpoints used for relative effectiveness assessment of pharmaceuticals: Clinical endpoints [Internet]. February 2013 [cited 2019 May 08]. Available from: https://www.eunethta.eu/wp-content/uploads/2018/01/Clinical-endpoints.pdf
Clinical study	Any investigation in relation to humans intended: (a) to discover or verify the clinical, pharmacological or other pharmacodynamic effects of one or more medicinal products; (b) to identify any adverse reactions to one or more medicinal products; or (c) to study the absorption, distribution, metabolism and excretion of one or more medicinal products; with the objective of ascertaining the safety and/or efficacy of those medicinal products. **Citation:** European Commission. Regulation No. 726/2004 of the European Parliament and of the Council of 31 March 2004 laying down community procedures for the authorisation and supervision of medicinal products for human and veterinary use and establishing a European Medicines Agency [Internet]. Official Journal of the European Union 2004; L136:1-33 [cited 2019 May 08]. Available from: https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-1/reg_2004_726/reg_2004_726_en.pdf
Clinical trial	A clinical study which fulfils any of the following conditions: (a) the assignment of the subject to a particular therapeutic strategy is decided in advance and does not fall within normal clinical practice of the Member State concerned; (b) the decision to prescribe the investigational medicinal products is taken together with the decision to include the subject in the clinical study; or (c) diagnostic or monitoring procedures in addition to normal clinical practice are applied to the subjects. **Citation:** European Commission. Regulation No. 726/2004 of the European Parliament and of the Council of 31 March 2004 laying down community procedures for the authorisation and supervision of medicinal products for human and veterinary use and establishing a European Medicines Agency [Internet]. Official Journal of the European Union 2004; L136:1-33 [cited 2019 May 08]. Available from: https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-1/reg_2004_726/reg_2004_726_en.pdf
Conditional marketing authorisation	A marketing authorisation granted in case of certain categories of medicinal products in order to meet unmet medical needs of patients and in the interests of public health, on the basis of less complete data than is normally the case and subject to specific obligations. **Citation**: European Commission. Regulation No. 507/2006 of 29 March 2006 on the conditional marketing authorisation for medicinal products for human use falling within the scope of Regulation (EC) No. 726/2004 of the European Parliament and of the Council [Internet]. Official Journal of the European Union 2006; L92:1-4 [cited 2019 May 08]. Available from: https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-1/reg_2006_507/reg_2006_507_en.pdf

Content validity	The degree to which an instrument has an appropriate sample of items for the construct being measured. **Citation**: Polit DF, Beck CT. Nursing research: Principles and methods. 7th edition. Philadelphia: Lippincott, Williams, & Wilkins; 2004. p. 423.
Core outcome set	A list of outcomes that should be measured and reported, as a minimum, in all clinical trials in specific areas of health or healthcare. **Citation**: Gargon E, Williamson PR, Young B. Improving core outcome set development: qualitative interviews with developers provided pointers to inform guidance. J Clin Epidemiol. 2017;86:140-152.
Effectiveness	The extent to which an intervention does more good than harm when provided under the usual circumstances of health care practice. **Citation**: High Level Pharmaceutical Forum. High Level Pharmaceutical Forum 2005-2008: Final Report [Internet]. October 2008 [cited 2019 May 08]. Available from: http://www.rees-france.com/en/IMG/pdf/2008_High_level_Pharma_forum__en_final_report.pdf
Efficacy	The extent to which an intervention does more good than harm under ideal circumstances. **Citation**: High Level Pharmaceutical Forum. High Level Pharmaceutical Forum 2005-2008: Final Report [Internet]. October 2008 [cited 2019 May 08]. Available from: http://www.rees-france.com/en/IMG/pdf/2008_High_level_Pharma_forum__en_final_report.pdf
European Public Assessment Report	A public assessment report prepared at the end of every centralised evaluation process to provide a summary of the grounds for the opinion in favour of a marketing authorisation as taken by the Committee for Human Medicinal Products. **Citation**: European Medicines Agency. Reflection paper on European public assessment report summary for the public [Internet]. January 2006 [cited 2019 May 08]. Available from: https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/reflection-paper-european-public-assessment-report-summary-public_en.pdf
Exceptional marketing authorisation	An authorisation granted in exceptional circumstances and following consultation with the applicant which is subject to a requirement for the applicant to introduce specific procedures, in particular concerning the safety of the medicinal product, notification to the to the competent authorities of any incident relating to its use, and action to be taken. **Citation**: European Commission. Regulation No. 726/2004 of the European Parliament and of the Council of 31 March 2004 laying down community procedures for the authorisation and supervision of medicinal products for human and veterinary use and establishing a European Medicines Agency [Internet]. Official Journal of the European Union 2004; L136:1-33 [cited 2019 May 08]. Available from: https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-1/reg_2004_726/reg_2004_726_en.pdf

Health technology assessment	A multidisciplinary comparative assessment process, based on clinical and non-clinical assessment domains, which compiles and evaluates the available evidence about the clinical and non-clinical issues related to the use of a health technology. **Citation**: European Commission. Proposal for a Regulation of the European Parliament and of the Council on health technology assessment and amending Directive 2011/24/EU [Internet]. January 2018 [cited 2019 May 08]. Available from: https://ec.europa.eu/health/sites/health/files/technology_assessment/docs/com2018_51final_en.pdf
Interim analysis	Any analysis intended to compare treatment arms with respect to efficacy or safety at any time prior to the formal completion of a trial. **Citation**: International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). Statistical Principles for Clinical Trials (E9) [Internet]. February 1998 [cited 2019 May 08]. Available from: https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E9/Step4/E9_Guideline.pdf
Intra-rater/subject reliability	The property of yielding equivalent results when used by the same rater on different occasions. **Citation**: International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). Statistical Principles for Clinical Trials (E9) [Internet]. February 1998 [cited 2019 May 08]. Available from: https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E9/Step4/E9_Guideline.pdf
Investigational medicinal product	A pharmaceutical form of an active ingredient or placebo being tested or used as a reference in a clinical trial, including a product with a marketing authorization when used or assembled (formulated or packaged) in a way different from the approved form, or when used for an unapproved indication, or when used to gain further information about an approved use. **Citation**: European Medicines Agency. Guideline for good clinical practice E6 (R2) [Internet]. December 2016 [cited 2019 May 08]. Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e-6-r1-guideline-good-clinical-practice-step-5_en.pdf
Marketing authorisation	An authorisation required for a medicinal product to be placed on the market of a Member State which has been issued by the competent authorities of that Member State in accordance with Directive 2001/83/EC. **Citation**: European Commission. Directive 2001/83/EC of the European Parliament and of the Council of 6 November 2001 on the Community code relating to medicinal products for human use [Internet]. Official Journal of the European Communities 2001; L311:1-62 [cited 2019 May 08]. Available from: https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2001:311:0067:0128:en:PD

Medicinal product	Any substance or combination of substances which may be administered to human beings with a view to making a medical diagnosis or to restoring, correcting or modifying physiological functions in human beings is likewise considered a medicinal product. **Citation**: European Commission. Directive 2001/83/EC of the European Parliament and of the Council of 6 November 2001 on the Community code relating to medicinal products for human use [Internet]. Official Journal of the European Communities 2001; L311:1-62 [cited 2019 May 08]. Available from: https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2001:311:0067:0128:en:PD
Meta-analysis	The formal evaluation of the quantitative evidence from two or more trials bearing on the same question. This most commonly involves the statistical combination of summary statistics from the various trials, but the term is sometimes also used to refer to the combination of the raw data. **Citation**: International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). Statistical Principles for Clinical Trials (E9) [Internet]. February 1998 [cited 2019 May 08]. Available from: https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E9/Step4/E9_Guideline.pdf
Orphan medicinal product	A medicinal product shall be designated as an orphan medicinal product if its sponsor can establish: (a) that it is intended for the diagnosis, prevention or treatment of a life-threatening or chronically debilitating condition affecting not more than five in 10 thousand persons in the Community when the application is made, or that it is intended for the diagnosis, prevention or treatment of a life-threatening, seriously debilitating or serious and chronic condition in the Community and that without incentives it is unlikely that the marketing of the medicinal product in the Community would generate sufficient return to justify the necessary investment; (b) that there exists no satisfactory method of diagnosis, prevention or treatment of the condition in question that has been authorised in the Community or, if such method exists, that the medicinal product will be of significant benefit to those affected by that condition. **Citation:** European Commission. Regulation No. 141/2000 of the European Parliament and of the Council of 16 December 1999 on orphan medicinal products [Internet]. Official Journal of the European Communities 2000; L18:1-5 [cited 2019 May 08]. Available from: https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2000:018:0001:0005:en:PDF

Parallel consultation procedure	A single gateway for requests for parallel discussions before the start of pivotal clinical trials on initial evidence generation for Marketing Authorisation Application/Reimbursement and Post- Licensing Evidence Generation (PLEG) involving the European Medicines Agency, European Network for Health Technology Assessment and health technology assessment bodies. **Citation:** European Medicines Agency. Guidance for parallel consultation [Internet]. June 2017 [cited 2019 May 08]. Available from: https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/guidance-parallel-consultation_en.pdf
Patient registry	Organised systems that use observational methods to collect uniform data on a population defined by a particular disease, condition, or exposure, and that is followed over time. **Citation**: European Medicines Agency. Patient registry initiative – Strategy and mandate of the cross-committee task force [Internet]. May 2017 [cited 2019 May 08]. Available from: https://www.ema.europa.eu/en/documents/other/patient-registry-initiative-strategy-mandate-cross-committee-task-force_en.pdf
Patient reported outcome	A range of measurement types, encompassing simple symptom measures (such as pain measured by Likert scale), more complex measures (such as activities of daily living or function), multidimensional measures (such as health-related quality of life) and satisfaction with treatment. **Citation**: European Network for Health Technology Assessment. Endpoints used for relative effectiveness assessment of pharmaceuticals: Clinical endpoints [Internet]. February 2013 [cited 2019 May 08]. Available from: https://www.eunethta.eu/wp-content/uploads/2018/01/Clinical-endpoints.pdf
Post-authorisation efficacy study	Studies conducted on medicinal products within the authorised therapeutic indication to address well-reasoned scientific uncertainties identified by EU regulators on aspects of the evidence of benefits that should be, or can only be, addressed post-authorisation. **Citation:** European Medicines Agency. Scientific guidance on post-authorisation efficacy studies [Internet]. October 2016 [cited 2019 May 08]. Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/scientific-guidance-post-authorisation-efficacy-studies-first-version_en.pdf

Pragmatic clinical trial	Pragmatic trials are used in situations in which randomisation is needed, and where there is a need to explore the use of the intervention in settings that are less restricted than the pivotal trials and whether that difference in usage, if any, affects the reported efficacy of the intervention. Pragmatic trials may also be used where non-adherence to treatment or relaxing of behaviours could be an issue or where population characteristics, co-morbidities and co-medications could have an impact on the benefits of the medicine. **Citation:** European Medicines Agency. Scientific guidance on post-authorisation efficacy studies [Internet]. October 2016 [cited 2019 May 08]. Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/scientific-guidance-post-authorisation-efficacy-studies-first-version_en.pdf
Primary endpoint	An endpoint that should reflect clinically relevant effects and is typically selected based on the principal objective of the study. **Citation**: European Medicines Agency. General considerations for clinical trials [Internet]. March 1998 [cited 2019 May 08]. Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e-8-general-considerations-clinical-trials-step-5_en.pdf
Relative efficacy/ effectiveness assessment	An assessment to determine the extent to which an intervention does more good than harm compared to one or more intervention alternatives for achieving the desired results when provided under ideal circumstances (efficacy) or under the usual circumstances of health care practice (effectiveness). **Citation**: European Network for Health Technology Assessment. Relative Effectiveness Assessment (REA) of Pharmaceuticals in Europe [Internet]. November 2010 [cited 2019 May 08]. Available from: https://www.eunethta.eu/wp-content/uploads/2018/01/Relative-Effectiveness-Assessment-of-Pharmaceuticals-A-EUnetHTA-Joint-Action-Initiative.pdf
Secondary endpoint	An endpoint that assesses other drug effects that may or may not be related to the primary endpoint. **Citation**: European Medicines Agency. General considerations for clinical trials [Internet]. March 1998 [cited 2019 May 08]. Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e-8-general-considerations-clinical-trials-step-5_en.pdf
Subgroup analysis	Grouping together patients with similar characteristics to explore variability of response to treatment between different groups of patients within a clinical trial dataset. **Citation**: European Medicines Agency. Guidelines on the investigation of subgroups in confirmatory clinical trials [Internet]. January 2019 [cited 2019 May 08]. Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-investigation-subgroups-confirmatory-clinical-trials_en.pdf

Surrogate endpoint	A variable that provides an indirect measurement of effect in situations where direct measurement of clinical effect is not feasible or practical. **Citation:** International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). Statistical Principles for Clinical Trials (E9) [Internet]. February 1998 [cited 2019 May 08]. Available from: https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E9/Step4/E9_Guideline.pdf
Translational research	The process of translating findings derived from basic science to the development of a new understanding of the disease mechanism, diagnosis and therapeutics. **Citation**: Bast RC Jr, Mills GB, Young RC. Translational research – traffic on the bridge. Biomed Pharmacother. 2001;55(9-10):5650571.

List of Acronyms

AIFA	Italian Medicines Agency, Italy
ALL	Acute Lymphoblastic Leukaemia
AML	Acute Myeloid Leukaemia
AOTMiT	Agencja Oceny Technologii Medycznych i Taryfikacji, Poland
APL	Acute Promyelocytic Leukaemia
ASCO	American Society of Clinical Oncology
ASCT	Allogeneic Stem Cell Transplantation
ASR	Age-standardised Mortality Rate
ATMP	Advanced Therapy Medicinal Product
ATV	Added Therapeutic Value
BOR	Best Overall Remission/Response Rate
CDP	Clinical Development Programme
CED	Coverage with Evidence Development
CEO	Chief Executive Officer
CHMP	Committee for Medicinal Products for Human Use
CLL	Chronic Lymphocytic Leukaemia
CMA	Conditional Marketing Authorisation
CML	Chronic Myeloid Leukaemia
CMML	Chronic Myelomonocytic Leukaemia
COMET	Core Outcome Measures in Effectiveness Trials

COS	Core Outcome Set
COS-STAD	Core Outcome Set – Standards for Development
COST-STAR	Core Outcome Set – Standards for Reporting
CR	Complete Remission/Response Rate
CRc	Combined Complete Remission/Response Rate
CRi/CRh	Complete Remission/Response Rate with Incomplete Blood Count Recovery
CRp	Complete Remission/Response Rate with Incomplete Platelet Recovery
CVI	Content Validity Index
CyR	Cytogenetic Response Rate
DFS	Disease-free Survival
DOR	Duration of Remission/Response
DPA	Directorate for Pharmaceutical Affairs, Malta
DPO	Data Protection Officer
e-Delphi	Electronic Delphi
EEA	European Economic Area
EFS	Event-free Survival
EMA	European Medicines Agency
ENVI	The European Parliament Committee on Environment, Public Health and Food Safety
EPAR	European Public Assessment Report

ESMO-MCBS	European Society for Medical Oncology – Magnitude of Clinical Benefit Scale
EU	European Union
EudraCT	The European Union Clinical Trials Register
EUnetHTA	European Network for Health Technology Assessment
FDA	Food and Drug Administration
FIMEA	Finnish Medicines Agency, Finland
FREC	Faculty of Medicine and Surgery Research Ethics Committee
GVHD	Graft Versus Host Disease
HAS	Haute Autorité de Santé, France
HEN	Health Evidence Network
HR	Haematologic Response Rate
HRQoL	Health-related Quality of Life
HSCT	Haematopoietic Stem Cell Transplantation
HTA	Health Technology Assessment
HTAB	Health Technology Assessment Body
HTAi	Health Technology Assessment International
IARC	International Agency for Research on Cancer
ICH	International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
I-CVI	Item-level Content Validity Index
IMP	Investigational Medicinal Product

INAHTA	International Network of Agencies for Health Technology Assessment
INFARMED	National Authority for Medicines and Health Products, Portugal
IP	Internet Protocol
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, Germany
LBI-HTA	Ludwig Boltzmann Institute for Health Technology Assessment, Austria
MA	Marketing Authorisation
MAH	Marketing Authorisation Holder
MCU	Medical Care Utilisation
MDH	Mater Dei Hospital
MDS	Myelodysplastic Syndromes
MedDRA	Medical Dictionary for Regulatory Activities
MLFS	Morphologic Leukaemia-free State
MPD	Myeloproliferative Disorders
MR	Molecular Remission/Response
MRD	Minimal Residual Disease
NCPE	National Centre for Pharmacoeconomics, Ireland
NGT	Nominal Group Technique
NHS	National Healthcare System
NICE	National Institute for Health and Care Excellence, England

NKI	Netherlands Cancer Institute
OMERACT	Outcome Measures in Rheumatology
ORR	Objective Remission/Response Rate
OS	Overall Survival
PAES	Post-authorisation Efficacy Study
PD	Rate of Progressive Disease
PFS	Progression-free Survival
PIP	Paediatric Investigation Plan
PR	Partial Remission/Response Rate
PRO	Patient Reported Outcome
REA	Relative Efficacy/Effectiveness Assessment
RECIST	Response Evaluation Criteria in Solid Tumours
ReVALeu	Response Evaluation in Leukaemia
RR	Relapse Rate
RTC	Return to Chronic Phase
SAMOC	Sir Anthony Mamo Oncology Centre
SAWP	Scientific Advice Working Party
S-CVI	Scale-level Content Validity Index
SEED	Shaping European Early Dialogues
SMC	Scottish Medicines Consortium, Scotland
SUKL	State Institute for Drug Control, Czech Republic
TFR	Treatment-free Remission/Response

TFS	Treatment-free Survival
TLV	Dental and Pharmaceutical Benefits Agency, Sweden
TTF	Time to Treatment Failure
TTP	Time to Progression
TTR	Time to Remission/Response
TTT	Time to Treatment
WHO	World Health Organization
WHO/Europe	World Health Organization Regional Office for Europe
ZINL	National Health Care Institute, The Netherlands

Chapter 1

Introduction

1.1 The Cancer Burden in the European Union

One in four deaths in the European Union (EU) is caused by cancer.[1] In 2014, cancer was the cause of death for over 1.3 million EU citizens.[2] Data from 2009 indicates that drug expenditure accounted for 27% of cancer-related healthcare costs in the EU, equivalent to more than €13.5 billion (Luengo-Fernandez et al, 2013). Soaring prices are predicted for treatment interventions with novel mechanistic pathways including targeted therapy such as immunotherapy, small molecule therapy and monoclonal antibodies (Prasad et al, 2017). The magnitude of cancer-related mortality coupled with constraints in health care budgets signal the onus on pharmaceutical sponsors to robustly design clinical development programmes (CDPs) for antineoplastic therapies that demonstrate tangible clinical benefit capable of convincing decision-makers.

1.2 Regulatory and Health Technology Assessments of Oncology Medicines

Patient access to innovative oncological therapy is preceded by highly complex, multi-stakeholder procedures comprising of regulatory, health technology assessment (HTA) and coverage bodies as key decision-makers. Figure 1.1 outlines the trajectory taken by an anticancer medicinal product from stages of translational research to authorisation and post-marketing approvals until it is ultimately taken up in clinical practice. Pharmaceutical companies requesting a marketing authorisation (MA) for anticancer medicines in the EU are obliged to apply for a centralised assessment by the European

[1] Eurostat. World Cancer Day: 1 in 4 deaths caused by cancer in the EU [Internet]. March 2017 [cited 2019 Apr 10]. Available from: http://ec.europa.eu/eurostat/web/products-eurostat-news/-/EDN-20170203-1?inheritRedirect=true&redirect=%2Feurostat%2F

[2] European Medicines Agency (EMA). Patient Registries Workshop, 28 October 2016: Observations and Recommendations Arising from the Workshop [Internet]. February 2017 [cited 2019 Apr 10]. Available from: https://www.ema.europa.eu/documents/report/report-patient-registries-workshop_en.pdf

Medicines Agency (EMA).[3] The EMA Committee for Medicinal Products for Human Use (CHMP) is tasked with the assessment of safety, quality and efficacy data submitted in medicinal product dossiers and to ensure a positive benefit-risk balance for the intervention under evaluation. Evolvement of the EU medicines regulatory framework introduced early access instruments to fast-track the development and regulatory review of applications, including the conditional marketing authorisation (CMA), authorisation under exceptional circumstances, accelerated assessments and adaptive pathways (Martinalbo et al, 2016).

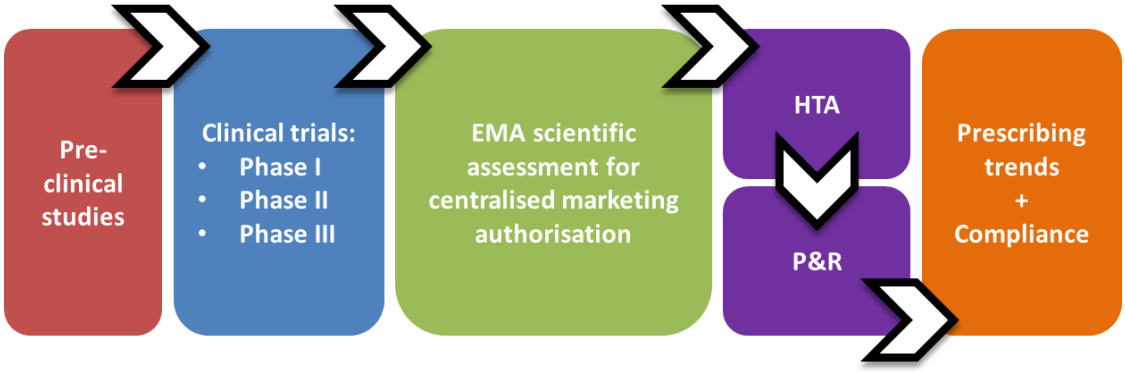


Figure 1.1 Pathway for patient accessibility to oncological therapy

Adapted from: Ades F, Zardavas D, Senterre C, De Azambuja E, Eniu A, Popescu R, et al. Hurdles and delays in access to anti-cancer drugs in Europe. Ecancermedicalscience. 2014;8:11 pages.

Acronyms: European Medicines Agency (EMA); Health Technology Assessment (HTA); Pricing and Reimbursement (P&R)

Hoekman and colleagues[4] (2014) report that when the first 11 conditional approvals in oncology during 2006-2013 were compared with 31 regular MA approvals for cancer medicines in the same period, paradoxically those granted a CMA had more frequent involvement of the external scientific advisory group, prolonged review times and lower instances of consensual votes. Similarly, more than half (n=12, 52%) of the applications

[3] European Commission. Regulation No. 726/2004 of the European Parliament and of the Council of 31 March 2004 laying down community procedures for the authorisation and supervision of medicinal products for human and veterinary use and establishing a European Medicines Agency [Internet]. Official Journal of the European Union 2004; L136:1-33 [cited 2019 Apr 20]. Available from: https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-1/reg_2004_726/reg_2004_726_en.pdf

[4] Hoekman JDB, Marie L, Boon Wouter PC. Improving the EU system for the marketing authorisation of medicines [Internet]. September 2014 [cited 2019 Apr 20]. Available from: http://lygature.c1.s3.aegirhost.nl/sites/lygature.c1.s3.aegirhost.nl/files/atoms/files/Escher_report_IA.pdf

for oncological products requesting accelerated assessments between 2006-2014 were rejected due to uncertain clinical relevance (Martinalbo et al, 2016). These scenarios demonstrate the challenges faced by pharmaceutical developers in providing persuasive evidence to the EMA. Irrespective of the chosen regulatory route, unconvincing clinical data may result in negative opinions or extended regulatory review timeframes, in turn defeating the principles of early access mechanisms.

The grant of a centralised European MA does not necessarily imply effective market access to national healthcare systems (NHS). HTA is described as a "key tool for Member States to ensure the accessibility, quality and sustainability of health care."[5] HTA is composed of two main domains targeting clinical (added therapeutic value (ATV), clinical efficacy/effectiveness and safety) and non-clinical (economic, ethical, organisational, social and legal) criteria.[5] The notion of ATV for medicinal products refers to the incremental therapeutic value brought by a new medicine compared with the best available treatment options already on the market and is estimated using relative efficacy or effectiveness assessments (REAs).[5] ATV secures return on investment by steering resources towards highest value medicines and rewarding interventions with added value. The generation of pre- and post-authorisation evidence underpins REAs.[5]

In contrast with the centralised European MA decisions, each 28 Member State performs HTAs independently at a national or regional level which are subsequently used to inform pricing and/or reimbursement decisions (Kleijnen et al, 2016). For some novel antineoplastic medicines, the steep acquisition prices are thought to be

[5] European Commission. Inception Impact Assessment: Strengthening of the EU cooperation on Health Technology Assessment (HTA) [Internet]. September 2016 [cited 2019 Apr 20]. Available from: http://ec.europa.eu/smartregulation/roadmaps/docs/2016_sante_144_health_technology_assessments_en.pdf

disproportionate to the modest clinical benefit gained (Sorenson, 2012). This presents dilemmas for HTA bodies (HTABs) on the true therapeutic value brought about by innovative products on the market.

1.2.1 Divergent Decisions and Associated Risks

European HTA jurisdictions are heterogeneous in terms of their organisational and methodological frameworks (Sorenson and Chalkidou, 2012; Allen et al, 2017a), assessment criteria (Kleijnen et al, 2016) and stage of appraisal (Martinalbo et al, 2016). This has manifested in large variability within and across states on the outcome and timing of recommendations for the reimbursement status of medicines, including anticancer therapies. For example, the time taken from an oncological agent receiving an MA to effectively access the market varied from 0 to 334 days in average across only four different European countries.[6] Such incongruence has also been reported beyond European confines, including the North American, Asian and Australian continents (Neumann et al, 2012; Chabot and Rocchi, 2014; Lim et al, 2014; Dranitsaris and Papadopoulos, 2015; Rocchi et al, 2015; Allen et al, 2017b).

An international study by Kanavos and colleagues (2010) reveal that more than half of the outcomes for six different HTA agencies between 2007 and 2009 had conflicting views. Factors driving the differences in recommendations included divergent clinical and economic evidence requirements, preferred clinical endpoints, data interpretation, choice of comparator and use of cost-effectiveness thresholds. In another study (Kleijnen et al, 2016), outcomes from HTAs for anticancer medicines granted an MA between 2011 and 2013 were reviewed for six European jurisdictions; the non-

[6] Hofmarcher T, Jonsson B, Wilking N. Access to high-quality oncology care across Europe, IHE Report [Internet]. 2014 [cited 2019 Apr 02]. Available from: http://portal.research.lu.se/ws/files/29186857/IHE_Report_2014_2.pdf

uniformity in recommendations to inform pricing and/or reimbursement decisions is highlighted in Table 1.1.

Table 1.1 Outcomes of recommendations that inform pricing and/or reimbursement decisions for anticancer agents

Adapted from: Kleijnen S, Lipska I, Leonardo Alves T, Meijboom K, Elsada A, Vervolgyi V, et al. Relative effectiveness assessments of oncology medicines for pricing and reimbursements decisions in European countries. Ann Oncol. 2016;27(9):1768-1775.

Jurisdiction		England	France	Germany	The Netherlands	Poland	Scotland
HTAB		NICE	HAS	IQWiG	ZIN	AOTMiT	SMC
Indication	Medicine						
Bone metastasis from solid tumours	Denosumab	(+)	(+)[a], (±)[a]	Not assessed	(±)	(-)	Not assessed
Breast cancer	Eribulin	(-)	(+)	(±)[b], (±)[b]	(+)	(-)	(-)
	Pertuzumab	Not assessed	(+)	(+)	Not assessed	(+)	(-)
Colorectal cancer	Aflibercept	(-)	(±)	(+)	Not assessed	(+)	(-)
Gastric cancer	Tegafur/gimeracil/oteracil	Not assessed	(-)	Not assessed	(-)	(-)	(+)
Melanoma	Ipilimumab (2nd line Tx)	(+)	(+)	(+)	(+)	(+)	(-)
	Vemurafenib	(+)	(+)	(+)	(+)	(+)	(-)
	Dabrafenib	(+)	(±)	(+)	Not assessed	(+)	(+)
NSCLC	Afatinib	(+)	(±)	(+)[c], (+)[c], (±)[c], (-)[c]	Not assessed	(+)	(+)
	Crizotinib	(-)	(+)	(±)	Not assessed	(-)	(-)
Prostate cancer	Cabazitaxel	(-)	(+)	(+)[d], (+)[d]	(+)	(-)	(-)
	Abiraterone (after Tx with taxane)	(+)	(+)	(+)	(±)	(+)	(-)
	Enzalutamide	(+)	(+)	(+)[e], (+)[e]	Not assessed	(+)	(+)
Renal cell carcinoma	Axitinib	(+)	(+)	(+)	Not assessed	(+)	(-)

Acronyms: Agencja Oceny Technologii Medycznych i Taryfikacji (AOTMiT); Haute Autorité de Santé (HAS); Health technology assessment body (HTAB); Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG); National Institute for Health and Care Excellence (NICE); Non-small cell lung cancer (NSCLC); Scottish Medicines Consortium (SMC); Treatment (Tx); Zorginstituut Nederland (ZIN)

Key: (+) recommended/added benefit; (-) not recommended/lesser benefit; (±) no added benefit proven/similar therapeutic value;
[a]HAS recommended that denosumab provides a minor improvement in actual benefit (level IV) in patients with breast cancer or prostate cancer with bone metastasis but does not provide an improvement in actual benefit (level V) in patients with other types of solid tumours with bone metastases;
[b]Eribulin was assessed by IQWiG for two patient subgroups;
[c]Afatinib was assessed by IQWiG for six different sub-populations, of which four were included in the analysis;
[d]Cabazitaxel was assessed by IQWiG for two different sub-populations.
[e]Enzalutamide was assessed by IQWiG for two different sub-populations

Regulatory early access instruments have increased flexibility in the authorisation process and premature clinical data has been accepted as basis for approvals of medicines used in malignancy (Cressman et al, 2015; Leyens and Brand, 2016). HTABs tend to be more rigid in their evaluations by requesting mature clinical datasets for economic modelling and REAs (Leyens and Brand, 2016; Martinalbo et al, 2016; Wang et al, 2018). Clinical endpoints, amongst other clinical trial design elements, considered valid and reliable by regulators are questioned and sometimes not accepted by HTABs and payers (Ruof et al, 2014; Martinalbo et al 2016). In a comparative analysis of parallel scientific advice outputs between HTABs and the EMA (Tafuri et al, 2016), 41% of the opinions on trial endpoints were not in full agreement (Figure 1.2).

The implications of regulatory and HTA discordance are of particular concern in rare malignancies which constitute approximately 40% of the orphan designations (Berggren et al, 2012; Pauwels et al, 2017). Orphan medicines tend to be associated with insufficient evidence at time of regulatory approval secondary to difficulties in patient

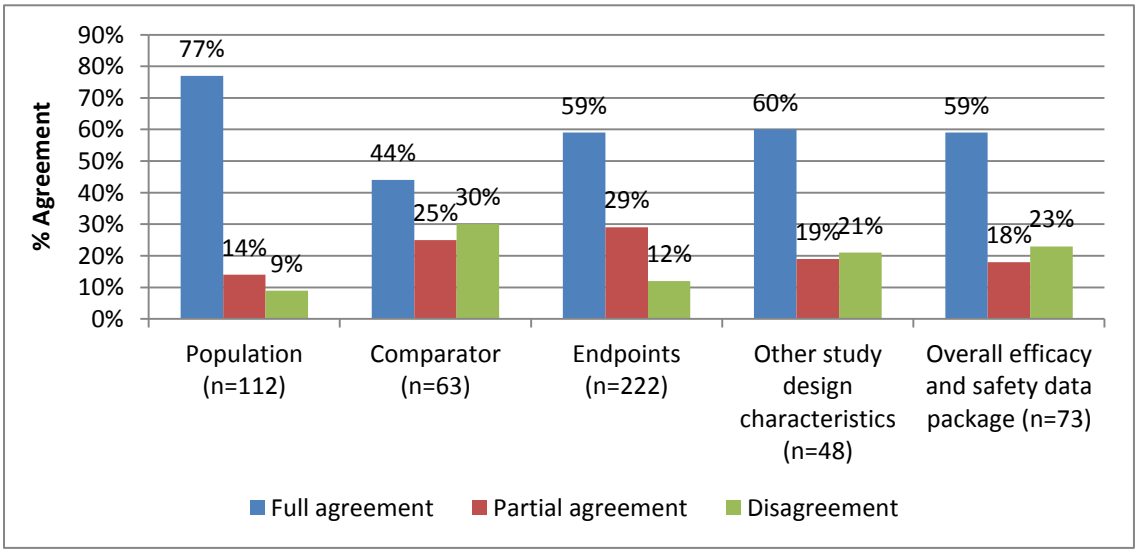


Figure 1.2 Level of agreement for different assessment domains between health technology assessment bodies and regulators for 31 parallel scientific advice procedures

Adapted from: Tafuri G, Pagnini M, Moseley J, Massari M, Petavy F, Behring A, et al. How aligned are the perspectives of EU regulators and HTA bodies? A comparative analysis of regulatory-HTA parallel scientific advice. Br J Clin Pharmacol. 2016;82:965-973.

Key: 'n' represents the total number of HTABs expressing an opinion for each domain

recruitment due to disease epidemiology, the use of surrogate endpoints and inadequate follow-up periods (Joppi et al, 2016). Disparities on how clinical uncertainty is addressed by HTABs may prove to be a barrier to access (Denis et al, 2010; Adkins et al, 2017). In a multi-stakeholder study, industry representatives have confirmed these observations by indicating that regulatory and HTA divergences have been mostly evident in areas of high clinical uncertainty such as oncology and orphan medicines (Wang et al, 2018).

This body of data presents the risks associated with inconsistent evidence needs between decision-makers which ensue in delays or failure for European citizens to effectively access new oncology agents. To this end, a paradigm shift towards streamlining HTA and regulatory clinical assessments should be attempted for innovations in cancer care to reach patients in a timely manner.

1.2.2 Blurring Boundaries between Decision-makers

Initiatives have been undertaken to increase interactions between decision-makers and avoid mismatches in evidence requested and that generated during a product's clinical development. Programmes promoting early dialogue during drug development stages between regulatory, HTA, coverage bodies and industry stakeholders have been successfully implemented. The European Commission pilot project 'Shaping European Early Dialogues' (SEED) (Martinalbo et al, 2016) and the Tapestry Networks (Fronsdal et al, 2012; Cuche et al, 2014) are prime examples of these multi-stakeholders platforms. The provision of joint technical advice from the European Network for Health Technology Assessment (EUnetHTA) and the EMA in the parallel scientific advice procedure (Elvidge, 2014), which has evolved to the current parallel consultation

procedure[7], serves as a beneficial instrument for pharmaceutical developers to bolster their data packages. The EUnetHTA and the EMA since 2010 have also collaborated to identify opportunities for amendments to the European Public Assessment Report (EPAR) template to facilitate subsequent REAs by HTABs (Fronsdal et al, 2012; Bergmann et al, 2014; Berntgen et al, 2014). In 2013[8], the EMA and EUnetHTA launched a three-year work plan on collaboration which has been renewed for the years 2017-2020[9] setting out a strategy to enhance dialogue on evidence needs in order to facilitate access to medicines for patients in the EU. The European Parliament Committee on the Environment, Public Health and Food Safety (ENVI) published a report[10] on the centralisation of ATV assessments proposing the common adoption of outputs from REAs by European national and regional HTABs. This report has triggered a legislative process by the European Commission for a Regulation to consolidate HTA in the EU with the objective of circumventing duplication of work and distorted market access by harmonising HTA procedures and methodologies.[11]

[7] European Network for Health Technology Assessment (EUnetHTA) and European Medicines Agency (EMA). Guidance for parallel consultation [Internet]. June 2017 [cited 2019 Apr 10]. Available from: https://www.ema.europa.eu/documents/regulatory-procedural-guideline/guidance-parallel-consultation_en.pdf

[8] European Network for Health Technology Assessment (EUnetHTA) and European Medicines Agency (EMA). EMA-EUnetHTA three-year work plan 2013-2015. November 2013 [cited 2019 Apr 10]. Available from: https://www.ema.europa.eu/documents/other/european-medicines-agency-eunethta-three-year-work-plan_en.pdf

[9] European Network for Health Technology Assessment (EUnetHTA) and European Medicines Agency (EMA). EMA-EUnetHTA three-year work plan 2017-2020. November 2017 [cited 2019 Apr 10]. Available from: https://www.ema.europa.eu/documents/other/ema-eunethta-three-year-work-plan-2017-2020_en.pdf

[10] European Parliament: Environment, Public Health and Food Safety Committee (ENVI). Towards a harmonised EU assessment of added therapeutic value of medicines [Internet]. June 2015 [cited 2019 Apr 10]. Available from: http://www.europarl.europa.eu/RegData/etudes/STUD/2015/542219/IPOL_STU(2015)542219_EN.pdf

[11] European Commission. Proposal for a Regulation of the European Parliament and of the Council on health technology assessment and amending Directive 2011/24/EU [Internet]. January 2018 [cited 2019 Apr 10]. Available from: https://ec.europa.eu/health/sites/health/files/technology_assessment/docs/com2018_51final_en.pdf

1.3 Evidence Generation in Oncology: A Continuum in Clinical Development

Considerations to evidence-generating strategies are given by pharmaceutical companies throughout the stages of medicinal product development; from initial research investment decisions to designing pre-marketing clinical trial protocols and devising mechanisms to collect effectiveness data through real-world use. Proceedings from the Health Technology Assessment International (HTAi) Policy Forum meeting in February 2011 underscore the need for industry stakeholders to generate evidence that is able to support regulatory and coverage approvals, therefore having a direct interest in ensuring clarity and predictability in those evidence requirements (Henshall et al, 2011). Evidence is collated along the product's lifecycle with efficacy data produced during pre-marketing phases being corroborated by outcome data recorded in patient registries, pragmatic trials, observational studies, digital health platforms and claims data when the intervention is used in routine clinical practice (Martinalbo et al, 2016).

1.3.1 Synthesis of Efficacy Data in Phase II and Phase III Clinical Trials

Assessment of clinical benefit in the regulatory review process and early HTAs at product launch is based on efficacy profiles generated in phase II and phase III clinical trials. The primary objective of phase II clinical trials, also termed "therapeutic exploratory" trials, is to explore the therapeutic efficacy of an investigated treatment in patients by evaluating preliminary evidence with the use of endpoints that measure drug activity and therapeutic benefit[12] (Umscheid et al, 2011). Phase III trials, referred to as "therapeutic confirmatory" or pivotal trials, are designed to establish the preliminary efficacy accumulated in Phase II.

[12] International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). ICH Harmonised Tripartite Guideline – General Considerations for Clinical Trials E8 [Internet]. July 1997 [cited 2019 Apr 10]. Available from: https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E8/Step4/E8_Guideline.pdf

"Patient-centred" endpoints, including overall survival (OS) and health-related quality of life (HRQoL), are widely accepted measures among decision-makers in trials for many cancer types given their clinical relevance and direct measurement of therapeutic benefit (Fiteni et al, 2014). The weight given to evidence generation of OS is unique to neoplastic disease secondary to the poor prognosis and high risk of mortality related with certain malignancies when compared to other therapeutic areas such as neurology, rheumatology and endocrinology (Wilson et al, 2015a). The gold standard OS may be associated with much longer follow-up periods, especially when the investigated therapy is studied against an effective standard of care in the control group (Fiteni et al, 2014) and/or in indolent malignancies (Wilson et al, 2015b). This prolonged interval in generating evidence can potentially delay access for patients to innovative treatments and significantly augment costs for clinical trial conduct. Promising agents showing high activity in early stages of clinical development may compromise clinical equipoise which necessitates the cross-over of patients to the experimental arm. Cross-over may confound the demonstration in OS improvement and requires that the trial investigates alternative endpoints (Blumenthal et al, 2017).

For these reasons, surrogate measures involving biological markers, such as progression-free survival (PFS) and disease-free survival (DFS), are increasingly substituting clinical endpoints in oncology clinical trials since they allow an expedited evaluation of efficacy (Fiteni et al, 2014). The strength of correlation of surrogate measures with clinical endpoints is however not always sufficiently demonstrated which may compromise therapeutic gain by the patient (Booth and Eisenhauer, 2012; Kim and Prasad, 2015; Carlson, 2016).

Regulatory bodies have issued guidance documents specifically on the design of clinical trials for medicines investigated in malignancy.[13,14] Both the Food and Drug Administration (FDA) and EMA guidelines consider OS as the cornerstone for approvals but surrogate markers reflecting tumour burden are considered acceptable if carefully and rigorously validated or in cases where there is lack of or limited alternative treatment options in rare or life-threatening cancer states. Apolone and colleagues (2005) have identified that response rate was selected as a primary measure for efficacy in over 60% of the clinical studies for a centralised EMA approval of 14 anticancer agents indicated in solid tumours between 1995 and 2004. This trend is also observed by the FDA where endpoints, other than survival, were used for the granting of regular marketing approvals of 68% of applications for oncology medicines (Johnson et al, 2003).

Evidence from case studies reveals that efficacy endpoints studied for anticancer medicines are valued differently between HTABs (Shah et al, 2013; Kleijnen et al, 2015). Akin to the regulatory position, outcome data in terms of OS is the most relevant measure for policymakers but statistically powering a clinical study to capture the effects on mortality may not be feasible for pharmaceutical developers (Jönsson, 2015). A comparative analysis on the inclusion of endpoints in REAs and their impact on HTA recommendations for solid tumour therapy shows that while OS consistently had an impact in 94% of the decisions, PFS impact significantly differed from no influence in Germany and up to 85% in other jurisdictions (Kleijnen et al, 2016).

[13] European Medicines Agency (EMA). Guidelines on the evaluation of anticancer medicinal products in man [Internet]. 2017 Sep [cited 2019 Apr 12]. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2017/11/WC500238764.pdf

[14] Food and Drug Administration (FDA). Guidance for industry clinical trial endpoints for the approval of cancer drugs and biologics [Internet]. 2007 May [cited 2019 Apr 12]. Available from: https://www.fda.gov/downloads/drugsGuidanceComplianceRegulatoyInformation/Guidance/UCM071590.pdf

1.3.2 Real-World Surveillance of Clinical Performance

Data garnered in clinical trials for marketing purposes involves a narrowly-defined patient cohort being investigated under controlled and ideal conditions. The clinical trial data elements collated and analysed for the product profile are used for both regulatory and health technology assessments. Decision-makers are also interested in the performance of the intervention in the real-world setting and how novel therapeutic modalities compare to established standard-of-care treatments in clinical practice.

Observational data, such as electronic health records and population surveillance registries, together with pragmatic trials complement clinical trials designed for regulatory purposes and contribute to knowledge required to bridge the efficacy to effectiveness gap by providing valuable information on long-term clinical outcomes that are not feasible to study in phases of clinical trials (Pignatti et al, 2015; Ording et al, 2016). This information is used by HTABs to revisit early positions about interventions evaluated at point of product launch whilst allowing for the continued assessment of added value. From a regulatory perspective, outcome data generated post-approval, for example in post-authorisation efficacy studies (PAES)[15], closes the loop for ancillary evidence required for a CMA to be converted to a standard MA and forms the basis for a more complete clinical picture on the benefits and risks of authorised medicines.

In oncology, one major pitfall for real-world evidence generation is the lack of standardisation of endpoints to be captured (Ording et al, 2016). Evidence criteria have been successfully defined and harmonised by registry networks such as the UK Renal Registry, the Big Multiple Sclerosis Data Registry, the Parkinson's Disease group and

[15] European Medicines Agency (EMA). Scientific guidance on post-authorisation efficacy studies [Internet]. 2016 October [cited 2019 Apr 12]. Available from: https://www.ema.europa.eu/documents/scientific-guideline/scientific-guidance-post-authorisation-efficacy-studies-first-version_en.pdf

the neuro-muscular TREAT-NMD Global Registries but only 18% of cancer registries collect sufficiently detailed information.[16] Coverage with evidence development (CED) is a concept adopted by public and private healthcare payers which is characterised by a defined period of formal introduction of the innovative product coupled with progressive accrual of patient outcomes from use of the intervention. The CED strategy has been successfully implemented by the Netherlands Cancer Institute (NKI) for novel therapies in melanoma, breast cancer and gastric carcinoma (Van Harten and Retel, 2016). Oncology is rapidly evolving from a system of cancer subtypes classified according to histology to the molecular identification of malignancies, leading to a collection of rare tumours being studied in smaller clinical trials (Eichler et al, 2012; Wilson et al, 2015a). The era of personalised medicine will put further pressure on healthcare payers to introduce CED or similar schemes that are based on large-scale real-world outcome studies and databases.

1.4 Rationale for Standardising Response Criteria

Heterogeneity in outcomes measured in clinical studies for cancer therapies presents a major barrier to evidence synthesis, increasing the complexity of regulatory, policy and healthcare decision-making. The magnitude of inconsistency is striking, with sources reporting over 25,000 of outcomes in cancer trials only having been used once or twice (Hirsch et al, 2013). Another relevant factor is outcome reporting bias where the investigator tends to capitalise on the opportunity of studying multiple response parameters by reporting findings that are only positive and statistically significant (Chan and Altman, 2005; Williamson et al, 2005). Misleading the end-user by presenting an overly optimistic effect of an intervention can have dangerous ramifications when this

[16] European Medicines Agency (EMA). Patient Registries Workshop, 28 October 2016 [Internet]. 2017 February [cited 2019 Apr 14]. Available from: https://www.ema.europa.eu/documents/report/report-patient-registries-workshop_en.pdf

evidence is used as foundation for decision-making (Clarke, 2007). Additionally, marked diversity in response criteria contributes to difficulty in interpreting findings between studies which in turn hinders potential systematic reviews and meta-analyses (Harman et al, 2013; Waters et al, 2014; Clarke and Williamson, 2016).

A proposed solution to address these issues is the development of core outcome sets (COSs). A COS is defined as "a list of outcomes that should be measured and reported, as a minimum, in all clinical trials in specific areas of health or healthcare" and developed by involving relevant stakeholders (Gargon et al, 2017). Clarke (2007) goes one step beyond by mentioning that the identified core outcomes should not be restricted to clinical trials but also adopted in systematic reviews and routine clinical practice. Standardising reporting outcomes will harness greater consistency in the body of evidence contributable to health research, for the ultimate benefit of the patient.

1.5 The Case of Leukaemia

This section describes the high incidence and mortality rates of leukaemia in Europe and the need to invest in outcomes research and initiatives that drive the clinical development of innovative anti-leukaemic agents.

1.5.1 Epidemiology and Mortality

Haematological malignancies, taken together, rank third in terms of cancer-related age-adjusted mortality in the European Economic Area (EEA).[17] According to 2018 statistics published by the World Health Organisation (WHO) International Agency for

[17] Eurostat. Causes of death, 2015 [Internet]. [cited 2019 Apr 13]. Available from: https://ec.europa.eu/eurostat/web/health/causes-death/data/database

Research on Cancer (IARC)[18], leukaemic disorders account for the highest age-standardised mortality rate (ASR) among haematological neoplasms for both sexes in Europe, exceeding the lymphoma and myeloma groups of blood cancers (Table 1.2). Leukaemia also ranks as the ninth highest cause of death among all cancer types in Europe for 2018. At a significant age-standardised rate of 7.4%, leukaemia comes in after Non-Hodgkin lymphoma for new cases reported in Europe in 2018 (Table 1.3).

Table 1.2 Estimated number of deaths in 2018 for haematological malignancies in Europe for both sexes and all ages[18]

Haematological Malignancy	Mortality		
	Number	Crude Rate (%)*	ASR (World) (%)*
Leukaemia	61,476	8.3	3.6
Non-Hodgkin lymphoma	48,096	6.5	2.6
Multiple myeloma	30,860	4.1	1.6
Hodgkin lymphoma	4,307	0.58	0.33

*Crude and age-standardised rates per 100,000. Acronyms: Age-standardised mortality rate (ASR)

Table 1.3 Estimated number of new cases in 2018 for haematological malignancies in Europe for both sexes and all ages[18]

Haematological Malignancy	Incidence		
	Number	Crude Rate (%)*	ASR (World) (%)*
Non-Hodgkin lymphoma	115,118	15.5	8.1
Leukaemia	94,780	12.7	7.4
Multiple myeloma	48,297	6.5	2.9
Hodgkin lymphoma	19,193	2.6	2.4

*Crude and age-standardised rates per 100,000. Acronyms: Age-standardised mortality rate (ASR)

[18] World Health Organisation (WHO) – International Agency for Research on Cancer (IARC). GLOBOCAN 2018 database: The Global Cancer Observatory (GCO) [Internet]. 2018 September [cited 2019 Apr 13]. Available from: http://gco.iarc.fr/today/home

These statistics make leukaemia, particularly certain acute and chronic variants, a disease category which is not optimally managed, therefore attracting investment by pharmaceutical companies. This is confirmed by the number of innovative treatment options in the pipeline, some of which are in the advanced phases of clinical development (Starr, 2016).

1.5.2 A Paucity in Outcomes Research

A number of cancer value frameworks, such as the European Society for Medical Oncology Magnitude of Clinical Benefit Scale (ESMO-MCBS) and the American Society of Clinical Oncology (ASCO) matrix, have been developed to determine the relative value of novel therapeutic agents in oncology (Cherny et al, 2015; Schnipper et al, 2015). The Response Evaluation Criteria In Solid Tumours (RECIST) is a set of criteria designed to provide an objective and uniform assessment of tumour burden between different trials (Wilson et al, 2015b). These initiatives are highly beneficial in determining the clinical value of treatments and for the identification of outcomes to measure in clinical studies, however they are either only available for solid tumour therapy or apply imperfectly for medicines indicated in haematologic cancers (Cheung et al, 2016). The Core Outcome Measures in Effectiveness Trials (COMET) database[19] is a publicly accessible repository of published and ongoing COS studies. Only five leukaemia-related COS studies feature in the database and are primarily concerned with the definition not prioritisation of outcomes. The opinions of researchers, clinicians and patient groups were sought in devising these core outcome sets. A lacuna still exists for regulatory and HTA perspectives on preferred parameters measured in efficacy studies for leukaemia clinical trials.

[19] Core Outcome Measures in Effectiveness Trials (COMET) Initiative. COMET Database [Internet]. [cited 2019 Apr 16]. Available from: http://www.comet-initiative.org/studies/search/

1.6 Aims and Objectives of the Study

The aims of this research were to determine core efficacy outcomes prioritised by decision-makers in the European Union (EU) for leukaemia clinical trials and to analyse the evolvement of evidence-generating strategies in clinical development programmes (CDPs) for medicinal products indicated in leukaemia.

The primary objective of the study was to:

- Identify a set of core efficacy outcomes prioritised by regulatory and health technology assessment (HTA) experts in the European Union (EU) for clinical trials investigating medicinal products in leukaemia.

The secondary objectives of the study were to:

- Analyse trends in efficacy parameters studied in leukaemia clinical trials and to determine whether they align with preferences of decision-makers;
- Explore the opinions of EU regulatory and HTA experts on the quality of evidence generated and alignment of clinical assessments for antineoplastic therapies.

Chapter 2

Methodology

This chapter outlines the methodological framework adopted to reach the research objectives described in the first chapter. The methodological design was divided into two parts:

- Part I presents the process undertaken to select the clinical trials included in the study and the methods chosen to perform an analysis of trends for the parameters investigated in leukaemia clinical trial efficacy studies.
- Part II details the steps followed in the development of an online surveying tool, the conduct of a two-round Delphi process with regulatory and HTA experts and the evaluation of decision-maker opinions in general and technical aspects related to clinical trial design and clinical assessments of antineoplastic agents.

2.1 Approvals and Consents

The following approvals and consents were granted prior to the study initiation (Appendix 1):

- Approval by the University of Malta Faculty of Medicine and Surgery Research Ethics Committee (FREC) for the proposed research protocol;
- Consents needed to engage oncologists and haematologists in the validation process (section 2.3.2), which included:
    i. Approval by the Data Protection Officer (DPO) at Mater Dei Hospital (MDH);
    ii. Approval by the Chief Executive Officer (CEO) at MDH;
    iii. Approval of the research proposal by the clinical chairperson of the Oncology and Haematology Department at Sir Anthony Mamo Oncology Centre (SAMOC) and MDH.

2.2 Part I: Trends in Leukaemia Clinical Trial Efficacy Studies

2.2.1 Identification of Clinical Trials

The European Union (EU) Clinical Trials Register[20] (EudraCT), an online repository of interventional clinical trials that are conducted in the EU and European Economic Area (EEA), was accessed to retrieve clinical trials investigating anti-leukaemic therapy. The specific clinical trial phases and the period when the trials were first entered into the EudraCT by a national competent authority (or a third country data provider in case of clinical trials conducted outside the EU/EEA) were defined in the search strategy. An eleven-year date range, spanning from 1 January 2007 to 31 December 2017, clinical trial Phases II, III and IV and the key term 'leukaemia' were inputted as search criteria. The definitions of each clinical trial phase are found in Table 2.1.

Table 2.1 Definitions of clinical trial phases II-IV[21]

Clinical trial phase	Definition
Phase II	Phase II aims to investigate the safety and effectiveness of a potential therapy. Usually between 100 and 300 people will be enlisted to take part with the aim of determining whether the treatment will be safe and effective to treat a condition.
Phase III	If previous trials have indicated a treatment is safe and that it also shows promise in being able to treat a condition, phase III clinical trials begin. These involve large numbers of participants, usually from several hundred to several thousand subjects, and are often spread between different hospitals and countries. If these trials show that a drug is safe and effective, the manufacturers can apply for a marketing authorisation.
Phase IV	Post-marketing studies to delineate additional information including the drug's risks, benefits, and optimal use. These studies are designed to monitor effectiveness of the approved intervention in the general population and to collect information about any adverse effects associated with widespread use.

[20] EU Clinical Trials Register [Internet]. [cited 2019 Apr 17]. Accessed from: https://www.clinicaltrialsregister.eu

[21] European Medicines Agency (EMA). How to search the EU Clinical Trials Register [Internet]. April 2014 [cited 2019 Apr 17]. Accessed from: https://www.clinicaltrialsregister.eu/doc/How_to_Search_EU_CTR.pdf

Results generated from the EudraCT were exported using the full trial details option so as to capture the entire record of each study. The raw data content was converted to tabular format in spreadsheet.

2.2.2 Clinical Trials Selection Process

The eligibility criteria for clinical trials that were included in this analysis are specified in Figure 2.1. To avoid limiting the scope of outcomes identified, a study design or age filter was not applied. Clinical trials were screened for the following main characteristics:

- *Investigational medicinal product (IMP)* – IMP consists of a chemical, biological or biotechnological agent;
- *International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) Medical Dictionary for Regulatory Activities (MedDRA) classification of medical condition(s) or disease(s) under investigation* – clinical indication under investigation is a variant of leukaemia;
- *Scope of trial* – efficacy study;
- *Selection of endpoints* – primary and/or secondary efficacy endpoints were measured.

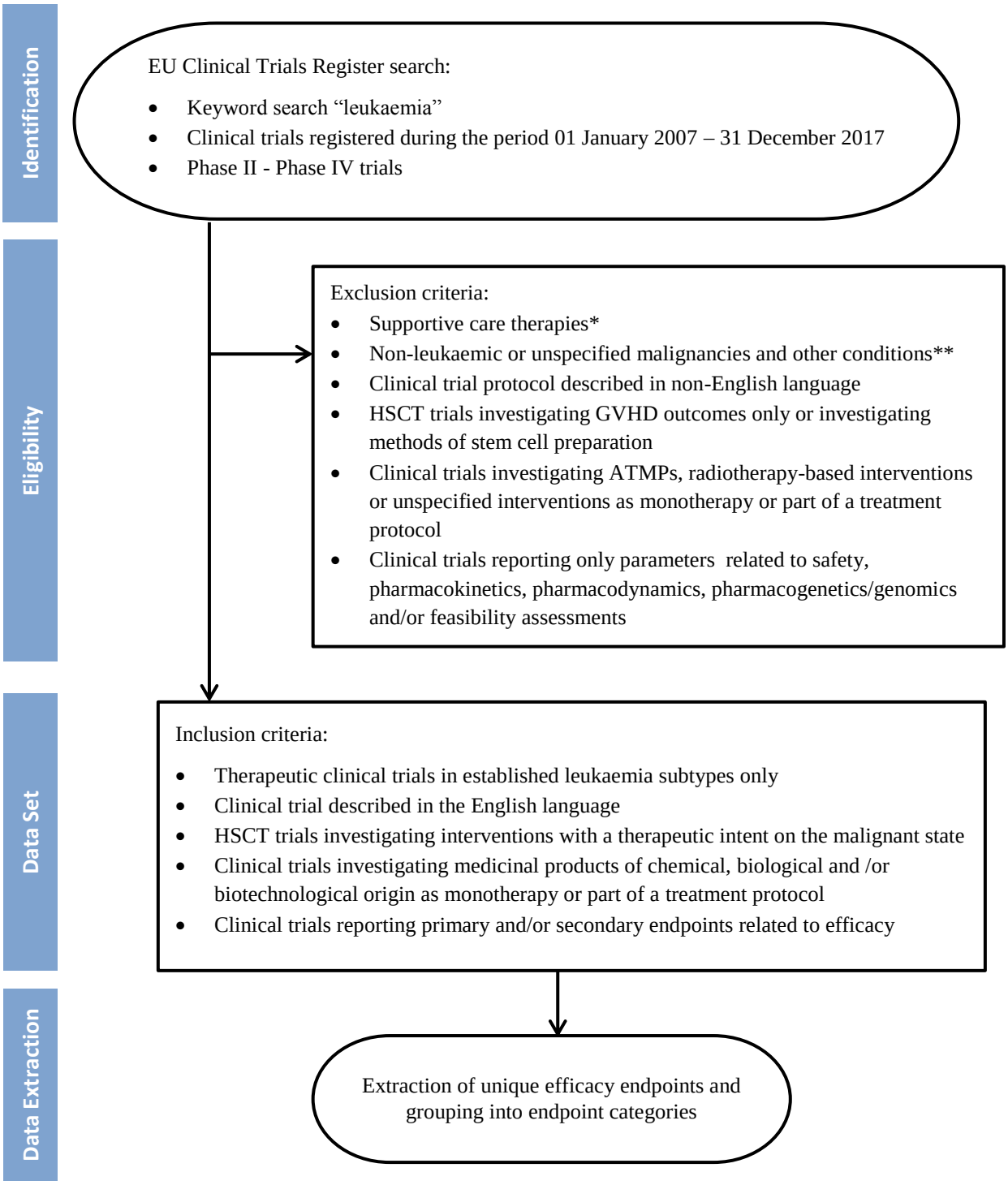


Figure 2.1 Clinical trials selection process and extraction of efficacy data

*Bacterial, fungal and viral infections, mucositis, anaemia, neutropenia, thrombocytopenia, nausea and vomiting, neurotoxicity, ototoxicity, coagulopathy, hormonal deficiency, reproductive toxicity and other iatrogenic effects, vaccination, adoptive immunotherapy

**Solid tumours, myelodysplastic syndromes (MDS), certain myeloproliferative disorders (MPD) (including myelofibrosis, polycythaemia vera, essential thrombocythaemia), lymphoma, multiple myeloma, amyloidosis

Acronyms: Advanced therapy medicinal product (ATMP); European Union (EU); Graft versus host disease (GVHD); Haematopoietic stem cell transplantation (HSCT)

2.2.3 Extraction and Categorisation of Efficacy Data

Primary and secondary endpoints reported in the EudraCT clinical trial protocol section on efficacy were reviewed for each selected trial. Efficacy endpoints were extracted and filtered for duplication to ensure uniqueness of each endpoint. Literature, including established institutional guidelines and published manuscripts relating to response assessment in leukaemic disease states, was accessed to differentiate between the identified endpoints. Variants of the same endpoint were considered as independent measures during compilation; for example, complete response rate, complete response rate with incomplete blood count recovery and complete response rate with incomplete platelet recovery were recorded as three distinct measures. Unique endpoints were then grouped into principal categories according to their type of measurement.

2.2.4 Data Mining of Trends

The selected clinical trials were given a study identification code, segregated according to year of registration in EudraCT and cross-tabulated against the compiled list of endpoints. Each endpoint could have been studied at a primary level, secondary level or both (in cases of subgroup or interim analyses) and this was reflected in a coding system adopted.

To understand the evolvement of efficacy assessment over time, a data mining exercise was performed on the data set by studying trends in the selection of endpoints. Descriptive statistics in the form of percentage frequency graphs and distribution time curves were applied to analyse patterns at an individual endpoint and category level over the eleven-year period studied. The difference of two proportions inferential test was operated to detect statistically significant changes in the choice of primary efficacy endpoints selected between set time frames.

2.3 Part II: Scientific Expert Opinions on Clinical Trial Design in Leukaemia

Twenty-two published study protocols (Table 2.2) applying the Delphi technique in the development of a COS for diverse clinical specialties were identified from PubMed and the Core Outcome Measures in Effectiveness Trials (COMET)[22] databases and reviewed. The review process was aimed at facilitating planning aspects and ensuring rigour in the methodological framework.

In addition to published study protocols, recommendations given by the following sources were reviewed for methodological guidance:

- The COMET handbook version 1.0 and four-step process in COS development (Williamson et al, 2017);
- The Outcome Measures in Rheumatology (OMERACT) filter 2.0 conceptual framework (Boers et al, 2014);
- Guidelines provided by Sinha and colleagues (Sinha et al, 2011) on the conduct of studies to elucidate core outcomes to measure in clinical trials and
- The Core Outcome Set – Standards for Development (COS-STAD) (Kirkham et al, 2017) and Core Outcome Set – Standards for Reporting (COS-STAR) (Kirkham et al, 2016) documents.

The protocol for this study was modelled using the above mentioned resources and tailored to the context of this research. A stepwise approach is described under this section outlining the methods undertaken to measure the validity, reliability and feasibility of a tool developed as part of the study protocol.

[22] Core Outcome Measures in Effectiveness Trials (COMET). The COMET database [Internet]. [cited 2019 Apr 10]. Accessed from: http://www.comet-initiative.org/studies/search/

Table 2.2 Core outcome set development protocols for specific medical conditions that were reviewed during methodological design

Medical condition	Citation
Cardiovascular disease	
Lower limb amputation for peripheral arterial disease	Ambler et al, 2017
Dermatology	
Eczema	Schmitt et al, 2015
Rosacea	Iyengar et al, 2016
Incontinence-associated dermatitis	Van den Bussche et al, 2017
Endocrinology	
Type 2 diabetes mellitus	Harman et al, 2018
Gastroenterology	
Inflammatory bowel disease	Ma et al, 2017
Haemorrhoidal disease	Van Tol et al, 2017
Malignant disease	
Oropharyngeal cancer	Waters et al, 2014
Localised prostatic cancer	MacLennan et al, 2015
Anal cancer	Fish et al, 2017
Basal cell carcinoma	Schlessinger et al, 2017a
Squamous cell carcinoma	Schlessinger et al, 2017b
Nephrology	
Haemodialysis	Tong et al, 2015
Polycystic kidney disease	Cho et al, 2017
Gynaecology	
Endometriosis	Hirsch et al, 2016
Ophthalmology	
Posterior segment-involving uveitis	Tallouzi et al, 2017
Paediatrics	
Childhood epilepsy	Morris et al, 2017
Acute uncomplicated appendicitis in children	Sherratt et al, 2017
Pain management	
Non-specific lower back pain	Chiarotto et al, 2014
Shoulder pain	Gagnier et al, 2017
Other disease states	
Otitis media with effusion in cleft palate	Harman et al, 2013
Critical illness	Connolly et al, 2018

2.3.1 Design of the Response Evaluation in Leukaemia (ReVALeu) Tool

2.3.1.1 Sections and Components of the Tool

The Response Evaluation in Leukaemia (ReVALeu) is an online surveying tool (Appendix 2) consisting of two sections which was developed to gauge the opinions of regulatory and HTA experts on generic and technical aspects of oncology clinical trials, mostly related to leukaemia. An introductory page explaining the research objectives and implications of findings, together with participant instructions, preceded the questionnaire items. The first section of the tool carried a broad scope where participants were asked for their perspectives on clinical assessments and evidence generation in the therapeutic fields of oncology and haematology. Weighted agreement, frequency and quality rating scales and rank-type questions were used in this section.

The second section constituted the predominant part of the tool and tackled technical questions on efficacy endpoints studied in clinical trials for medicines in leukaemia. The identified list of endpoints was presented according to the categories specified under section 3.1.2. Three multiple choice fields were asked for each unique efficacy measure (Table 2.3). The optional use of free text entry in comment boxes at the end of each set of statements allowed respondents to justify the ratings given.

Table 2.3 Multiple choice fields for section two questionnaire items of the ReVALeu tool

Multiple choice field	Options for selection	Type of variable
Classification of endpoint	Primary or secondary	Nominal, categorical
Clinical trial phase-specificity for endpoint to be studied	Phase II, Phase III and/or Phase IV	Nominal, categorical
Importance of endpoint in demonstrating efficacy	5-point likert scale of weighted importance ratings, ranging from 'Not important at all' (weighting of 1) to 'Very important' (weighting of 5)	Ordinal, categorical

The tool consisted of 124 questionnaire items grouped under 21 main questions: Section 1 carried 9 questions (4 HTA-specific and 5 common to both groups of decision-makers) and Section 2 featured 12 questions (3 questions for each endpoint category, common to both panels). The potential for 'opinion-molding' was avoided by formulating questions and statements in an objective and non-leading manner.

The ReVALeu tool was constructed to incorporate both conventional survey and Delphi components. Only section 2 items of the tool requesting the importance ratings for endpoints in demonstrating efficacy were selected to undergo the Delphi process and this was communicated with participants in the preface. The remaining questions and statements were included as one-time response survey items to gain deeper insight.

2.3.1.2 Features of the Online Surveying Platform

Web-based technologies have been broadly cited in literature as means to facilitate the conduct of surveying research methods (Donohoe et al, 2012; Holloway, 2012; Helms et al, 2017). Given its advanced interface and global recognition, SurveyMonkey was selected as the online platform for the design, piloting and dissemination of the ReVALeu tool and for tracking and exporting responses.

The platform boasts several features that offer convenient and efficient data collation, translating to an enhanced survey experience for the researcher and participants. The use of custom variables as a form of embedded data provided the possibility to individualise survey profiles through personalised web links which correlated survey responses to participants. This information in conjunction with the internet protocol (IP) address generated as a metadata element enabled participant follow-up in attempt to amplify the response rate. Rather than requiring participants to skip questions manually, the skip logic function permitted adaptive questioning by showing questions that are

only relevant to the expert group, therefore customising the process. This function was applied to the first section of the ReVALeu tool where certain question sets for HTA and regulatory experts differed. Paradata refers to data generated passively by the respondent during survey completion, such as survey completion time, which was used to inform same round or subsequent participants on the expected time to complete the respective questionnaire rounds. The platform allowed responses to be exported in data files compatible with data management and statistical software packages (SPSS and Excel) for subsequent manipulation.

2.3.1.3 Appendix of Definitions

An exhaustive list of online published sources relating to response assessment in oncology and haematology was reviewed to compile a table of definitions for each efficacy endpoint (Appendix 3) to assist the respondent in distinguishing between the identified efficacy endpoints for the four main leukaemia subtypes: Acute Lymphoblastic Leukaemia (ALL), Acute Myeloid Leukaemia (AML), Chronic Lymphocytic Leukaemia (CLL) and Chronic Myeloid Leukaemia (CML). This guidance document was developed to be used in parallel with the ReVALeu online surveying tool during its completion. More than one definition for the same endpoint was provided if definitions from different sources were considered to contrast.

2.3.2 Psychometric Evaluation of the Tool

The following section outlines the steps undertaken to conduct a comprehensive psychometric evaluation of the ReVALeu tool to ensure its validity, reliability and ease of completion.

2.3.2.1 Selection of the Validation Panel Members

On the premise that a reactive type of Delphi technique was planned to be executed, it was not possible for the tool content to be validated across successive rounds. Validation was conducted a priori to the start of the Delphi process. A multi-disciplinary validation panel was assembled composed of the following eight members:

- Two resident consultants in paediatric and adolescent oncology at MDH and SAMOC, Malta;
- One resident consultant in haematology at MDH, Malta;
- One resident specialist in haematology at MDH, Malta;
- A joint review from two health technology assessors of oncology and haematology medicinal products at the Directorate for Pharmaceutical Affairs (DPA), Malta
- One pharmaceutical and HTA policy pharmacist at the DPA, Malta
- One regulatory clinical assessor at the Malta Medicines Authority, Malta and alternate member of the EMA Committee for Medicinal Products for Human Use (CHMP) Scientific Advice Working Party (SAWP)
- One head of the Life Science Informatics Laboratory, Oncology Department, Mario Negri Institute for Pharmacological Research, Italy

The validation members were purposively selected to comprehensively cover the expertise required in validating the sections of the ReVALeu tool. Validation responses were collected electronically or through face-to-face interviews when requested. The number of participating experts falls within the prescribed range of three to ten as recommended by Lynn (1986) for the content validity index (CVI) method, the validation method employed, to provide a sufficient level of control for chance agreement.

2.3.2.2 Development of the Expert Validation Template

An expert validation template (Appendix 4) was developed to test the questionnaire items and the overall tool (including the appendix of definitions) for the three domains of relevance, clarity, and structure and layout. Four-point likert scales were used as advocated by Lynn (1986) to avoid having an ambivalent midpoint. A rating of '1' and '4' indicate the lowest and highest scores respectively for the mentioned validation domains. For items receiving a rating of '1' or '2', the expert was requested to justify the score in the comments field, proposing improvements or omission. Besides its purpose to challenge the content of the ReVALeu tool, the template was developed to ensure user-friendliness in completion and readability of the tool. The validation template was provided to the panel as both an electronic fillable form and print version.

2.3.2.3 The Content Validity Index (CVI) Method

The CVI method was developed by Lynn (1986) and is a widely recognised rigorous scale development procedure to measure the content validity of an instrument (Polit and Beck, 2006). Content validity is described as "the degree to which a sample of items, taken together, constitute an adequate operational definition of a construct" (Polit and Beck, 2006). Two indices at an item (I-CVI) and scale (S-CVI) level were computed in the validation process (Appendix 5). The I-CVI represents the proportion of members from the panel rating items for a domain as '3' or '4'. The I-CVI threshold for revision or omission was taken as 88% at a 0.05 level of significance using the matrix in Figure 2.2 proposed by Lynn (1986). The S-CVI was calculated by averaging the I-CVI values by the number of items for each validation domain and adopting a 90% acceptance criterion (Waltz et al, 2005). Items that failed to achieve the specified threshold were considered for revision or omission (Appendix 6).

Number of Experts	Number of Experts Endorsing Item or Instrument as Content Valid								
	2	3	4	5	6	7	8	9	10
2	1.00								
3	.67	1.00							
4	.50	.75	1.00						
5	.40	.60	.80	1.00					
6	.33	.50	.67	.83	1.00				
7	.29	.43	.57	.71	.86	1.00			
8	.25	.38	.50	.63	.75	.88	1.00		
9	.22	.33	.44	.56	.67	.78	.89	1.00	
10	.20	.30	.40	.50	.60	.70	.80	.90	1.00

Figure 2.2 Proportion of experts (above the line) whose endorsement is required to establish content validity beyond the 0.05 level of significance

Reproduced from: Lynn MR. Determination and quantification of content validity. Nurs Res. 1986;35(6):382-385.

2.3.2.4 Reliability Testing and Piloting

The intra-subject reliability of the tool was confirmed using the test-retest approach with 3 HTA experts and 2 regulatory clinical assessors, allowing a two-week window between the first and second questionnaire administrations. Statistical tests operated included the Kendall Tau test for ordinal scales and Kappa values for nominal scales using IBM SPSS® Statistics version 23. Pilot testing was necessary to assess the feasibility and ease of completion of the online tool. The tool was piloted by the same five experts on the most recent versions of common web browsers using desktop, tablet and smartphone devices.

2.3.3 Eligibility and Recruitment of Study Participants

Main considerations in the expert selection for the Delphi study included the sample size, level of expertise and the homogeneity of the panel. The main success driver for a Delphi technique is underpinned by the pooled expertise of the panellists. There is no universal agreement on a sampling method in Delphi processes and a pragmatic approach was taken in line with the context of this research. Representativeness in the Delphi is governed by the qualities of the participants making up the expert panel and does not require the sample to be statistically representative of the study population (Hasson et al, 2000; Powell, 2003; Du Plessis and Human, 2007). The sample of experts

required is dependent on the scope of the research objective and available resources. Multiple sources suggest that a panel of 10 to 15 subjects could be sufficient if the background of the Delphi subjects is homogeneous (Delbecq et al, 1975; Linstone and Turoff, 1975; Skulmoski et al, 2007). Larger panels have been significantly associated with lower response rates and greater potential for attrition bias in Delphi studies (Gargon et al, 2019).

Experts were invited to participate using a non-probability, purposive sampling approach. Subjects were eligible to participate if they are involved in health technology and regulatory clinical assessments for onco-haematology medicinal products, either directly or through expert scientific advice. Contact details for HTABs in the EU were retrieved from online platforms of the EUnetHTA[23], World Health Organization Regional Office for Europe (WHO/Europe) Health Evidence Network (HEN)[24] and International Network of Agencies for Health Technology Assessment (INAHTA)[25]. HTABs were asked to identify possible candidates as experts to the Delphi panel based on the eligibility criteria previously outlined.

[23] European Network for Health Technology Assessment (EUnetHTA). EUnetHTA Partner Organisations and Institutions [Internet]. [cited 2019 Apr 15]. Available from: https://www.eunethta.eu/about-eunethta/eunethtanetwork/

[24] World Health Organization (WHO) Europe. Health Evidence Network (HEN) Current Technical Members [Internet]. [cited 2019 Apr 15] Available from: http://www.euro.who.int/en/data-and-evidence/evidence-informed-policy-making/health-evidence-network-hen/technical-members/current-technical-members

[25] International Network of Agencies for Health Technology Assessment (INAHTA). INAHTA Members List [Internet]. [cited 2019 Apr 15]. Available from: http://www.inahta.org/members/members_list/

EMA CHMP and SAWP members and alternates and clinical experts on blood cancers were identified from the EMA experts database[26] and constituted the regulatory scientific personnel invited to the Delphi. A recruitment letter was disseminated to potential participants or their institutions detailing the objectives, scope and methodology of the research and their expected input (Appendix 7). Only experts that expressed willingness and interest to participate were recruited in the first Delphi round.

Two independent Delphi panels were assembled, one for each of the HTA and regulatory expert groups. Sinha et al (2011) comments on the benefits of multiple Delphi panels since potentially conflicting views that are included in the same panel may be diluted and underrepresented in the final consensus. Separate expert panels also facilitate the detection of divergent opinions, therefore allowing comparative analyses between the results reached in the respective groups. This viewpoint is also supported by Williamson and colleagues (2012) stating that "separate panels for different stakeholder groups followed by work to integrate the multiple perspectives may be more appropriate."

2.3.4 Anonymity of Participants

An encoding system was used when exporting data into statistical software packages, with each code corresponding to the expert identification details, enabling follow-up. This study is considered as pseudo-anonymous when it comes to identification of participants with the research team since the principal investigator was able to link the participants to their survey responses. Sinha et al (2011) describes full anonymity of a Delphi process when "participants should not know the identities of the other

[26] European Medicines Agency (EMA). The European Experts Database [Internet]. [cited 2019 Apr 15]. Available from: https://www.ema.europa.eu/about-us/how-we-work/european-medicines-regulatory-network/european-experts/az

individuals in the group, nor should they know the specific answers that any other individual gave." Full anonymity was sustained among the panellists as experts could not identify the participants' identities and their corresponding responses. The web-survey software of choice additionally incorporates a number of physical and electronic security measures to safeguard data protection.[27]

2.3.5 Data Collection

The Delphi technique is an established consensus-building method employed in decision-making. Through a series of iterative rounds and controlled feedback, opinions from a group of experts are harnessed and aligned. Predicated on the notion that "two heads are better than one, or…n heads are better than one" (Dalkey, 1972), the Delphi approach conceptually pushes for a majority-based perspective on the specific topic under study. Uncertainty and considerable variability is documented for features of the Delphi such as the definition of consensus, number of rounds and acceptable response rate (Boulkedid et al, 2011; Sinha et al, 2011).

2.3.5.1 Selection of the Consensus Method

The Delphi technique was selected over other consensus methods such as the nominal group technique (NGT), focus groups and face-to-face interviews, since it offers superiority for prioritisation of complex and technical parameters among experts (McMillan et al, 2016). The Delphi approach also allows for a virtual and anonymous rating system which is not possible with other consensus methods having face-to-face interaction as a fundamental criterion throughout most of the process. A modified, electronic variant of the Delphi technique (e-Delphi) is conducted in this study

[27] SurveyMonkey. SurveyMonkey Security Statement [Internet]. [cited 2019 Apr 18]. Accessed from: https://www.surveymonkey.com/mp/legal/security/

according to the definitions provided by Keeney and colleagues (2011) since the first round items were pre-determined from clinical trial data and less iterations than the classical three-round Delphi were planned. The electronic execution of the Delphi process, by means of an online surveying platform and emailing system, is reported to have the additional advantages of wider outreach to participants from diverse geographical locations, shorter turnaround times, augmented response rates and improved data quality (Gill et al, 2013; Helms et al, 2017).

2.3.5.2 Number of Rounds and Interval Period

The factors that determined the number of iterations included homogeneity of the expert groups, participant fatigue, resources and the type of Delphi selected. A two-round e-Delphi process was conducted for both panels and this was justified by the homogeneity of expertise in each group (Skulmoski et al, 2007) and with the intention of abating the risk of attrition bias resulting from dropouts (McMillan et al, 2016).

A fixed interval period between successive rounds was assigned in pursue of keeping the expert panel members on track. A maximum four-week interval between Delphi rounds was adhered to; three weeks for the collation of responses, as implemented in published COS development Delphi studies (Harman et al, 2013; Chiarotto et al, 2014; MacLennon et al, 2015; Iyengar et al, 2016; Ma et al, 2017; Schlessinger et al, 2017(a); Schlessinger et al, 2017 (b); Sherratt et al, 2017; Connolly et al, 2018) and a further week to analyse responses, compile the group feedback and create the second round questionnaire. Up to three follow-up reminder emails were sent to non-responders per round.

2.3.5.3 Rating Scale and Consensus Analysis

A five-point likert scale of weighted importance ratings was used for the Delphi items. This type of scale has been implemented by previous Delphi studies (Smail-Faugeron et al, 2013; Giannarou and Zervas, 2014) since it allows the respondent to unambiguously differentiate the level of importance between one point on the scale and another.

The Delphi consensus threshold has been regarded as "one of the most contentious components of the Delphi method" (Von der Gracht, 2012). The mean as a measure of central tendency or metrics of dispersion (range, standard deviation, interquartile range and coefficient of variation) should be considered as stopping criteria for consensus when purely quantitative scales are used. Questionnaire items being tested in this Delphi consist of likert scales capturing ordinal-type data, therefore percentage agreement falling within a range on a scale was the selected criteria for consensus. The mean and standard deviation were reserved to assess stability of responses across rounds as described in section 2.3.6.3. The consensus threshold was determined a priori to avoid any post-hoc bias of results. The rating scale was dichotomised into consensus achieved or not and a two-sided consensus was possible on the scale. Consensus for inclusion of the endpoint as an important outcome was reached if greater than or equal to 75% of the experts have rated statements as "Important" or "Very important". Similarly, consensus for exclusion of an endpoint as an important parameter was achieved if a frequency of greater than or equal to 75% of the responses selected "Not important at all" or "Not important" on the scale. The consensus level was set at 75% based on findings from a systematic review that investigated consensus measurement in 98 Delphi studies (Diamond et al, 2014). The majority of these studies (n=25) used percent agreement to determine consensus, with 75% being the median threshold to define consensus. This consensus threshold has also been recommended by Keeney and colleagues (2006).

2.3.5.4 Targeted Response Rate and Strategies to Reduce Attrition

A 70% response rate is required to maintain rigour (Keeney et al, 2011) and strategies were followed to reduce panel fatigue and retain participants across rounds. In order to engage the experts, a personalised approach was used to create buy-in and a clear explanation of the e-Delphi process was communicated. A string of reminders was sent to non-responders emphasising the value of their contributions in the research outcomes. Gargon et al (2019) report that the number of questionnaire items is inversely related to response rates in Delphi studies. The risk of participant fatigue was further reduced by carrying forward to the subsequent round only items that failed to reach consensus.

Delphi experts are expected to have appropriate knowledge about the topic of concern (McMillan et al, 2016). Mismatches between the nature or level of knowledge being requested in questionnaire items and the expertise of panellists are considered as a contributing factor to attrition. This risk was offset by developing and validating a tool that matches the expertise sought for, to produce core outcome data in leukaemic disease that is of quality and reliable.

2.3.5.5 Delphi Round One

The full set of questions in the ReVALeu tool, that is both Delphi and conventional survey items, was provided to experts recruited in the first round. The type of Delphi that was operated is termed as a 'reactive Delphi' since panel members in the first survey round were asked to rate pre-determined statements generated from clinical trial efficacy data, rather than providing responses to open questions (Gill et al, 2013). The e-Delphi first round responses were collected in September 2018 for HTA experts and between October and November 2018 for regulatory representatives.

2.3.5.6 Delphi Round Two

A new survey was developed for the second Delphi round retracting all the questionnaire items except for scales of importance ratings for items that have not reached consensus in the first round (Appendix 8). Only first round panellists submitting complete responses were invited to the second round. Feedback between rounds has been described as "an essential component of the Delphi procedure" (Boulkedid et al, 2011). Between the first and second iterations, quantitative and qualitative data on importance ratings given by the panel members was presented to the respondents completing the preceding round (Appendix 9). Quantitative data included the weighted mean rating and median as a measure of central tendency, standard deviation as an indication of disagreement and frequency distribution charts. The qualitative component comprised a summary of comments provided by the panel.

A hallmark feature of a successful Delphi is reminding participants of their individual scores selected in the previous round (Powell, 2003; Keeney et al, 2006). In promoting the intended converging effect of the Delphi, the experts were given visibility to their previous round responses by providing them with their individual ratings profile. This enabled the participants to revisit their first round positions relative to the common group opinions.

The e-Delphi second round survey data was collected in October 2018 for HTA experts and in November 2018 for regulatory representatives. An overview of the Part II data collection framework is presented in Figure 2.3. A letter of participation was sent to the panel experts at the end of the Delphi process (Appendix 10) together with a detailed report highlighting the main findings.

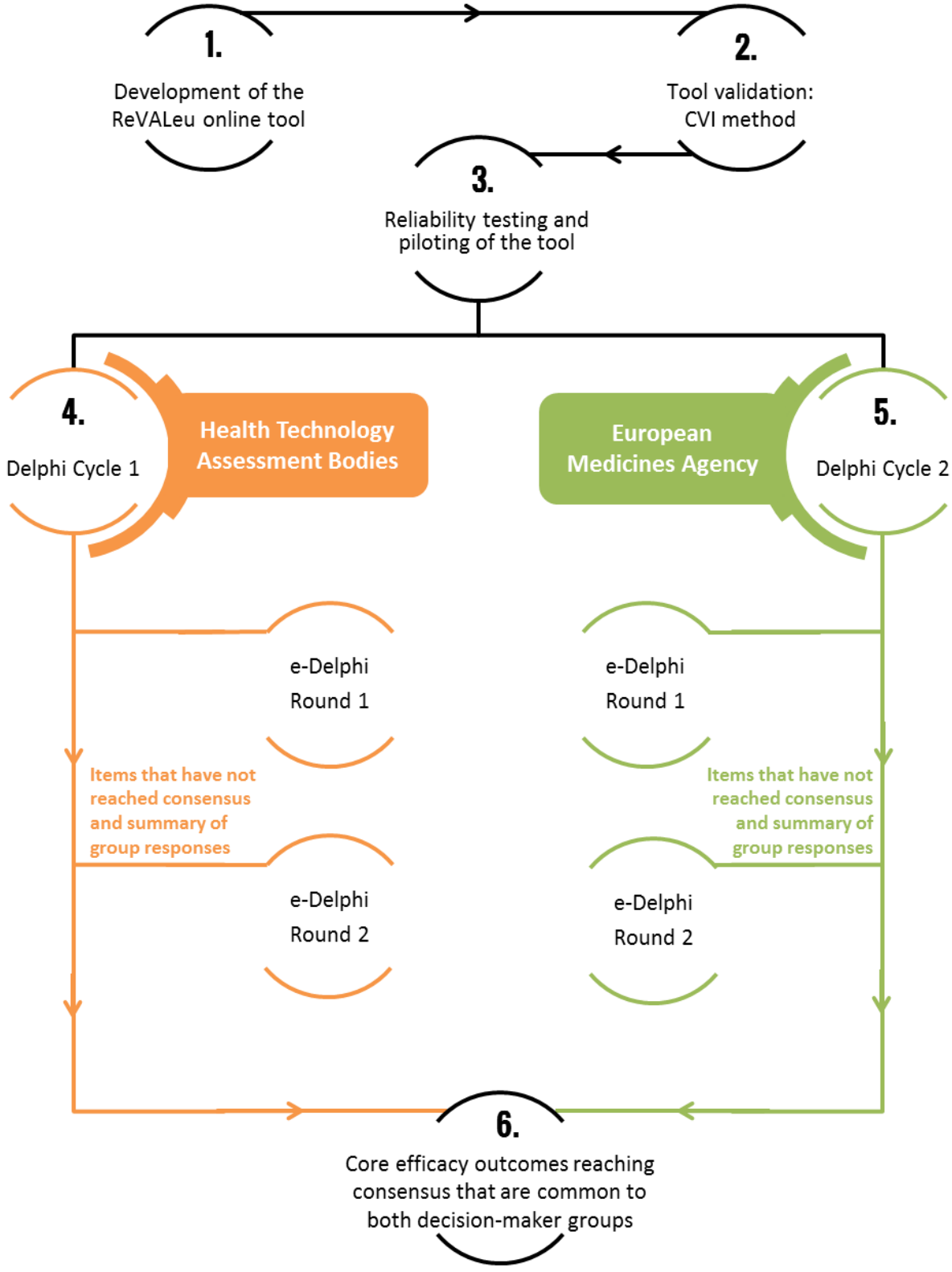


Figure 2.3 Flow diagram depicting the data collection framework for Part II of the study

Acronyms: Clinical Trial (CT); Content Validity Index (CVI); European Union (EU); Health Technology Assessment (HTA); Response Evaluation in Leukaemia (ReVALeu)

2.3.6 Data Processing

This section provides an overview of the statistical and qualitative methods adopted to analyse the expert responses across both Delphi rounds and explains how the inter-rater stability testing was conducted.

2.3.6.1 HTA and Regulatory Opinions

Inferential statistics were operated for non-Delphi items to determine statistically significant differences between HTA and regulatory opinions. Weighted mean ratings for ordinal scales were compared between the two expert panels and analysed using the non-parametric Mann-Whitney U Test in IBM SPSS® Statistics version 23. A statistically significant difference in opinion was obtained for statements if a p-value was found to be less than the 0.05 level of significance.

Descriptive statistics were reported for both Delphi and non-Delphi items. 'Don't know' answers in Section 1 of the ReVALeu tool were not anchored to a weighting and excluded from descriptive and inferential analyses.

Free text comments inputted by regulatory and HTA experts in both Delphi rounds were qualitatively explored using a thematic analysis approach. Trending arguments were identified by grouping the comments into overarching themes, with features for each theme being synthesised.

2.3.6.2 Determination of the Regulatory-HTA Core Efficacy Outcomes

The efficacy elements that reached consensus after the two rounds of both Delphi cycles were compared and endpoints common to both decision-maker groups were identified as the regulatory-HTA core outcomes for efficacy studies in leukaemia clinical trials. The resulting list of regulatory-HTA core efficacy outcomes were assessed against the

ideal core outcome areas of death, life impact, resource use and pathophysiological manifestations that should be represented in every COS set as recommended by Boers et al (2014).

2.3.6.3 Inter-rater Stability Testing

Stability is defined as "the consistency of responses between successive rounds of a study" (Dajani et al, 1979). Behaviour in responses among panel experts is reported to be complex and unpredictable across iterative rounds (Greatorax and Dexter, 2000). Analysis of response stability goes beyond the mere determination of items reaching consensus at the terminal Delphi round since it seeks to understand the process leading to the consensus reached in a Delphi study. The stability of responses was quantified for items not reaching consensus in the first Delphi round and featuring in the second. To assess the stability pattern for inter-rater opinion and level of agreement, the mean as a measure of the aggregate judgement and the standard deviation to determine the degree of dispersion around the central opinion were calculated for both rounds and compared.

Chapter 3

Results

The results chapter reports on the following:

- Characteristics of clinical trials investigating medicines in leukaemia that were included in the study;
- The identification and evolvement of endpoints used in efficacy studies as part of leukaemia clinical trial protocols;
- Perspectives of regulatory and HTA experts on the quality of evidence generated and alignment of clinical assessments for antineoplastic agents and
- Technical opinions of decision-makers on the design of efficacy studies in leukaemia clinical trials.

3.1 Part I: Trends in Leukaemia Clinical Trial Efficacy Studies

3.1.1 Characteristics of Clinical Trials

The EU Clinical Trials Register search generated 666 trials with 235 trials being omitted on the basis of exclusion criteria described. The screening process leading to the data set is outlined in Figure 3.1. Four hundred and thirty-one (431) leukaemia clinical trials were eligible in the study, comprising a total of nearly 110,000 patients. Peak trial registration was observed in the years 2009 and 2012 with 51 and 58 trials registered respectively (Figure 3.2). Acute lymphoblastic leukaemia trials recruited the highest percentage of paediatric patients (46%, n=31) whilst trials investigating chronic lymphocytic leukaemia enrolled patients exclusively from the adult and elderly populations (Table 3.1). Only 1% (n=4) of the selected trials were conducted in investigator sites outside the EU/EEA and were included if they formed part of a paediatric investigation plan (PIP) or were sponsored by a MAH and involved the use of a medicinal product covered by an EU MA for paediatric use.

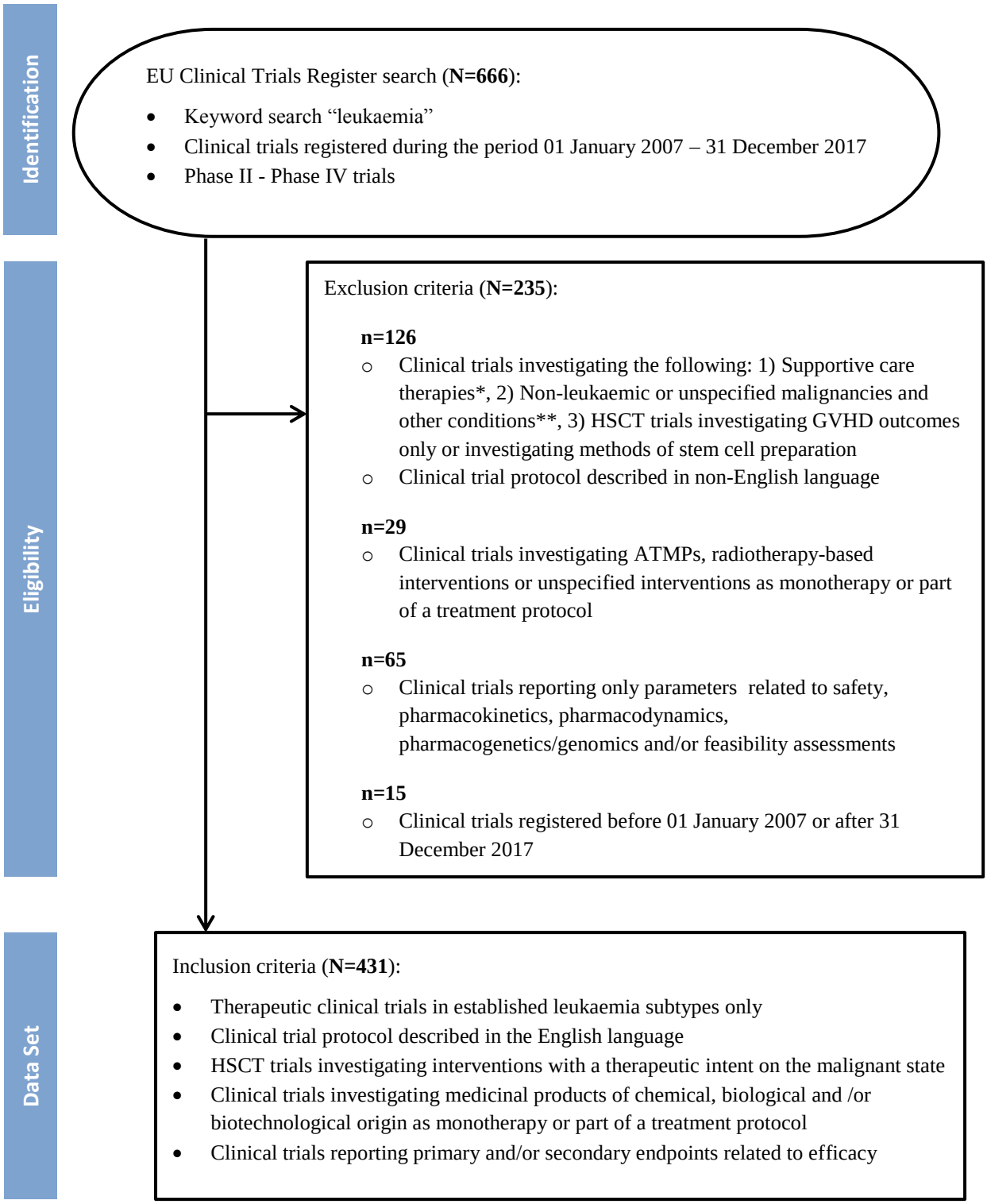


Figure 3.1 Data set of clinical trials included in the study

*Bacterial, fungal and viral infections, mucositis, anaemia, neutropenia, thrombocytopenia, nausea and vomiting, neurotoxicity, ototoxicity, coagulopathy, hormonal deficiency, reproductive toxicity and other iatrogenic effects, vaccination, adoptive immunotherapy

**Solid tumours, myelodysplastic syndromes (MDS), certain myeloproliferative disorders (MPD) (including myelofibrosis, polycythaemia vera, essential thrombocythaemia), lymphoma, multiple myeloma, amyloidosis

Acronyms: Advanced therapy medicinal product (ATMP); European Union (EU); Graft versus host disease (GVHD); Haematopoietic stem cell transplantation (HSCT)

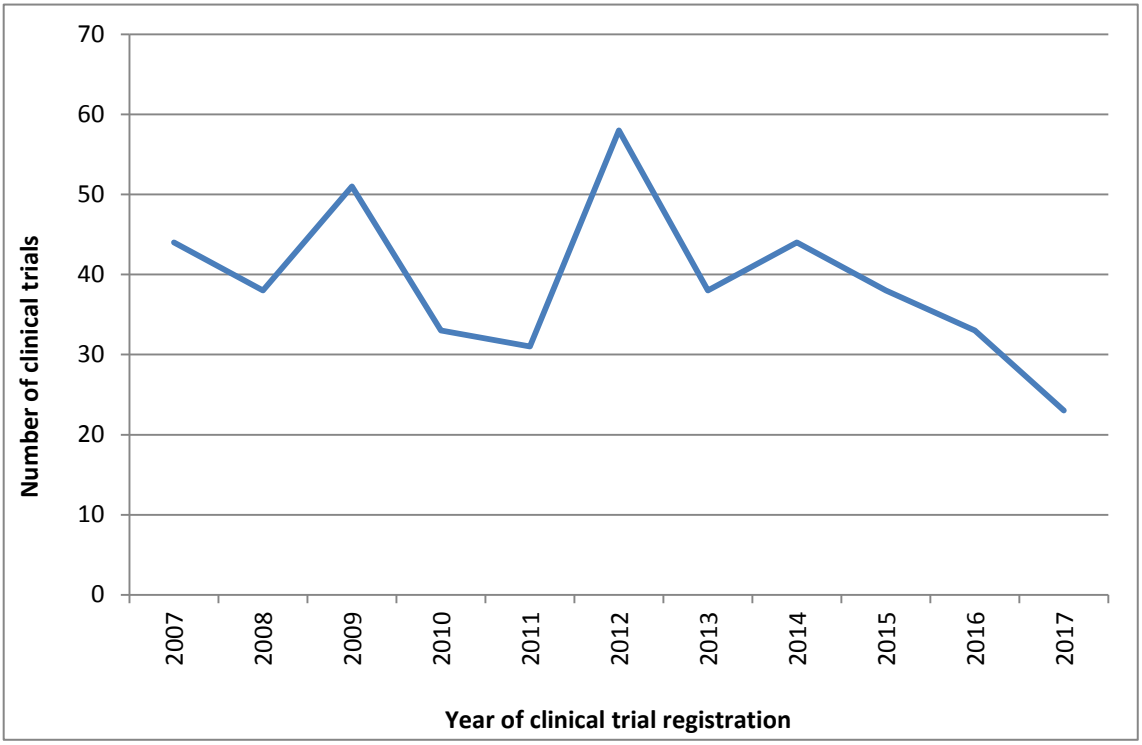


Figure 3.2 Number of clinical trials registered in EudraCT by year, 2007-2017 (N=431)

Table 3.1 Number and characteristics of patients recruited in the clinical trials (N=431)

Clinical indication	Total number of subjects planned to be included in the clinical trials	Trial subjects under 18 (%)	Trial subjects adults (18-64 years) (%)	Trial subjects elderly (>=65 years) (%)
All leukaemia subtypes	108,633*	15	86	77
Acute lymphoblastic leukaemia (ALL)	19,709	46	69	38
Acute myeloid leukaemia (AML)	46,914	11	83	76
Chronic lymphocytic leukaemia (CLL)	20,114	0	98	93
Chronic myeloid leukaemia (CML)	16,206	7	93	90
Other leukaemia subtypes	6,031	28	79	77

*Total number of patients does not tally since more than one leukaemia subtype may be studied by the same clinical trial

Table 3.2 shows the frequency of leukaemia subtypes investigated by trials included in this study, stratified according to the WHO 2016 classification systems for lymphoid and myeloid neoplasms (Arber et al, 2016; Swerdlow et al, 2016). The most commonly studied leukaemia variants were acute myeloid leukaemia (37.4%, n=161) and chronic lymphocytic leukaemia (23.7%, n=102), collectively accounting for more than 60% of the trials conducted.

Table 3.2 Frequency of clinical trials investigating leukaemia subtypes using the World Health Organization (WHO) 2016 classification systems for lymphoid and myeloid neoplasms (N=431)

Leukaemia subtype	Frequency of clinical trials (n)*	Percentage from total clinical trials (%)*
Acute myeloid leukaemia (AML) and related neoplasms		
AML unspecified	161	37.4
Acute promyelocytic leukaemia (APL)	3	0.7
β-lymphoblastic leukaemia/T-lymphoblastic leukaemia/Acute leukaemia of ambiguous lineage		
Acute lymphoblastic leukaemia (ALL) unspecified	68	15.8
β-cell type acute leukaemia	3	0.7
Myeloproliferative neoplasms (MPN)		
Chronic myeloid leukaemia (CML)	67	15.5
Mature β-cell neoplasms		
Chronic lymphocytic leukaemia (CLL)	102	23.7
Hairy cell leukaemia	4	0.9
Burkitt leukaemia	2	0.5
Plasma cell leukaemia	2	0.5
Myelodysplastic/myeloproliferative neoplasms (MDS/MPN)		
Chronic myelomonocytic leukaemia (CMML)	17	3.9
Mature T and NK neoplasms		
T-cell leukaemia unspecified	5	1.2
Other		
Unspecified acute/chronic leukaemia	17	3.9

*Frequency (n) and percentage (%) do not tally since more than one leukaemia subtype may be studied by the same clinical trial

3.1.2 Analysis on Efficacy Parameters

Thirty-six unique efficacy endpoints were identified from protocols of the included trials and grouped into four principal categories comprising of survival (14%, n=5), time-to-event (17%, n=6), response rates and biomarkers (44%, n=16) and a miscellaneous group (25%, n=9) with parameters such as patient reported outcomes (PROs) and medical care utilisation (MCU) amongst others (Table 3.3).

Table 3.3 Categories of efficacy endpoints studied in leukaemia clinical trials (N=36)

Endpoint Category 1: Survival (n=5)	**Endpoint Category 2: Time-To-Event (n=6)**
Overall survival (OS)	Duration of remission/response (DOR)
Progression-free survival (PFS)	Time to remission/response (TTR)
Event-free survival (EFS)	Time to progression (TTP)
Disease-free survival (DFS)	Time to treatment (TTT)
Treatment-free survival (TFS)	Time to treatment failure (TTF)
	Treatment-free remission/response (TFR)
Endpoint Category 3: Response Rates and Biomarkers (n=16)	**Endpoint Category 4: Other (n=9)**
Relapse rate (RR)/Rate of Progressive Disease (PD)	Patient reported outcomes (PROs)
Complete remission/response rate (CR)	Number of patients with asparagine depletion
Complete remission/response rate with incomplete platelet recovery (CRp)	Secondary malignancies
Complete remission/response rate with incomplete blood count recovery (CRi/CRh)	Number of blood cell transfusions
Partial remission/response rate (PR)	Number of chemotherapy cycles received
Combined complete remission/response rate (CRc)	Incidence and duration of febrile neutropenia
Objective remission/response rate (ORR)	Medical care utilisation (MCU)
Best overall remission/response rate (BOR)	Incidence, duration and severity of opportunistic infections
Haematologic response rate (HR)	Proportion of patients undergoing subsequent allogeneic stem cell transplant (ASCT)
Cytogenetic response rate (CyR)	
Molecular remission/response (MR)	
Minimal residual disease (MRD)	
Cytokine serum concentrations	
Morphologic leukaemia-free state (MLFS)	
Splenomegaly and hepatomegaly response rates	
Return to chronic phase (RTC)	

When examining endpoints studied in clinical trials per year, the general trend at a category level is that response rates and biomarkers (42% mean rate of clinical trials) and survival parameters (35% mean rate of clinical trials) were consistently investigated to a greater extent throughout the eleven-year period when compared to the other endpoint categories (Figure 3.3).

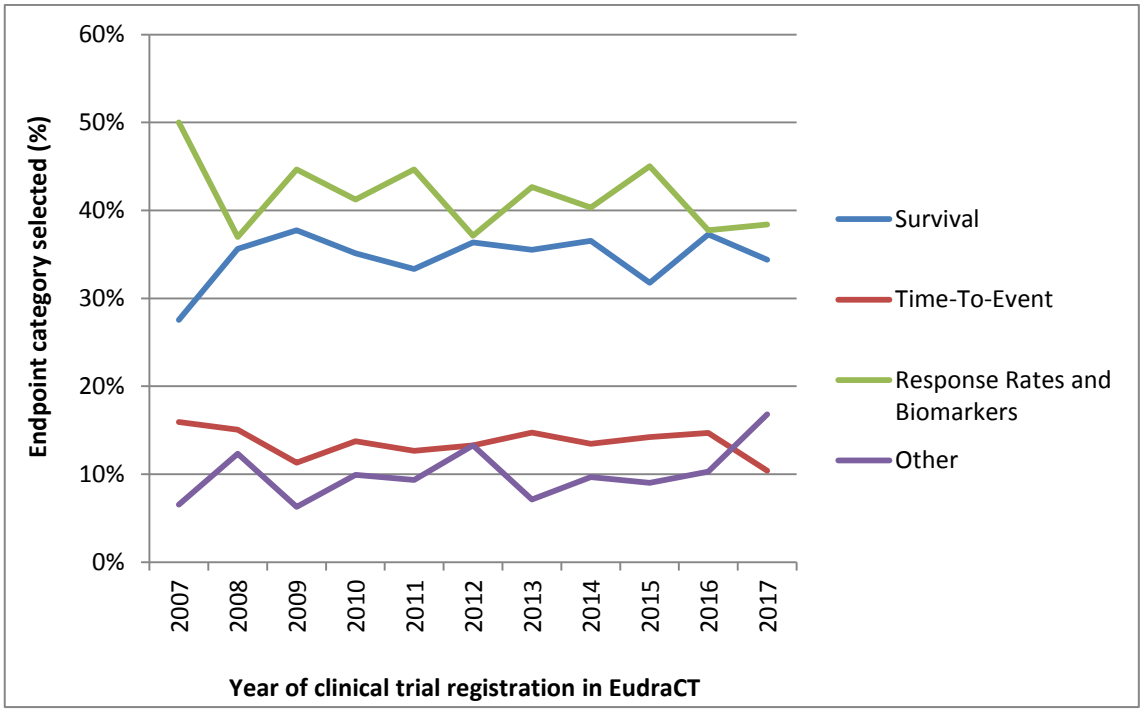


Figure 3.3 Time-trend analysis of selected efficacy endpoints by category in leukaemia clinical trials, 2007-2017 (N=431)

Overall survival (66%, n=285), duration of remission/response (35%, n=150), complete remission/response rate (39%, n=168) and patient reported outcomes (19%, n=81) were found to be the top ranking efficacy parameters in their respective categories (Figures 3.4-3.7). Figure 3.8 depicts how overall survival was the most studied endpoint for all leukaemia clinical trials, however complete response rate surpassed overall survival as the primary measure of choice (OS: 17%, n=72; CR: 19%, n=81). Objective remission/response rate (14%, n=62) and progression-free survival (12%, n=51) rank as the third and fourth preferred primary endpoints respectively.

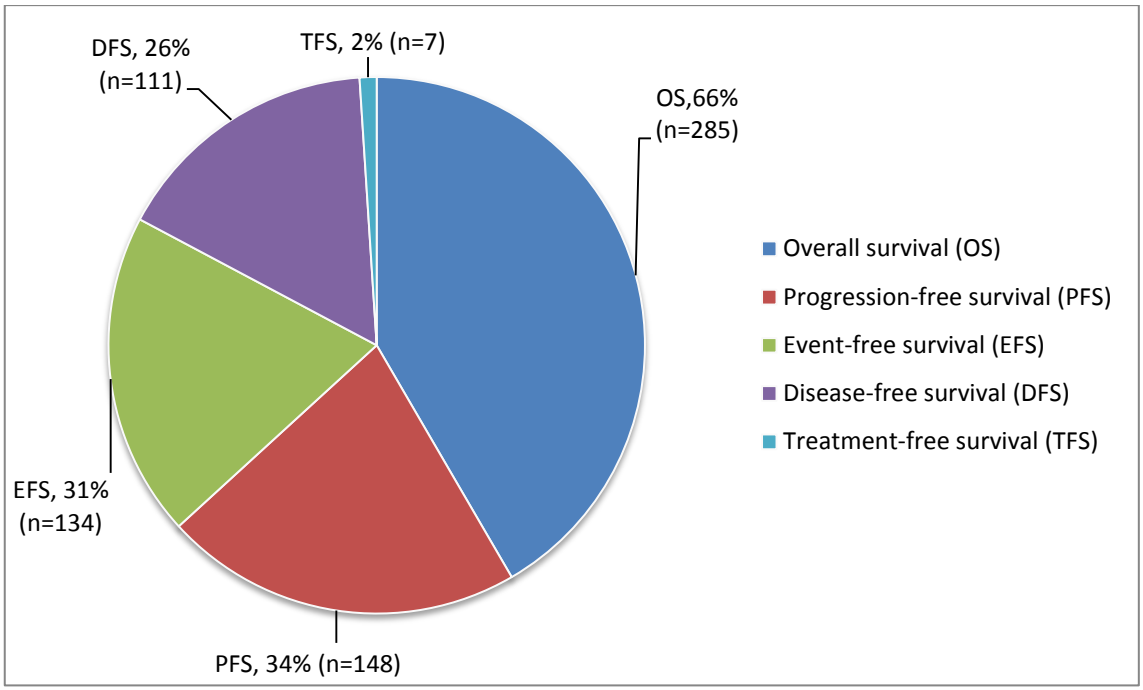


Figure 3.4 Top 5 efficacy endpoints studied for the category of survival (N=431)

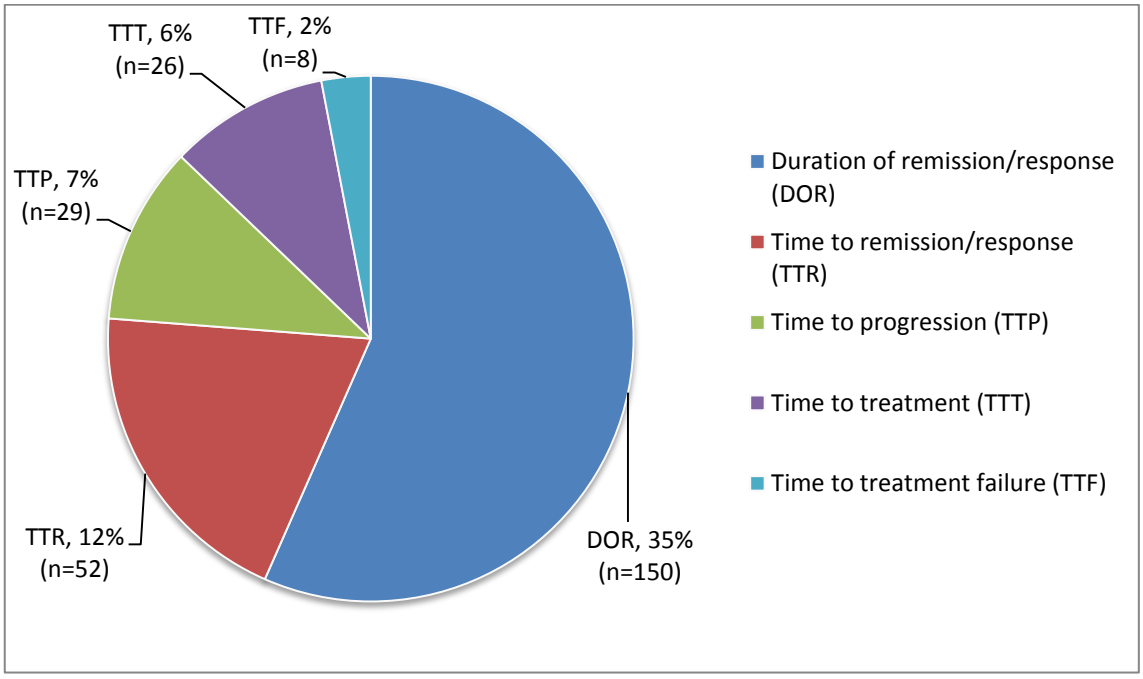


Figure 3.5 Top 5 efficacy endpoints studied for the category of time-to-event (N=431)

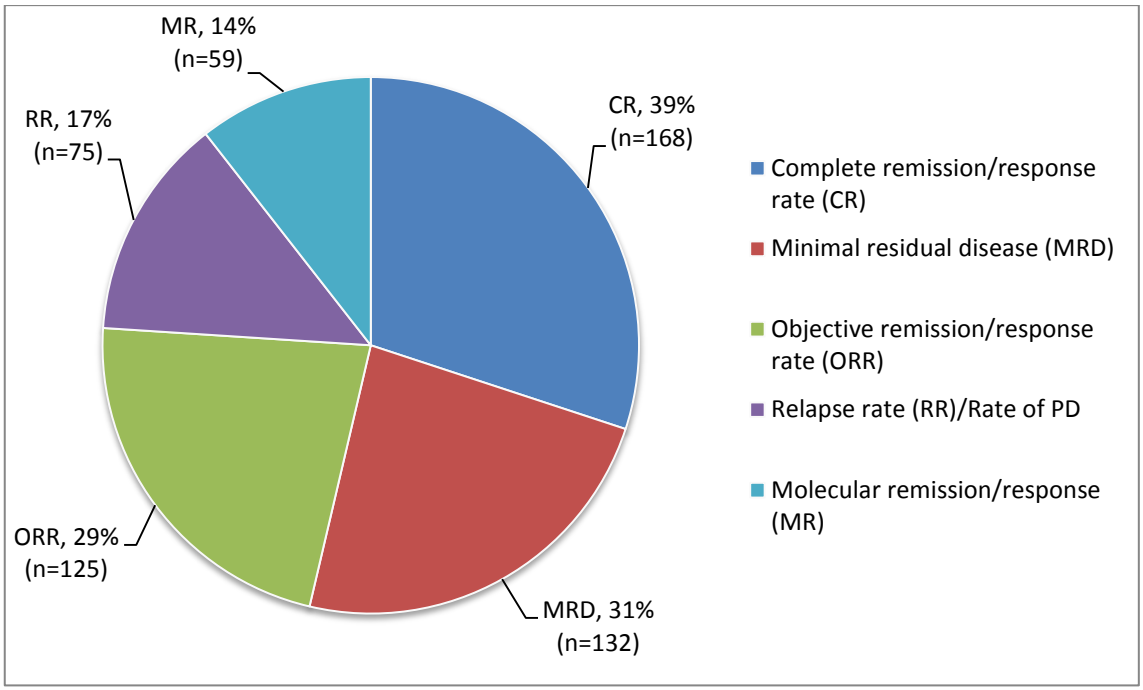


Figure 3.6 Top 5 efficacy endpoints studied for the category of response rates and biomarkers (N=431)

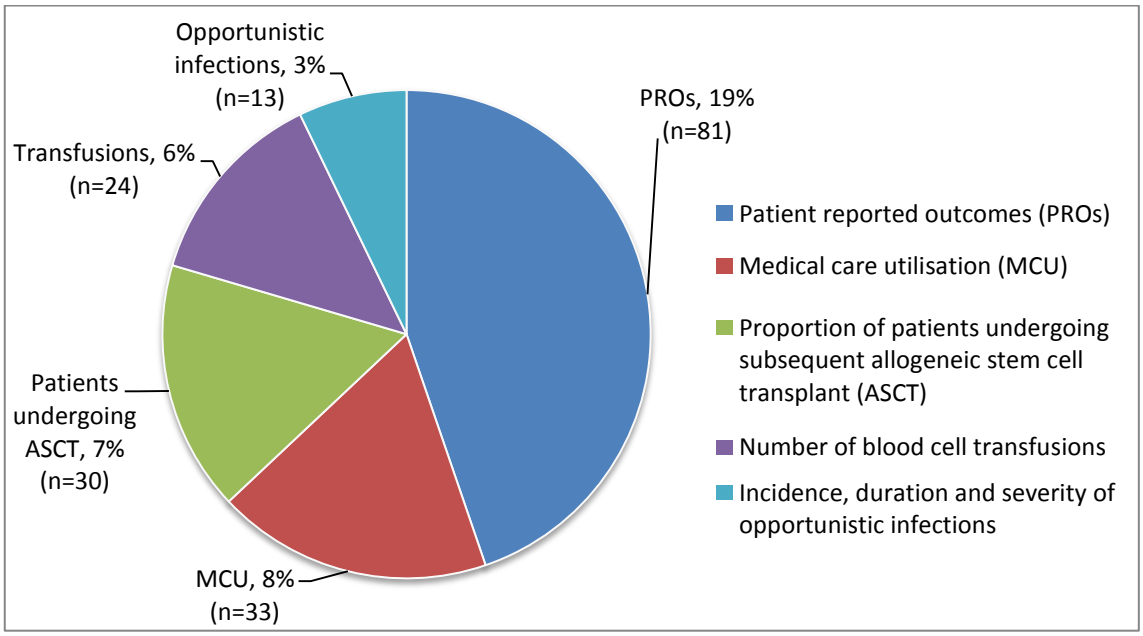


Figure 3.7 Top 5 efficacy endpoints studied for the category of other (N=431)

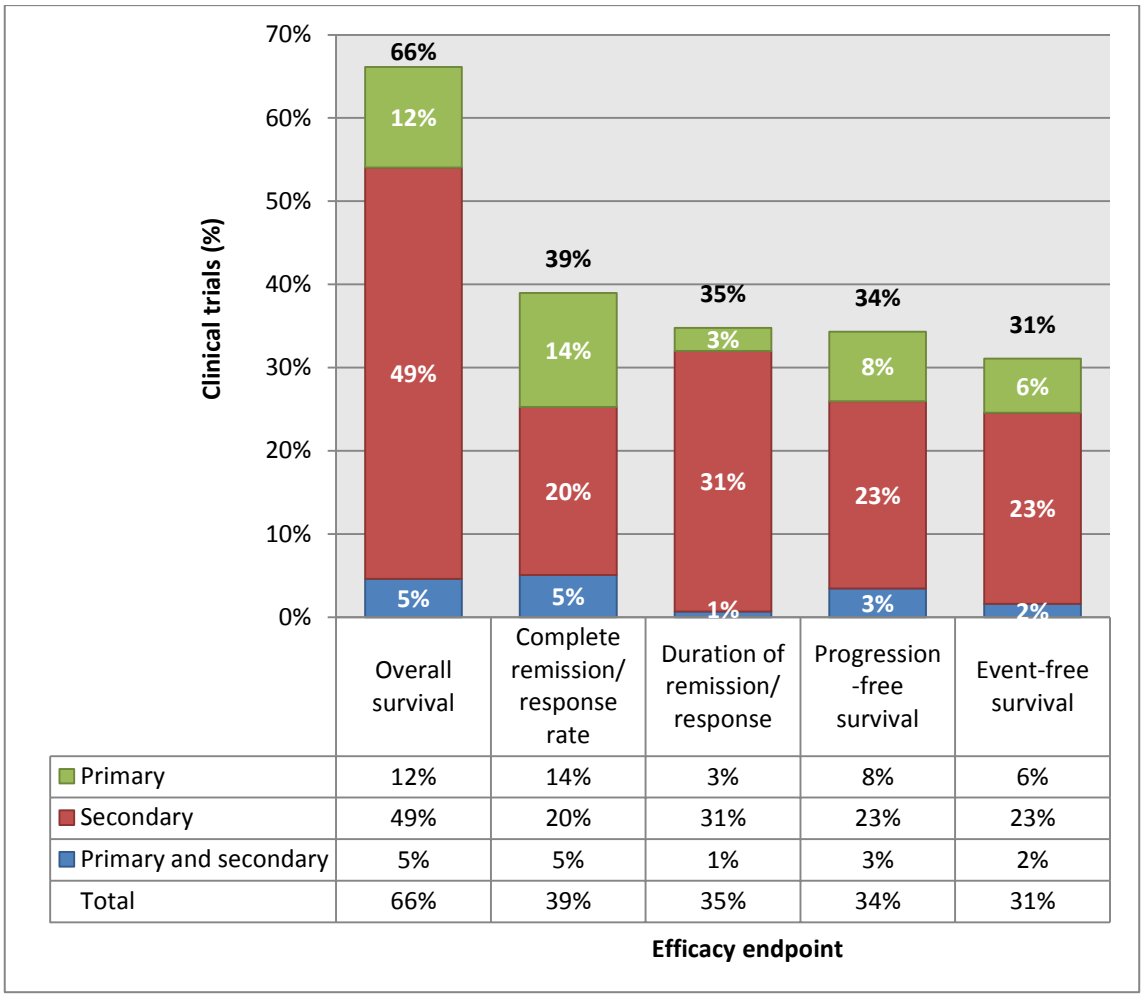
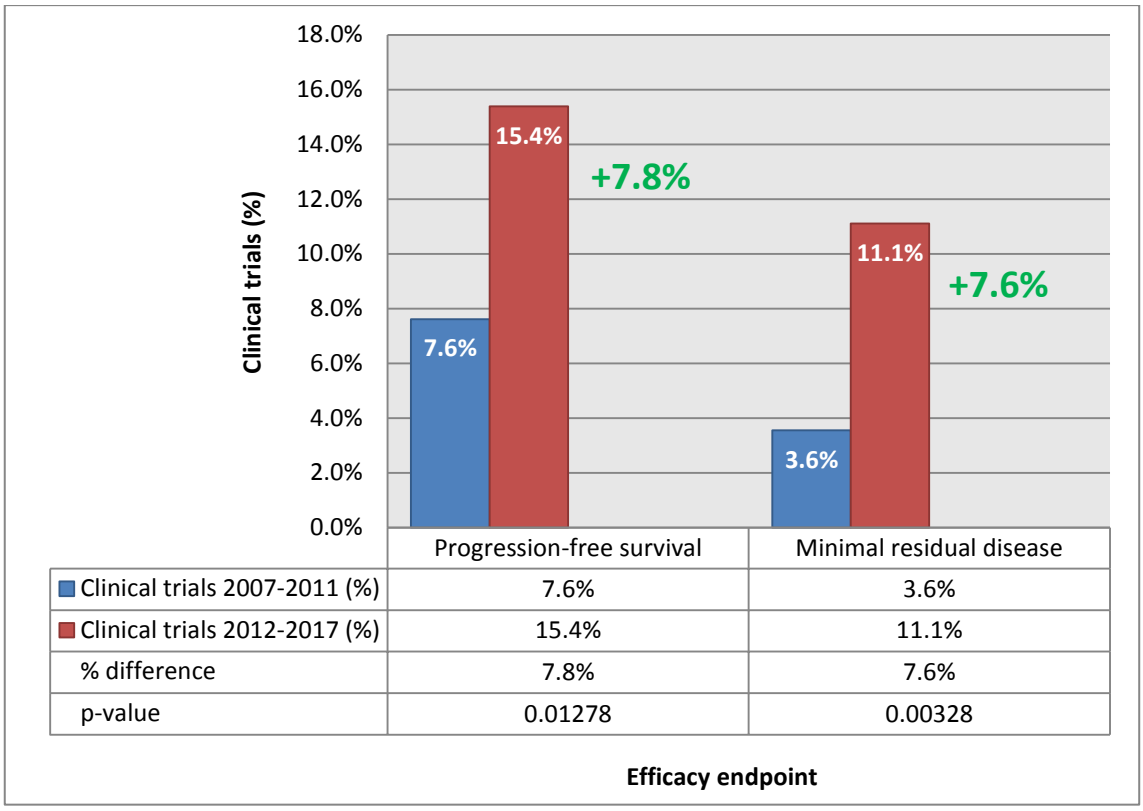


Figure 3.8 Most frequently reported endpoints in leukaemia clinical trial efficacy studies, 2007-2017 (N=431)

Patterns in the selection of primary efficacy endpoints over time were analysed by comparing the frequency of endpoint selection for the latter six years of the study period with the preceding five years. Figures 3.9 and 3.10 show that a statistically significant upward trend was detected for progression-free survival and the sensitive molecular marker minimal residual disease (PFS: p=0.013; MRD: p=0.003). The number of efficacy outcomes studied in trials investigating the same clinical indication was assessed. Efficacy studies in clinical trials for chronic lymphocytic leukaemia demonstrated the greatest variability with 69% (n=70) of the trials studying 4 or more parameters, having a mean of 4.9 outcomes per trial (Figure 3.11).



	Progression-free survival	Minimal residual disease
Clinical trials 2007-2011 (%)	7.6%	3.6%
Clinical trials 2012-2017 (%)	15.4%	11.1%
% difference	7.8%	7.6%
p-value	0.01278	0.00328

Figure 3.9 **Primary efficacy endpoints with a statistically significant change in uptake in leukaemia clinical trials pre- and post-2012 (N=431)**

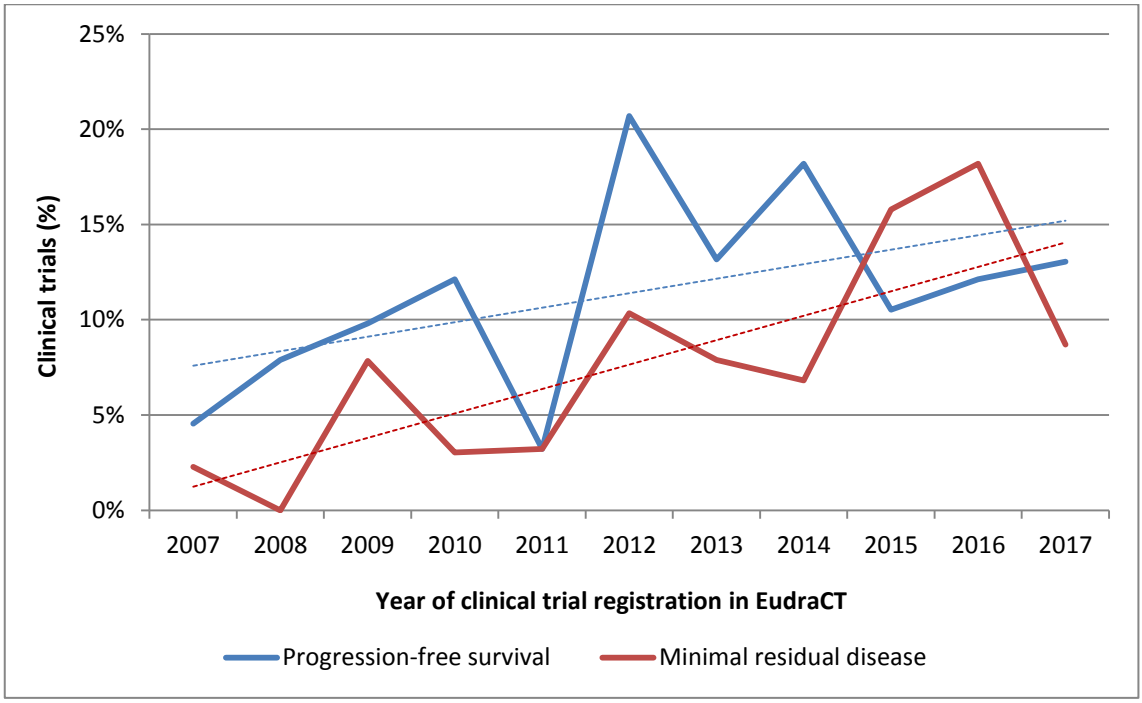


Figure 3.10 **Time-trend analysis of progression-free survival and minimal residual disease as primary efficacy endpoints in leukaemia clinical trials, 2007-2017 (N=431)**

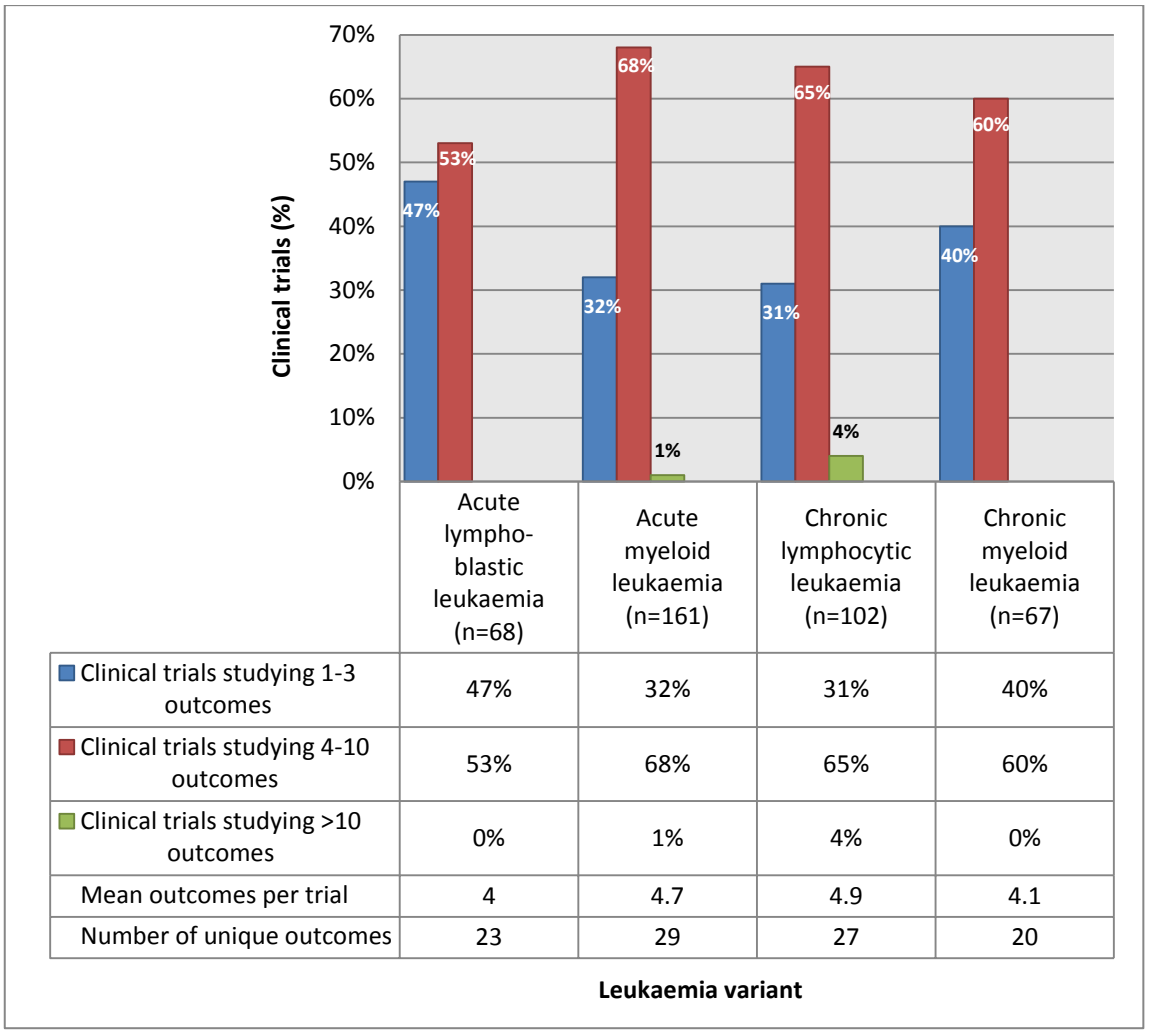


Figure 3.11 Number of efficacy outcomes studied in clinical trials for the main leukaemia variants (N=384)

The preferred primary endpoints for the four main leukaemia subtypes are presented in Figure 3.12. Results suggest that the majority of clinical trials for three of the main leukaemia variants reported surrogate endpoints as their primary measure of choice throughout the study period; event-free survival in acute lymphoblastic leukaemia (16%, n=11), progression-free survival in chronic lymphocytic leukaemia (44%, n=45) and molecular response rate in chronic myeloid leukaemia (40%, n=27) were most extensively studied for their respective clinical indication.

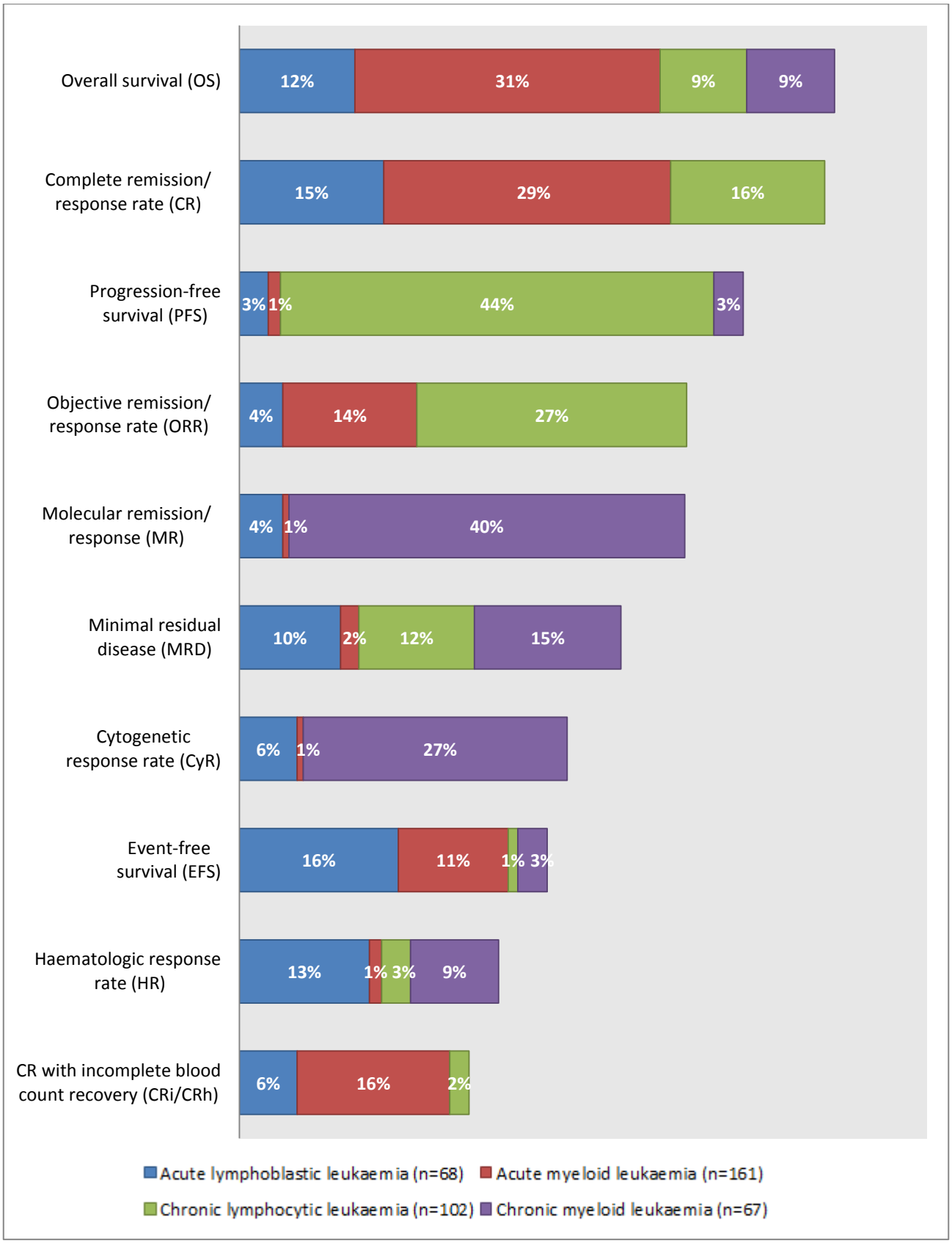


Figure 3.12 Preferred primary efficacy endpoints studied in clinical trials for the main leukaemia subtypes (N=384)

Table 3.4 highlights the clinical trial phase characteristics for the data set. Efficacy studies were predominantly conducted in exploratory Phase II clinical trials (65%, n=281) followed by pivotal Phase III trials (35%, n=152). Only 3% (n=15) of clinical trials investigated the efficacy of leukaemia treatments in Phase IV post-authorisation studies. Overall survival was the highest selected primary efficacy measure for Phase III and Phase IV trials, with complete response rate being the most studied endpoint under Phase II conditions.

Table 3.4 Clinical trial phase characteristics for the included studies

Clinical trial phase	Number of trials (n)*	Percentage from total trials (N=431) (%)	Preferred primary efficacy measure
Phase II	281	65	Complete remission/response rate (22%, n=61)
Phase III	152	35	Overall survival (30%, n=46)
Phase IV	15	3	Overall survival (20%, n=3)

*Total number of trials does not tally to N=431 since more than one clinical trial phase may be studied by the same trial

3.2 Part II: Scientific Expert Opinions on Clinical Trial Design in Leukaemia

This section describes the results of the psychometric evaluation of the ReVALeu tool, the analyses of scientific expert opinions across both Delphi rounds and the identification of core efficacy outcomes for leukaemia trials.

3.2.1 Validation, Reliability Testing and Piloting of the Tool

The 8 validation members completed the validation process of the ReVALeu tool by rating the questions, statements and appendix of endpoint definitions for the 3 domains

of relevance, clarity, and structure and layout (Appendix 5). Six questions (26%), all pertaining to first section of the tool, obtained an I-CVI score below the 88% threshold and were flagged for revision or omission. Five out of these questions were modified and only one was omitted. For seven questions and the appendix of endpoint definitions, ancillary comments were provided; despite these tool components obtaining scores higher than the established criterion, feedback from the expert validation panel was deemed important and thus incorporated (Appendix 6). S-CVI outputs were calculated at 95%, 91% and 93% for the validation domains of relevance, clarity, and structure and layout respectively, further affirming the robustness of the ReVALeu tool.

Five members completed the two rounds of reliability testing for the validated ReVALeu tool. First round responses were statistically compared for concordance with second round responses. Intra-subject reliability was upheld across both sections of the tool, with 97% (n=120) of the questionnaire items obtaining a p-value of less than the 0.05 criterion, rejecting the null hypothesis. Only 3% (n=4) of the tool items did not demonstrate sufficient intra-subject reliability (p-value <0.05), however it is believed that statistical significance was not achieved due to the limited number of subjects participating in reliability testing and these items were still retained.

The estimated time generated by the online survey platform in completing the first Delphi round was observed to approach the simulated time taken by the piloting members. Approximately 20 minutes for completion of Delphi round one was recorded. The ReVALeu tool was found to be compatible with the online browsers tested on smart devices.

3.2.2 Recruitment of Delphi Experts and Response Rates

Thirty-six experts agreed to participate in the Delphi study (Table 3.5). Twelve experts were recruited from 8 public and 1 private HTABs in 9 different EU countries with optimal geographical distribution, representing 25% of the EU-28 population[28]. These included Austria (n=1), Czech Republic (n=2), Finland (n=1), Ireland (n=2), Malta (n=2), Portugal (n=1), The Netherlands (n=1), Italy (n=1) and Sweden (n=1). The 9 participating HTABs encompassed the following:

1. Ludwig Boltzmann Institute for Health Technology Assessment (LBI-HTA), Austria
2. State Institute for Drug Control (SUKL), Czech Republic
3. Finnish Medicines Agency (FIMEA), Finland
4. National Centre for Pharmacoeconomics (NCPE), Ireland
5. Directorate for Pharmaceutical Affairs (DPA), Malta
6. National Authority of Medicines and Health Products (INFARMED), Portugal
7. National Health Care Institute (ZINL), The Netherlands
8. Italian Medicines Agency (AIFA), Italy
9. Dental and Pharmaceutical Benefits Agency (TLV), Sweden

Twenty-four regulatory experts from the EMA confirmed participation: 3 from the Committee for Medicinal Products for Human Use (CHMP), 5 from the Scientific Advice Working Party (SAWP) and 16 external experts.

[28] Eurostat. Population as a percentage of EU28 population [Internet]. [cited 2019 Apr 20]. Available from: https://ec.europa.eu/eurostat/tgm/table.do?tab=table&init=1&language=en&pcode=tps00005&plugin=1

Table 3.5 Recruitment of experts in the Delphi study

Decision-maker group	Number of EMA experts/HTABs invited to the study	No response (%)	Rejected participation (%)	Accepted participation (%)
Regulatory	270 experts	235 experts (87)	11 experts (4)	24 experts (9): 3 CHMP members 4 SAWP members 1 SAWP alternate 16 External experts
HTA	74 HTABs	61 HTABs (82)	4 HTABs (5)	12 experts from 9 HTABs (12)

Acronyms: European Medicines Agency (EMA); Committee for Medicinal Products for Human Use (CHMP); Health Technology Assessment (HTA); Health Technology Assessment Body (HTAB); Scientific Advice Working Party (SAWP)

A response rate of 83% was achieved for both survey rounds, exceeding the 70% limit required between rounds for the Delphi technique. A total of 55 questionnaires were completed in 4 iterations with a mean completion rate of 92% (Table 3.6).

Table 3.6 Response and completion rates for each survey round

Decision-maker group	Recruited (n)	Delphi round one complete responses (n)	Delphi round two complete responses (n)
HTA	12	12	10
Regulatory	24	18	15
Total participants	36	30	25
Response rate (%)		**83**	**83**
Completion rate (%)		**87.5**	**97**

Acronyms: Health Technology Assessment (HTA)

3.2.3 First Delphi Iteration

3.2.3.1 General Opinions on the Clinical Development and Assessments of Antineoplastic Agents

Figure 3.13 shows that HTA experts expressed stronger dissent than their regulatory counterparts when asked if there is currently sufficient cooperation between both decision-makers throughout the life cycle of medicinal products (weighted mean ratings: 3.0 (regulatory); 2.5 (HTA)). This trend is replicated to a greater extent in Figure 3.14 where participants were asked on their perception of alignment between regulatory and HTA assessment procedures with regards to the clinical evidence needs for oncology medicines (weighted mean ratings: 3.1 (regulatory); 2.4 (HTA)).

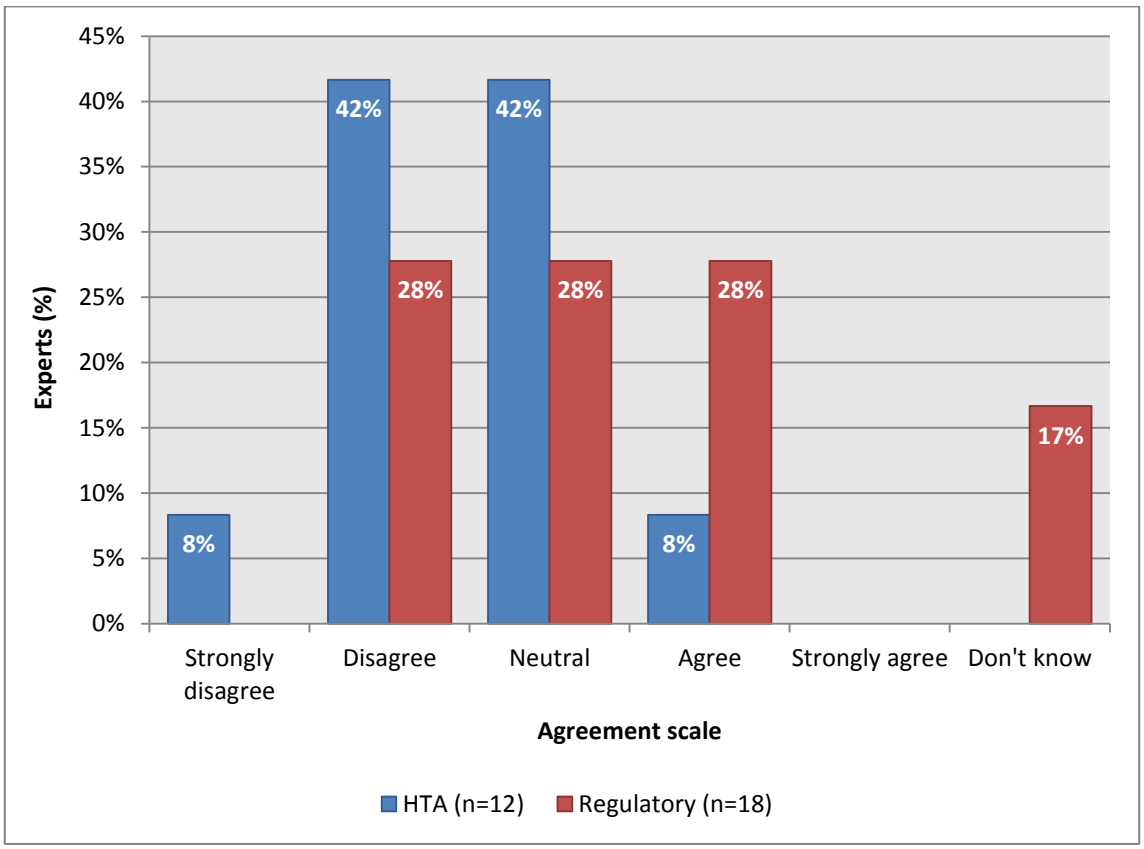


Figure 3.13 Agreement on whether the current level of cooperation between regulatory and HTA bodies is sufficient (N=30)

Acronyms: Health Technology Assessment (HTA)

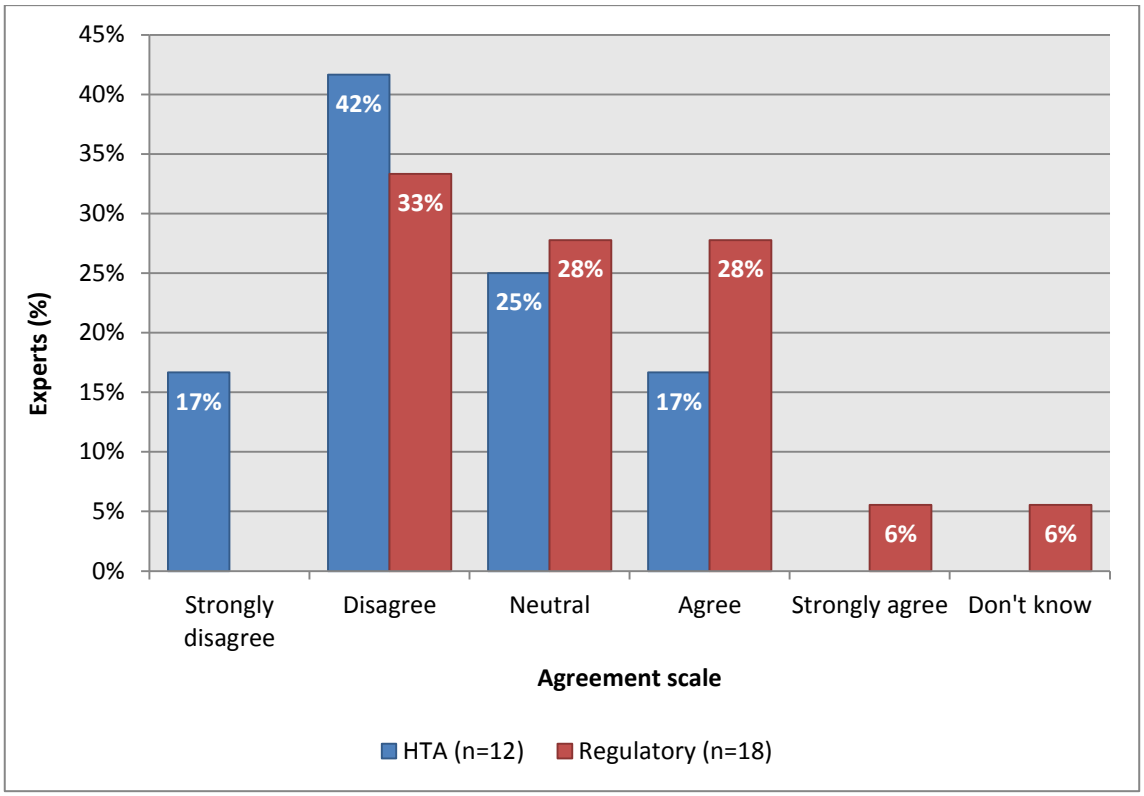


Figure 3.14 Agreement on whether the clinical evidence needs for oncology medicines between regulatory and HTA bodies are aligned (N=30)

Acronyms: Health Technology Assessment (HTA)

When asked on their perception regarding the quality of evidence generated for antineoplastic agents in the pre- and post-authorisation phases, decision-makers had significantly conflicting views (Table 3.7). The radar graphs in Figures 3.15 and 3.16 demonstrate that regulatory opinions are skewed towards higher quality data. A mean of 73% versus 21% for the regulatory and HTA opinions respectively rated the evidence as good, very good or excellent.

Table 3.7 Regulatory and HTA expert ratings on the quality of evidence generated for antineoplastic agents in the pre- and post-authorisation phases (N=30)

Decision-maker group	Pre-authorisation (weighted mean rating)	Post-authorisation (weighted mean rating)
Regulatory (n=18)	3.3	2.9
HTA (n=12)	2.2	2.1
p-value	0.01	0.04

Acronyms: Health Technology Assessment (HTA)

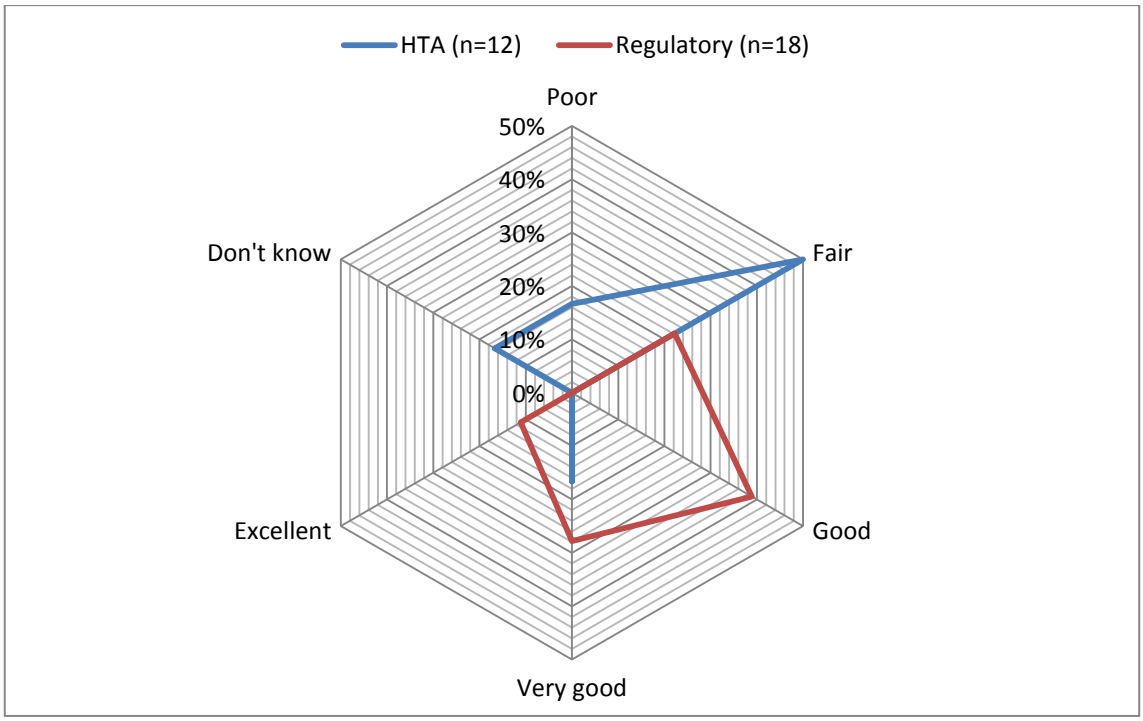


Figure 3.15 Decision-maker opinions on the quality of evidence generated for antineoplastic agents in the pre-authorisation phase (N=30)

Acronyms: Health Technology Assessment (HTA)

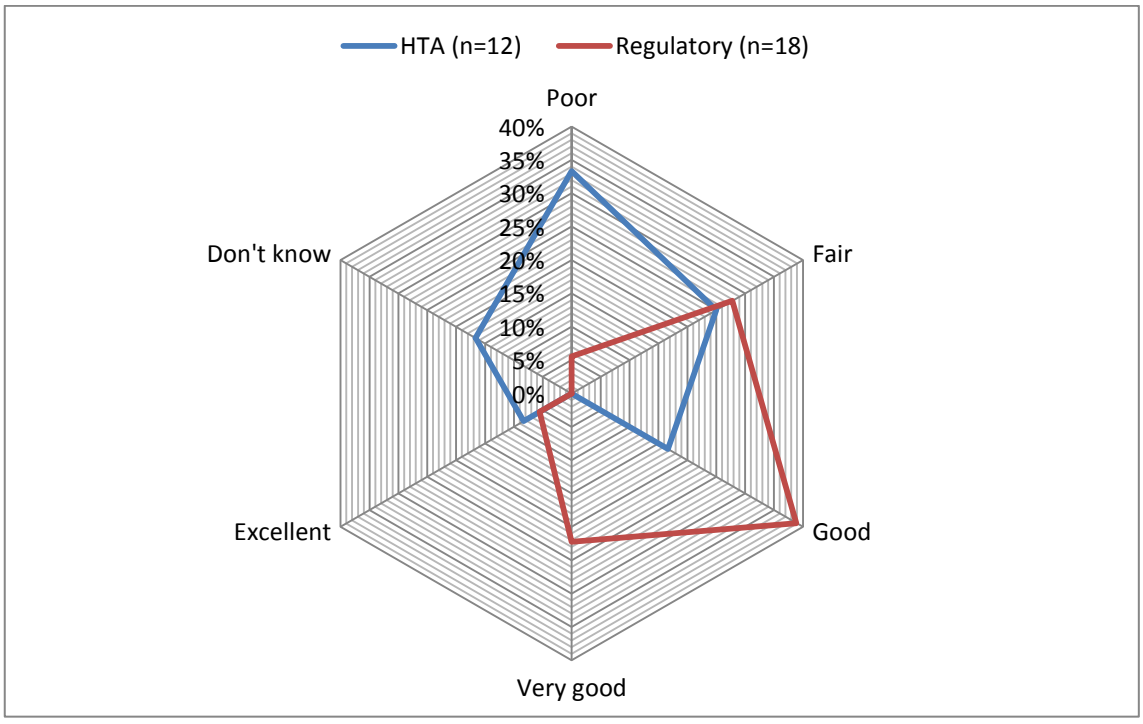


Figure 3.16 Decision-maker opinions on the quality of evidence generated for antineoplastic agents in the post-authorisation phase (N=30)

Acronyms: Health Technology Assessment (HTA)

Attitudes of panel members towards the potential impact of harmonising regulatory and HTA clinical assessments on patient access to innovations were examined. Regulatory experts (weighted mean rating of 3.8) expressed a firmer standpoint than HTA respondents (weighted mean rating of 3.5) that divergences in clinical evidence requirements negatively impacts patient access to novel anticancer treatments. Notwithstanding this, both groups of decision-makers shared a common position by ranking patients as the top stakeholder to benefit from enhanced interactions between regulatory and HTA procedures (Table 3.8).

Table 3.8 Stakeholder groups ranked by panellists according to who is most likely to benefit from enhanced cooperation between regulatory and HTA decision-makers (N=30)

Stakeholder group	**Regulatory (n=18) (weighted mean rank scores)**
Patients	4.2
Clinicians	4.2
Industry	3.9
Payers	3.6
HTABs	2.8
EMA	2.3
Stakeholder group	**HTA (n=12) (weighted mean rank scores)**
Patients	4.4
Payers	3.9
HTABs	3.5
Industry	3.5
EMA	3.5
Clinicians	2.9

Acronyms: European Medicines Agency (EMA); Health Technology Assessment (HTA); Health Technology Assessment Body (HTAB)

Figure 3.17 illustrates that the majority of HTA members consider the clinical evidence requested by their respective HTAB for anticancer medicines is akin to that of other HTABs (weighted mean rating of 3.8). HTA experts were of the opinion that their HTAB evidentiary requirements are not on par to those being requested by the EMA (weighted mean rating of 2.4). This difference was found to be statistically significant with a p-value of 0.006.

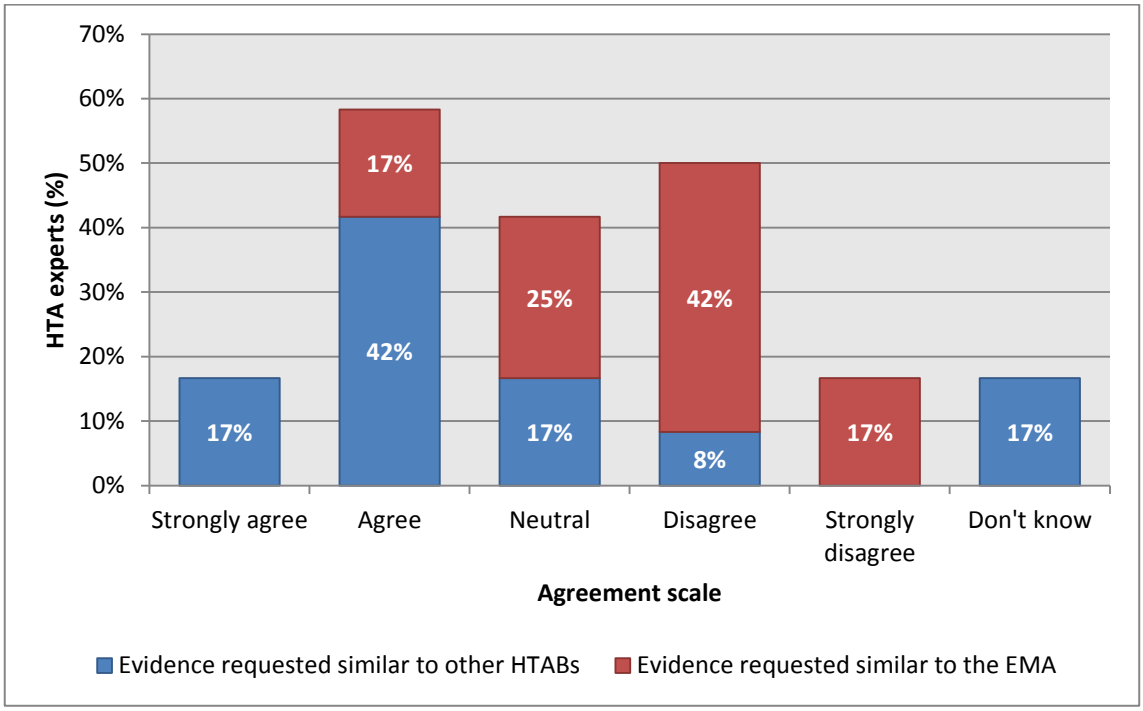


Figure 3.17 Comparability of clinical evidence requested by HTABs for anticancer medicines to that of other HTABs and the EMA (N=12)

Acronyms: European Medicines Agency (EMA); Health Technology Assessment (HTA); Health Technology Assessment Body (HTAB)

Five HTA respondents representing five distinct HTABs namely AIFA in Italy, NCPE in Ireland, INFARMED in Portugal, TLV in Sweden and ZINL in The Netherlands indicated that they had been involved in parallel scientific advice/consultation procedures with the EMA SAWP (Figure 3.18). Seventy-five per cent (75%) of the participating HTA experts responded that they always consult the EMA EPAR when conducting HTA clinical assessments of antineoplastic agents (Figure 3.19).

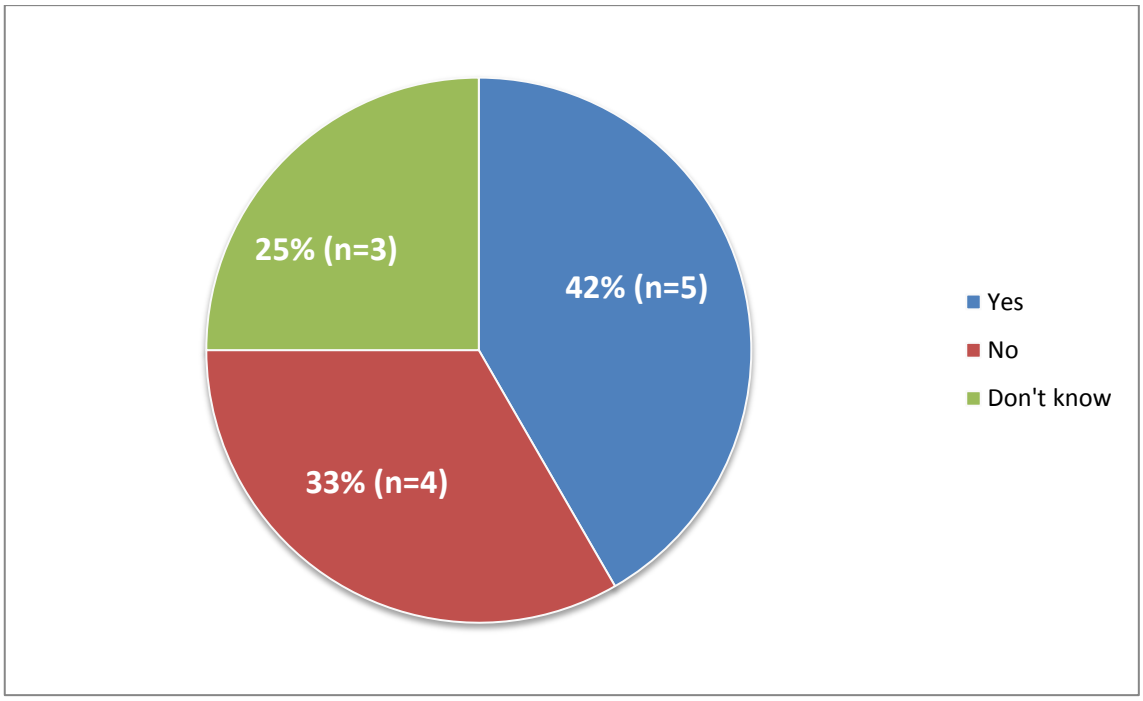


Figure 3.18 HTAB involvement in parallel scientific advice/consultation procedures with the EMA SAWP (N=12)

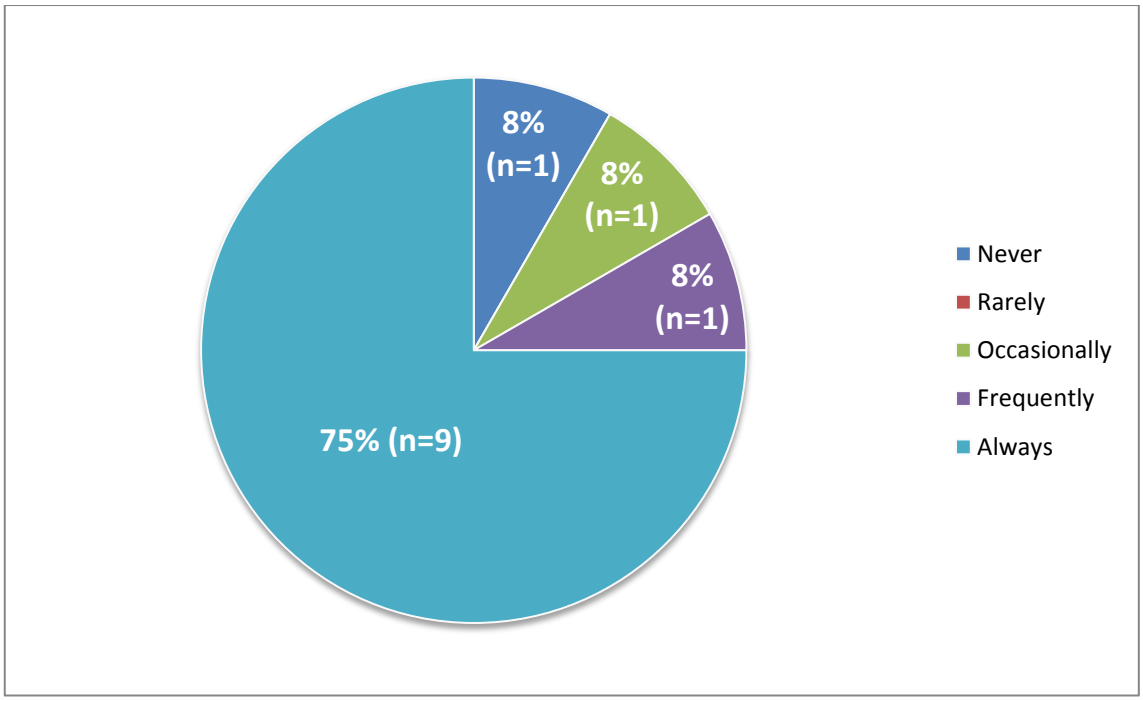


Figure 3.19 Frequency at which the EMA EPAR is consulted in HTA clinical evaluations of antineoplastic agents (N=12)

3.2.3.2 Technical Opinions on Efficacy Studies in Leukaemia Clinical Trials

For each of the 36 efficacy measures identified, the panel members selected the classification (primary or secondary) and most suitable clinical trial phase for investigation (Phase II, Phase III and/or Phase IV) according to their scientific judgement. Figure 3.20 depicts the efficacy parameters that have been chosen by at least 50% of the experts in each decision-maker group to be studied at a primary level. The only parameters that met this cut-off point were overall survival, complete remission/response rate, progression-free survival and objective remission/response rate with a mean percentage frequency of 92%, 68%, 61% and 50% respectively.

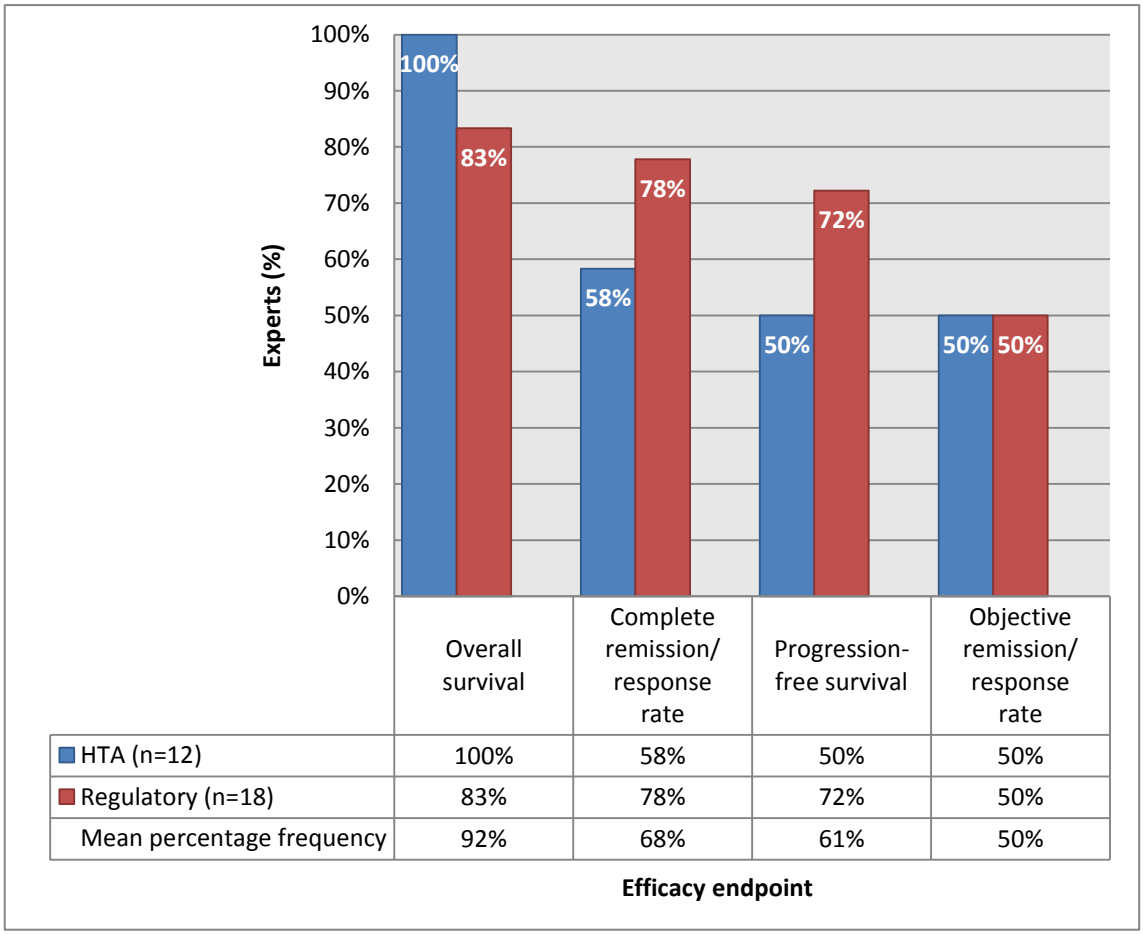


Figure 3.20 Preferred primary efficacy endpoints by regulatory and HTA experts for leukaemia clinical trials (N=30)

Acronyms: Health Technology Assessment (HTA)

Figure 3.21 shows that regulators are much more amenable to Phase II clinical studies than HTABs (mean percentage frequency: 73% (regulatory); 18% (HTA)). HTA experts express a preference for Phase III as the clinical trial phase of choice for the efficacy endpoints under review (mean percentage frequency: 65% (regulatory); 95% (HTA)). This finding underpins the inclination of HTABs towards more robust evidence generating strategies under clinical trial Phase III conditions rather than Phase II studies with smaller patient cohorts.

Both groups of decision-makers were in agreement that post-authorisation Phase IV studies are mostly ideal to measure longer-term outcomes such as patient reported outcomes, medical care utilisation, number of treatment cycles or blood transfusions required, the onset of infectious disease or secondary malignancies and the rate of progression to stem cell transplantation, collectively categorised under the 'Other' category (Table 3.9).

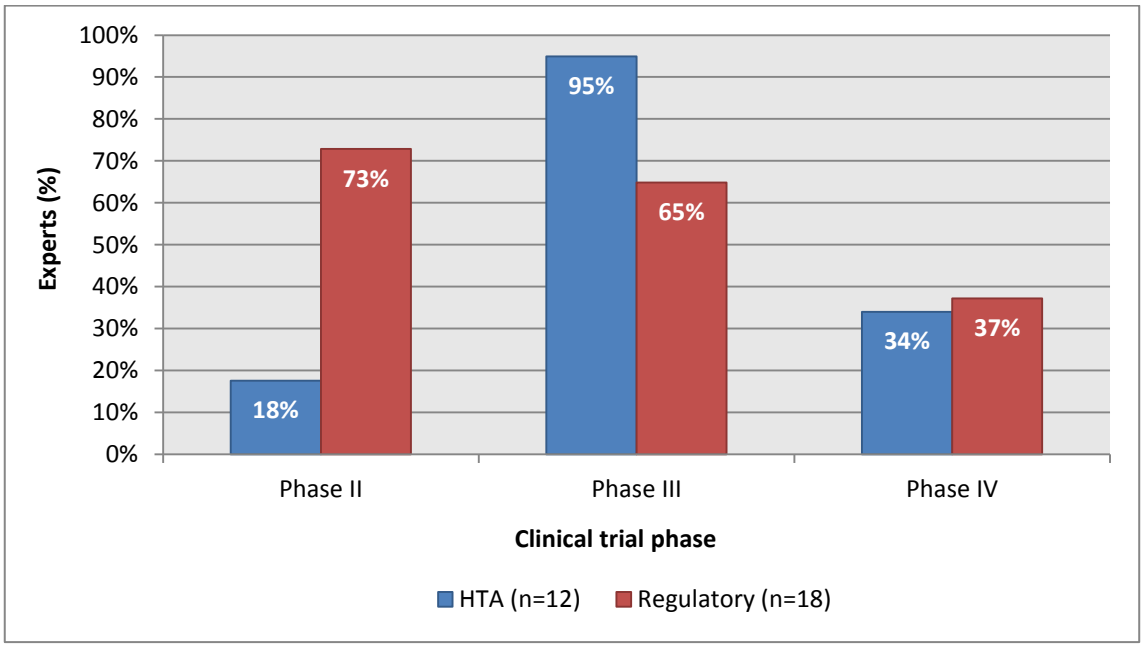


Figure 3.21 Decision-maker clinical trial phase of choice for efficacy studies in leukaemia clinical trials (N=30)

Acronyms: Health Technology Assessment (HTA)

Table 3.9 Preferred clinical trial phase for decision-makers by efficacy endpoint category (N=30)

	HTA (n=12)			Regulatory (n=18)		
Endpoint category	**Phase II**	**Phase III**	**Phase IV**	**Phase II**	**Phase III**	**Phase IV**
Survival (%)	17	97	35	60	70	36
Time-To-Event (%)	24	100	24	69	56	35
Response Rates and Biomarkers (%)	22	97	26	82	58	27
Other (%)	6	86	55	66	79	58

The results for the first round Delphi items are summarised in Tables 3.10 and 3.11. Twenty-five per cent (25%, n=9) of the efficacy parameters satisfied the consensus threshold for inclusion by the HTA panel experts, with the remaining items (75%, n=27) not reaching consensus and progressing to Delphi round two. The regulatory cohort achieved higher consensus rates in the first round with 36% (n=13) of the efficacy outcomes reaching consensus for inclusion and 64% (n=23) of the endpoints being carried forward to the next round. None of the Delphi items met the stopping criteria for exclusion in round one.

Table 3.10 Delphi round one consensus rates for efficacy outcomes (N=36)

	HTA	**Regulatory**
Number of outcomes reaching consensus for inclusion as 'important' (%)	9 (25)	13 (36)
Number of outcomes reaching consensus for exclusion as 'important' (%)	0 (0)	0 (0)
Number of outcomes not reaching consensus (%)	27 (75)	23 (64)

Acronyms: Health Technology Assessment (HTA)

Table 3.11 Efficacy outcomes reaching consensus in the first Delphi round

HTA (N=12)		
Efficacy Outcome	**Percentage of experts rating outcome as 'Important' or 'Very important' (n)**	**Weighted mean rating of importance**
Overall survival (OS)	100 (12)	4.9
Progression-free survival (PFS)	83 (10)	4.3
Event-free survival (EFS)	75 (9)	3.9
Disease-free survival (DFS)	83 (10)	4.1
Time to progression (TTP)	75 (9)	4.0
Complete remission/response rate (CR)	83 (10)	4.1
Partial remission/response rate (PR)	75 (9)	3.8
Patient-reported outcomes (PROs)	100 (12)	4.3
Proportion of patients undergoing subsequent allogeneic stem cell transplant (ASCT)	83 (10)	4.0
Regulatory (N=18)		
Efficacy Outcome	**Percentage of experts rating outcome as 'Important' or 'Very important' (%) (n)**	**Weighted mean rating of importance**
Overall survival (OS)	94 (17)	4.7
Progression-free survival (PFS)	100 (18)	4.6
Event-free survival (EFS)	83 (15)	3.9
Disease-free survival (DFS)	89 (16)	4.0
Duration of remission/response (DOR)	100 (18)	4.5
Time to progression (TTP)	89 (16)	4.1
Time to treatment failure (TTF)	83 (15)	4.0
Treatment-free remission/response (TFR)	78 (14)	3.4
Complete remission/response rate (CR)	94 (17)	4.6
Objective remission/response rate (ORR)	78 (14)	3.8
Cytogenetic response rate (CyR)	83 (15)	4.0
Molecular remission/response (MR)	83 (15)	4.2
Minimal residual disease (MRD)	83 (15)	3.9

Acronyms: Health Technology Assessment (HTA)

3.2.4 Second Delphi Iteration

A summary of results achieved for the second round Delphi items is presented in Tables 3.12 and 3.13. No further items reached consensus among the HTA panel for neither inclusion nor exclusion. Iteration and visibility of the group opinions led the regulatory experts to reach consensus on 3 more parameters. At the end of both Delphi rounds, a total of 9 (25%) items constituted the final list of outcomes reaching consensus for HTA members whilst 16 (44%) endpoints were accrued for regulatory experts. A set of 6 efficacy measures, representing 17% of the 36 endpoints identified, were common to both decision-maker consensus lists and were determined as the core efficacy outcomes (Figure 3.22). Five out of the 6 core outcomes were found to be applicable to the 4 main leukaemia subtypes (Table 3.14).

Table 3.12 Delphi round two consensus rates for efficacy outcomes (HTA: N=27; Regulatory: N=23)

	HTA	Regulatory
Number of outcomes reaching consensus for inclusion as 'important' (%)	0 (0)	3 (13)
Number of outcomes reaching consensus for exclusion as 'important' (%)	0 (0)	0 (0)
Number of outcomes not reaching consensus (%)	27 (100)	20 (87)

Table 3.13 Efficacy outcomes reaching consensus in the second Delphi round

Regulatory (N=15)		
Efficacy Outcome	**Percentage of experts rating outcome as 'Important' or 'Very important' (n)**	**Weighted mean rating of importance**
Treatment-free survival (TFS)	80 (12)	3.8
Relapse rate (RR)/Rate of progressive disease (PD)	87 (13)	4.1
Complete remission/response rate with incomplete blood count recovery (CRi/CRh)	80 (12)	4.0

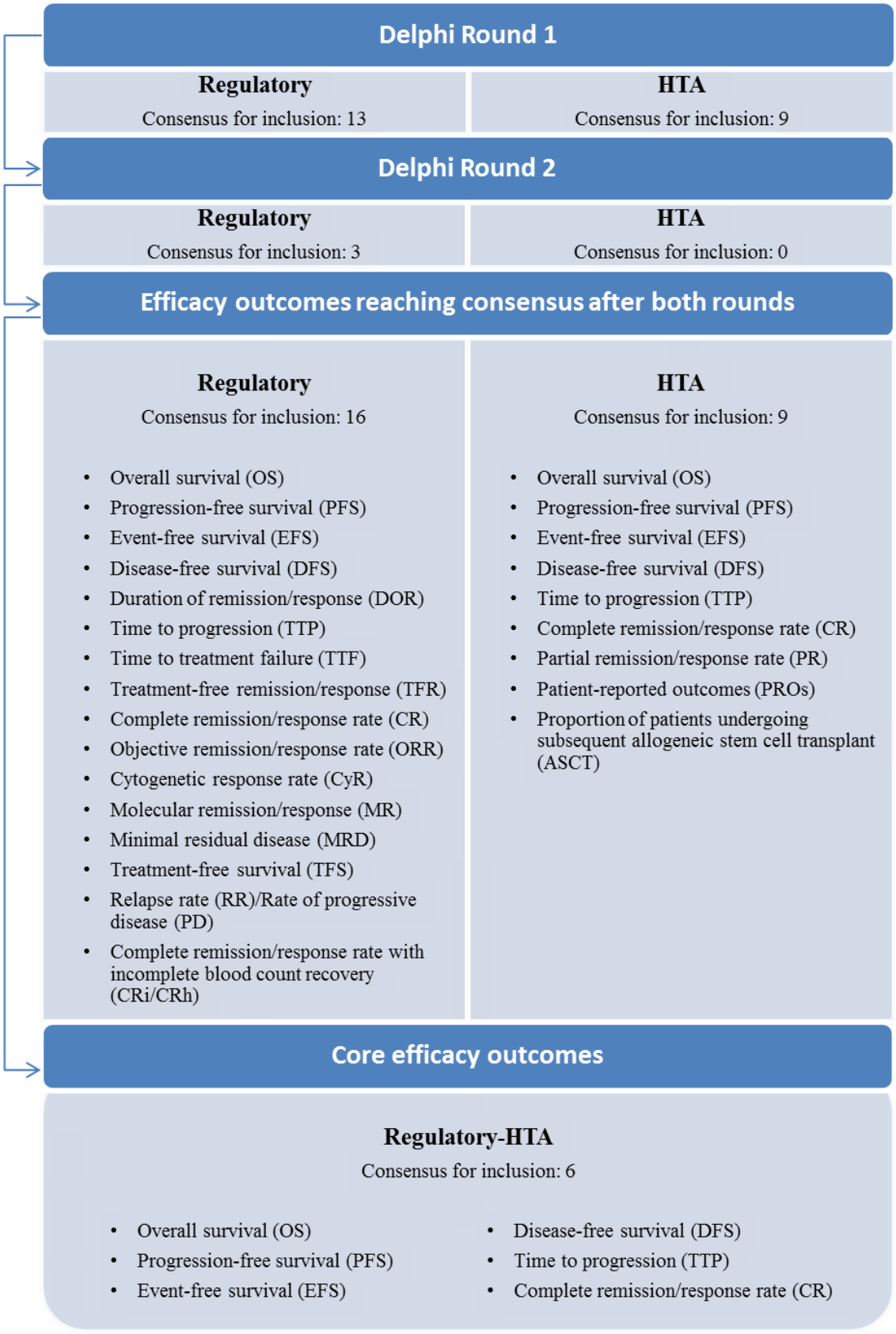


Figure 3.22 Results of the Delphi process leading to the identification of the regulatory-HTA core efficacy outcomes

Acronyms: Health Technology Assessment (HTA)

Table 3.14 Applicability of the core efficacy outcomes to the main leukaemia variants

Core efficacy outcome	Main leukaemia variants			
	Acute lymphoblastic leukaemia	**Acute myeloid leukaemia**	**Chronic lymphocytic leukaemia**	**Chronic myeloid leukaemia**
Overall survival (OS)	✓	✓	✓	✓
Progression-free survival (PFS)	✓	✓	✓	✓
Event-free survival (EFS)	✓	✓	✓	✓
Disease-free survival (DFS)	✓	✓	✓	✓
Time to progression (TTP)	✓	✓	✓	✓
Complete remission/response rate (CR)	✓	✓	✓	

3.2.5 Thematic Analysis of Expert Comments

In total 22 comments were inputted by the panellists throughout the Delphi process, ten from regulatory experts and twelve from their HTA counterparts. As delineated in Table 3.15, comments were structured into three main themes representing patterns in scientific opinions between or within the expert groups. The first theme highlights the importance given by decision-makers to adapting efficacy endpoints according to stage of disease and clinical trial phase. The second theme shows concern expressed commonly among HTA respondents on the use of patient-reported, infection-related and resource utilisation outcomes as indicators of efficacy. The third theme further confirms the weighting given to long-term survival measures by HTABs as opposed to surrogate markers.

Table 3.15 Thematic analysis of expert comments

Qualitative themes	Regulatory expert comments	HTA expert comments
Theme 1: Importance of endpoint depends on stage of disease and clinical trial phase	**n=1:** "All of them* are very important according to the phase of clinical trials." **referring to endpoints grouped under category 1*	**n=2:** "…Progression-free survival and Disease-free survival are important and generally applicable endpoints (but depending on the stage of disease and indication, when these should be selected)…" "The importance depends on the study Phase (II vs. III) and the stage of the disease."
Theme 2 : Endpoint(s) does not merely assess efficacy but is important for safety and risk-benefit evaluations		**n=4:** "Most of these* are important from safety point of view, and also for consideration of risk-benefit ration, but not from mere efficacy point of view…" "Secondary malignancies, incidence/duration of febrile neutropenia and incidence, duration and severity of opportunistic infections are relevant and informative for safety conclusions, yet of lesser utility for efficacy conclusions." "Most of the listed items* are safety-related, not efficacy and very important safety variables." "Some of these* endpoints are relevant in determining overall risk benefit, but not for demonstrating efficacy specifically." **referring to endpoints grouped under category 4*
Theme 3: Preference for long-term outcomes		**n=4:** "OS the most solid and un-biased endpoint." "Long-term efficacy endpoints are usually appreciated in haematological patients." "…not as important as OS, or PFS etc." "Unvalidated endpoints such as MRD unimportant…would be considered more important if there was a definitive link to a hard endpoint such as OS."

3.2.6 Stability of Responses Across Rounds

Delphi responses from both panels were considered to be relatively stable between the first and second iterations. For HTA experts, differences between the first and second round weighted mean ratings of importance, representing the group aggregate judgements, ranged from 0-0.6. Between round group opinions for regulators were considered more stable with changes in the weighted mean ratings ranging from 0-0.3. No major variation was noted for values of standard deviation as a measure of dispersion around the central group opinions.

3.3 Dissemination of Results

1) An abstract submission titled 'Analysis of efficacy parameters used in clinical trials for anti-leukaemic therapy' was accepted as an oral presentation at the 10th Malta Medical School Conference (MMSC) held in Malta, 29 November-1 December 2018 (Appendix 11).

2) An abstract submission titled 'Decision-maker perspectives on outcomes studied in leukaemia clinical trials: The Response Evaluation in Leukaemia (REVALEU) study protocol' was accepted as a poster presentation at The Professional Society for Health Economics and Outcomes Research (ISPOR) 2019 conference held in New Orleans, USA, 18-22 May 2019 (Appendix 11).

3) An abstract submission titled 'Decision-maker perspectives on outcomes studied in leukaemia clinical trials: The Response Evaluation in Leukaemia (REVALEU) study protocol' was accepted as a publication in the 'Value in Health' Journal (Appendix 11).

4) A study protocol titled ‘Regulatory and health technology assessment perspectives on outcomes studied in leukaemia clinical trials’ was registered in the Core Outcome Measures in Effectiveness Trials (COMET) online database (registration number:1234), available from: http://www.comet-initiative.org/studies/details/1234 (Appendix 11).

5) An abstract submission titled ‘Opinions of decision-makers on the clinical development and assessment of antineoplastic agents’ was accepted as a poster presentation at the European Association of Faculties of Pharmacies (EAFP) 2019 conference held in Krakow, Poland, 15-17 May 2019 (Appendix 11).

6) An abstract submission titled ‘Outcomes studied in leukaemia clinical trials: A need for harmonisation?’ was accepted as a poster presentation at the 79th FIP World Congress of Pharmacy and Pharmaceutical Sciences held in Abu Dhabi, United Arab Emirates, 22-26 September 2019 (Appendix 11).

7) An abstract titled ‘Efficacy endpoints studied in clinical trials for early-onset leukaemia’ was accepted as a poster presentation at the European Society for Medical Oncology (ESMO) 2019 congress held in Barcelona, Spain, 27 September – 1 October 2019 (Appendix 11).

8) An abstract titled ‘Trends in efficacy parameters studied in clinical trials for acute myeloid leukaemia’ was accepted as a poster presentation at The Professional Society for Health Economics and Outcomes Research (ISPOR) Europe 2019 conference held in Copenhagen, Denmark, 2-6 November 2019 (Appendix 11).

Chapter 4

Discussion

This chapter discusses the following aspects:

- The characteristics and variability of efficacy outcomes studied in leukaemia clinical trials;
- The level of congruence between the type of evidence produced in leukaemia clinical trials versus that expected by regulatory and HTA bodies;
- Areas of agreement and mismatches between regulatory and HTA perspectives on clinical assessment parameters;
- The features and multi-stakeholder impact of the core outcomes identified as common to regulatory and HTA decision-makers;
- Strengths and potential improvements of the study methodological design, including recommendations for future work;
- The general conclusions drawn from this research and its implications on the landscape for clinical trials, regulatory-HTA interactions and medicines access.

4.1 Evolvement and Heterogeneity of Outcomes in Leukaemia Clinical Trials

Increased clinical research costs associated with lengthy trial durations may lead to economic constraints that curb drug development and access to innovative cancer therapies (Stewart et al, 2010). Results from Part I of this research are in line with recommendations made by Wilson and colleagues (2015a) stating that pharmaceutical companies should challenge trial designs by integrating new endpoints to traditional ones that allow earlier assessments of efficacy throughout the course of development.

An upsurge in the number of oncology trials designating surrogate outcomes as primary endpoints has been reported in the last decade (Cressman et al, 2015). This trend was mirrored in this study where surrogate efficacy parameters such as progression-free survival and minimal residual disease in leukaemia clinical trials were found to be significantly more frequently selected as primary measures post-2012 compared to previous years. This finding demonstrates that the assessment landscape for efficacy in clinical trial protocols for anti-leukaemic therapies is moving towards new paradigms that revolve around biomarker profiling.

The high uptake of surrogate endpoints observed in leukaemia clinical trials can potentially accelerate the production of efficacy data, shortening trial follow-up periods. The inference being abated costs for clinical trial conduct and expedited availability of innovative interventions. Evolvement in leukaemia clinical trial endpoints necessitates the recognition and action by regulatory, HTA and clinical stakeholders by adapting assessment procedures and developing or updating institutional guidelines.

Regulatory and HTA decision-makers are faced with the continuous balancing act of accelerating approvals to realise the unmet medical needs of patients affected by certain conditions versus having a comprehensive clinical data set that allows the adequate assessment of benefit-risk and relative effectiveness. Rapid market release of innovative products is desirable but for agents with an absence of firm evidence, mechanisms should be in place that secure real benefit is derived to avoid undue financial burden on European healthcare systems.

Clinical trials for three of the four main leukaemia variants were found to report a surrogate variable as the preferred primary endpoint for efficacy. Validation studies have correlated progression-free survival in chronic lymphocytic leukaemia

(Beauchemin et al, 2015) and molecular response rate in chronic myeloid leukaemia (Akwaa and Liesveld, 2013; Oriana et al, 2013) with measures of direct clinical benefit. There is a lack of data supporting the use of event-free survival as a surrogate endpoint for longer-term outcomes in acute lymphoblastic leukaemia.

The use of not fully validated surrogate endpoints and the conduct of non-randomised and/or early phase clinical trials can be regarded as measures to secure early access in order to protect public health or as dangerous precedents that give way to the marketing and prescribing of ineffective treatments. This dilemma must be tackled on a case by case basis depending on the ethics and feasibility of the situation. The acceptance of advanced response criteria should be based on the premise of rigorous validation studies substantiating their correlation with patient-relevant outcomes. The Prentice criteria (Prentice, 1989) provide a stringent framework to ensure the inherent validity of surrogate markers in translating to clinically meaningful results. Prentice explains that a surrogate endpoint "should yield unambiguous information about differential treatment effects on the true endpoint".

Inconsistencies were observed in both the number and type of outcomes reported within trials for the main leukaemia subtypes, particularly in chronic lymphocytic leukaemia. In terms of scientific robustness, systematic reviews and meta-analyses are considered the pinnacle of methods for collating evidence in the hierarchy of evidence-generating processes (Murad et al, 2016). Figure 4.1 shows how these analyses would prove challenging to perform if contrasting elements are reported as primary and supporting evidentiary data between trials investigating the same leukaemia subtype. The identification of core outcomes to be reported commonly among clinical trials would standardise the approach of examining efficacy and consolidate the quality of review and combinatorial analyses. From the developer's perspective, the observed high

diversity in outcome reporting also places a risk on the clinical development strategy of generating data that is incompatible with the evidence requirements of decision-makers, negatively impacting on the company's resources.

The arguments presented establish a clear demand for cross-collaborative initiatives among multiple stakeholders to define trial outcomes that are collectively agreed upon in attempt of harmonising clinical outputs.

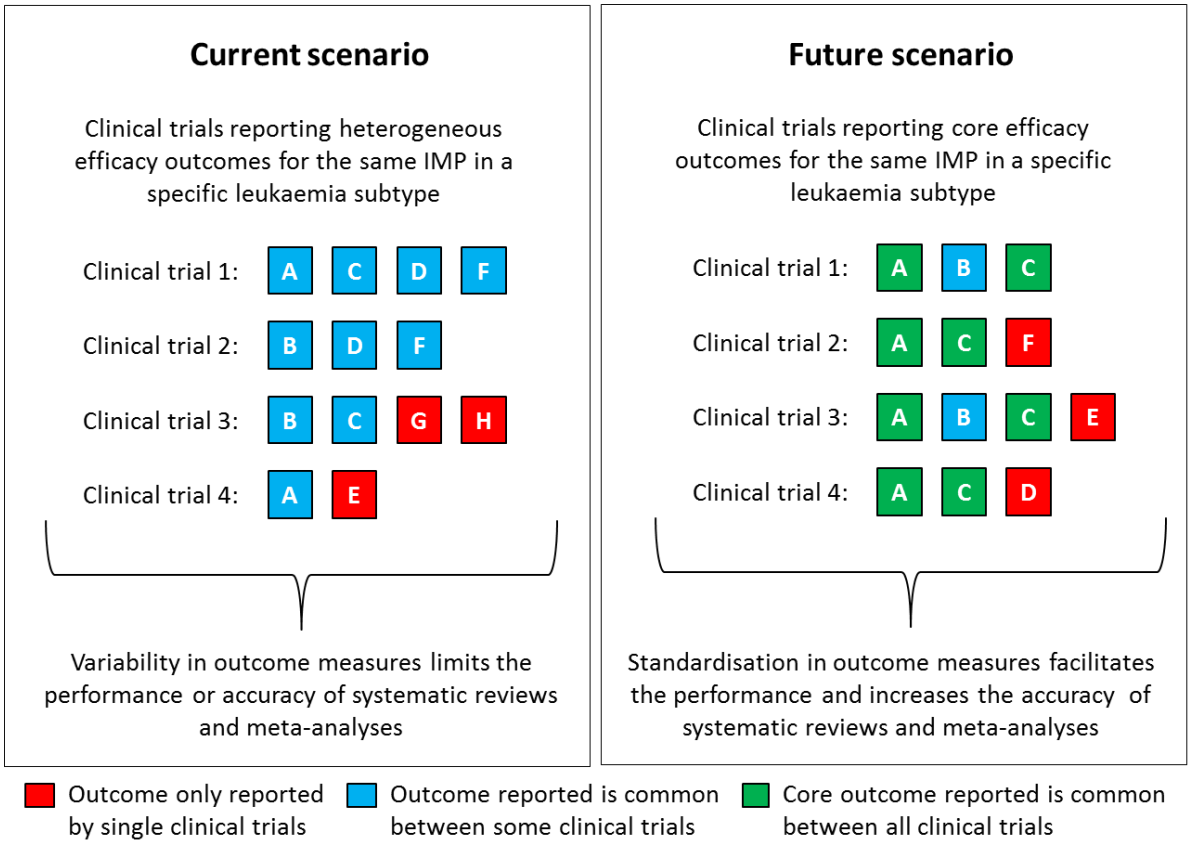


Figure 4.1 The impact of standardising clinical trial outcome selection on the validity of systematic reviews and meta-analyses

Key: Letters A-H represent unique efficacy outcomes reported in clinical trials

Acronyms: Investigational medicinal product (IMP)

4.2 Clinical Trial Design in Leukaemia: Is there Alignment with Expectations of Decision-makers?

For a medicinal product to be granted a marketing authorisation, the pharmaceutical company must demonstrate that the product fulfils the three regulatory 'hurdles' of safety, quality and efficacy. These requirements have been further supplemented by the 'fourth hurdle' which consists of comparative economic and clinical evaluations prior to coverage by the healthcare payer (Rawlins, 2012). Crossing all four 'hurdles' requires pharmaceutical sponsors to adopt a thought process in designing trials capable of generating evidence that satisfies not only the needs of the regulator but also for the treatment under investigation to show superiority in cost- and clinical-effectiveness analyses.

Patterns in the selection of endpoints in leukaemia clinical trial efficacy studies were compared to the preferences of regulatory and HTA experts. The resulting top four primary efficacy variables investigated across all trials were found to correspond with the parameters obtaining the highest ranking among the decision-makers to be studied at a primary level. A distinguishing feature was the order of importance given to these endpoints where, in contrast to the trial protocols, both expert groups expressed clear preference for overall survival to be selected as the gold standard over complete response rate. The Delphi process filtered the 36 unique efficacy endpoints to six core outcomes prioritised jointly by the participating experts. The remaining 30 endpoints are either deemed important to only one of the decision-maker groups or to none at all. The implication of this finding is that medicines developers in the field of leukaemia are assessing response criteria that are subordinate to decision-makers, at the expense of the applicant dossier not getting the regulatory green light or failure for the medicinal product to be enlisted as formulary-positive.

Genetic markers involving the measurement of chromosomal or transcript levels were observed to be at the forefront of assessing the efficacy of interventions in mutation-based leukaemias such as chronic myeloid leukaemia. Regulatory stakeholders have expressed willingness to accept this type of evidence, but this position is not shared with HTABs since none of the cytogenetic, molecular or minimal residual disease endpoints reached consensus among this expert group. The list of endpoints reaching consensus by HTA experts signal the weighting given by technology bodies to patient-relevant survival and quality of life metrics. Despite this demand, only 19% (n=81) of clinical trials investigated patient reported outcomes as an outcome of interest which raises concern on whether HTA evidence needs are being overlooked during trial design.

One final critical observation is the stage at which efficacy studies are conducted. Regulatory experts indicated that both Phase II and Phase III clinical trials may be appropriate to conduct efficacy studies. Conversely, their HTA counterparts are sceptic of early phase studies, greatly favouring late phase trials to demonstrate efficacy. The latter finding is in conflict with what is actually occurring in the great majority of trials since 65% (n=281) of the trials are in Phase II of clinical development. There is an underlying risk for negative HTA recommendations should these trials fail to progress to Phase III and present Phase II data to HTABs.

The clinical development and market access teams within a pharmaceutical company are historically distinct. Forward-looking enterprises are fostering greater cooperation and interactions between the two units in order to capture commercial insights in early product development to proactively support regulatory and HTA approvals.[29]

[29] Galante D. The evolving reimbursement landscape – Considerations for clinical trial design [Internet]. May 2018 [cited 2019 April 10]. Available from: https://www.clinicalleader.com/doc/the-evolving-reimbursement-landscape-considerations-for-clinical-trial-design-0001

4.3 Regulatory and HTA Perspectives on Clinical Assessment Parameters: Commonalities and Disparities

The views of regulatory and HTA participants were considered to have common ground in certain aspects but were also observed to contrast in other areas. Decision-makers held distinct opinions on the quality of evidence generated for antineoplastic agents in both pre- and post-authorisation phases, demonstrating how their perception of satisfactory evidence differs. An example to underscore this divide is the weight given to the respective clinical trial phases. HTABs had an overall low selection rate for Phase II as the ideal clinical trial phase for the efficacy endpoints under review, expressing clear preference for Phase III studies. Paradoxically, regulatory experts rated Phase II clinical trials as most suitable for the majority of the endpoints.

Both decision-maker groups agreed that survival endpoints such as overall survival and progression-free survival should be given utmost importance in efficacy studies. Regulatory experts were far more accepting of endpoints grouped under the category of response rates and biomarkers. An apparent difference was noted between the decision-maker groups on tumour-centric measures, based on cytogenetic and molecular response rates, reaching consensus as important outcomes. The HTA panel greatly favoured patient-centric outcomes over biomarkers, with patient-reported outcomes obtaining the second highest weighted mean rating after overall survival. In line with the observations reported by Wang et al (2018), this study identified the use of biomarkers, surrogate endpoints and patient-reported outcomes as areas to prioritise potential alignment between regulatory and HTA stakeholders.

The reluctance of HTA experts to accept molecular markers as valid outcomes can be traced to a study conducted by Liberti et al (2015) highlighting the lack of HTA

preparedness to implement adaptive pathways where a mean respondent rate of only 29% have rated five EU HTABs (TLV, NICE, AIFA, HAS, IQWiG) as committed in the successful implementation of adaptive pathways versus 75% for the EMA.

The risk of divergent opinions on clinical assessment parameters may potentially increase trial complexity at the expense of pharmaceutical companies curtailing investment initiatives in the development of innovative therapies. This necessitates the streamlining of scientific clinical judgments between the regulatory and HTA facets in order to facilitate the evidence-generating processes for industry stakeholders.

4.4 Core Efficacy Outcomes Identified: Characteristics and Potential Impact

The identified core efficacy outcomes were compared against the ideal core outcome areas of death, life impact, resource use and pathophysiological manifestations that should be represented in every COS set as recommended by Boers et al (2014). Mortality and pathophysiological changes are represented by the core list identified, however life impact, reflected in patient-reported outcomes, only achieved consensus with the HTA members. Neither of the expert groups reached consensus on endpoints related to resource use. The six core outcomes determined in this study for leukaemia trials were observed to approach those being requested by the National Institute for Health and Care Excellence (NICE) in scopes and technology appraisals for haematological malignancies published over the period 2001-2017.[30] NICE showed clear preference for overall survival, progression-free survival, response rates and HRQoL, with the former three also featuring in the core set identified in this research.

[30] HARMONY Alliance. 15 years of Hematological Malignancies outcome reporting to NICE: Data for core outcome sets [Internet]. [cited 2019 Mar 10]. Available from: https://www.harmony-alliance.eu/webimages/files/Harmony_A0_Poster_NICE_FINAL_newsletter.pdf

Kleijnen et al (2016) reported that overall survival and progression-free survival had the highest positive impact on HTAs, which aligns with the HTA preferences in this study.

As depicted in Figure 4.2, this research sought to overcome the heterogeneity of outcomes reported in leukaemia clinical trials by narrowing down the identified 36 unique efficacy endpoints to six core outcomes, jointly prioritised by both groups of decision-makers. The resulting core outcomes are intended to standardise the reporting of clinical trial efficacy data in order to meet the expectations of regulators and HTABs. Diminishing variability in the requests for clinical data creates congruence between EU Member States which promotes a sense of business predictability for pharmaceutical companies, in turn stimulating innovation for the benefit of the patient.

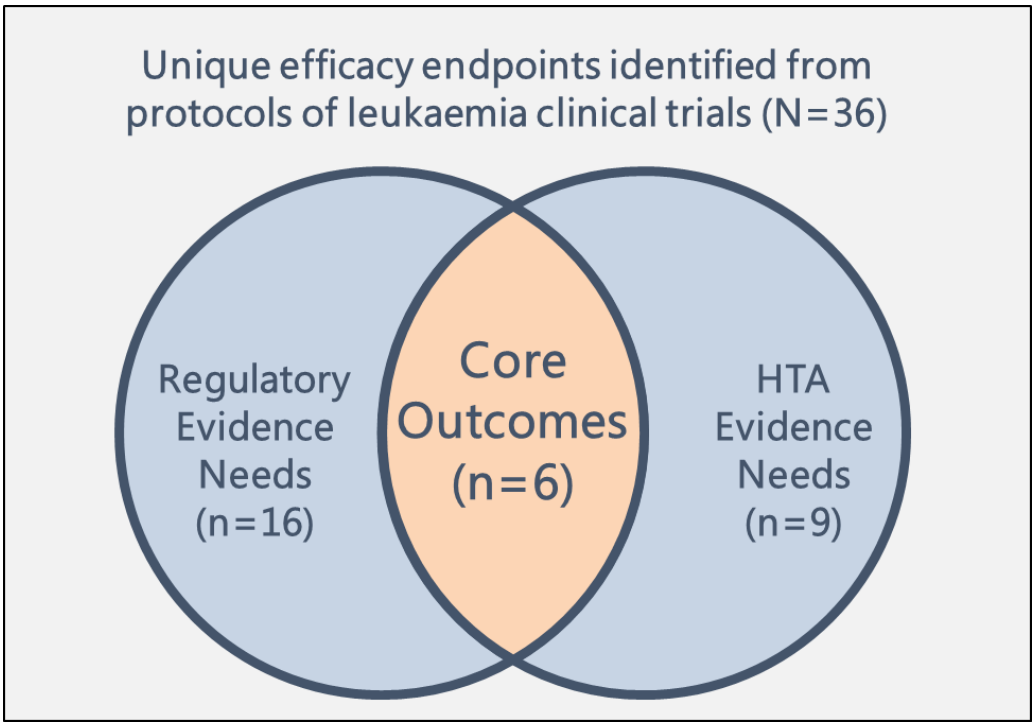


Figure 4.2 Graphical representation of the core outcome selection from the pool of efficacy endpoints identified

Key: 'n' represents the number of efficacy endpoints

Figure 4.3 illustrates how the harmonisation of clinical data reported in clinical trials and clinical studies creates synergy between multiple stakeholders in the medicines access chain. A common approach to HTA clinical assessments in the EU through consolidated expertise and centralised REAs will circumvent redundant duplication of assessments conducted at individual national or regional levels, leading to enhanced resource use and greater sustainability in HTA operations. Standardising clinical outputs translates to higher quality data being aggregated for the intervention under review which is free from investigator bias, therefore having payers secure greater return on investment for the treatments covered.

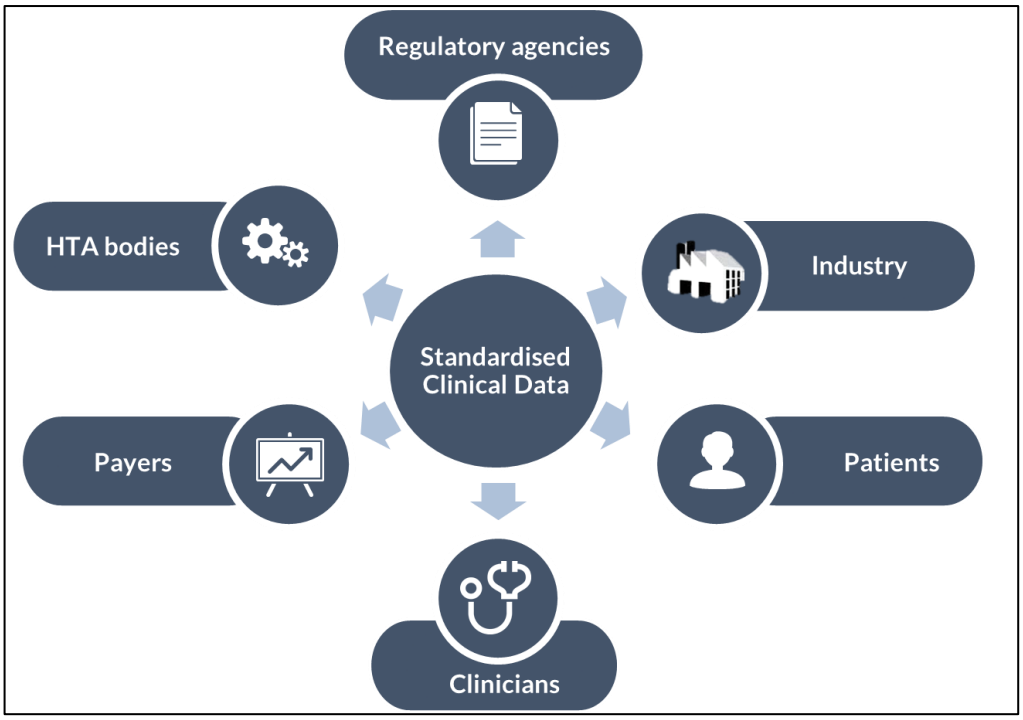


Figure 4.3 Multi-stakeholder impact of standardising clinical data

Similar to HTABs, regulatory and clinician decision-making is highly dependent on data and trends emerging from the use of the intervention in the real-world setting. Various challenges currently exist in the EU with regards to the development of patient

registries and protocols governing their data structures.[31] Identifying a standardised set of outcomes to be recorded in patient registries will facilitate the interoperability of repositories, the data of which is used as basis to inform post-authorisation regulatory, HTA and prescribing decisions.

For industry and patients, harmonising criteria amongst decision-makers is a win-win scenario. Through initiatives that streamline and increase visibility of regulatory and HTA evidence needs, pharmaceutical companies will be able to obviate costly parallel studies that are conducted to accommodate ancillary requests, therefore tapping earlier into the market with innovative products. Patients are the ultimate beneficiaries of such efforts towards cohesive clinical assessment decisions. A main factor contributing to improved longevity is the launch of new drugs, with an estimated 40% increase in life expectancy being attributable to the uptake of innovation (Lichtenberg, 2005). Rapid access to novel oncology and haematology medicines offers patients with therapeutic alternatives that provide enhanced disease management which results in better outcomes.

4.5 Strengths and Limitations of the Research

Pathological and molecular underpinnings of malignancies, together with studies confirming the validity of surrogate markers, are unravelled at a fast pace. For this reason, a prospective study design by means of the Delphi technique captured scientific opinions which are recent and reflecting current knowledge. The benefits of anonymity in the Delphi technique gives this method an edge compared to other consensus processes since the expression of opinions is free from domineering individuals and

[31] European Medicines Agency. Initiative for patient registries – Strategy and pilot phase [Internet]. September 2015 [cited 2019 Mar 10]. Available from: https://www.ema.europa.eu/en/documents/other/initiative-patient-registries-strategy-pilot-phase_en.pdf

socio-psychological pressures. The low attrition rate observed for the two rounds in both Delphi cycles was also considered a successful feature, with response rates being higher than published COS Delphi studies (Byrne et al, 2017; Kuizenga-Wessel et al, 2017; Meher et al, 2019). The broad time period selected for clinical trials analysed in this study allowed for the identification of patterns in clinical trial design over time.

This research is associated with some limitations. Efficacy in decision-making should be regarded as a means to an end and not itself the end since other elements such as safety aspects and economic factors also influence assessment outcomes. It has been stated that "it is important to ensure that views from all key stakeholder groups are included when making the final decision regarding the COS" (Williamson et al, 2012). Comprehensive stakeholder input with the involvement of patient groups, clinicians, industry representatives, payers and researchers, would have captured wider insight in selecting the core outcomes. Input from multiple stakeholder groups increases representativeness and thus the breadth of agreement among professionals in the supply chain, leading to an expedited bench-to-bedside trajectory for treatments under development. This study tackled the opinions of regulatory and HTA experts on efficacy endpoints only and therefore the final list of core outcomes cannot be regarded as a true COS. Another limitation is that the participation of HTABs from 9 out of 28 EU Member States may limit the generalisability of results to non-participating European HTA institutions.

Suboptimal data input by persons registering the clinical trial protocol details into the EudraCT may also introduce an element of inherent error in the study. "Consensus should not be seen as the "correct answer", and should be interpreted as the opinion of a specific group of experts on a given topic" (Du Plessis and Human, 2007). It is recommended that the results of a Delphi process are validated further in structured

discussions by means of workshops, focus groups, interviews or other ways (Hasson et al, 2000; Williamson et al, 2017). A face-to-face meeting to discuss the resulting COS list has however been recommended as an optional step (Williamson et al, 2012). A post-Delphi consensus meeting was not held in this study which might have raised certain issues for debate on the applicability of the identified core outcomes.

4.6 Recommendations for Future Work

This research has mainly focused on 'what' efficacy outcomes are prioritised by decision-makers for leukaemia trials yet future work is needed on 'how' and 'when' such parameters are to be measured, that is, what instruments should be used and at what intervals should measurements be taken (Gargon et al, 2014; Spargo et al, 2018). The development of core outcomes is not final since iterative revision by means of periodic reviews of the identified measures is warranted as a form of validation to ensure that the core outcomes are still relevant and important and in order to incorporate new emerging endpoints deemed critical by stakeholders (Williamson et al, 2012).

This study is in line with the priority areas for collaboration of the joint EMA and EUnetHTA 2017-2020 three-year work plan on pre- and post-authorisation evidence generation, with the aim of guiding medicines developers to generate evidence which is able to address both regulatory and HTA information needs.[32] "Achieving consistency is not something that can be left to serendipity. It will require consensus, guidelines and adherence." (Clarke, 2007). Benefits of the identified core outcomes can only be realised through successful mobilisation in clinical studies by engaging with relevant stakeholder groups and devising implementation strategies.

[32] European Network for Health Technology Assessment (EUnetHTA) and European Medicines Agency (EMA). EMA-EUnetHTA three-year work plan 2017-2020. November 2017 [cited 2019 May 10]. Available from: https://www.ema.europa.eu/documents/other/ema-eunethta-three-year-work-plan-2017-2020_en.pdf

Multiple factors such as therapeutic intent (induction, neo-adjuvant, consolidation, post-remission, maintenance or palliative therapy), stage of disease (acute, chronic, refractory, advanced or metastatic), patient's age, natural history of the disease, line of therapy, genetic factors and rarity of the disease all influence the choice of outcomes to be studied. This study sought to determine which outcomes, overarching and common to the main leukaemia subtypes, are prioritised by regulatory and HTA decision-makers. Further research is warranted to determine those criteria that are applicable to more specific therapeutic settings within the clinical area of leukaemia. Investigators should steer away from a 'one-size fits all' approach in endpoint selection and tailor evidence-generating strategies according to the specific disease characteristics.

Taking the natural history of the four main leukaemia subtypes as an example, chronic lymphocytic leukaemia, acute lymphoblastic leukaemia in children and adolescents under 15 years of age and chronic myeloid leukaemia are associated with a longer five-year relative survival rate compared with acute myeloid leukaemia and adult acute lymphoblastic leukaemia which present as more aggressive forms and are generally associated with a poor prognosis.[33] In the latter setting, overall survival would be the preferred parameter to detect any incremental benefits in survival brought about by the test intervention in comparison with the control agent. For leukaemias characterised by indolent disease development or extended longevity, surrogate markers proven to predict the effects on survival would be suitable primary efficacy endpoints to overcome excessively long trial durations.

[33] Leukaemia and Lymphoma Society. Facts and statistics: Leukaemia [Internet]. [cited 2019 Mar 10]. Available from: https://www.lls.org/http%3A/llsorg.prod.acquia-sites.com/facts-and-statistics/facts-and-statistics-overview/facts-and-statistics#Leukemia

4.7 Conclusions

Clinical trial design in leukaemia is moving away from traditional methods of assessing efficacy and heralding an era of "predictalytics". Evidence-generating strategies have evolved efficacy studies from being focused on measuring long-term efficacy outcomes to examining novel markers and composite endpoints that bypass the maturation of survival data, allowing for more rapid assessments. The lack of robustly validated surrogate endpoints presents as a caveat to this approach in clinical development, leading to the acceptance of surrogate markers if verified as true correlates to patient-relevant outcomes.

Features of efficacy studies in leukaemia trials were not fully aligned with expectations of regulatory and HTA experts. This was reflected by decision-makers expressing conflicting opinions on the quality of evidence being generated. Pharmaceutical companies should consider seeking to harness both regulatory and HTA positions along the course of clinical development to obtain authorisation and reimbursement approvals.

Incongruences were also observed between decision-makers for certain efficacy parameters being studied in leukaemia trials, which may heighten research costs and limit investment in clinical development by pharmaceutical companies. The elucidation of core outcomes shared by regulatory and HTA agencies is a step forward towards streamlining evidence requirements. Following this pathway facilitates the strategic planning by sponsors of clinical trials for medicines in leukaemia by integrating the scientific views of both decision-makers earlier in clinical development phases. Optimising clinical trial data packages can potentially translate to favourable clinical assessment decisions which in turn spur research initiatives at the benefit of leukaemia patients, public health and society at large.

References

Ades F, Zardavas D, Senterre C, De Azambuja E, Eniu A, Popescu R, et al. Hurdles and delays in access to anti-cancer drugs in Europe. Ecancermedicalscience. 2014;8:11 pages.

Adkins EM, Nicholson L, Floyd D, Ratcliffe M, Chevrou-Severac H. Oncology drugs for orphan indications: how are HTA processes evolving for this specific drug category? Clinicoecon Outcomes Res. 2017;9:327-342.

Akwaa F, Liesveld J. Surrogate end points for long-term outcomes in chronic myeloid leukemia. Leuk Lymphoma. 2013;54(10):2103-2111.

Allen N, Liberti L, Walker SR, Salek S. A comparison of reimbursement recommendations by European HTA agencies: Is there opportunity for further alignment? Front Pharmacol. 2017a;8:14 pages.

Allen N, Walker SR, Liberti L, Salek S. Health technology assessment (HTA) case studies: Factors influencing divergent HTA reimbursement recommendations in Australia, Canada, England and Scotland. Value Health. 2017b;20(3):320-328.

Ambler GK, Bosanquet DC, Brookes-Howell L, Thomas-Jones E, Waldron CA, Edwards AGK, et al. Development of a core outcome set for studies involving patients undergoing major lower limb amputation for peripheral arterial disease: study protocol for a systematic review and identification of a core outcome set using a Delphi survey. Trials. 2017;18(1):7 pages.

Apolone G, Joppi R, Bertele V, Garattini S. Ten years of marketing approvals of anticancer drugs in Europe: regulatory policy and guidance documents need to find a balance between different pressures. Br J Cancer. 2005;93(5):504-509.

Arber DA, Orazi A, Hasserjian R, Thiele J, Borowitz MJ, Le Beau MM, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukaemia. Blood. 2016;127(20):2391-2405.

Beauchemin C, Johnston JB, Lapierre ME, Aissa F, Lachaine J. Relationship between progression-free survival and overall survival in chronic lymphocytic leukemia: a literature-based analysis. Curr Oncol. 2015;22(3):e148-156.

Berggren R, Moller M, Moss R, Poda P, Smietana K. Outlook for the next 5 years in drug innovation. Nat Rev Drug Discov. 2012;11(6):435-436.

Bergmann L, Enzmann H, Broich K, Hebborn A, Marsoni S, Goh L, et al. Actual developments in European regulatory and health technology assessment of new cancer drugs: What does this mean to oncology? Ann Oncol. 2014;25:303-306.

Berntgen M, Gourvil A, Pavlovic M, Goettsch W, Eichler HG, Kristensen FB. Improving the contribution of regulatory assessment reports to health technology assessments – A collaboration between the European Medicines Agency and the European Network for Health Technology Assessment. Value Health. 2014;17:634-641.

Blumenthal GM, Kluetz PG, Schneider J, Goldberg KB, McKee AE, Pazdur R. Oncology drug approvals: Evaluating endpoints and evidence in an era of breakthrough therapies. Oncologist. 2017;22(7):762-767.

Boers M, Kirwan JR, Wells G, Beaton D, Gossec L, D'Agostino MA, et al. Developing core outcome measurement sets for clinical trials: OMERACT filter 2.0. J Clin Epidemiol. 2014;67(7):745-53.

Booth CM, Eisenhauer EA. Progression-free survival: Meaningful or simply measurable? J Clin Oncol. 2012;30(10):1030-1033.

Boulkedid R, Abdoul H, Loustau M, Sibony O, Alberti C. Using and reporting the Delphi method for selecting healthcare quality indicators: a systematic review. 2011;6(6):9 pages.

Byrne M, O'Connell A, Egan AM, Dinneen SF, Hynes L, O'Hara MC, et al. A core outcomes set for clinical trials of interventions for young adults with type 1 diabetes: an international, multi-perspective Delphi consensus study. Trials. 2017;18:8 pages.

Carlson RH. Stronger correlation of surrogate endpoints with survival needed. Lancel Oncol. 2016;17(7):e273.

Cassar Flores A, Marshall S, Cordina M. Use of the Delphi technique to determine safety features to be included in a neonatal and paediatric prescription chart. Int J Clin Pharm. 2014;36(6):1179–1189.

Chabot I, Rocchi A. Oncology drug health technology assessment recommendations: Canadian versus UK experiences. Clinicoecon Outcomes Res. 2014;6:357-367.

Chaffin WW, Talley WK. Individual stability in Delphi studies. Technol Forecast Soc Change. 1980;16:67-73.

Chan AW, Altman DG. Identifying outcome reporting bias in randomised trials on PubMed: review of publications and survey of authors. BMJ. 2005;330(7494):8 pages.

Cherny NI, Sullivan R, Dafni U, Kerst JM, SobreroA, Zielinski C, et al. A standardised, generic, validated approach to stratify the magnitude of clinical benefit that can be anticipated from anti-cancer therapies: the European Society for Medical Oncology Magnitude of Clinical Benefit Scale (ESMO-MCBS). Ann Oncol. 2015;26(8):1547-1573.

Cheung MC, Cheng S, McDonald E, Arciero V, Mahin Q, Chan K. Assessment of value using the American Society of Clinical Oncology (ASCO) and European Society of Medical Oncology (ESMO) frameworks for novel therapies for the hematologic malignancies. Blood. 2016;128:3578.

Chiarotto A, Terwee CB, Deyo RA, Boers M, Lin CW, Buchbinder R, et al. A core outcome set for clinical trials on non-specific low back pain: study protocol for the development of a core domain set. Trials. 2014;15:9 pages.

Cho Y, Sautenet B, Rangan G, Craig JC, Ong ACM, Chapman A, et al. Standardised Outcomes in Nephrology – Polycystic Kidney Disease (SONG-PKD): study protocol for eastablishing acore outcomes set in polycystic kidney disease. Trials. 2017;18(1):11 pages.

Clarke M, Williamson PR. Core outcome sets and systematic reviews. Syst Rev. 2016;5:4 pages.

Clarke M. Standardising outcomes for clinical trials and systematic reviews. Trials. 2007;8:3 pages.

Connolly B, Denehy L, Hart N, Pattison N, Williamson P, Blackwood B. Physical Rehabilitation Core Outcomes in Critical illness (PRACTICE): protocol for development of a core outcome set. Trials. 2018;19(1):8 pages.

Cressman S, Browman GP, Hoch JS, Kovacic L, Peacock SJ. A time-trend economic analysis of cancer drug trials. Oncologist. 2015;20(2):729-736.

Cuche M, Beckerman R, Chowdhury CA, Van Weelden MA. Early dialogue with health technology assessment bodies: a European perspective. Int J Technol Assess Health Care. 2014;30(6):571-580.

Dajani JS, Sincoff MZ, Talley WK. Stability and agreement criteria for the termination of Delphi studies. Technol Forecast Soc Chang. 1979;13:83-90.

Dalkey NC. Studies in the quality of life: Delphi and decision-making. Lexington, MA: Lexington Books; 1972. p.15.

Dean B, Barber N, Schachter M. What is a prescribing error? Qual Health Care. 2000;9:232–237.

Delbecq AL, Van de Ven AH, Gustafson DH. Group techniques for program planning: a guide to nominal group and Delphi processes. Glenview, Illinois: Scott, Foresman; 1975. p.52.

Denis A, Mergaert L, Fostier C, Cleemput I, Simoens S. A comparative study of European rare disease and orphan drug markets. Health policy. 2010;97(2-3):173-179.

Diamond IR, Grant RC, Feldman BM, Pencharz PB, Ling SC, Moore AM, et al. Defining consensus: a systematic review recommends methodologic criteria for reporting Delphi studies. J Clin Epidemiol. 2014;67(4):401-409.

Donohoe H, Stellefson M, Tennant B. Advantages and limitations of the e-Delphi technique: Implications for health education researchers. Am J Health Educ. 2012;43(1):38-46.

Dranitsaris G, Papadopoulos G. Health technology assessment of cancer drugs in Canada, the United Kingdom and Australia: Should the United States take notice? Appl Health Econ Health Policy. 2015;13:291-302.

Du Plessis E, Human SP. The art of the Delphi technique: Highlighting its scientific merit. Health SA Gesondheid. 2007;12(4):13-24.

Eichler HG, Oye K, Baird LG, Abadie E, Brown J, Drum CL, et al. Adaptive licensing: taking the next step in the evolution of drug approval. Clin Pharmacol Ther. 2012;91(3):426-437.

Elvidge S. EMA's parallel advice workshop bridges regulatory and reimbursement divide. Nat Rev Drug Discov. 2014;13(1):1 page.

Fish R, Sanders C, Williamson PR, Renehan AG. Core outcome research measures in anal cancer (CORMAC): protocol for systematic review, qualitative interviews and Delphi survey to develop a core outcome set in anal cancer. BMJ Open. 2017;7(11):8 pages.

Fiteni F, Westeel V, Pivot X, Borg C, Vernerey D, Bonnetain F. Endpoint in cancer clinical trials. J Visc Surg. 2014;151(1):17-22.

Fronsdal K, Pichler F, Marshani-Bayne L, Henshall C, Rottingen JA, Morland B, et al. Interaction initiatives between regulatory, health technology assessment and coverage bodies, and industry. Int J Technol Assess health Care. 2012;28(4):374-381.

Gagnier JJ, Page MJ, Huang H, Verhagen AP, Buchbinder R. Creation of a core outcome set for clinical trials of people with shoulder pain: a study protocol. Trials. 2017;18(1):9 pages.

Gargon E, Crew R, Burnside G, Williamson PR. Higher number of items associated with significantly lower response rates in COS Delphi surveys. J Clin Epidemiol. 2019;108:110-120.

Gargon E, Guring B, Medley N, Altman DG, Blazeby JM, Clarke M, et al. Choosing important health outcomes for comparative effectiveness research: A systematic review. PLoS One. 2014;9(6):12 pages.

Gargon E, Williamson PR, Young B. Improving core outcome set development: qualitative interviews with developers provided pointers to inform guidance. J Clin Epidemiol. 2017;86:140-152.

Giannarou L, Zervas E. Using Delphi technique to build consensus in practice. Int J Business and Applied Management. 2014;9(2):65-82.

Gill FJ, Leslie GD, Grech C, Latour JM. Using a web-based survey tool to undertake a Delphi study: Application for nurse education research. Nurse Educ Today. 2013;33(11):1322-1328.

Greatorax J, Dexter T. An accessible analytical approach for investigating what happens between the rounds of a Delphi study. J Adv Nurs. 2000;32(4):1016-1024.

Harman NL, Bruce IA, Callery P, Tierney S, Sharif MO, O'Brien K, et al. MOMENT-Management of Otitis Media with Effusion in Cleft Palate: protocol for a systematic review of the literature and identification of a core outcome set using a Delphi survey. Trials. 2013;14:8 pages.

Harman NL, Wilding J, Curry D, Harris J, Logue J, Pemberton RJ, et al. Selecting Core Outcomes for Randomised Effectiveness trials In Type 2 Diabetes (SCORE-IT): study protocol for the development of a core outcomes set. Trials. 2018;19(1):6 pages.

Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. J Adv Nurs. 2000;32(4):1008-1015.

Helms C, Gardner A, McInnes E. The use of advanced web-based survey design in Delphi research. J Adv Nurse. 2017;73(12):3168-3177.

Henshall C, Mardhani-Bayne L, Fronsdal KB, Klemp M. Interactions between health technology assessment, coverage, and regulatory processes: emerging issues, goals, and opportunities. Int J Technol Assess Health Care. 2011;27(3):253-260.

Hirsch BR, Califf RM, Cheng SK, Tasneem A, Horton J, Chiswell K, et al. Characteristics of oncology clinical trials: insights from a systematic analysis of ClinicalTrials.gov. JAMA Intern Med. 2013;173(11):972-979.

Hirsch M, Duffy JM, Barker C, Hummelshoj L, Johnson NP, Mol B, et al. Protocol for developing, disseminating and implementing a core outcome set for endometriosis. BMJ Open. 2016;6(12):e013998.

Holloway K. Doing the E-Delphi: using online survey tools. Comput Inform Nurs. 2012;30(7):347-350.

Iyengar S, Williamson PR, Schmitt J, Johannsen L, Maher IA, Sobanko JF, et al. Development of a core outcome set for clinical trials in rosacea: study protocol for a systematic review of the literature and identification of a core outcome set using a Delphi survey. Trials. 2016;17(1):6 pages.

Johnson JR, Williams G, Pazdur R. End points and United States Food and Drug Administration approval of oncology drugs. J Clin Oncol. 2003;21(7):1404-1411.

Joppi R, Gerardi C, Bertele V, Garattini S. Letting post-marketing bridge the evidence gap: the case of orphan drugs. BMJ. 2016;353:5 pages.

Jönsson B. Bringing in health technology assessment and cost-effectiveness considerations at an early stage of drug development. Mol Oncol. 2015;9(5):1025-1033.

Kanavos P, Nicod E, Vand den Aardweg S, Pomedli S. The impact of health technology assessments: an international comparison. Euro Observer. 2010;12(4):1-6.

Keeney S, Hasson F, McKenna H. Consulting the oracle: ten lessons from using the Delphi technique in nursing research. J Adv Nurs. 2006;53(2):205-212.

Keeney S, Hasson F, McKenna H. The Delphi technique in nursing and health research. West Sussex, UK: Wiley-Blackwell; 2011. p.53.

Kim C, Prasad V. Cancer drugs approved on the basis of a surrogate end point and subsequent overall survival: An analysis of 5 years of US Food and Drug Administration approvals. JAMA Intern Med. 2015;175(12):1992-1994.

Kirkham JJ, Davis K, Altman DG, Blazeby JM, Clarke M, Tunis S, et al. Core Outcomes Set-STAndards for Development: The COS-STAD recommendations. PLoS Med. 2017;14(11):10 pages.

Kirkham JJ, Gorst S, Altman DG, Blazeby Jm, Clarke M, Devane D, et al. Core Outcome Set-STAandards for Reporting: The COS-STAR Statement. PLoS Med. 2016;13(10):11 pages.

Kleijnen S, Fathallah M, Van dr Linden MW, Vancraeynest P, Dahmani B, Timoney A, et al. Can a joint assessment provide relevant information for national/local relative effectiveness assessments? An in-depth comparison of pazopanib assessmenets. Value Health. 2015;18(5):663-672.

Kleijnen S, Lipska I, Leonardo Alves T, Meijboom K, Elsada A, Vervolgyi V, et al. Relative effectiveness assessments of oncology medicines for pricing and reimbursements decisions in European countries. Ann Oncol. 2016;27(9):1768-1775.

Kuizenga-Wessel S, Steutel NF, Benninga MA, Devreker T, Scarpato E, Staiano A, et al. Development of a core outcome set for clinical trials in childhood constipation: a study using a Delphi technique. BMJ Paediatr Open. 2017;1(1):10 pages.

Leyens L, Brand A. Early patient access to medicines: Health technology assessment bodies need to catch up with new marketing authorization methods. Public Health Genomics. 2016;19:187-191.

Liberti L, Stolk P, McAuslane N, Somauroo A, Breckenridge AM, Leufkens HG. Adaptive licensing and facilitated regulatory pathways: A survey of stakeholder perceptions. Clin Pharmacol Ther. 2015;98(5):477-479.

Lichtenberg FR. The impact of new drug launches on longevity: Evidence from longitudinal, disease-level data from 52 countries, 1982-2001. Int J Health Care Finance Econ. 2005;5:47-73.

Lim CS, Lee YG, Koh Y, Heo DS. International comparison of the factors influencing reimbursement of targeted anti-cancer drugs. BMC Health Serv Res. 2014;14:10 pages.

Linstone HA, Turoff M. Introduction to the Delphi method: Techniques and applications. Reading, MA: Addison-Wesley Publishing Company; 1975. p.3-12.

Luengo-Fernandez R, Leal J, Gray A, Sullivan R. Economic burden of cancer across the European Union: a population-based cost analysis. Lancet Oncol. 2013;14(12):1165-1174.

Lynn MR. Determination and quantification of content validity. Nurs Res. 1986;35(6):382-385.

Ma C, Panaccione R, Fedorak RN, Parker CE, Khanna R, Levesque BG, et al. Development of a core outcome set for clinical trials in inflammatory bowel disease: study protocol for a systematic review of the literature and identification of a core outcome set using a Delphi survey. BMJ Open. 2017;7(6):7 pages.

MacLennan S, Bekema HJ, Williamson PR, Campbell MK, Stewart F, MacLennan SJ, et al. A core outcome set for localised prostate cancer effectiveness trials: protocol for a systematic review of the literature and stakeholder involvement through interviews and a Delphi survey. Trials. 2015;16:10 pages.

Martinalbo J, Bowen D, Camarero J, Chapelin M, Démolis P, Foggi P, et al. Early market access of cancer drugs in the EU. Ann Oncol. 2016;27:96-105.

McMillan SS, King M, Tully MP. How to use the nominal group and Delphi techniques. Int J Clin Pharm. 2016;38(3):655-662.

Meher S, Cuthbert A, Kirkham JJ, Williamson P, Abalos E, Aflaifel N, et al. Core outcome sets for prevention and treatment of postpartum haemorrhage: an international Delphi consensus study. BJOG. 2019;126(1):83-93.

Morris C, Dunkley C, Gibbon FM, Currier J, Roberts D, Rogers M, et al. Core Health Outcomes In Childhood Epilepsy (CHOICE): protocol for the selection of a core outcome set. Trials. 2017;18(1):8 pages.

Murad MH, Asi N, Alsawas M, Alahdab F. New evidence pyramid. Evid Based Med. 2016;21(4):125-127.

Neumann PJ, Bliss SK, Chambers JD. Therapies for advanced cancers pose a special challenge for health technology assessment organizations in many countries. Health Aff (Millwood). 2012;31(4):700-708.

Ording AG, Cronin-Fenton D, Ehrenstein V, Lash TL, Acquavella J, Rorth M, et al. Challenges in translating endpoints from trials to observational studies in oncology. Clin Epidemiol. 2016;8:195-200.

Oriana C, Martin H, Toby P, Chris C, Ruth G, Claudius R, et al. Complete cytogenetic response and major molecular response as surrogate outcomes for overall survival in first-live treatment of chronic myelogenous leukemia: A case study for technology appraisal on the basis of surrogate outcomes evidence. Value Health. 2013;16(6):1081-1090.

Pauwels K, Huys I, Casteels M, Larsson K, Voltz C, Penttila K, et al. Are products with an orphan designation for oncology indications different from products for other rare indications? A retrospective analysis of European orphan designations granted between 2002-2012. Orphanet J Rare Dis. 2017;12(1):7 pages.

Pignatti F, Jonsson B, Blumethal G, Justice R. Assessment of benefits and risks in development of targeted therapies for cancer – The view of regulatory authorities. Mol Oncol. 2015;9(5):1034-1041.

Polit DF, Beck CT. The content validity index: are you sure you know what's being reported? Critique and recommendations. Res Nurs Health. 2006;29(5):489-497.

Powell C. The Delphi technique: myths and realities. J Adv Nurs. 2003;41(4):376-382.

Prasad V, De Jesus K, Mailankody S. The high price of anticancer drugs: Origins, implications, barriers, solutions. Nat Rev Clin Oncol. 2017;14(6):381-390.

Prentice RL. Surrogate endpoints in clinical trials: definition and operational criteria. Stat Med. 1989;8(4):431-440.

Rawlins MD. Crossing the fourth hurdle. Br J Clin Pharmacol. 2012;73(6):855-860.

Rocchi A, Chabot I, Glennie J. Evolution of health technology assessment: Best practices of the pan-Canadian oncology drug review. Clinicoecon Outcomes Res. 2015;7:287-298.

Ruof J, Knoerzer D, Dunne AA, Dintsios CM, Staab T, Schwartz FW. Analysis of endpoints used in marketing authorisations versus value assessments of oncology medicines in Germany. Health Policy. 2014;118(2):242-254.

Schlessinger DI, Iyengar S, Yanes AF, Lazaroff JM, Godinez-Puig V, Chen BR, et al. Development of a core outcome set for clinical trials in basal cell carcinoma: study protocol for a systematic review of the literature and identification of a core outcome set using a Delphi survey. Trials. 2017a;18(1):6 pages.

Schlessinger DI, Iyengar S, Yanes AF, Chiren SG, Godinez-Puig V, Chen BR, et al. Development of a core outcome set for clinical trials in squamous cell carcinoma: study protocol for a systematic review of the literature and identification of a core outcome set using a Delphi survey. Trials. 2017b;18(1):6 pages.

Schmitt J, Apfelbacher C, Spuls PI, Thomas KS, Simpson EL, Furue M, et al. The Harmonizing Outcomes Measures for Eczema (HOME) roadmap: A methodological framework to develop core sets of outcome measurements in dermatology. J Invest Deramtol. 2015;135(1):24-30.

Schnipper LE, Davidson NE, Wollins DS, Tyne C, Blayney DW, Blum D, et al. American Society of Clinical Oncology statement: A conceptual framework to assess the value of cancer treatment options. J Clin Oncol. 2015;33(23):2563-2577.

Shah KK, Mestre-Ferrandix J, Towse A, Smyth EN. A review of health technology appraisals: case studies in oncology. Int J Technol Assess Health Care. 2013;29(1):101-109.

Sherratt FC, Eaton S, Walker E, Beasant L, Blazeby JM, Young B, et al. Development of a core outcome set to determine the overall treatment suc cess of acute uncomplicated appendicitis in children: a study protocol. BMJ Paediatr Open. 2017;1(1):5 pages.

Sinha IP, Smyth RL, Williamson PR. Using the Delphi technique to determine which outcomes to measure in clinical trials: Recommendations for the future based on a systematic review of existing studies. PLoS Med. 2011;8(1):5 pages.

Skulmoski GJ, Hartman FT, Krahn J. The Delphi method for graduate research. JITE: Research. 2007;6:1-21.

Smail-Faugeron V, Chabouis HF, Durieux P, Attal JP, Muller-Bolla M, Courson F. Development of a core set of outcomes for randomized controlled trials with multiple outcomes – Example of pulp treatments of primary teeth for extensive decay in children. PLoS One. 2013;8(1):9 pages.

Sorenson C, Chalkidou K. Reflections on the evolution of health technology assessment in Europe. Health Econ Policy Law. 2012;7:25-45.

Sorenson C. Valuing end-of-life care in the United States: the case of new cancer drugs. Health Econ Policy Law. 2012;7(4):411-430.

Spargo M, Ryan C, Downey D, Hughes C. Development of a core outcome set for trials investigating the long-term management of bronchiectasis. Chron Respir Dis. 2018;16:1-9.

Starr P. The haematologic drug pipeline: Exciting new treatment options looming. Am Health Drug Benefits. 2016;9:1-20.

Stewart DJ, Whitney SN, Kurzrock R. Equipose lost: ethics, costs, and the regulation of cancer clinical research. J Clin Oncol. 2010;28:2925-2935.

Swerdlow SH, Campo E, Pileri SA, Harris NL, Harald S, Siebert R, et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. Blood. 2016;127(20):2375-2390.

Tafuri G, Pagnini M, Moseley J, Massari M, Petavy F, Behring A, et al. How aligned are the perspectives of EU regulators and HTA bodies? A comparative analysis of regulatory-HTA parallel scientific advice. Br J Clin Pharmacol. 2016;82:965-973.

Tallouzi MO, Mathers JM, Moore DJ, Murray PI, Bucknall N, Blazeby JM, et al. COSUMO: study protocol for the development of a core outcome set for efficacy and effectiveness trials in posterior segment-involving uveitis. Trials. 2017;18(1):9 pages.

Tong A, Manns B, Hemmelgarn B, Wheeler DC, Tugwell P, Winkelmayer WC, et al. Standardised outcomes in nephrology – Haemodialysis (SONG-HD): study protocol for establishing a core outcome set in haemodialysis. Trials. 2015;16:9 pages.

Umscheid C, Margolis D, Grossman C. Key concepts of clinical trials: A narrative review. Postgrad Med. 2011;123(5):194-204.

Van den Bussche K, De Meyer D, Van Damme N, Kottner J, Beeckman D. CONSIDER – Core Outcome Set in IAD Research: study protocol for establishing a core set of outcomes and measurements in incontinence-associated dermatitis research. J Adv Nurs. 2017;73(10):2473-2483.

Van Harten WH, Retel VP. Innovations that reach the patient: early health technology assessment and improving the chances of coverage and implementation. Ecancermedicalscience. 2016;10:6 pages.

Van Rol RR, Melenhorst J, Dirksen CD, Stassen LPS, Breukink SO. Protocol for the development of a Core Outcome Set (COS) for hemorrhoidal disease: an international Delphi study. Int J Colorectal Dis. 2017;32(7):1091-1094.

Von der Gracht HA. Consensus measurement in Delphi studies: Review and implications for future quality assurance. Technol Forecast Soc Change. 2012;79(8):1525-1536.

Waltz CF, Strickland OL, Lenz ER. Measurement in nursing and health research. 3rd edition. 2005. p.178.

Wang T, McAuslane N, Liberti L, Leufkens H, Hovels A. Building synergy between regulatory and HTA agencies beyond processes and procedures – Can we effectively align the evidentiary requirements? A survey of stakeholder perceptions. Value Health. 2018;21(6):707-714.

Waters AM, Tudor Smith C, Young B, Jones TM. The CONSENSUS study: protocol for a mixed methods study to establish which outcomes should be included in a core outcome set for oropharyngeal cancer. Trials. 2014;15:9 pages.

Williamson PR, Altman DG, Bagley H, Barnes KL, Blazeby J, Brookes ST, et al. The COMET Handbook: version 1.0. Trials. 2017;18(3):50 pages.

Williamson PR, Altman DG, Blazeby JM, Clarke M, Devane D, Gargon E, et al. Developing core outcome sets for clinical trials: issues to consider. Trials. 2012;13:8 pages.

Williamson PR, Gamble C, Altman DG, Hutton JL. Outcome selection bias in meta-analysis. Stat Methods Med Res. 2005;14(5):515-524.

Wilson MK, Collyar D, Chingos DT, Friedlander M, Ho TW, Karakasis K, et al. Outcomes and endpoints in cancer trials: bridging the divide. Lancet Oncol. 2015a;16(1):e43-52.

Wilson MK, Karakasis K, Oza AM. Outcomes and endpoints in trials of cancer treatment: the past, present, and future. Lancet Oncol. 2015b;16(1):e32-e42.

Appendix 1

Approvals and consents

L-Università ta' Malta

Faculty of Medicine & Surgery

University of Malta
Msida MSD 2080, Malta

Tel: +356 2340 1879/1891/1167
umms@um.edu.mt

www.um.edu.mt/ms

Ref No: **FRECMDS_1718_044**

Monday 30th April 2018

Mr Dylan Said
Ta' San Tumas
Block 2, Flat 3
Triq il-Gandoffli
Marsascala MSK3021

Dear Mr Said,

Please refer to your application submitted to the Research Ethics Committee in connection with your research entitled:

Regulatory and Health Technology Assessment Perspectives on Measures of Clinical Benefit in Antineoplastic Therapy

The Faculty Research Ethics Committee granted ethical approval for the above mentioned protocol.

Yours sincerely,

Dr. Mario Vassallo
Chairman
Research Ethics Committee

L-Università ta' Malta

Dylan Said <dylan.said.11@um.edu.mt>

RE: MDH DPO Approval

1 message

Data Protection at MDH <datapro.mdh@gov.mt> 4 July 2018 at 06:52
To: Dylan Said <dylan.said.11@um.edu.mt>
Cc: Young Sharon at Health-MDH <sharon.young@gov.mt>, Aquilina Graziella at Health-MDH <graziella.aquilina@gov.mt>, Falzon Edward at Health-SAMOC <edward.falzon@gov.mt>

Dear Mr Said

On the basis of the documentation you submitted, from the MDH data protection point of view you have been cleared to proceed with your study provided that you obtain approval from MDH CEO.

Please contact Ms. Graziella Aquilina on 2545 5334 to present a copy of your approvals and fill in the appropriate Data Protection Form.

Remember that in no way should you retain any personal details you obtain from your research and this should be destroyed at the end of your study.

You are requested to submit a copy of your findings to this office at the end of your study.

NB: This email does not cover ethical approval.

Regards

Simon Caruana

A/Data Protection Officer

Health Informatics

SPTAR MATER DEI

T +356 25455340

E simon.caruana@gov.mt

Mater Dei Hospital, Triq tal-Qroqq, Msida, Malta MSD 2090 | Tel +356 2545 0000 | www.mdh.gov.mt

L-Università ta' Malta

Dylan Said <dylan.said.11@um.edu.mt>

RE: MDH CEO Approval

1 message

Falzon Ivan at Health-MDH <ivan.falzon@gov.mt> 4 July 2018 at 10:59
To: Dylan Said <dylan.said.11@um.edu.mt>
Cc: Data Protection at MDH <datapro.mdh@gov.mt>

proceed

Ivan Falzon
Chief Executive Officer | TeaMDH

SPTAR MATER DEI

T +356 2545 4102

M +356 9995 0393

E ivan.falzon@gov.mt

Mater Dei Hospital, Triq tal-Qroqq, Msida, Malta MSD 2090 | Tel +356 2545 0000 | www.careandcure.gov.mt

From: Dylan Said [mailto:dylan.said.11@um.edu.mt]
Sent: 04 July 2018 10:07
To: Falzon Ivan at Health-MDH <ivan.falzon@gov.mt>
Cc: Data Protection at MDH <datapro.mdh@gov.mt>
Subject: MDH CEO Approval

Dear Mr. Falzon,

As a second year Doctorate in Pharmacy (Pharm D) student undertaking my thesis project entitled 'Regulatory and Health Technology Assessment Perspectives on Measures of Clinical Benefit in Antineoplastic Therapy', I am hereby asking your kind permission to engage oncologists and haematologists in a preliminary validation phase of the study to review the data collection tool.

Clearance has already been obtained from the MDH Data Protection Office, Faculty of Medicine and Surgery Research Ethics Committee, Dr. Laspina as clinical chairman for the Department of Haematology and Oncology and the

SIR ANTHONY MAMO

FORM :	**Oncology Proposal/Approval Audit/ Research purposes**
Document Code: ONCO-GeFO-P/A-001. Ver.01	Reference SOP : ONCO-Ge-PD.AP--001.Ver.01

Heads of:
(Name, Surname and Section (in block letters) and Signature)

Radiotherapy Department: ***(if applicable)***

Radiography

Medical Physics

Nursing

Other SAMOC Departments/ Wards:

Clinical Chairperson (Haematology - Oncology):
Name and Surname (in block letters) and Signature:

Dr. Stefan Laspina
Clinical Chairman
Department of Haematology & Oncology

Quality Assurance Manager:
Name and Surname (in block letters) and Signature:

EDWARD FALZON

An approval is granted to carry out the study/audit at any SAMOC Department. Patient information can be accessible only by complying with the following data protection principles, which are set out in the Data Protection Act 2001. In summary these state that patient's data shall:

- ***Be obtained and processed fairly and lawfully and shall not be processed unless certain conditions are met. Therefore patient's information (including scans) should be made anonymous by an appointed radiotherapy staff (from the Head of section)***
- ***Be obtained for a specified and lawful purpose and shall not be processed in any manner incompatible with that purpose.***
- ***Be adequate, relevant and not excessive for those purposes (In the case of a study or audit).***
- ***Be accurate and kept up to date.***
- ***Not be kept longer than is necessary for that purpose***
- ***Be processed in accordance with the data subject's rights.***
- ***Be kept safe from unauthorised access, accidental loss or destruction.***
- ***Not be transferred to any third party unlawfully.***

Appendix 2

Pre- and post-validation versions of the Response Evaluation in Leukaemia (ReVALeu) online surveying tool

Pre-validation version of the ReVALeu tool:

L-Università ta' Malta | Department of Pharmacy
Faculty of Medicine & Surgery

Response Evaluation in Leukaemia (ReVALeu)

Introduction

Dear Participant,

Thank you for accepting to participate as an expert in the Delphi panel and for allocating time to complete the questionnaires. The Response Evaluation in Leukaemia (ReVALeu) online surveying tool has been developed and validated to determine the prioritised efficacy endpoints in clinical assessments of anti-leukaemic therapies among Health Technology Assessment (HTA) Bodies (HTABs) and between HTABs and regulators. The Delphi technique has been selected as the method for data collection which provides iteration and potentially directs respondents towards consensus.

Efficacy endpoints were extracted from clinical trials of anticancer agents investigated in established leukaemia indications for all populations as a representative case study for haematological malignancies. Characteristics of the selected clinical trials are detailed in Section 2.

This research is congruent with the priority areas for collaboration of the joint European Medicines Agency (EMA) and European Network for Health Technology Assessment (EUnetHTA) 2017-2020 three-year work plan on pre- and post-authorisation evidence generation, with the aim of guiding medicines developers to generate evidence which is able to address both regulatory and HTA information needs. The EMA-EUnetHTA three-year work plan may be accessed here.

Research ethics approval has been granted by the Faculty of Medicine and Surgery Research Ethics Committee at the University of Malta.

L-Università ta' Malta
Faculty of Medicine & Surgery
Department of Pharmacy

Response Evaluation in Leukaemia (ReVALeu)

Delphi Round 1 Instructions

This questionnaire is the first of two (2) rounds, in which experts are requested to rate statements or provide comments on both generic and technical aspects of assessments. The first round questionnaire is divided into two main sections as follows:

Section 1 (General Questions): Questions asked in this section carry a broad scope and relate to the dynamics of regulatory and HTA decision-makers, including their respective data needs, current state of play and the impact of aligning evidence requirements between both facets.

Section 2 (Technical Questions): This section tackles technical questions on efficacy endpoints used in leukaemia clinical trials. For each unique efficacy measure, there are three (3) multiple choice fields, including classification (primary or secondary), trial phase-specificity (Phase II-IV) and clinical importance, and one (1) optional field for comments justifying answers on the endpoint in question.

For both sections, fields that require an answer are marked with an asterisk (*) and must be answered before progressing with the questionnaire.

Only Section 2 statements on the importance of the endpoint in demonstrating efficacy that have not reached consensus will progress to the second Delphi round. A feedback loop reporting the panel's distribution of responses and a summary of comments will be provided to the experts after each round.

The estimated time to complete the questionnaire is approximately 20 minutes and each Delphi round will have a **deadline of three weeks** for participants to respond.

L-Università ta' Malta
Faculty of Medicine & Surgery
Department of Pharmacy

Response Evaluation in Leukaemia (ReVALeu)

*** 1. Indicate your domain:**

- ○ Regulatory (European Medicines Agency)
- ○ Health Technology Assessment

L-Università ta' Malta
Faculty of Medicine & Surgery
Department of Pharmacy

Response Evaluation in Leukaemia (ReVALeu)

Section 1: General Questions

*** 2. Do you agree that there is sufficient interaction between the regulatory and HTA facets in the clinical assessment of applications?**

To be answered by both HTABs and regulators.

(HTA: Health Technology Assessment; HTAB: Health Technology Assessment Body)

☐ Strongly disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly agree

*** 3. Do you agree that clinical evidence needs in assessments of oncology/haematology medicines are aligned between HTABs in EU Member States?**

To be answered by HTABs only.

(EU: European Union; HTAB: Health Technology Assessment Body)

☐ Strongly disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly agree

*** 4. Do you agree that clinical evidence needs in assessments of oncology/haematology medicines are aligned between regulators and HTABs in the EU?**

To be answered by both HTABs and regulators.

(EU: European Union; HTAB: Health Technology Assessment Body)

☐ Strongly disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly agree

*** 5. What is the perceived level of agreement of your HTAB on the acceptance of clinical data for oncology/haematology medicines relative to:**
i) Other HTABs?
ii) Regulators?

To be answered by HTABs only.

(HTAB: Health Technology Assessment Body)

	Other HTABs	Regulators
Strongly opposed	☐	☐
Opposed	☐	☐
Neutral	☐	☐
Favoured	☐	☐
Strongly favoured	☐	☐
Do not know	☐	☐

*** 6. Is the EMA European Public Assessment Report (EPAR) consulted during the HTA clinical evaluation of oncology/haematology medicines?**

To be answered by HTABs only.

(EMA: European Medicines Agency; HTA: Health Technology Assessment; HTAB: Health Technology Assessment Body)

☐ Never ☐ Rarely ☐ Occasionally ☐ Frequently ☐ Always

*** 7. Has your HTAB been involved in parallel scientific advice/consultation procedures with the EMA Scientific Advice Working Party (SAWP)?**

To be answered by HTABs only.

(EMA: European Medicines Agency; HTAB: Health Technology Assessment Body)

☐ Yes ☐ No

If yes, please provide an indication on the number of procedures involved related to oncology/haematology medicines:

*** 8. How would you rate the quality of evidence generated in the field of oncology/haematology during:**

i) Pre-authorisation/reimbursement phases?

ii) Post-authorisation/reimbursement phases?

To be answered by both HTABs and regulators.

	Pre-authorisation/reimbursement phases	Post-authorisation/reimbursement phases
Poor	☐	☐
Fair	☐	☐
Good	☐	☐
Very good	☐	☐
Excellent	☐	☐

*** 9. Do you agree that the risk of clinical assessment disharmony between regulatory and HTA procedures impacts patient access to novelties in anticancer therapy?**

To be answered by both HTABs and regulators.

(HTA: Health Technology Assessment; HTAB: Health Technology Assessment Body)

◯ Strongly disagree ◯ Disagree ◯ Neutral ◯ Agree ◯ Strongly agree

*** 10. From the following list, rank stakeholders from 1 to 6 (1 being highest and 6 being lowest) according to who is most likely to benefit from enhanced collaboration in clinical assessments between the regulatory and HTA facets:**

To be answered by both HTABs and regulators.

(HTA: Health Technology Assessment; HTAB: Health Technology Assessment Body)

⠿ ▲▼ Regulators

⠿ ▲▼ HTABs

⠿ ▲▼ Clinicians

⠿ ▲▼ Industry

⠿ ▲▼ Patients

⠿ ▲▼ Payers

*** 11. What factor(s) influence your HTAB recommendation outcomes (*more than one option may be selected*)?**

To be answered by HTABs only.

(HTAB: Health Technology Assessment Body)

- [ ] Cost-effectiveness
- [ ] Budgetary impact
- [ ] Relative effectiveness assessment (REA)
- [ ] Other (please specify)

Response Evaluation in Leukaemia (ReVALeu)

Section 2: Technical Questions

This set of questions relate to efficacy endpoints extracted from clinical trials in the EU Clinical Trials Register database over an 11-year period (January 2007 - December 2017), that have been selected after the application of inclusion criteria. In total, 431 clinical trials satisfied acceptance criteria and qualified for the study, having the following characteristics:

Clinical Trial Characteristic	% **Clinical Trials (N=431)**
Ongoing	61 (n=261)
Completed	23 (n=100)
Prematurely Ended/Restarted/Temporarily Halted/Not Specified	16 (n=70)
Trial Subjects Under 18 (birth-17 years)	15 (n=64)
Trial Subjects Adults (18-64 years)	86 (n=369)
Trial Subjects Elderly (≥65 years)	77 (n=330)

Thirty-six (36) unique efficacy endpoints have been identified and grouped into four principle categories, comprising:

- **Endpoint Category 1: Survival** (n=5)
- **Endpoint Category 2: Time-To-Event** (n=6)
- **Endpoint Category 3: Response Rates and Biomarkers** (n=16)
- **Endpoint Category 4: Other** (n=9)

Please refer to Appendix 1 attached for guidance on endpoint definitions.

L-Università ta' Malta
Faculty of Medicine & Surgery
Department of Pharmacy

Response Evaluation in Leukaemia (ReVALeu)

Section 2(a): Endpoint Category 1 (Survival)

*** 12. How would you classify this endpoint?**

	Primary	Secondary
Overall survival (OS)	○	○
Progression-free survival (PFS)	○	○
Event-free survival (EFS)	○	○
Disease-free survival (DFS)	○	○
Treatment-free survival (TFS)	○	○

Comments:

* **13. Indicate the most suitable clinical trial phase(s) for this endpoint (more *than one option may be selected*):**

	Phase II	Phase III	Phase IV
Overall survival (OS)	☐	☐	☐
Progression-free survival (PFS)	☐	☐	☐
Event-free survival (EFS)	☐	☐	☐
Disease-free survival (DFS)	☐	☐	☐
Treatment-free survival (TFS)	☐	☐	☐

Comments:

*** 14. Indicate the importance of this endpoint in demonstrating efficacy:**

	Not important at all	Not important	Neutral	Important	Very important
Overall survival (OS)	○	○	○	○	○
Progression-free survival (PFS)	○	○	○	○	○
Event-free survival (EFS)	○	○	○	○	○
Disease-free survival (DFS)	○	○	○	○	○
Treatment-free survival (TFS)	○	○	○	○	○

Comments:

L-Università ta' Malta
Faculty of Medicine & Surgery
Department of Pharmacy

Response Evaluation in Leukaemia (ReVALeu)

Section 2(b): Endpoint Category 2 (Time-To-Event)

*** 15. How would you classify this endpoint?**

	Primary	Secondary
Duration of remission/response (DOR)	☐	☐
Time to remission/response (TTR)	☐	☐
Time to progression (TTP)	☐	☐
Time to treatment (TTT)	☐	☐
Time to treatment failure (TTF)	☐	☐
Treatment-free remission/response (TFR)	☐	☐

Comments:

* **16. Indicate the most suitable clinical trial phase(s) for this endpoint *(more than one option may be selected)*:**

	Phase II	Phase III	Phase IV
Duration of remission/response (DOR)	☐	☐	☐
Time to remission/response (TTR)	☐	☐	☐
Time to progression (TTP)	☐	☐	☐
Time to treatment (TTT)	☐	☐	☐
Time to treatment failure (TTF)	☐	☐	☐
Treatment-free remission/response (TFR)	☐	☐	☐

Comments:

*** 17. Indicate the importance of this endpoint in demonstrating efficacy:**

	Not important at all	Not important	Neutral	Important	Very important
Duration of remission /response (DOR)	○	○	○	○	○
Time to remission/response (TTR)	○	○	○	○	○
Time to progression (TTP)	○	○	○	○	○
Time to treatment (TTT)	○	○	○	○	○
Time to treatment failure (TTF)	○	○	○	○	○
Treatment-free remission /response (TFR)	○	○	○	○	○

Comments:

L-Università ta' Malta
Faculty of Medicine & Surgery
Department of Pharmacy

Response Evaluation in Leukaemia (ReVALeu)

Section 2(c): Endpoint Category 3 (Response Rates and Biomarkers)

*** 18. How would you classify this endpoint?**

	Primary	Secondary
Relapse rate (RR)/Rate of progressive disease (PD)	☐	☐
Complete remission/response rate (CR)	☐	☐
Complete remission/response rate with incomplete platelet recovery (CRp)	☐	☐
Complete remission/response rate with incomplete blood count recovery (CRi/CRh)	☐	☐
Partial remission/response rate (PR)	☐	☐
Combined complete remission/response rate (CRc)	☐	☐
Objective remission/response rate (ORR)	☐	☐
Best overall response rate (BOR)	☐	☐
Haematologic response rate (HR)	☐	☐
Cytogenetic response rate (CyR)	☐	☐

	Primary	Secondary
Molecular remission/response (MR)	○	○
Minimal residual disease (MRD)	○	○
Cytokine serum concentrations	○	○
Morphologic leukaemia-free state (MLFS)	○	○
Splenomegaly and hepatomegaly response rates	○	○
Return to chronic phase (RTC)	○	○

Comments:

* **19. Indicate the most suitable clinical trial phase(s) for this endpoint *(more than one option may be selected)*:**

	Phase II	Phase III	Phase IV
Relapse rate (RR)/Rate of progressive disease (PD)	☐	☐	☐
Complete remission/response rate (CR)	☐	☐	☐
Complete remission/response rate with incomplete platelet recovery (CRp)	☐	☐	☐
Complete remission/response rate with incomplete blood count recovery (CRi/CRh)	☐	☐	☐
Partial remission/response rate (PR)	☐	☐	☐

	Phase II	Phase III	Phase IV
Combined complete remission/response rate (CRc)	☐	☐	☐
Objective remission/response rate (ORR)	☐	☐	☐
Best overall response rate (BOR)	☐	☐	☐
Haematologic response rate (HR)	☐	☐	☐
Cytogenetic response rate (CyR)	☐	☐	☐
Molecular remission/response (MR)	☐	☐	☐
Minimal residual disease (MRD)	☐	☐	☐
Cytokine serum concentrations	☐	☐	☐
Morphologic leukaemia-free state (MLFS)	☐	☐	☐
Splenomegaly and hepatomegaly response rates	☐	☐	☐
Return to chronic phase (RTC)	☐	☐	☐

Comments:

*** 20. Indicate the importance of this endpoint in demonstrating efficacy:**

	Not important at all	Not important	Neutral	Important	Very important
Relapse rate (RR)/Rate of progressive disease (PD)	○	○	○	○	○
Complete remission/response rate (CR)	○	○	○	○	○

Complete remission/response rate with incomplete platelet recovery (CRp)	☐	☐	☐	☐	☐
Complete remission/response rate with incomplete blood count recovery (CRi/CRh)	☐	☐	☐	☐	☐
Partial remission/response rate (PR)	☐	☐	☐	☐	☐
Combined complete remission/response rate (CRc)	☐	☐	☐	☐	☐
Objective remission/response rate (ORR)	☐	☐	☐	☐	☐
Best overall response rate (BOR)	☐	☐	☐	☐	☐
Haematologic response rate (HR)	☐	☐	☐	☐	☐
Cytogenetic response rate (CyR)	☐	☐	☐	☐	☐
Molecular remission/response (MR)	☐	☐	☐	☐	☐
Minimal residual disease (MRD)	☐	☐	☐	☐	☐
Cytokine serum concentrations	☐	☐	☐	☐	☐
Morphologic leukaemia-free state (MLFS)	☐	☐	☐	☐	☐
Splenomegaly and hepatomegaly response rates	☐	☐	☐	☐	☐
Return to chronic phase (RTC)	☐	☐	☐	☐	☐

Comments:

L-Università ta' Malta
Faculty of Medicine & Surgery
Department of Pharmacy

Response Evaluation in Leukaemia (ReVALeu)

Section 2(d): Endpoint Category 4 (Other)

*** 21. How would you classify this endpoint?**

	Primary	Secondary
Patient reported outcomes (PROs)	☐	☐
Number of patients with asparagine depletion	☐	☐
Secondary malignancies	☐	☐
Number of blood cell transfusions	☐	☐
Number of chemotherapy cycles received	☐	☐
Incidence and duration of febrile neutropenia	☐	☐
Medical care utilisation (MCU)	☐	☐
Incidence, duration and severity of opportunistic infections	☐	☐
Proportion of patients undergoing subsequent allogeneic stem cell transplant (ASCT)	☐	☐

Comments:

* **22. Indicate the most suitable clinical trial phase(s) for this endpoint *(more than one option may be selected)*:**

	Phase II	Phase III	Phase IV
Patient reported outcomes (PROs)	☐	☐	☐
Number of patients with asparagine depletion	☐	☐	☐
Secondary malignancies	☐	☐	☐
Number of blood cell transfusions	☐	☐	☐
Number of chemotherapy cycles received	☐	☐	☐
Incidence and duration of febrile neutropenia	☐	☐	☐
Medical care utilisation (MCU)	☐	☐	☐
Incidence, duration and severity of opportunistic infections	☐	☐	☐
Proportion of patients undergoing subsequent allogeneic stem cell transplant (ASCT)	☐	☐	☐

Comments:

*** 23. Indicate the importance of this endpoint in demonstrating efficacy:**

	Not important at all	Not important	Neutral	Important	Very important
Patient reported outcomes (PROs)	○	○	○	○	○
Number of patients with asparagine depletion	○	○	○	○	○
Secondary malignancies	○	○	○	○	○
Number of blood cell transfusions	○	○	○	○	○
Number of chemotherapy cycles received	○	○	○	○	○
Incidence and duration of febrile neutropenia	○	○	○	○	○
Medical care utilisation (MCU)	○	○	○	○	○
Incidence, duration and severity of opportunistic infections	○	○	○	○	○
Proportion of patients undergoing subsequent allogeneic stem cell transplant (ASCT)	○	○	○	○	○

Comments:

Thank you for completing the first Delphi round. A summary of the expert panel responses will be provided in the second (final) round.

Post-validation version of the ReVALeu tool:

L-Università ta' Malta | Faculty of Medicine & Surgery | Department of Pharmacy

Response Evaluation in Leukaemia (ReVALeu)

Introduction

Dear Participant,

Thank you for accepting to participate as an expert in the Delphi panel and for allocating time to complete the questionnaires. The Response Evaluation in Leukaemia (ReVALeu) online surveying tool has been developed and validated to determine the prioritised efficacy endpoints in clinical assessments of anti-leukaemic therapies among Health Technology Assessment (HTA) Bodies (HTABs) and between HTABs and the European Medicines Agency (EMA). The Delphi technique has been selected as the method for data collection which provides iteration and potentially directs respondents towards consensus.

Efficacy endpoints were extracted from clinical trials of anticancer agents investigated in established leukaemia indications for all age groups as a representative case study for haematological malignancies. The case of leukaemia has been chosen since it accounts for the highest age-standardised mortality rate from haematological malignancies in Europe [1]. Characteristics of the selected clinical trials are detailed in Section 2.

This research is congruent with the priority areas for collaboration of the joint EMA and European Network for Health Technology Assessment (EUnetHTA) 2017-2020 three-year work plan on pre- and post-authorisation evidence generation, with the aim of guiding medicines developers to generate evidence which is able to address both regulatory and HTA information needs. The EMA-EUnetHTA three-year work plan may be accessed here.

Research ethics approval has been granted by the Faculty of Medicine and Surgery Research Ethics Committee at the University of Malta.

References:

[1] World Health Organisation (WHO) – International Agency for Research on Cancer (IARC). GLOBOCAN 2018 database: The Global Cancer Observatory (GCO) [Internet]. 2018 September [cited 2018 Sep 13]. Available from: http://gco.iarc.fr/today/home

L-Università ta' Malta | Department of Pharmacy
Faculty of Medicine & Surgery

Response Evaluation in Leukaemia (ReVALeu)

Delphi Round 1 Instructions

This questionnaire is the first of two (2) rounds, in which HTA and EMA assessors and experts are requested to rate statements or provide comments on both generic and technical aspects of assessments. The first round questionnaire is divided into two main sections as follows:

Section 1 (General Questions): Questions asked in this section carry a broad scope and relate to the current state of play and impact of cooperation among HTABs and between the regulatory and HTA domains in assessment and scientific advice procedures.

Section 2 (Technical Questions): This section tackles technical questions on efficacy endpoints used in leukaemia clinical trials. For each unique efficacy measure, there are three (3) multiple choice fields, including classification (primary or secondary), trial phase-specificity (Phase II-IV) and clinical importance, and one (1) optional field for comments justifying answers on the endpoint in question.

For both sections, fields that require an answer are marked with an asterisk (*) and must be answered before progressing with the questionnaire.

Only Section 2 statements on the importance of the endpoint in demonstrating efficacy that have not reached consensus will progress to the second Delphi round. A feedback loop reporting the panel's distribution of responses and a summary of comments will be provided to the experts after each round.

The estimated time to complete the questionnaire is approximately 20 minutes and each Delphi round will have a **deadline of three weeks** for participants to respond.

L-Università ta' Malta
Faculty of Medicine & Surgery
Department of Pharmacy

Response Evaluation in Leukaemia (ReVALeu)

*** 1. Indicate the agency type you are representing or provide professional services to:**

- ○ European Medicines Agency (EMA)
- ○ Health Technology Assessment Body (HTAB)

L-Università ta' Malta | Department of Pharmacy
Faculty of Medicine & Surgery

Response Evaluation in Leukaemia (ReVALeu)

Section 1: General Questions (HTA)

*** 1. Do you agree that there is sufficient cooperation between the EMA and HTABs throughout the life cycle of medicinal products?**

(EMA: European Medicines Agency; HTAB: Health Technology Assessment Body)

◯ Strongly disagree ◯ Disagree ◯ Neutral ◯ Agree ◯ Strongly agree ◯ Don't know

*** 2. Do you agree that clinical evidence requirements in HTAs of oncology/haematology medicines are aligned between HTABs in EU Member States?**

(EU: European Union; HTA: Health Technology Assessment; HTAB: Health Technology Assessment Body)

◯ Strongly disagree ◯ Disagree ◯ Neutral ◯ Agree ◯ Strongly agree ◯ Don't know

*** 3. Do you agree that clinical evidence requirements are aligned between regulatory and health technology assessments of oncology/haematology medicines in the EU?**

(EU: European Union)

◯ Strongly disagree ◯ Disagree ◯ Neutral ◯ Agree ◯ Strongly agree ◯ Don't know

*** 4. Do you agree that the clinical evidence requested by your HTAB for oncology/haematology medicines is similar to:**
i) Other HTABs?
ii) The EMA?

(EMA: European Medicines Agency; HTAB: Health Technology Assessment Body)

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree	Don't know
Other HTABs	◯	◯	◯	◯	◯	◯
EMA	◯	◯	◯	◯	◯	◯

*** 5. Is the EMA European Public Assessment Report (EPAR) consulted during the HTA clinical evaluation of oncology/haematology medicines?**

(EMA: European Medicines Agency; HTA: Health Technology Assessment)

◯ Never ◯ Rarely ◯ Occasionally ◯ Frequently ◯ Always

*** 6. Has your HTAB been involved in parallel scientific advice/consultation procedures with the EMA Scientific Advice Working Party (SAWP)?**

(EMA: European Medicines Agency; HTAB: Health Technology Assessment Body)

- ◯ Yes
- ◯ No
- ◯ Don't know

If yes, please provide an indication on the number of procedures involved related to oncology/haematology medicines:

*** 7. How would you rate the quality of evidence generated in the field of oncology/haematology during:**

i) Pre-authorisation phases?

ii) Post-authorisation phases?

	Poor	Fair	Good	Very good	Excellent	Don't know
Pre-authorisation phases	◯	◯	◯	◯	◯	◯
Post-authorisation phases	◯	◯	◯	◯	◯	◯

Comments:

*** 8. Do you agree that the risk of divergent clinical evidence requirements between regulatory and health technology assessments impacts patient access to novelties in anticancer therapy?**

◯ Strongly disagree ◯ Disagree ◯ Neutral ◯ Agree ◯ Strongly agree

Comments:

*** 9. From the following list, rank stakeholders from 1 to 6 (1 being highest and 6 being lowest) according to who is most likely to benefit from enhanced cooperation between the EMA and HTABs throughout the medicinal product life cycle:**

(EMA: European Medicines Agency; HTAB: Health Technology Assessment Body)

	EMA	N/A
	HTABs	N/A
	Clinicians	N/A
	Industry	N/A
	Patients	N/A
	Payers	N/A

L-Università ta' Malta
Faculty of Medicine & Surgery
Department of Pharmacy

Response Evaluation in Leukaemia (ReVALeu)

Section 1: General Questions (Regulatory)

*** 1. Do you agree that there is sufficient cooperation between the EMA and HTABs throughout the life cycle of medicinal products?**

(EMA: European Medicines Agency; HTAB: Health Technology Assessment Body)

☐ Strongly disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly agree ☐ Don't know

*** 2. Do you agree that clinical evidence requirements are aligned between regulatory and health technology assessments of oncology/haematology medicines in the EU?**

(EU: European Union)

☐ Strongly disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly agree ☐ Don't know

*** 3. How would you rate the quality of evidence generated in the field of oncology/haematology during:**

i) Pre-authorisation phases?

ii) Post-authorisation phases?

	Poor	Fair	Good	Very good	Excellent	Don't know
Pre-authorisation phases	☐	☐	☐	☐	☐	☐
Post-authorisation phases	☐	☐	☐	☐	☐	☐

Comments:

*** 4. Do you agree that the risk of divergent clinical evidence requirements between regulatory and health technology assessments impacts patient access to novelties in anticancer therapy?**

☐ Strongly disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly agree

Comments:

*** 5. From the following list, rank stakeholders from 1 to 6 (1 being highest and 6 being lowest) according to who is most likely to benefit from enhanced cooperation between the EMA and HTABs throughout the medicinal product life cycle:**

(EMA: European Medicines Agency; HTAB: Health Technology Assessment Body)

	EMA	N/A
	HTABs	N/A
	Clinicians	N/A
	Industry	N/A
	Patients	N/A
	Payers	N/A

Response Evaluation in Leukaemia (ReVALeu)

Section 2: Technical Questions

This set of questions relate to efficacy endpoints extracted from clinical trials in the EU Clinical Trials Register database over an 11-year period (January 2007 - December 2017), that have been selected after the application of inclusion criteria. In total, 431 clinical trials satisfied acceptance criteria and qualified for the study, having the following characteristics:

Clinical Trial Characteristic	% **Clinical Trials (N=431)**
Ongoing	61 (n=261)
Completed	23 (n=100)
Prematurely Ended/Restarted/Temporarily Halted/Not Specified	16 (n=70)
Trial Subjects Under 18 (birth-17 years)	15 (n=64)
Trial Subjects Adults (18-64 years)	86 (n=369)
Trial Subjects Elderly (≥65 years)	77 (n=330)

Thirty-six (36) unique efficacy endpoints have been identified and grouped into four principal categories, comprising:

- **Endpoint Category 1: Survival** (n=5)
- **Endpoint Category 2: Time-To-Event** (n=6)
- **Endpoint Category 3: Response Rates and Biomarkers** (n=16)
- **Endpoint Category 4: Other** (n=9)

Please refer to Appendix 1 attached for guidance on endpoint definitions.

L-Università ta' Malta | Faculty of Medicine & Surgery | Department of Pharmacy

Response Evaluation in Leukaemia (ReVALeu)

Section 2(a): Endpoint Category 1 (Survival)

*** 1. How would you classify this endpoint?**

	Primary	Secondary
Overall survival (OS)	○	○
Progression-free survival (PFS)	○	○
Event-free survival (EFS)	○	○
Disease-free survival (DFS)	○	○
Treatment-free survival (TFS)	○	○

Comments:

*** 2. Indicate the most suitable clinical trial phase(s) for this endpoint (*more than one option may be selected*):**

	Phase II	Phase III	Phase IV
Overall survival (OS)	☐	☐	☐
Progression-free survival (PFS)	☐	☐	☐
Event-free survival (EFS)	☐	☐	☐
Disease-free survival (DFS)	☐	☐	☐
Treatment-free survival (TFS)	☐	☐	☐

Comments:

*** 3. Indicate the importance of this endpoint in demonstrating efficacy:**

	Not important at all	Not important	Neutral	Important	Very important
Overall survival (OS)	○	○	○	○	○
Progression-free survival (PFS)	○	○	○	○	○
Event-free survival (EFS)	○	○	○	○	○
Disease-free survival (DFS)	○	○	○	○	○
Treatment-free survival (TFS)	○	○	○	○	○

Comments:

L-Università ta' Malta
Faculty of Medicine & Surgery
Department of Pharmacy

Response Evaluation in Leukaemia (ReVALeu)

Section 2(b): Endpoint Category 2 (Time-To-Event)

*** 1. How would you classify this endpoint?**

	Primary	Secondary
Duration of remission/response (DOR)	☐	☐
Time to remission/response (TTR)	☐	☐
Time to progression (TTP)	☐	☐
Time to treatment (TTT)	☐	☐
Time to treatment failure (TTF)	☐	☐
Treatment-free remission/response (TFR)	☐	☐

Comments:

*** 2. Indicate the most suitable clinical trial phase(s) for this endpoint (*more than one option may be selected*):**

	Phase II	Phase III	Phase IV
Duration of remission/response (DOR)	☐	☐	☐
Time to remission/response (TTR)	☐	☐	☐
Time to progression (TTP)	☐	☐	☐
Time to treatment (TTT)	☐	☐	☐
Time to treatment failure (TTF)	☐	☐	☐
Treatment-free remission/response (TFR)	☐	☐	☐

Comments:

*** 3. Indicate the importance of this endpoint in demonstrating efficacy:**

	Not important at all	Not important	Neutral	Important	Very important
Duration of remission/response (DOR)	○	○	○	○	○
Time to remission/response (TTR)	○	○	○	○	○
Time to progression (TTP)	○	○	○	○	○
Time to treatment (TTT)	○	○	○	○	○
Time to treatment failure (TTF)	○	○	○	○	○
Treatment-free remission/response (TFR)	○	○	○	○	○

Comments:

L-Università ta' Malta | Department of Pharmacy
Faculty of Medicine & Surgery

Response Evaluation in Leukaemia (ReVALeu)

Section 2(c): Endpoint Category 3 (Response Rates and Biomarkers)

*** 1. How would you classify this endpoint?**

	Primary	Secondary
Relapse rate (RR)/Rate of progressive disease (PD)	☐	☐
Complete remission/response rate (CR)	☐	☐
Complete remission/response rate with incomplete platelet recovery (CRp)	☐	☐
Complete remission/response rate with incomplete blood count recovery (CRi/CRh)	☐	☐
Partial remission/response rate (PR)	☐	☐
Combined complete remission/response rate (CRc)	☐	☐
Objective remission/response rate (ORR)	☐	☐
Best overall response rate (BOR)	☐	☐
Haematologic response rate (HR)	☐	☐
Cytogenetic response rate (CyR)	☐	☐
Molecular remission/response (MR)	☐	☐

	Primary	Secondary
Minimal residual disease (MRD)	○	○
Cytokine serum concentrations	○	○
Morphologic leukaemia-free state (MLFS)	○	○
Splenomegaly and hepatomegaly response rates	○	○
Return to chronic phase (RTC)	○	○

Comments:

*** 2. Indicate the most suitable clinical trial phase(s) for this endpoint (*more than one option may be selected*):**

	Phase II	Phase III	Phase IV
Relapse rate (RR)/Rate of progressive disease (PD)	☐	☐	☐
Complete remission/response rate (CR)	☐	☐	☐
Complete remission/response rate with incomplete platelet recovery (CRp)	☐	☐	☐
Complete remission/response rate with incomplete blood count recovery (CRi/CRh)	☐	☐	☐
Partial remission/response rate (PR)	☐	☐	☐
Combined complete remission/response rate (CRc)	☐	☐	☐
Objective remission/response rate (ORR)	☐	☐	☐

	Phase II	Phase III	Phase IV
Best overall response rate (BOR)	☐	☐	☐
Haematologic response rate (HR)	☐	☐	☐
Cytogenetic response rate (CyR)	☐	☐	☐
Molecular remission/response (MR)	☐	☐	☐
Minimal residual disease (MRD)	☐	☐	☐
Cytokine serum concentrations	☐	☐	☐
Morphologic leukaemia-free state (MLFS)	☐	☐	☐
Splenomegaly and hepatomegaly response rates	☐	☐	☐
Return to chronic phase (RTC)	☐	☐	☐

Comments:

*** 3. Indicate the importance of this endpoint in demonstrating efficacy:**

	Not important at all	Not important	Neutral	Important	Very important
Relapse rate (RR)/Rate of progressive disease (PD)	○	○	○	○	○
Complete remission/response rate (CR)	○	○	○	○	○
Complete remission/response rate with incomplete platelet recovery (CRp)	○	○	○	○	○

	Not important at all	Not important	Neutral	Important	Very important
Complete remission/response rate with incomplete blood count recovery (CRi/CRh)	○	○	○	○	○
Partial remission/response rate (PR)	○	○	○	○	○
Combined complete remission/response rate (CRc)	○	○	○	○	○
Objective remission/response rate (ORR)	○	○	○	○	○
Best overall response rate (BOR)	○	○	○	○	○
Haematologic response rate (HR)	○	○	○	○	○
Cytogenetic response rate (CyR)	○	○	○	○	○
Molecular remission/response (MR)	○	○	○	○	○
Minimal residual disease (MRD)	○	○	○	○	○
Cytokine serum concentrations	○	○	○	○	○
Morphologic leukaemia-free state (MLFS)	○	○	○	○	○
Splenomegaly and hepatomegaly response rates	○	○	○	○	○
Return to chronic phase (RTC)	○	○	○	○	○

Comments:

L-Università ta' Malta | Faculty of Medicine & Surgery | Department of Pharmacy

Response Evaluation in Leukaemia (ReVALeu)

Section 2(d): Endpoint Category 4 (Other)

*** 1. How would you classify this endpoint?**

	Primary	Secondary
Patient reported outcomes (PROs)	☐	☐
Number of patients with asparagine depletion	☐	☐
Secondary malignancies	☐	☐
Number of blood cell transfusions	☐	☐
Number of chemotherapy cycles received	☐	☐
Incidence and duration of febrile neutropenia	☐	☐
Medical care utilisation (MCU)	☐	☐
Incidence, duration and severity of opportunistic infections	☐	☐
Proportion of patients undergoing subsequent allogeneic stem cell transplant (ASCT)	☐	☐

Comments:

*** 2. Indicate the most suitable clinical trial phase(s) for this endpoint (*more than one option may be selected*):**

	Phase II	Phase III	Phase IV
Patient reported outcomes (PROs)	☐	☐	☐
Number of patients with asparagine depletion	☐	☐	☐
Secondary malignancies	☐	☐	☐
Number of blood cell transfusions	☐	☐	☐
Number of chemotherapy cycles received	☐	☐	☐
Incidence and duration of febrile neutropenia	☐	☐	☐
Medical care utilisation (MCU)	☐	☐	☐
Incidence, duration and severity of opportunistic infections	☐	☐	☐
Proportion of patients undergoing subsequent allogeneic stem cell transplant (ASCT)	☐	☐	☐

Comments:

Thank you for completing the first Delphi round.

A summary of the panel responses together with your own individual ratings will be provided in the second (final) round.

Appendix 3

Definitions of efficacy endpoints in leukaemia clinical trials (Appendix 1 of the ReVALeu tool)

Response Evaluation in Leukaemia (ReVALeu)

Appendix 1: Definitions of clinical trial efficacy endpoints

The below description tables have been compiled to assist the respondent in differentiating between the identified efficacy endpoints for the four main leukaemia subtypes; Acute Lymphoblastic Leukaemia (ALL), Acute Myeloid Leukaemia (AML), Chronic Lymphocytic Leukaemia (CLL) and Chronic Myeloid Leukaemia (CML). This guidance document is to be used in conjunction with the Response Evaluation in Leukaemia (ReVALeu) online surveying tool during its completion. More than one definition for the same endpoint may be provided if descriptions from different sources are considered to contrast.

Endpoint Category 1: Survival

Efficacy Endpoint	***Alternative Nomenclature/ Variations of Endpoint***	***Definition of Endpoint***			
		ALL	***AML***	***CLL***	***CML***
Overall survival (OS)	Mortality rate; Early mortality rate	Time from randomisation to death from any cause.[1]			
Progression-free survival (PFS)	Transformation-free survival	Time from randomisation (or registration, in non-randomised trials) to objective tumour progression, or death from any cause, whichever first.[1]			
Event-free survival (EFS)	Failure-free survival (FFS)	Lack of achievement of complete response (CR), relapse and death without relapse are counted as events in an EFS analysis. Those patients who did not reach CR during the pre-specified induction phase will be considered as having an event at time zero.[1] Time from randomisation to disease progression, death, or discontinuation of treatment for any reason (eg, toxicity, patient preference, or initiation of a new treatment without documented progression).[2] FFS includes the addition of systemic therapy in its definition.[16]			
Disease-free survival (DFS)	Leukaemia-free survival (LFS); Recurrence-free survival; Relapse-free survival (RFS);	Time from randomisation to objective recurrence or death from any cause.[1]			
Treatment-free survival (TFS)		Survival without the need for treatment of recurrent or persistent cancer.[3]			

Endpoint Category 2: Time-To-Event

Efficacy Endpoint	*Alternative Nomenclature/ Variations of Endpoint*	*Definition of Endpoint*			
		ALL	***AML***	***CLL***	***CML***
Duration of remission/response (DOR)	Continuous complete remission/response (CCR); Time to relapse	Time from initial response until documented tumour progression.[2]			
Time to remission/response (TTR)		Time from the start of treatment to the first objective tumour response.[20]			
Time to progression (TTP)	Time to transformation (TTT)	Time from randomisation to observed tumour progression, censoring for death not related to the underlying malignancy.[1]			
Time to treatment (TTT)	Time to next treatment (TTNT)	Time from end of primary treatment to institution of next therapy.[2]			
Time to treatment failure (TTF)		Time from randomisation to discontinuation of therapy for any reason including death, progression, toxicity or add-on of new anti-cancer therapy.[1]			
Treatment-free remission/response (TFR)		A stable deep molecular response (DMR) without the need for ongoing treatment after having been in deep molecular residual disease (MRD) over a long period of time.[19]			

Endpoint Category 3: Response Rates and Biomarkers

Efficacy Endpoint	*Alternative Nomenclature/ Variations of Endpoint*	*Definition of Endpoint*			
		ALL	***AML***	***CLL***	***CML***
Relapse rate (RR)/Rate of progressive disease (PD)	Cumulative incidence of relapse (CIR); Relapse-free incidence (RFI); Cumulative progression	• Relapse rate (RR): The proportion of patients relapsing from complete remission (CR) or molecular remission. • Rate of PD: The proportion of patients progressing from stable disease (SD) to accelerated phase (AP) or blast phase (BP).			
Complete remission/response rate (CR)	Morphologic complete remission/response rate (CR); Noncytopenic complete remission/response rate (CR)	Response criteria[4,5,6,7]: • Neutrophils (μL): >1,000 • Platelets(μL): >100,000 • Bone marrow (BM) blasts (%): <5%		Response criteria[8]: • Symptoms: None • Peripheral blood lymphocytes: $<4x10^9$/L • Lymphadenopathy: None >1.5cm in diameter • No hepatomegaly and splenomegaly • Neutrophils: $>1.5x10^9$/L • Platelets: $>100x10^9$/L • Haemoglobin: >11g/dL (untransfused) • BM lymphocytes: <30%; no nodules	
Complete remission/response rate with incomplete platelet recovery (CRp)	Morphologic complete remission/response rate with incomplete platelet recovery (CRp)	Response criteria[4,5,6,7]: • CR in ALL/AML • Platelets(μL): <100,000		Response criteria[8]: • CR in CLL • Platelets: $<100x10^9$/L	
Complete remission/response rate with incomplete blood count recovery (CRi/CRh)	Morphologic complete remission/response rate with incomplete blood count recovery (CRi/CRh)	Response criteria[4,5,6,7]: • CR in ALL/AML • Neutrophils (μL): <1,000		Response criteria[8]: • CR in CLL • Neutrophils: $<1.5x10^9$/L • Haemoglobin: <11g/dL	
Partial remission/response rate (PR)	Morphologic partial remission/response rate (PR)	Response criteria[4,5,6,7]: • CR in ALL/AML • BM blasts (%): ≤50% ↓ to a value of 5-25%		Response criteria[8]: • Lymphadenopathy, hepatomegaly, splenomegaly, blood lymphocytes, BM lymphocytes: ≥50% ↓ • Platelets: >100,00/μL or ≥50% ↑ • Haemoglobin: >11g/dL or ≥50% ↑ • Neutrophils: >1500/μL or ≥50% ↑	
Combined complete remission/response rate (CRc)	Composite complete remission/response rate (CRc)	The sum of complete responses, including those with incomplete platelet or blood count recovery[12]: CRc = CR + CRp + (CRi/CRh)			
Objective remission/response rate (ORR)	Overall remission/response rate (ORR)	The proportion of patients in whom a complete response or partial response was observed[1]: ORR = CR + CRp + (CRi/CRh) + PR			

Best overall response rate (BOR)	Best observed response (BOR)	The best response recorded from the start of the study treatment until the disease progression/recurrence, including CR, PR, SD, PD.[21]	
Haematologic response rate (HR)	Complete haematologic response rate (CHR), Major haematologic response rate (MaHR), Minor haematologic response rate (MiHR), No evidence of leukaemia (NEL), Overall haematologic response rate (OHR); Haematological improvement (HI)	HI[17,18]: • Erythroid (HI-E): Haemoglobin ≥ 1.5g/dL, reductions in red blood cell (RBC) transfusions by at least 4 RBC transfusions/8 week • Platelet (HI-P): ↑ ≥30x10^9/L (pre-treatment >20x10^9/L) or >20x10^9/L and at least 100% ↑ (pre-treatment <20x10^9/L) • Neutrophil (HI-N): At least 100% ↑ and >0.5x10^9/L CHR[9,10,11]: • Complete normalisation of peripheral blood counts with leukocyte count <10x10^9/L • Platelets: <450x10^9/L • No immature granulocytes, basophils <5% • Non-palpable spleen MiHR[15]: • Blasts less than 15% and blasts promyelocytes less than 30% • Peripheral blood basophils less than 20%, blasts less than 15%, and blasts promyelocytes less than 30% • No extramedullary disease other than in spleen and liver NEL[15]: • CHR without full recovery of platelets and neutrophils MaHR[15]: • CHR + NEL OHR[15]: • CHR + NEL + MiHR	
Cytogenetic response rate (CyR)	Complete cytogenetic response rate (CCyR), Major partial cytogenetic response rate (MCyR), Partial cytogenetic response rate (PCyR)	* *	Response criteria[9,10,11]: • CyR: Any Philadelphia (Ph+) chromosome value less than baseline reading • PCyR: 1–35% Ph+ metaphases • MCyR: 0–35% Ph+ metaphases (CCyR + PCyR) • CCyR: No Ph+ metaphases
Molecular remission/response (MR)	Complete molecular remission/response (CMR), Major molecular remission/response (MMR)	* *	Response criteria[9,10,11]: • MMR: Ratio of BCR-ABL1 to ABL (or other housekeeping genes) ≤ 0.1% on the International Scale or ≥3-log reduction on International Scale of BCR-ABL1 mRNA • CMR: Undetectable BCR-ABL1 mRNA transcripts
Minimal residual disease (MRD)	Deep molecular response (DMR)	Any residual disease after suboptimal induction chemotherapy, but at the same time refers to the lowest levels of disease potentially compatible with cure or to molecularly defined relapse after long-term remission. MRD is detected only by laboratory techniques more sensitive than morphology, such as flow cytometry (immunologic MRD) or polymerase chain reaction (PCR) (molecular MRD).[13]	
Cytokine serum concentrations		Abnormalities of cytokine signalling pathways are characteristic of all forms of leukaemia: lymphoid and myeloid, acute and chronic.[14]	
Morphologic leukaemia-free state (MLFS)		<5% blasts in the bone marrow with no blasts with Auer rods and no extramedullary disease. Does not require count recovery.[12]	

*CML definitions for cytogenetic (CyR) and molecular (MR) response criteria are also applicable to Philadelphia chromosome-positive (Ph+) ALL and AML.

Splenomegaly and hepatomegaly response rates				• Splenomegaly response rate[8]: The proportion of subjects with a 50% ↓ from baseline or no splenomegaly. • Hepatomegaly response rate[8]: The proportion of subjects with a 50% ↓ from baseline or no hepatomegaly.	
Return to chronic phase (RTC)				Response criteria[22]: • <15% blasts in peripheral blood and bone marrow • <30% blasts plus promyelocytes in the peripheral blood and bone marrow • <20% basophils in peripheral blood • No extramedullary disease except liver or spleen enlargement	

Endpoint Category 4: Other

Efficacy Endpoint	***Alternative Nomenclature/ Variations of Endpoint***	***Definition of Endpoint***			
		ALL	***AML***	***CLL***	***CML***
Patient reported outcomes (PROs)	Quality of life (QoL)/Health-related quality of life (HRQoL/HRQL)	PRO is an umbrella term covering both single dimension and multi-dimension measures of symptoms, HRQL, health status, adherence to treatment and satisfaction with treatment.[1] PRO measures (PROMs) are the tools and/or instruments that have been developed to ensure both a valid and reliable measurement of these PROs[1] eg. EORTC QLQ-C30, ECOG and Karnofsky performance status, Charlson Comorbidity Index (CCI), Global Health Status/QOL, EQ-5D, Functional Assessment of Cancer Therapy-Leukemia (FACT-Leu), (PROMIS) Cancer Fatigue Short Form, Instrumental activities of daily living (iADLs), FACIT-Fatigue Scale, etc.			
Number of patients with asparagine depletion	Nadir serum asparaginase activity (NSAA)	Serum asparaginase activity level of ≥0.1 IU/ml since complete asparagine depletion is observed less consistently with asparaginase activity levels below this cut-off.[23,24]			
Secondary malignancies					
Number of blood cell transfusions	Last day of transfusion after each cycle; Transfusion independence				
Number of chemotherapy cycles received	Number of consolidation courses actually given				
Incidence and duration of febrile neutropenia					
Medical care utilisation (MCU)	Number of days alive and out of the hospital (NDAOH)	Resource use includes antibiotics, number and duration of hospitalisations and number of days of attendance at the day hospital.			
Incidence, duration and severity of opportunistic infections					
Proportion of patients undergoing subsequent allogeneic stem cell transplant (ASCT)					

References:

1. European Medicines Agency. Guidelines on the evaluation of anticancer medicinal products in man [Internet]. 2017 Sep [cited 2018 Jul 05]. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2017/11/WC500238764.pdf

2. Food and Drug Administration (FDA). Guidance for industry clinical trial endpoints for the approval of cancer drugs and biologics [Internet]. 2007 May [cited 2018 Jul 05]. Available from: https://www.fda.gov/downloads/drugsGuidanceComplianceRegulatoyInformation/Guidance/UCM071590.pdf

3. Banerjee M, Reyes-Gastelum D, Haymart M. Treatment-free survival in patients with differentiated thyroid cancer. J Clin Endocrinol Metab. 2018;103(7):2720-2727.

4. Cheson BD, Bennet JM, Kopecky KJ, Buchner T, Willman CL, Estey EH, et al. Revised recommendations of the International Working Group for Diagnosis, Standardisation of Response Criteria, Treatment Outcomes and Reporting Standards for Therapeutic Trials in Acute Myeloid Leukaemia. J Clin Oncol. 2003;21:4642-4649.

5. Appelbaum FR, Rosenblum D, Arceci R, Carroll WL, Breitfeld PP, Forman SJ, et al. End points to establish the efficacy of new agents in the treatment of acute leukaemia. Blood. 2007;109:1810-1815.

6. Douglas Smith B, Karp JE. What are the endpoints of therapy for acute leukaemias? Old definitions and new challenges. Clin Lymphoma Myeloma. 2009;9(3):S296-S301.

7. Gokbuget N, Dombret H, Bassan R, Wadleigh M, Doubek M, Ribera J. Inclusion and response criteria for clinical trials in relapsed/refractory acute lymphoblastic leukemia and usefulness of historical control trials. Haematologica. 2017;102:e118.

8. Hallek M, Cheson BD, Catovsky D, Caligaris-Cappio F, Dighiero G, Dohner H, et al. Guidelines for the diagnosis and treatment of chronic lymphocytic leukaemia: a report from the International Workshop on Chronic Lymphocytic Leukaemia updating the National Cancer Institute – Working Group 1996 guidelines. Blood. 2008;111:5446-5456.

9. Baccarani M, Castagnetti F, Gugliotta G, Palandri F, Soverini S. Response definitions and European Leukemianet Management recommendations. Best Pract Res Clin Haematol 2009;22:331–41.

10. O'Brien S, Berman E, Borghaei H, DeAngelo DJ, Devetten MP, Devine S, et al. The NCCN chronic myelogenous leukemia clinical practice guidelines in oncology. J Natl Compr Canc Netw. 2009;7(9):984-1023.

11. Tibes R, Mesa RA. Evolution of clinical trial endpoints in chronic myeloid leukaemia: Efficacious therapies require sensitive monitoring techniques. Leuk Res. 2012:36(6):664-671.

12. Medeiros BC. Interpretation of clinical endpoints in trials of acute myeloid leukemia. Leuk Res. 2018;68:32-39.

13. Paietta E. Assessing minimal residual disease (MRD) in leukaemia: a changing definition and concept? Bone Marrow Transplant. 2002;29(6):459-465.

14. Van Etten RA. Aberrant cytokine signaling in leukemia. Oncogene. 2007;26(47):6738-6749.

15. Cortes J, Rousselot P, Kim DW, Ritchie E, Hamerschlak N, Coutre S, et al. Dasatinib induces complete hematologic and cytogenetic responses in patients with imatinib-resistant or –intolerant chronic myeloid leukemia in blast crisis. Blood. 2007;1093207-3213.

16. Palmer J, Chai X, Martin PJ, Weisdorf D, Inamoto Y, Pidala J, et al. Failure-free survival in a prospective cohort of patients with chronic graft-versus-host disease. Haematologica. 2015;100(5):690-695.

17. Cheson BD, Greenberg PL, Bennett JM, Lowenberg B, Wijermans PW, Nimer SD, et al. Clinical application and proposal for modification of the International Working Group (IWG) response criteria in myelodysplasia. Blood. 2006;108:419-425.

18. Yanada M, Huang X, Garcia-Manero G, O'Brien S, Ravandi F, Borthakur G, et al. Effect of haematological improvement on survival in patients given targeted therapy as initial treatment of acute myeloid leukaemia or high-risk myelodysplastic syndrome. Br J Haematol. 2007;138:555-560.

19. Rea D, Etienne G, Mahon FX. What is treatment free remission in chronic myeloid leukemia? Oncotarget. 2018;9(4):4279.

20. Takeda M, Okamoto I, Nakagawa K. Survival outcome assessed according to tumor responsive and shrinkage pattern in patients with EGFR mutation-positive non-small-cell lung cancer treated with gefitinib or erlotinib. J Thorac Oncol. 2014;9(2):200-204.

21. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation in solid tumours: revised RECIST guidelines (version 1.1). Eur J Cancer. 2009;45(2):228-247.

22. Sawyers CL, Hochhaus A, Feldman E, Goldman JM, Miller CB, Ottmann OG, et al. Imatinib induces hematologic and cytogenetic responses in patients with chronic myelogenous leukemia in myeloid blast crisis: results of a phase II study. Blood. 2002;99(10):3530-3590.

23. Van der Sluis IM, Vrooman LM, Pieters R, Baruchel A, Escherish G, Goulden N, et al. Consensus expert recommendations for identification and management of asparaginase hypersensitivity and silent inactivation. Haematologica. 2016;101(3):279-285.

24. Salzer W, Bostrom B, Messinger Y, Perissinotti AJ, Marini B. Asparaginase activity levels and monitoring in patients with acute lymphoblastic leukemia. Leuk Lymphoma. 2018;59(8):1797-1806.

Appendix 4

Expert validation template for the Response Evaluation in Leukaemia (ReVALeu) tool

Expert validation of the Response Evaluation in Leukaemia (ReVALeu) online surveying tool

Introduction and Instructions

Dear Participant,

Thank you for accepting to participate as a member of the expert validation process for the data collection tool that will be used as part of the research project entitled 'Evidence Generation in the Clinical Development of Medicines for Leukaemia' being undertaken by Doctorate in Pharmacy student Dylan Said.

Please indicate your area of expertise or practice:

☐ Academia

☐ Clinical data management

☐ Clinical oncology and haematology

☐ Health technology assessment (HTA)

☐ Regulatory sciences

☐ Research

☐ Other, please specify:Click here to enter text.

The Response Evaluation in Leukaemia (ReVALeu) online surveying tool has been developed to determine the prioritised efficacy endpoints in clinical assessments of anti-leukaemic therapies among HTA bodies (HTABs) and between HTABs and regulators. Efficacy endpoints were extracted from clinical trials of anticancer agents investigated in established leukaemia indications for all populations as a representative case study for haematological malignancies.

You are kindly requested to rate fields of the tool using the following three validation domains on a Likert scale of 1 to 4:

i) **Relevance** to research construct: 1 (not relevant) – 4 (very relevant)

ii) **Clarity** of questions and statements: 1 (not clear) – 4 (very clear)

iii) **Structure and Layout** of the questionnaire: 1 (not well structured) – 4 (very well structured)

For questions or statements rated 2 or less, please suggest possible improvements in the comments section.

Section 1: General Questions

Question or Statement Number	Relevance	Clarity	Structure and Layout
	Rate each question or statement from 1-4		
2	Select rating	Select rating	Select rating
3	Select rating	Select rating	Select rating
4	Select rating	Select rating	Select rating
5	Select rating	Select rating	Select rating
6	Select rating	Select rating	Select rating
7	Select rating	Select rating	Select rating
8	Select rating	Select rating	Select rating
9	Select rating	Select rating	Select rating
10	Select rating	Select rating	Select rating
11	Select rating	Select rating	Select rating

Question or Statement Number	Comments
	Please suggest possible improvements for questions or statements rated as 2 or less.
Question number	Click here to insert comments
Question number	Click here to insert comments
Question number	Click here to insert comments
Question number	Click here to insert comments
Question number	Click here to insert comments
Question number	Click here to insert comments
Question number	Click here to insert comments
Question number	Click here to insert comments
Question number	Click here to insert comments
Question number	Click here to insert comments

Section 2: Technical Questions

Section 2(a): Endpoint Category 1 (Survival)

Question or Statement Number	Relevance	Clarity	Structure and Layout
	Rate each question or statement from 1-4		
12	Select rating	Select rating	Select rating
13	Select rating	Select rating	Select rating
14	Select rating	Select rating	Select rating
Question or Statement Number	Comments		
	Please suggest possible improvements for questions or statements rated as 2 or less.		
Question number	Click here to insert comments		
Question number	Click here to insert comments		
Question number	Click here to insert comments		

Section 2(b): Endpoint Category 2 (Time-To-Event)

Question or Statement Number	Relevance	Clarity	Structure and Layout
	Rate each question or statement from 1-4		
15	Select rating	Select rating	Select rating
16	Select rating	Select rating	Select rating
17	Select rating	Select rating	Select rating
Question or Statement Number	Comments		
	Please suggest possible improvements for questions or statements rated as 2 or less.		
Question number	Click here to insert comments		
Question number	Click here to insert comments		
Question number	Click here to insert comments		

Section 2(c): Endpoint Category 3 (Response Rates and Biomarkers)

Question or Statement Number	Relevance	Clarity	Structure and Layout
	Rate each question or statement from 1-4		
18	Select rating	Select rating	Select rating
19	Select rating	Select rating	Select rating
20	Select rating	Select rating	Select rating
Question or Statement Number	Comments		
	Please suggest possible improvements for questions or statements rated as 2 or less.		
Question number	Click here to insert comments		
Question number	Click here to insert comments		
Question number	Click here to insert comments		

Section 2(d): Endpoint Category 4 (Other)

Question or Statement Number	Relevance	Clarity	Structure and Layout
	Rate each question or statement from 1-4		
21	Select rating	Select rating	Select rating
22	Select rating	Select rating	Select rating
23	Select rating	Select rating	Select rating
Question or Statement Number	Comments		
	Please suggest possible improvements for questions or statements rated as 2 or less.		
Question number	Click here to insert comments		
Question number	Click here to insert comments		
Question number	Click here to insert comments		

Appendix 1: Definitions of clinical trial efficacy endpoints

Relevance	Clarity	Structure and Layout
Rate the overall Appendix from 1-4		
Select rating	Select rating	Select rating

Appendix Definition	Comments
	Please suggest possible improvements to the Appendix for domains rated as 2 or less.
Definition	Click here to insert comments
Definition	Click here to insert comments
Definition	Click here to insert comments
Definition	Click here to insert comments
Definition	Click here to insert comments
Definition	Click here to insert comments
Definition	Click here to insert comments
Definition	Click here to insert comments
Definition	Click here to insert comments
Definition	Click here to insert comments
Definition	Click here to insert comments
Definition	Click here to insert comments
Definition	Click here to insert comments

Thank you for your participation

Appendix 5

The Content Validity Index (CVI) validation method: I-CVI and S-CVI values

Expert Validation of the ReVALeu Tool: Content Validity Index (CVI)									
Validation Domain: Relevance to research construct									
Question or Statement Number	**Expert 1**	**Expert 2**	**Expert 3**	**Expert 4**	**Expert 5**	**Expert 6**	**Expert 7**	**Expert 8**	**I-CVI (%)**
Section 1: General Questions									
2	X	X	X	X	X	X	X	X	100%
3	X	X	X	X	X	X	X	X	100%
4	X	X	X	X	X	X	X	X	100%
5	X	X	X	X	-	X	-	X	75%
6	X	X	X	X	X	X	X	X	100%
7	X	X	X	X	X	X	X	X	100%
8	X	X	X	X	-	-	X	X	75%
9	X	X	X	X	-	-	X	X	75%
10	X	X	X	X	X	X	X	X	100%
11	X	X	X	X	-	-	X	X	75%
Section 2: Technical Questions									
Section 2(a): Endpoint Category 1 (Survival)									
12	X	X	X	X	X	X	X	X	100%
13	X	X	X	X	X	X	X	X	100%
14	X	X	X	X	X	X	X	X	100%
Section 2(b): Endpoint Category 2 (Time-To-Event)									
15	X	X	X	X	X	X	X	X	100%
16	X	X	X	X	X	X	X	X	100%
17	X	X	X	X	X	X	X	X	100%
Section 2(c): Endpoint Category 3 (Response Rates and Biomarkers)									
18	X	X	X	X	X	X	X	X	100%
19	X	X	-	X	X	X	X	X	88%
20	X	X	X	X	X	X	X	X	100%
Section 2(d): Endpoint Category 4 (Other)									
21	X	X	X	X	X	X	X	X	100%
22	X	X	-	X	X	X	X	X	88%
23	X	X	X	X	X	X	X	X	100%
Appendix 1: Definitions of clinical trial efficacy endpoints	X	X	X	X	X	X	X	X	100%
									S-CVI/Ave (%)
									95%

Expert Validation of the ReVALeu Tool: Content Validity Index (CVI)									
Validation Domain: Clarity of questions and statements									
Question or Statement Number	**Expert 1**	**Expert 2**	**Expert 3**	**Expert 4**	**Expert 5**	**Expert 6**	**Expert 7**	**Expert 8**	**I-CVI (%)**
Section 1: General Questions									
2	X	X	X	X	-	-	X	X	75%
3	X	X	X	X	-	X	X	X	88%
4	X	X	X	X	-	-	X	X	75%
5	X	X	-	X	-	X	-	X	63%
6	X	X	X	X	X	X	X	X	100%
7	X	X	X	X	X	X	X	X	100%
8	X	X	X	X	-	-	X	X	75%
9	X	X	X	X	-	-	X	X	75%
10	X	X	X	X	X	-	X	X	88%
11	X	X	X	X	-	-	X	X	75%
Section 2: Technical Questions									
Section 2(a): Endpoint Category 1 (Survival)									
12	X	X	X	X	X	X	X	X	100%
13	X	X	X	X	X	X	X	X	100%
14	X	X	X	X	X	X	X	X	100%
Section 2(b): Endpoint Category 2 (Time-To-Event)									
15	X	X	X	X	X	X	X	X	100%
16	X	X	X	X	X	X	X	X	100%
17	X	X	X	X	X	X	X	X	100%
Section 2(c): Endpoint Category 3 (Response Rates and Biomarkers)									
18	X	X	X	X	X	X	X	X	100%
19	X	X	-	X	X	X	X	X	88%
20	X	X	X	X	X	X	X	X	100%
Section 2(d): Endpoint Category 4 (Other)									
21	X	X	X	X	X	X	X	X	100%
22	X	X	-	X	X	X	X	X	88%
23	X	X	X	X	X	X	X	X	100%
Appendix 1: Definitions of clinical trial efficacy endpoints	X	X	X	X	X	X	X	X	100%
									S-CVI/Ave (%)
									91%

Expert Validation of the ReVALeu Tool: Content Validity Index (CVI)

Validation Domain: Structure and layout of questionnaire

Question or Statement Number	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Expert 6	Expert 7	Expert 8	I-CVI (%)
Section 1: General Questions									
2	X	X	X	X	X	-	X	X	88%
3	X	X	X	X	X	X	X	X	100%
4	X	X	X	X	X	-	X	X	88%
5	X	X	X	X	-	-	-	X	63%
6	X	X	X	X	X	X	X	X	100%
7	X	X	X	X	X	X	X	X	100%
8	X	X	X	X	-	-	X	X	75%
9	X	X	X	X	-	X	X	X	88%
10	X	X	X	X	X	X	X	X	100%
11	X	X	X	X	-	-	X	X	75%
Section 2: Technical Questions									
Section 2(a): Endpoint Category 1 (Survival)									
12	X	X	X	X	X	X	X	X	100%
13	X	X	X	X	X	X	X	X	100%
14	X	X	X	X	X	X	X	X	100%
Section 2(b): Endpoint Category 2 (Time-To-Event)									
15	X	X	X	X	X	X	X	X	100%
16	X	X	X	X	X	X	X	X	100%
17	X	X	X	X	X	X	X	X	100%
Section 2(c): Endpoint Category 3 (Response Rates and Biomarkers)									
18	X	X	X	X	X	X	X	X	100%
19	X	X	-	X	X	X	X	X	88%
20	X	X	X	X	X	X	X	X	100%
Section 2(d): Endpoint Category 4 (Other)									
21	X	X	X	X	X	X	X	X	100%
22	X	X	-	X	X	X	X	X	88%
23	X	X	X	X	X	X	X	X	100%
Appendix 1: Definitions of clinical trial efficacy endpoints	X	X	X	X	X	X	X	X	100%
									S-CVI/Ave (%)
									93%

Appendix 6

Questionnaire items revised or omitted from the Response Evaluation in Leukaemia (ReVALeu) tool

Question or statement number in the pre-validation version of the tool (Appendix 2)	Modification/Omission
2	**Modification(s)**: 'cooperation' instead of 'interaction'; 'European Medicines Agency' instead of 'regulatory'; deletion of the word 'facets'; focus on the life cycle of medicinal products instead of 'clinical assessments'.
4	**Modification(s)**: 'requirements' instead of 'needs'; sentence re-structured to differentiate more clearly between regulatory and HTA assessments.
5	**Modification(s)**: re-worded comprehensively to increase clarity of question objective; description for tick boxes optimised to include agreement scale instead; change in question layout.
8	**Modification(s)**: removal of any reference to reimbursement since targeted study subjects are concerned with clinical assessments; addition of comments section since 'quality of evidence' may be interpreted subjectively; change in question layout.
9	**Modification(s)**: question re-worded in view of the separate remits of EMA and HTABs; comments section added.
11	**Omission**: this question was determined to be out of scope for this study.

Question or statement number in the pre-validation version of the tool (Appendix 2)	Other comments
3	**Modification(s)**: 'requirements' instead of 'needs'; 'health technology assessments' instead of 'assessments'.
2,3,4,5,7,8	**Modification(s):** Addition of 'Don't know' option.
10	**Modification(s)**: Minor amendment in sentence structure.
Appendix of definitions	**Modification(s)**: Definition of 'asparagine depletion' specified in endpoint category 4 (Other).

Appendix 7

Recruitment letter for participation to the Delphi study

Recruitment letter to health technology assessment bodies:

L-Università ta' Malta
Faculty of Medicine & Surgery
Department of Pharmacy

22 August 2018

Dear Sir/Madam,

As a second year Doctorate in Pharmacy student at the University of Malta, I would like to invite expert(s) in the clinical assessment of onco-haematology medicinal products from your institution to participate in a Delphi study. This comprises part of the methodology for my thesis project titled 'Evidence generation in the clinical development of medicines for leukaemia' conducted in partial fulfilment of the requirements for the Doctorate in Pharmacy degree.

A robust clinical development strategy of antineoplastic agents, investigating parameters that satisfy regulatory and health technology assessment (HTA) requirements, can potentially contribute to enhanced patient access to innovative treatments. Harmonisation of evidence needs eliminates duplication of assessments and promotes data uniformity along the life cycle of a medicine, with inferred benefits extending to all stakeholders concerned. The purpose of the study is to determine a set of core efficacy outcomes prioritised by EU medicines regulators and Health Technology Assessment Bodies (HTABs) for clinical trials investigating medicines in leukaemia.

The study will employ a web-based platform to disseminate a validated online questionnaire in a two-round e-Delphi method, with respondent instructions preceding each survey round. Duration of completion for each questionnaire is approximately 20 minutes and participants will be asked questions relating to generic and technical aspects of clinical assessments, evidence generation and efficacy studies. Respondents will have a three-week window to answer each survey round. Subsequent to each round, participants will receive feedback consisting of the distribution of responses and a summary of comments, followed by a detailed report and a letter of participation at the end of the study. Participants and their views are representative of their respective institution and agreement to participate is fully compliant with institutional research requirements. Each participant will have the right to withdraw from the study at any point. Data collection is planned between September-December 2018.

A coding system will secure anonymity of study participants among the panel experts, with responses being identifiable to the principal investigator only to allow follow-up. Two independent panels will be convened: one panel composed of experts from HTABs across EU Member States, while the second panel will constitute regulatory representatives from the European Medicines Agency (EMA). This research is congruent with the priority areas for collaboration of the joint EMA-EUnetHTA 2017-2020 three-year work plan on pre- and post-authorisation evidence generation, with the aim of optimising applicant data packages for authorisation and reimbursement assessments.

The participation of expert(s) in the study on behalf of your institution is highly appreciated and will significantly contribute towards the existing technical cooperation in assessments. You are kindly requested to confirm participation, including relevant contact details of participant(s), or to forward any queries to the undersigned via email on dylan.said.11@um.edu.mt.

Thank you in advance for your time and consideration.

Best regards,

Dylan Said (Principal investigator)
Department of Pharmacy,
Faculty of Medicine and Surgery,
University of Malta,
Msida,
Malta

Recruitment letter to experts from the European Medicines Agency:

L-Università ta' Malta
Faculty of Medicine & Surgery
Department of Pharmacy

22 August 2018

Dear Sir/Madam,

As a second year Doctorate in Pharmacy student at the University of Malta, I would like to invite you to participate in a Delphi study. This comprises part of the methodology for my thesis project titled 'Evidence generation in the clinical development of medicines for leukaemia' conducted in partial fulfilment of the requirements for the Doctorate in Pharmacy degree.

A robust clinical development strategy of antineoplastic agents, investigating parameters that satisfy regulatory and health technology assessment (HTA) requirements, can potentially contribute to enhanced patient access to innovative treatments. Harmonisation of evidence needs eliminates duplication of assessments and promotes data uniformity along the life cycle of a medicine, with inferred benefits extending to all stakeholders concerned. The purpose of the study is to determine a set of core efficacy outcomes prioritised by EU medicines regulators and Health Technology Assessment Bodies (HTABs) for clinical trials investigating medicines in leukaemia.

The study will employ a web-based platform to disseminate a validated online questionnaire in a two-round e-Delphi method, with respondent instructions preceding each survey round. Duration of completion for each questionnaire is approximately 20 minutes and participants will be asked questions relating to generic and technical aspects of clinical assessments, evidence generation and efficacy studies. Respondents will have a three-week window to answer each survey round. Subsequent to each round, participants will receive feedback consisting of the distribution of responses and a summary of comments, followed by a detailed report and a letter of participation at the end of the study. Participants and their views are representative of their respective institution and agreement to participate is fully compliant with institutional research requirements. Each participant will have the right to withdraw from the study at any point. Data collection is planned between September-December 2018.

A coding system will secure anonymity of study participants among the panel experts, with responses being identifiable to the principal investigator only to allow follow-up. Two independent panels will be convened: one panel composed of experts from HTABs across EU Member States, while the second panel will constitute regulatory representatives from the European Medicines Agency (EMA). This research is congruent with the priority areas for collaboration of the joint EMA-EUnetHTA 2017-2020 three-year work plan on pre- and post-authorisation evidence generation, with the aim of optimising applicant data packages for authorisation and reimbursement assessments.

Your participation would be highly appreciated and will significantly contribute towards the existing technical cooperation in assessments. You are kindly requested to confirm participation or to forward any queries to the undersigned via email on dylan.said.11@um.edu.mt.

Thank you in advance for your time and consideration.

Best regards,

Dylan Said (Principal investigator)
Department of Pharmacy,
Faculty of Medicine and Surgery,
University of Malta,
Msida,
Malta

Appendix 8

Delphi round two questionnaires

Delphi cycle 1 (HTA) round 2:

L-Università ta' Malta
Faculty of Medicine & Surgery
Department of Pharmacy

Response Evaluation in Leukaemia (ReVALeu)

Delphi Round 2 Instructions

Dear Participant,

Thank you for allocating the necessary time to complete the first Delphi round, your expertise is considered of great value in reaching the objectives of this research.

Descriptive statistics and individual comments have been provided as feedback on the group opinions for the first Delphi round Section 2 items requesting the importance ratings of survival, time-to-event, response rates/biomarkers and other endpoints in demonstrating the efficacy of anti-leukaemic therapy.

A consensus analysis was performed on the above items adopting a pre-determined approach and threshold. A two-sided consensus on the rating scale is possible and reached if ≥75% of the experts have rated statements as 4 ("Important") or 5 ("Very important") or if a frequency of ≥75% of the responses selected 1 ("Not important at all") or 2 ("Not important") on the scale.

From the 36 efficacy endpoints identified, 25% (n=9) satisfied the criteria for consensus and have been retracted from the tool.

To promote the convergence of opinions, you are kindly requested to revisit your individual ratings profile, compare your responses to the overall scores of the group and re-rate the 27 endpoints that have not reached consensus. The appendix with endpoint definitions is once again provided as a reference document. I would like to take this opportunity to encourage you to complete the second and final round of the Delphi study.

Fields that require an answer are marked with an asterisk (*) and must be answered before progressing with the questionnaire.

A summary of the panel's ratings and comments will be reported back to the experts, indicating the final list of efficacy endpoints that have reached consensus after both rounds.

The estimated time to complete the second Delphi round is approximately 5-10 minutes and experts will have a **deadline of three weeks** to respond.

L-Università ta' Malta
Faculty of Medicine & Surgery
Department of Pharmacy

Response Evaluation in Leukaemia (ReVALeu)

Section 2(a): Endpoint Category 1 (Survival)

*** 1. Indicate the importance of this endpoint in demonstrating efficacy:**

	Not important at all	Not important	Neutral	Important	Very important
Treatment-free survival (TFS)	○	○	○	○	○

Comments:

L-Università ta' Malta | Faculty of Medicine & Surgery | Department of Pharmacy

Response Evaluation in Leukaemia (ReVALeu)

Section 2(b): Endpoint Category 2 (Time-To-Event)

*** 2. Indicate the importance of this endpoint in demonstrating efficacy:**

	Not important at all	Not important	Neutral	Important	Very important
Duration of remission/response (DOR)	○	○	○	○	○
Time to remission/response (TTR)	○	○	○	○	○
Time to treatment (TTT)	○	○	○	○	○
Time to treatment failure (TTF)	○	○	○	○	○
Treatment-free remission/response (TFR)	○	○	○	○	○

Comments:

L-Università ta' Malta | Department of Pharmacy
Faculty of Medicine & Surgery

Response Evaluation in Leukaemia (ReVALeu)

Section 2(c): Endpoint Category 3 (Response Rates and Biomarkers)

*** 3. Indicate the importance of this endpoint in demonstrating efficacy:**

	Not important at all	Not important	Neutral	Important	Very important
Relapse rate (RR)/Rate of progressive disease (PD)	○	○	○	○	○
Complete remission/response rate with incomplete platelet recovery (CRp)	○	○	○	○	○
Complete remission/response rate with incomplete blood count recovery (CRi/CRh)	○	○	○	○	○
Combined complete remission/response rate (CRc)	○	○	○	○	○
Objective remission/response rate (ORR)	○	○	○	○	○
Best overall response rate (BOR)	○	○	○	○	○
Haematologic response rate (HR)	○	○	○	○	○
Cytogenetic response rate (CyR)	○	○	○	○	○
Molecular remission/response (MR)	○	○	○	○	○
Minimal residual disease (MRD)	○	○	○	○	○
Cytokine serum concentrations	○	○	○	○	○

	Not important at all	Not important	Neutral	Important	Very important
Morphologic leukaemia-free state (MLFS)	○	○	○	○	○
Splenomegaly and hepatomegaly response rates	○	○	○	○	○
Return to chronic phase (RTC)	○	○	○	○	○

Comments:

L-Università ta' Malta
Faculty of Medicine & Surgery
Department of Pharmacy

Response Evaluation in Leukaemia (ReVALeu)

Section 2(d): Endpoint Category 4 (Other)

*** 4. Indicate the importance of this endpoint in demonstrating efficacy:**

	Not important at all	Not important	Neutral	Important	Very important
Number of patients with asparagine depletion	○	○	○	○	○
Secondary malignancies	○	○	○	○	○
Number of blood cell transfusions	○	○	○	○	○
Number of chemotherapy cycles received	○	○	○	○	○
Incidence and duration of febrile neutropenia	○	○	○	○	○
Medical care utilisation (MCU)	○	○	○	○	○
Incidence, duration and severity of opportunistic infections	○	○	○	○	○

Comments:

Delphi cycle 2 (regulatory) round 2:

L-Università ta' Malta | Department of Pharmacy
Faculty of Medicine & Surgery

Response Evaluation in Leukaemia (ReVALeu)

Delphi Round 2 Instructions

Dear Participant,

Thank you for allocating the necessary time to complete the first Delphi round, your expertise is considered of great value in reaching the objectives of this research.

Descriptive statistics and individual comments have been provided as feedback on the group opinions for the first Delphi round Section 2 items requesting the importance ratings of survival, time-to-event, response rates/biomarkers and other endpoints in demonstrating the efficacy of anti-leukaemic therapy.

A consensus analysis was performed on the above items adopting a pre-determined approach and threshold. A two-sided consensus on the rating scale is possible and reached if ≥75% of the experts have rated statements as 4 ("Important") or 5 ("Very important") or if a frequency of ≥75% of the responses selected 1 ("Not important at all") or 2 ("Not important") on the scale.

From the 36 efficacy endpoints identified, 36% (n=13) satisfied the criteria for consensus and have been retracted from the tool.

To promote the convergence of opinions, you are kindly requested to revisit your individual ratings profile, compare your responses to the overall scores of the group and re-rate the 23 endpoints that have not reached consensus. The appendix with endpoint definitions is once again provided as a reference document. I would like to take this opportunity to encourage you to complete the second and final round of the Delphi study.

Fields that require an answer are marked with an asterisk (*) and must be answered before progressing with the questionnaire.

A summary of the panel's ratings and comments will be reported back to the experts, indicating the final list of efficacy endpoints that have reached consensus after both rounds.

The estimated time to complete the second Delphi round is approximately 5-10 minutes and experts will have a **deadline of three weeks** to respond.

L-Università ta' Malta
Faculty of Medicine & Surgery
Department of Pharmacy

Response Evaluation in Leukaemia (ReVALeu)

Section 2(a): Endpoint Category 1 (Survival)

*** 1. Indicate the importance of this endpoint in demonstrating efficacy:**

	Not important at all	Not important	Neutral	Important	Very important
Treatment-free survival (TFS)	○	○	○	○	○

Comments:

L-Università ta' Malta
Faculty of Medicine & Surgery
Department of Pharmacy

Response Evaluation in Leukaemia (ReVALeu)

Section 2(b): Endpoint Category 2 (Time-To-Event)

*** 2. Indicate the importance of this endpoint in demonstrating efficacy:**

	Not important at all	Not important	Neutral	Important	Very important
Time to remission/response (TTR)	☐	☐	☐	☐	☐
Time to treatment (TTT)	☐	☐	☐	☐	☐

Comments:

L-Università ta' Malta | Department of Pharmacy
Faculty of Medicine & Surgery

Response Evaluation in Leukaemia (ReVALeu)

Section 2(c): Endpoint Category 3 (Response Rates and Biomarkers)

*** 3. Indicate the importance of this endpoint in demonstrating efficacy:**

	Not important at all	Not important	Neutral	Important	Very important
Relapse rate (RR)/Rate of progressive disease (PD)	☐	☐	☐	☐	☐
Complete remission/response rate with incomplete platelet recovery (CRp)	☐	☐	☐	☐	☐
Complete remission/response rate with incomplete blood count recovery (CRi/CRh)	☐	☐	☐	☐	☐
Partial remission/response rate (PR)	☐	☐	☐	☐	☐
Combined complete remission/response rate (CRc)	☐	☐	☐	☐	☐
Best overall response rate (BOR)	☐	☐	☐	☐	☐
Haematologic response rate (HR)	☐	☐	☐	☐	☐
Cytokine serum concentrations	☐	☐	☐	☐	☐
Morphologic leukaemia-free state (MLFS)	☐	☐	☐	☐	☐
Splenomegaly and hepatomegaly response rates	☐	☐	☐	☐	☐
Return to chronic phase (RTC)	☐	☐	☐	☐	☐

Comments:

L-Università ta' Malta
Faculty of Medicine & Surgery
Department of Pharmacy

Response Evaluation in Leukaemia (ReVALeu)

Section 2(d): Endpoint Category 4 (Other)

*** 4. Indicate the importance of this endpoint in demonstrating efficacy:**

	Not important at all	Not important	Neutral	Important	Very important
Patient reported outcomes (PROs)	☐	☐	☐	☐	☐
Number of patients with asparagine depletion	☐	☐	☐	☐	☐
Secondary malignancies	☐	☐	☐	☐	☐
Number of blood cell transfusions	☐	☐	☐	☐	☐
Number of chemotherapy cycles received	☐	☐	☐	☐	☐
Incidence and duration of febrile neutropenia	☐	☐	☐	☐	☐
Medical care utilisation (MCU)	☐	☐	☐	☐	☐
Incidence, duration and severity of opportunistic infections	☐	☐	☐	☐	☐
Proportion of patients undergoing subsequent allogeneic stem cell transplant (ASCT)	☐	☐	☐	☐	☐

Comments:

Appendix 9

Controlled feedback provided to the expert panels between the first and second Delphi rounds

Controlled feedback provided to Delphi cycle 1 panellists (HTA):

Response Evaluation in Leukaemia (ReVALeu) online surveying tool (HTA)

Delphi Round 1: Distribution of Responses and Summary of Comments

Dear Participant,

Thank you for allocating the necessary time to complete the first Delphi round, your expertise is considered of great value in reaching the objectives of this research.

Descriptive statistics and individual comments are provided herein as feedback on the group opinions for the first Delphi round Section 2 items requesting the importance ratings of Survival, Time-To-Event, Response Rates/Biomarkers and Other endpoints in demonstrating the efficacy of anti-leukaemic therapy.

A consensus analysis was performed on the above items adopting a pre-determined approach and threshold. A two-sided consensus on the rating scale is possible and reached if ≥75% of the experts have rated statements as 4 ("Important") or 5 ("Very important") or if a frequency of ≥75% of the responses selected 1 ("Not important at all") or 2 ("Not important") on the scale. Only statements not reaching consensus have progressed to the second Delphi round.

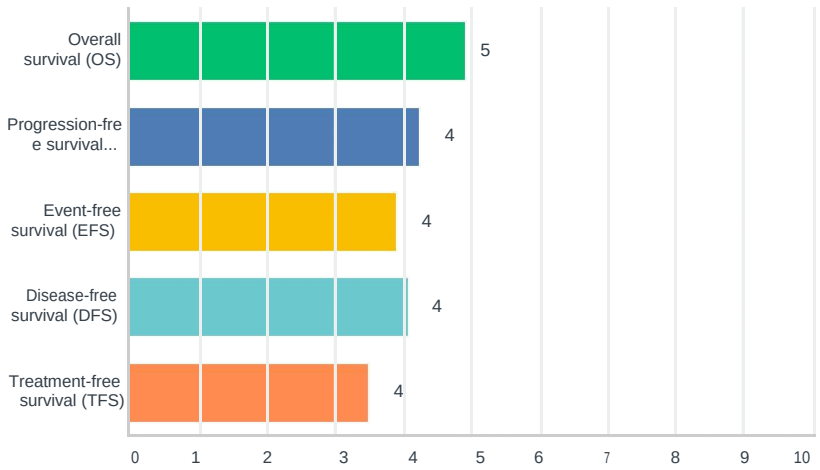
From the 36 efficacy endpoints identified, 25% (n=9) satisfied the criteria for consensus and have been retracted from the tool. The 27 endpoints that have not reached consensus are highlighted below for ease of reference.

To promote the convergence of opinions, you are kindly requested to revisit your individual ratings profile, compare your responses to the overall scores of the group and re-rate the 27 endpoints that have not reached consensus. I would like to take this opportunity to encourage you to complete the final round of the Delphi study by following the web link provided in the email.

Endpoint Category 1: Survival

Indicate the importance of this endpoint in demonstrating efficacy:

Answered: 12 Skipped: 0



	NOT IMPORTANT AT ALL (1)	NOT IMPORTANT (2)	NEUTRAL (3)	IMPORTANT (4)	VERY IMPORTANT (5)	TOTAL	WEIGHTED AVERAGE
Overall survival (OS)	0.00% 0	0.00% 0	0.00% 0	8.33% 1	91.67% 11	12	4.92
Progression-free survival (PFS)	0.00% 0	8.33% 1	8.33% 1	33.33% 4	50.00% 6	12	4.25
Event-free survival (EFS)	0.00% 0	0.00% 0	25.00% 3	58.33% 7	16.67% 2	12	3.92
Disease-free survival (DFS)	0.00% 0	0.00% 0	16.67% 2	58.33% 7	25.00% 3	12	4.08
Treatment-free survival (TFS)	0.00% 0	8.33% 1	41.67% 5	41.67% 5	8.33% 1	12	3.50

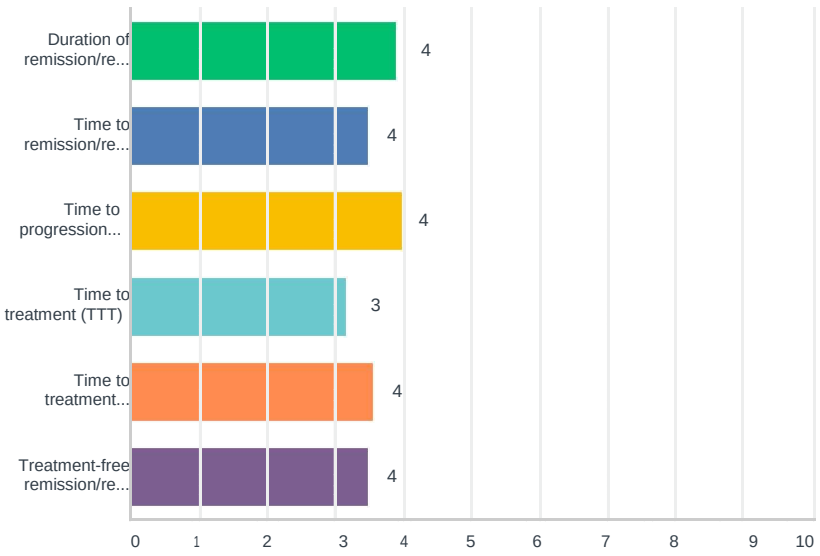
BASIC STATISTICS	MINIMUM	MAXIMUM	MEDIAN	MEAN	STANDARD DEVIATION
Overall survival (OS)	4.00	5.00	5.00	4.92	0.28
Progression-free survival (PFS)	2.00	5.00	4.50	4.25	0.92
Event-free survival (EFS)	3.00	5.00	4.00	3.92	0.64
Disease-free survival (DFS)	3.00	5.00	4.00	4.08	0.64
Treatment-free survival (TFS)	2.00	5.00	3.50	3.50	0.76

#	COMMENTS:	DATE
1	OS the most solid and un-biased endpoint. Progression-free survival and Disease-free survival are important and generally applicable endpoints (but depending on stage of disease and indication, when these should be selected, as indicated above).	9/21/2018 1:50 PM

Endpoint Category 2: Time-To-Event

Indicate the importance of this endpoint in demonstrating efficacy:

Answered: 12 Skipped: 0



	NOT IMPORTANT AT ALL (1)	NOT IMPORTANT (2)	NEUTRAL (3)	IMPORTANT (4)	VERY IMPORTANT (5)	TOTAL	WEIGHTED AVERAGE
Duration of remission/response (DOR)	0.00% 0	0.00% 0	33.33% 4	41.67% 5	25.00% 3	12	3.92
Time to remission/response (TTR)	0.00% 0	8.33% 1	41.67% 5	41.67% 5	8.33% 1	12	3.50
Time to progression (TTP)	0.00% 0	0.00% 0	25.00% 3	50.00% 6	25.00% 3	12	4.00
Time to treatment (TTT)	0.00% 0	8.33% 1	66.67% 8	25.00% 3	0.00% 0	12	3.17
Time to treatment failure (TTF)	0.00% 0	0.00% 0	41.67% 5	58.33% 7	0.00% 0	12	3.58
Treatment-free remission/response (TFR)	0.00% 0	8.33% 1	41.67% 5	41.67% 5	8.33% 1	12	3.50

BASIC STATISTICS

	MINIMUM	MAXIMUM	MEDIAN	MEAN	STANDARD DEVIATION
Duration of remission/response (DOR)	3.00	5.00	4.00	3.92	0.76
Time to remission/response (TTR)	2.00	5.00	3.50	3.50	0.76
Time to progression (TTP)	3.00	5.00	4.00	4.00	0.71
Time to treatment (TTT)	2.00	4.00	3.00	3.17	0.55

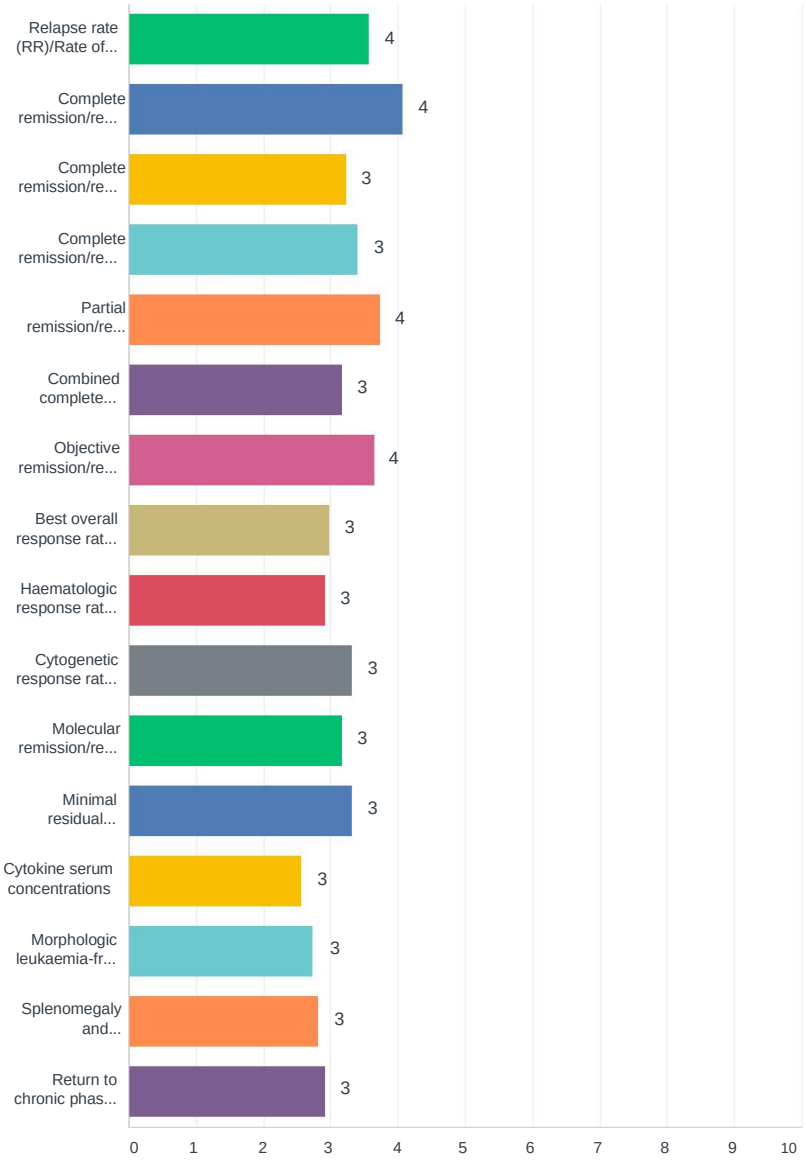
Time to treatment failure (TTF)	3.00	4.00	4.00	3.58	0.49
Treatment-free remission/response (TFR)	2.00	5.00	3.50	3.50	0.76

#	COMMENTS:	DATE
1	Long-term efficacy endpoints are usually appreciated in hematological patients .	9/21/2018 1:53 PM
2	Ii is quite difficult to indicate the importance as it may differ across treatment modalities (e.g. early but short-lived responses versus later but durable ones).	9/10/2018 9:00 AM

Endpoint Category 3: Response Rates and Biomarkers

Indicate the importance of this endpoint in demonstrating efficacy:

Answered: 12 Skipped: 0



	NOT IMPORTANT AT ALL (1)	NOT IMPORTANT (2)	NEUTRAL (3)	IMPORTANT (4)	VERY IMPORTANT (5)	TOTAL	WEIGHTED AVERAGE
Relapse rate (RR)/Rate of progressive disease (PD)	0.00% 0	0.00% 0	50.00% 6	41.67% 5	8.33% 1	12	3.58

Complete remission/response rate (CR)	0.00% 0	0.00% 0	16.67% 2	58.33% 7	25.00% 3	12	4.08
Complete remission/response rate with incomplete platelet recovery (CRp)	0.00% 0	16.67% 2	41.67% 5	41.67% 5	0.00% 0	12	3.25
Complete remission/response rate with incomplete blood count recovery (CRi/CRh)	0.00% 0	8.33% 1	41.67% 5	50.00% 6	0.00% 0	12	3.42
Partial remission/response rate (PR)	8.33% 1	0.00% 0	16.67% 2	58.33% 7	16.67% 2	12	3.75
Combined complete remission/response rate (CRc)	0.00% 0	25.00% 3	41.67% 5	25.00% 3	8.33% 1	12	3.17
Objective remission/response rate (ORR)	0.00% 0	16.67% 2	25.00% 3	33.33% 4	25.00% 3	12	3.67
Best overall response rate (BOR)	0.00% 0	25.00% 3	50.00% 6	25.00% 3	0.00% 0	12	3.00
Haematologic response rate (HR)	8.33% 1	16.67% 2	50.00% 6	25.00% 3	0.00% 0	12	2.92
Cytogenetic response rate (CyR)	0.00% 0	8.33% 1	50.00% 6	41.67% 5	0.00% 0	12	3.33
Molecular remission/response (MR)	0.00% 0	16.67% 2	50.00% 6	33.33% 4	0.00% 0	12	3.17
Minimal residual disease (MRD)	0.00% 0	16.67% 2	33.33% 4	50.00% 6	0.00% 0	12	3.33
Cytokine serum concentrations	8.33% 1	33.33% 4	50.00% 6	8.33% 1	0.00% 0	12	2.58
Morphologic leukaemia-free state (MLFS)	8.33% 1	25.00% 3	50.00% 6	16.67% 2	0.00% 0	12	2.75
Splenomegaly and hepatomegaly response rates	0.00% 0	33.33% 4	50.00% 6	16.67% 2	0.00% 0	12	2.83
Return to chronic phase (RTC)	0.00% 0	25.00% 3	58.33% 7	16.67% 2	0.00% 0	12	2.92

BASIC STATISTICS

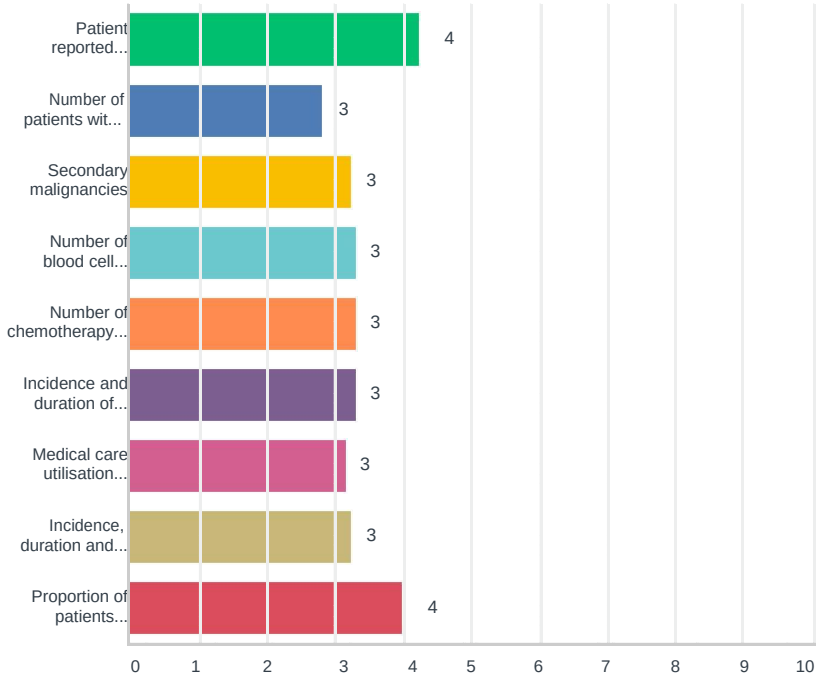
	MINIMUM	MAXIMUM	MEDIAN	MEAN	STANDARD DEVIATION
Relapse rate (RR)/Rate of progressive disease (PD)	3.00	5.00	3.50	3.58	0.64
Complete remission/response rate (CR)	3.00	5.00	4.00	4.08	0.64
Complete remission/response rate with incomplete platelet recovery (CRp)	2.00	4.00	3.00	3.25	0.72

Complete remission/response rate with incomplete blood count recovery (CRi/CRh)	2.00	4.00	3.50	3.42	0.64
Partial remission/response rate (PR)	1.00	5.00	4.00	3.75	1.01
Combined complete remission/response rate (CRc)	2.00	5.00	3.00	3.17	0.90
Objective remission/response rate (ORR)	2.00	5.00	4.00	3.67	1.03
Best overall response rate (BOR)	2.00	4.00	3.00	3.00	0.71
Haematologic response rate (HR)	1.00	4.00	3.00	2.92	0.86
Cytogenetic response rate (CyR)	2.00	4.00	3.00	3.33	0.62
Molecular remission/response (MR)	2.00	4.00	3.00	3.17	0.69
Minimal residual disease (MRD)	2.00	4.00	3.50	3.33	0.75
Cytokine serum concentrations	1.00	4.00	3.00	2.58	0.76
Morphologic leukaemia-free state (MLFS)	1.00	4.00	3.00	2.75	0.83
Splenomegaly and hepatomegaly response rates	2.00	4.00	3.00	2.83	0.69
Return to chronic phase (RTC)	2.00	4.00	3.00	2.92	0.64

#	COMMENTS:	DATE
1	As stated above, depending on the stage of the disease etc.... These all are important, but not as important as OS, or PFS etc.	9/21/2018 2:01 PM

Endpoint Category 4: Other

Indicate the importance of this endpoint in demonstrating efficacy:



	NOT IMPORTANT AT ALL (1)	NOT IMPORTANT (2)	NEUTRAL (3)	IMPORTANT (4)	VERY IMPORTANT (5)	TOTAL	WEIGHTED AVERAGE
Patient reported outcomes (PROs)	0.00% 0	0.00% 0	0.00% 0	75.00% 9	25.00% 3	12	4.25
Number of patients with asparagine depletion	8.33% 1	8.33% 1	75.00% 9	8.33% 1	0.00% 0	12	2.83
Secondary malignancies	8.33% 1	16.67% 2	16.67% 2	58.33% 7	0.00% 0	12	3.25
Number of blood cell transfusions (platelets or red blood cells)	8.33% 1	0.00% 0	50.00% 6	33.33% 4	8.33% 1	12	3.33
Number of chemotherapy cycles received	8.33% 1	16.67% 2	33.33% 4	16.67% 2	25.00% 3	12	3.33
Incidence and duration of febrile neutropenia	8.33% 1	8.33% 1	25.00% 3	58.33% 7	0.00% 0	12	3.33
Medical care utilisation (MCU)	8.33% 1	16.67% 2	41.67% 5	16.67% 2	16.67% 2	12	3.17

Incidence, duration and severity of opportunistic infections	8.33% 1	8.33% 1	33.33% 4	50.00% 6	0.00% 0	12	3.25
Proportion of patients undergoing subsequent allogeneic stem cell transplant (ASCT)	0.00% 0	0.00% 0	16.67% 2	66.67% 8	16.67% 2	12	4.00

BASIC STATISTICS

	MINIMUM	MAXIMUM	MEDIAN	MEAN	STANDARD DEVIATION
Patient reported outcomes (PROs)	4.00	5.00	4.00	4.25	0.43
Number of patients with asparagine depletion	1.00	4.00	3.00	2.83	0.69
Secondary malignancies	1.00	4.00	4.00	3.25	1.01
Number of blood cell transfusions	1.00	5.00	3.00	3.33	0.94
Number of chemotherapy cycles received	1.00	5.00	3.00	3.33	1.25
Incidence and duration of febrile neutropenia	1.00	4.00	4.00	3.33	0.94
Medical care utilisation (MCU)	1.00	5.00	3.00	3.17	1.14
Incidence, duration and severity of opportunistic infections	1.00	4.00	3.50	3.25	0.92
Proportion of patients undergoing subsequent allogeneic stem cell transplant (ASCT)	3.00	5.00	4.00	4.00	0.58

#	COMMENTS:	DATE
1	Most of these are important from safety point of view, and also for consideration of risk-benefit-ration, but not from mere EFFICACY point of view as asked in this question.	9/21/2018 2:07 PM

Controlled feedback provided to Delphi cycle 2 panellists (regulatory):

Response Evaluation in Leukaemia (ReVALeu) online surveying tool (Regulatory)

Delphi Round 1: Distribution of Responses and Summary of Comments

Dear Participant,

Thank you for allocating the necessary time to complete the first Delphi round, your expertise is considered of great value in reaching the objectives of this research.

Descriptive statistics and individual comments are provided herein as feedback on the group opinions for the first Delphi round Section 2 items requesting the importance ratings of Survival, Time-To-Event, Response Rates/Biomarkers and Other endpoints in demonstrating the efficacy of anti-leukaemic therapy.

A consensus analysis was performed on the above items adopting a pre-determined approach and threshold. A two-sided consensus on the rating scale is possible and reached if ≥75% of the experts have rated statements as 4 ("Important") or 5 ("Very important") or if a frequency of ≥75% of the responses selected 1 ("Not important at all") or 2 ("Not important") on the scale. Only statements not reaching consensus have progressed to the second Delphi round.

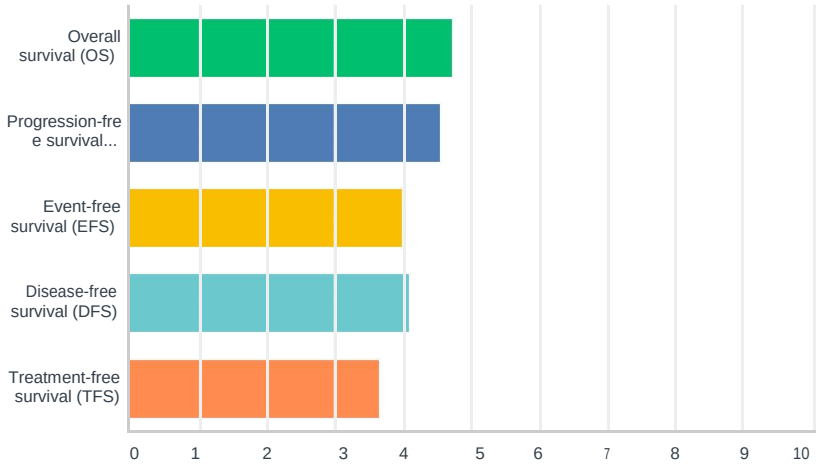
From the 36 efficacy endpoints identified, 36% (n=13) satisfied the criteria for consensus and have been retracted from the tool. The 23 endpoints that have not reached consensus are highlighted below for ease of reference.

To promote the convergence of opinions, you are kindly requested to revisit your individual ratings profile, compare your responses to the overall scores of the group and re-rate the 23 endpoints that have not reached consensus. I would like to take this opportunity to encourage you to complete the final round of the Delphi study by following the web link provided in the email.

Endpoint Category 1: Survival

Indicate the importance of this endpoint in demonstrating efficacy:

Answered: 18 Skipped: 0



	NOT IMPORTANT AT ALL (1)	NOT IMPORTANT (2)	NEUTRAL (3)	IMPORTANT (4)	VERY IMPORTANT (5)	TOTAL	WEIGHTED AVERAGE
Overall survival (OS)	0.00% 0	0.00% 0	5.56% 1	16.67% 3	77.78% 14	18	4.72
Progression-free survival (PFS)	0.00% 0	0.00% 0	0.00% 0	44.44% 8	55.56% 10	18	4.56
Event-free survival (EFS)	0.00% 0	0.00% 0	16.67% 3	66.67% 12	16.67% 3	18	4.00
Disease-free survival (DFS)	0.00% 0	0.00% 0	11.11% 2	66.67% 12	22.22% 4	18	4.11
Treatment-free survival (TFS)	0.00% 0	11.11% 2	16.67% 3	66.67% 12	5.56% 1	18	3.67

BASIC STATISTICS

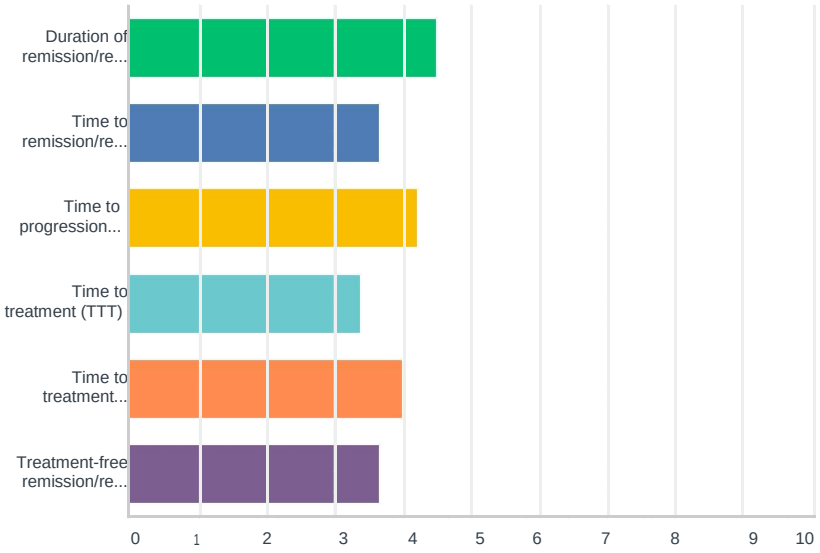
	MINIMUM	MAXIMUM	MEDIAN	MEAN	STANDARD DEVIATION
Overall survival (OS)	3.00	5.00	5.00	4.72	0.56
Progression-free survival (PFS)	4.00	5.00	5.00	4.56	0.50
Event-free survival (EFS)	3.00	5.00	4.00	4.00	0.58
Disease-free survival (DFS)	3.00	5.00	4.00	4.11	0.57
Treatment-free survival (TFS)	2.00	5.00	4.00	3.67	0.75

#	COMMENTS:	DATE
1	All of them are very important according to the phase of clinical trials.	11/2/2018 7:19 AM

Endpoint Category 2: Time-To-Event

Indicate the importance of this endpoint in demonstrating efficacy:

Answered: 18 Skipped: 0

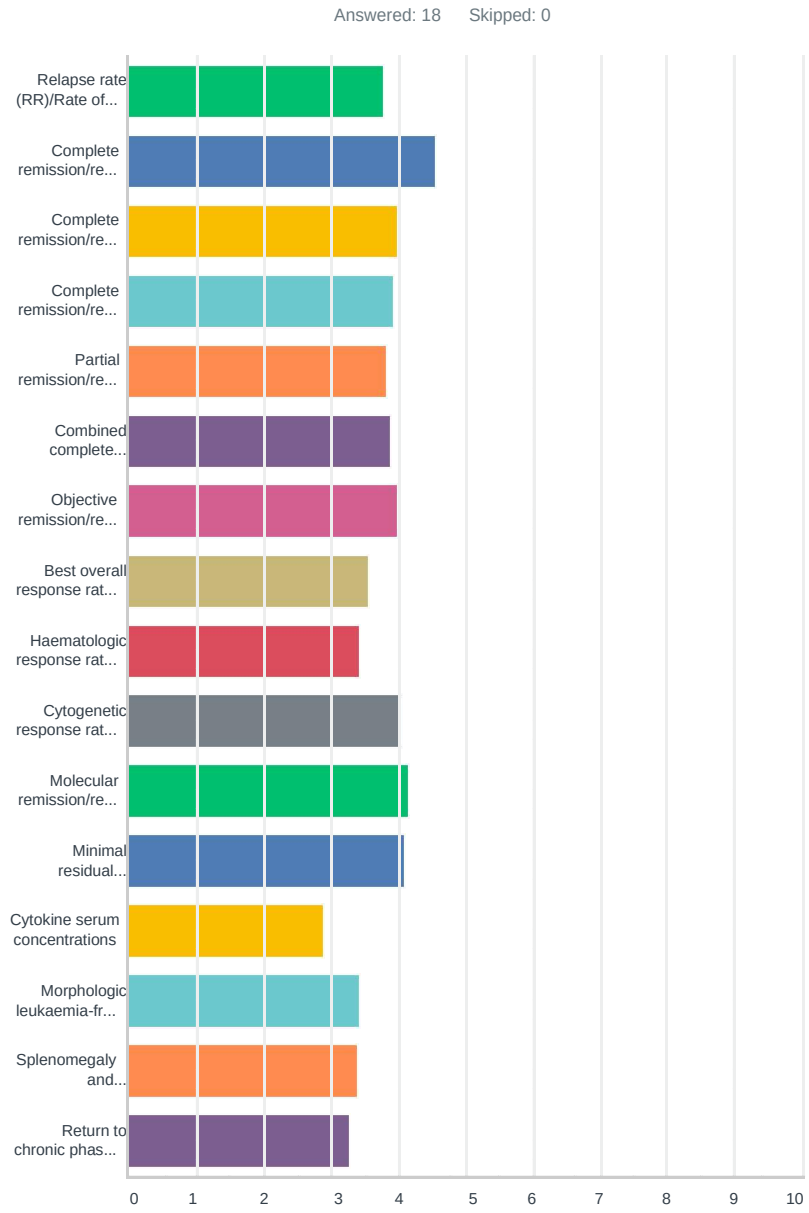


	NOT IMPORTANT AT ALL (1)	NOT IMPORTANT (2)	NEUTRAL (3)	IMPORTANT (4)	VERY IMPORTANT (5)	TOTAL	WEIGHTED AVERAGE
Duration of remission/response (DOR)	0.00% 0	0.00% 0	0.00% 0	50.00% 9	50.00% 9	18	4.50
Time to remission/response (TTR)	0.00% 0	5.56% 1	33.33% 6	50.00% 9	11.11% 2	18	3.67
Time to progression (TTP)	0.00% 0	0.00% 0	11.11% 2	55.56% 10	33.33% 6	18	4.22
Time to treatment (TTT)	5.56% 1	11.11% 2	27.78% 5	50.00% 9	5.56% 1	18	3.39
Time to treatment failure (TTF)	0.00% 0	0.00% 0	16.67% 3	66.67% 12	16.67% 3	18	4.00
Treatment-free remission/response (TFR)	5.56% 1	5.56% 1	11.11% 2	72.22% 13	5.56% 1	18	3.67

BASIC STATISTICS	MINIMUM	MAXIMUM	MEDIAN	MEAN	STANDARD DEVIATION
Duration of remission/response (DOR)	4.00	5.00	4.50	4.50	0.50
Time to remission/response (TTR)	2.00	5.00	4.00	3.67	0.75
Time to progression (TTP)	3.00	5.00	4.00	4.22	0.63
Time to treatment (TTT)	1.00	5.00	4.00	3.39	0.95
Time to treatment failure (TTF)	3.00	5.00	4.00	4.00	0.58
Treatment-free remission/response (TFR)	1.00	5.00	4.00	3.67	0.88

Endpoint Category 3: Response Rates and Biomarkers

Indicate the importance of this endpoint in demonstrating efficacy:



	NOT IMPORTANT AT ALL (1)	NOT IMPORTANT (2)	NEUTRAL (3)	IMPORTANT (4)	VERY IMPORTANT (5)	TOTAL	WEIGHTED AVERAGE
Relapse rate (RR)/Rate of progressive disease (PD)	0.00% 0	5.56% 1	38.89% 7	27.78% 5	27.78% 5	18	3.78

Complete remission/response rate (CR)	0.00% 0	5.56% 1	0.00% 0	27.78% 5	66.67% 12	18	4.56
Complete remission/response rate with incomplete platelet recovery (CRp)	0.00% 0	5.56% 1	22.22% 4	38.89% 7	33.33% 6	18	4.00
Complete remission/response rate with incomplete blood count recovery (CRi/CRh)	0.00% 0	5.56% 1	22.22% 4	44.44% 8	27.78% 5	18	3.94
Partial remission/response rate (PR)	0.00% 0	5.56% 1	22.22% 4	55.56% 10	16.67% 3	18	3.83
Combined complete remission/response rate (CRc)	0.00% 0	5.56% 1	27.78% 5	38.89% 7	27.78% 5	18	3.89
Objective remission/response rate (ORR)	0.00% 0	5.56% 1	16.67% 3	50.00% 9	27.78% 5	18	4.00
Best overall response rate (BOR)	0.00% 0	5.56% 1	44.44% 8	38.89% 7	11.11% 2	18	3.56
Haematologic response rate (HR)	0.00% 0	11.11% 2	44.44% 8	33.33% 6	11.11% 2	18	3.44
Cytogenetic response rate (CyR)	0.00% 0	5.56% 1	11.11% 2	55.56% 10	27.78% 5	18	4.06
Molecular remission/response (MR)	0.00% 0	5.56% 1	11.11% 2	44.44% 8	38.89% 7	18	4.17
Minimal residual disease (MRD)	0.00% 0	5.56% 1	11.11% 2	50.00% 9	33.33% 6	18	4.11
Cytokine serum concentrations	11.11% 2	33.33% 6	22.22% 4	22.22% 4	11.11% 2	18	2.89
Morphologic leukaemia-free state (MLFS)	0.00% 0	22.22% 4	22.22% 4	44.44% 8	11.11% 2	18	3.44
Splenomegaly and hepatomegaly response rates	0.00% 0	16.67% 3	33.33% 6	44.44% 8	5.56% 1	18	3.39
Return to chronic phase (RTC)	0.00% 0	22.22% 4	33.33% 6	38.89% 7	5.56% 1	18	3.28

BASIC STATISTICS

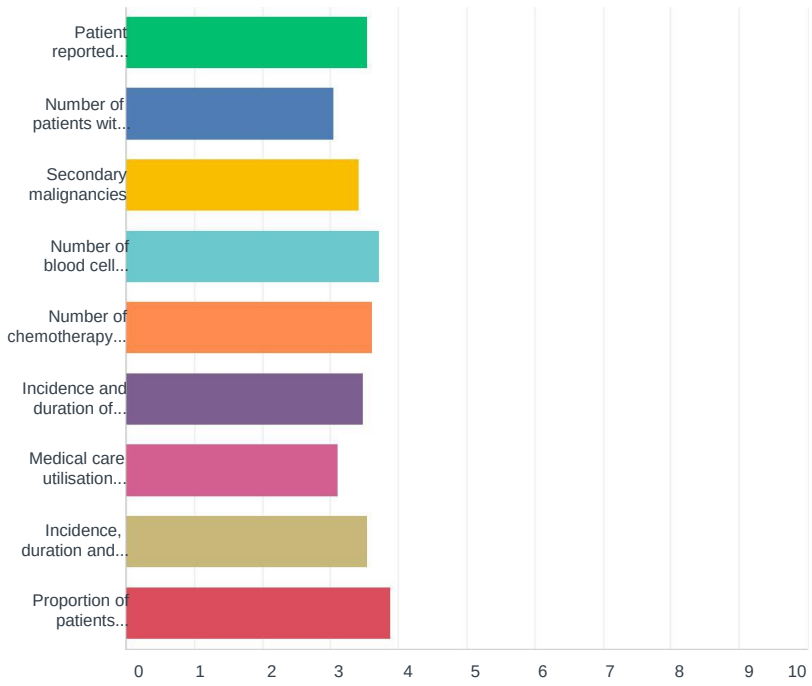
	MINIMUM	MAXIMUM	MEDIAN	MEAN	STANDARD DEVIATION
Relapse rate (RR)/Rate of progressive disease (PD)	2.00	5.00	4.00	3.78	0.92
Complete remission/response rate (CR)	2.00	5.00	5.00	4.56	0.76
Complete remission/response rate with incomplete platelet recovery (CRp)	2.00	5.00	4.00	4.00	0.88
Complete remission/response rate with incomplete blood count recovery (CRi/CRh)	2.00	5.00	4.00	3.94	0.85

Partial remission/response rate (PR)	2.00	5.00	4.00	3.83	0.76
Combined complete remission/response rate (CRc)	2.00	5.00	4.00	3.89	0.87
Objective remission/response rate (ORR)	2.00	5.00	4.00	4.00	0.82
Best overall response rate (BOR)	2.00	5.00	3.50	3.56	0.76
Haematologic response rate (HR)	2.00	5.00	3.00	3.44	0.83
Cytogenetic response rate (CyR)	2.00	5.00	4.00	4.06	0.78
Molecular remission/response (MR)	2.00	5.00	4.00	4.17	0.83
Minimal residual disease (MRD)	2.00	5.00	4.00	4.11	0.81
Cytokine serum concentrations	1.00	5.00	3.00	2.89	1.20
Morphologic leukaemia-free state (MLFS)	2.00	5.00	4.00	3.44	0.96
Splenomegaly and hepatomegaly response rates	2.00	5.00	3.50	3.39	0.83
Return to chronic phase (RTC)	2.00	5.00	3.00	3.28	0.87

Endpoint Category 4: Other

Indicate the importance of this endpoint in demonstrating efficacy:

Answered: 18 Skipped: 0



	NOT IMPORTANT AT ALL (1)	NOT IMPORTANT (2)	NEUTRAL (3)	IMPORTANT (4)	VERY IMPORTANT (5)	TOTAL	WEIGHTED AVERAGE
Patient reported outcomes (PROs)	0.00% 0	22.22% 4	22.22% 4	33.33% 6	22.22% 4	18	3.56
Number of patients with asparagine depletion	0.00% 0	38.89% 7	27.78% 5	22.22% 4	11.11% 2	18	3.06
Secondary malignancies	0.00% 0	33.33% 6	5.56% 1	44.44% 8	16.67% 3	18	3.44
Number of blood cell transfusions	0.00% 0	11.11% 2	27.78% 5	38.89% 7	22.22% 4	18	3.72
Number of chemotherapy cycles received	0.00% 0	11.11% 2	44.44% 8	16.67% 3	27.78% 5	18	3.61
Incidence and duration of febrile neutropenia	0.00% 0	16.67% 3	33.33% 6	33.33% 6	16.67% 3	18	3.50
Medical care utilisation (MCU)	0.00% 0	33.33% 6	27.78% 5	33.33% 6	5.56% 1	18	3.11

Incidence, duration and severity of opportunistic infections	0.00% 0	22.22% 4	22.22% 4	33.33% 6	22.22% 4	18	3.56
Proportion of patients undergoing subsequent allogeneic stem cell transplant (ASCT)	0.00% 0	11.11% 2	16.67% 3	44.44% 8	27.78% 5	18	3.89

BASIC STATISTICS

	MINIMUM	MAXIMUM	MEDIAN	MEAN	STANDARD DEVIATION
Patient reported outcomes (PROs)	2.00	5.00	4.00	3.56	1.07
Number of patients with asparagine depletion	2.00	5.00	3.00	3.06	1.03
Secondary malignancies	2.00	5.00	4.00	3.44	1.12
Number of blood cell transfusions	2.00	5.00	4.00	3.72	0.93
Number of chemotherapy cycles received	2.00	5.00	3.00	3.61	1.01
Incidence and duration of febrile neutropenia	2.00	5.00	3.50	3.50	0.96
Medical care utilisation (MCU)	2.00	5.00	3.00	3.11	0.94
Incidence, duration and severity of opportunistic infections	2.00	5.00	4.00	3.56	1.07
Proportion of patients undergoing subsequent allogeneic stem cell transplant (ASCT)	2.00	5.00	4.00	3.89	0.94

#	COMMENTS:	DATE
1	ultimately no detrimental effect on PROs should be achieved, a derived benefit in PROs can be gathered as supportive in case of primary and secondary achieved EPs. Isolated PRO benefit is of limited value. Rather have incidence, severity and duration of neutropenia AND febrile neutropenia.	10/23/2018 8:08 AM

Appendix 10

Letter of participation to experts completing the Delphi process

L-Università ta' Malta
Faculty of
Medicine & Surgery

Department
of Pharmacy

dd/mm/yyyy

Dear <Participant's name>,

This letter is being sent to extend my gratitude for accepting to participate in the study and successfully completing the two Delphi rounds pertaining to the thesis entitled 'Evidence generation in the clinical development of medicines for leukaemia' held during September – November 2018.

Kindly find attached a detailed report on the main findings from the Delphi study.

Thanking you once again.

Best regards,

Dylan Said (Principal Investigator)
Department of Pharmacy,
Faculty of Medicine and Surgery,
University of Malta,
Msida,
Malta

Appendix 11

Dissemination of results in international fora

Abstract accepted as an oral presentation at the 10th Malta Medical School Conference (MMSC) held in Malta, 29 November-1 December 2018:

Analysis of efficacy parameters used in clinical trials for anti-leukaemic therapy

Dylan Said, John Joseph Borg, Maresca Attard Pizzuto, Anthony Serracino-Inglott

Introduction: Clinical trials (CTs) in leukaemia have validated biological markers as surrogates for measures of direct clinical benefit, such as overall survival (OS), contributing to shorter CT durations. The aim of the study was to analyse the choice of efficacy endpoints over time for leukaemia CTs conducted in the European Union (EU).

Methods: Interventional Phase II to Phase IV CTs investigating therapies in subtypes of leukaemia for all ages over an 11-year period (2007-2017) were identified from the EU Clinical Trials Register database. CTs reporting efficacy data for medicinal products of chemical, biological and biotechnological origin were included in the study. To understand the shift in response assessment, efficacy endpoints were extracted from the selected CTs, categorised and trends in the frequency of selection evaluated.

Results: Thirty-six unique efficacy endpoints were identified from the final data set of CTs (*N*=431) and categorised into four domains consisting of survival (*n*=5*)*, response rates and biomarkers (*n*=16), time-to-event (*n*=6), and other (*n*=9) parameters. OS (66%, *n*=285) was the most studied endpoint across the defined period at primary and/or secondary level, with complete response (CR) rate (19%, *n*=81) surpassing OS (17%, *n*=72) as a primary measure. The proportion of CTs reporting minimal residual disease (MRD) as an endpoint registered the highest frequency change from 33% (2012-2014) to 50% (2015-2017).

Conclusion: Regulatory authorities consider OS as a robust endpoint which is in line with the findings observed. CR is a strong predictor of OS whilst MRD is emerging as a surrogate endpoint, both bypassing the maturation of survival data.

Abstract accepted as a poster presentation at The Professional Society for Health Economics and Outcomes Research (ISPOR) 2019 conference held in New Orleans, USA, 18-22 May 2019:

Decision-maker perspectives on outcomes studied in leukaemia clinical trials: The Response Evaluation in Leukaemia (REVALEU) study protocol

Said D, Borg JJ, Attard Pizzuto M, Serracino-Inglott A

Objectives: Evidence of efficacy for antineoplastic agents may be valued differently by regulatory and health technology assessment (HTA) bodies in the European Union (EU), impacting decision-making and access to novel medicines. A study protocol was developed to identify and harmonise core outcomes prioritised by European regulatory and HTA experts for clinical trials (CTs) investigating leukaemic disorders.

Methods: The protocol was developed as follows: (1) identification of Phase II to Phase IV CTs registered in the EU Clinical Trials Register database throughout an 11-year period (2007-2017); (2) screening of CTs against inclusion criteria; (3) extraction of efficacy endpoints from selected trials and grouping according to type of measurement; (4) design of the Response Evaluation in Leukaemia (REVALEU) online surveying tool; (5) testing of the tool for content validity by means of the content validity index (CVI) method and for reliability using the test-retest approach; (6) recruitment of regulatory and HTA onco-haematology experts in a two-round e-Delphi process with two independent panels; (7) determination of the outcomes that have reached consensus.

Results: Thirty-six unique efficacy measures were identified from the final data set of CTs (N=431) and grouped into the endpoint clusters of survival (n=5), response rates and biomarkers (n=16), time-to-event (n=6), and other (n=9). The REVALEU tool demonstrated high content validity as shown from the mean scale-level CVI (S-CVI) score of 93% for the assessed domains. Intra-subject reliability was upheld across the tool as confirmed from the Kendall-Tau and Kappa statistical test values ($p<0.05$). Thirty-six experts were recruited in the e-Delphi process; 24 from configurations and external expertise of the European Medicines Agency (EMA) and 12 from HTA bodies in 9 EU countries.

Conclusions: The protocol developed will form the basis for the identification of core outcomes, overarching and leukaemia subtype-specific, that are able to support industry stakeholders in obtaining regulatory and reimbursement approvals.

Abstract accepted as a publication in the 'Value in Health' Journal:

Decision-maker perspectives on outcomes studied in leukaemia clinical trials: The Response Evaluation in Leukaemia (REVALEU) study protocol

Said D, Borg JJ, Attard Pizzuto M, Serracino-Inglott A

Objectives: Evidence of efficacy for antineoplastic agents may be valued differently by regulatory and health technology assessment (HTA) bodies in the European Union (EU), impacting decision-making and access to novel medicines. A study protocol was developed to identify and harmonise core outcomes prioritised by European regulatory and HTA experts for clinical trials (CTs) investigating leukaemic disorders.

Methods: The protocol was developed as follows: (1) identification of Phase II to Phase IV CTs registered in the EU Clinical Trials Register database throughout an 11-year period (2007-2017); (2) screening of CTs against inclusion criteria; (3) extraction of efficacy endpoints from selected trials and grouping according to type of measurement; (4) design of the Response Evaluation in Leukaemia (REVALEU) online surveying tool; (5) testing of the tool for content validity by means of the content validity index (CVI) method and for reliability using the test-retest approach; (6) recruitment of regulatory and HTA onco-haematology experts in a two-round e-Delphi process with two independent panels; (7) determination of the outcomes that have reached consensus.

Results: Thirty-six unique efficacy measures were identified from the final data set of CTs (N=431) and grouped into the endpoint clusters of survival (n=5), response rates and biomarkers (n=16), time-to-event (n=6), and other (n=9). The REVALEU tool demonstrated high content validity as shown from the mean scale-level CVI (S-CVI) score of 93% for the assessed domains. Intra-subject reliability was upheld across the tool as confirmed from the Kendall-Tau and Kappa statistical test values ($p<0.05$). Thirty-six experts were recruited in the e-Delphi process; 24 from configurations and external expertise of the European Medicines Agency (EMA) and 12 from HTA bodies in 9 EU countries.

Conclusions: The protocol developed will form the basis for the identification of core outcomes, overarching and leukaemia subtype-specific, that are able to support industry stakeholders in obtaining regulatory and reimbursement approvals.

A study protocol was registered in the Core Outcome Measures in Effectiveness Trials (COMET) online database (registration number:1234), available from: http://www.comet-initiative.org/studies/details/1234

COMET - New Project

Contact Name	Dylan Said
Email address	dylan.said.11@um.edu.mt
Study title	Regulatory and Health Technology Assessment Perspectives on Outcomes Studied in Leukaemia Clinical Trials
Brief Summary	A robust clinical development strategy for medicines indicated in neoplastic disease, gauging metrics of health benefits that satisfy regulatory and health technology assessment (HTA) requirements, potentially favours drug approvals and enhances patient accessibility. Harmonisation of decision-maker evidence needs promotes data uniformity collated along the lifecycle of a medicine. The purpose of the study is to determine a core set of efficacy endpoints prioritised by EU regulators and Health Technology Assessment Bodies (HTABs) in the assessment of haematology anticancer agents, specifically those investigated in leukaemia subtypes. Focus of the study is placed on 'what' efficacy parameters should be measured, with additional insights being captured on preferred clinical trial design elements.
Contributors and affiliations/organisations	Principal investigator: Dylan Said, Department of Pharmacy, University of Malta Supervisor: Prof Anthony Serracino-Inglott, Department of Pharmacy, University of Malta Co-supervisor: Dr Maresca Attard Pizzuto, Department of Pharmacy, University of Malta Advisor: Prof John Joseph Borg, Malta Medicines Authority
Funding source(s)	None.
Health area: category	Blood disorders
Health area: disease name	Leukaemia
Target population: minimum age of population	Birth
Target population: maximum age of population	>65 years of age
Target population: sex	Either
Nature/type of intervention e.g. drug, surgery, any	Medicinal products of chemical, biological and biotechnological origin

Other details about the population within the health area	
Setting for intended use	Research
Other	
Methods	Delphi process,Survey
Other	
Brief description of methods	Interventional Phase II to Phase IV clinical trials investigating therapies in subtypes of leukaemia for all ages over an eleven-year period (2007-2017) were identified from the EU Clinical Trials Register. Unique efficacy endpoints were extracted from the selected clinical trials and categorised into four major categories. Clinical trial efficacy data was used to design the Response Evaluation in Leukaemia (ReVALeu) online surveying tool. The tool was validated, tested for reliability and subsequently disseminated in a two-round, e-Delphi technique with two independent expert panels having recognised expertise in the therapeutic area of onco-haematology. The first Delphi panel recruited experts from HTA bodies across EU Member States, while the second Delphi panel was composed of regulatory representatives from the European Medicines Agency (EMA). Participants were asked to rate the importance of identified endpoints in demonstrating efficacy on a weighted likert scale. Parameters reaching consensus after both rounds and common to both groups of decision-makers were identified as the core outcomes.
Stakeholder group(s) involved	Policy makers,Regulatory agency representatives
Additional or other	Onco-haematology experts from the European Medicines Agency (EMA) and Health Technology Assessment Bodies (HTABs) in the European Union
Study start date	October 2017
Study end date	June 2019
Protocol	N/A

Abstract accepted as a poster presentation at the European Association of Faculties of Pharmacies (EAFP) 2019 conference held in Krakow, Poland, 15-17 May 2019:

Opinions of decision-makers on the clinical development and assessment of antineoplastic agents

Said D, Borg JJ, Attard-Pizzuto M, Serracino-Inglott A

INTRODUCTION

Regulatory early access routes have increased flexibility in the authorisation process by accepting less comprehensive data as basis for approvals [1,2]. Health technology assessment (HTA) bodies tend to be more rigid in their evaluations by requesting mature clinical datasets for economic modelling and comparative efficacy assessments [2,3,4]. An analysis of parallel scientific advice outputs between HTA bodies and the European Medicines Agency (EMA) reports that both decision-maker groups failed to fully agree on multiple clinical trial design elements [5]. Industry stakeholders have indicated that oncology medicinal products are significantly associated with divergences in HTA and regulatory positions [4]. The objective of this study was to capture and compare the perspectives of regulatory and HTA decision-makers on aspects related to the clinical development and assessment of antineoplastic agents.

MATERIALS AND METHODS

Development, Validation and Reliability Testing of Survey

An online survey was developed to gauge opinions of regulatory and HTA experts on the quality of evidence generated and alignment of clinical assessments in the therapeutic fields of oncology and haematology. Attitudes towards the potential impact of harmonising HTA and regulatory clinical assessments on patient access to innovations were also examined. Survey items consisted of weighted agreement and quality rating scales and rank-type questions. The survey was validated using the content validity index (CVI) method [6] by a multi-disciplinary 8-member validation panel composed of clinical (n=4), regulatory (n=1), HTA (n=2) and informatics (n=1) specialists. The intra-subject reliability of the survey items was confirmed by means of the test-retest approach.

Recruitment of Study Participants

Oncology experts were invited to participate using non-probability, purposive sampling. Contact details for HTA bodies were retrieved from the online platforms of the European Network for Health Technology Assessment (EUnetHTA), World Health Organization Regional Office for Europe (WHO/Europe) Health Evidence Network (HEN) and the International Network of Agencies for Health Technology Assessment (INAHTA). Members and alternate members of the EMA Committee for Medicinal Products for Human Use (CHMP), Scientific Advice Working Party (SAWP) and external experts in clinical oncology were identified from the EMA experts database and constituted the regulatory scientific personnel invited to participate in the study.

Statistical Analysis of Opinions

Descriptive and inferential statistics were operated to report the distribution of regulatory and HTA opinions. Weighted mean ratings for ordinal scales were compared between the two expert groups and analysed for statistical significance using the non-parametric Mann-Whitney U Test.

RESULTS AND DISCUSSION

Twelve (12) HTA experts from 9 different EU countries and 18 regulatory representatives completed the survey questions. HTA experts expressed stronger dissent than their regulatory counterparts when asked on their level of agreement to the current alignment between regulatory and HTA clinical evidence needs in assessment procedures of oncology medicines (weighted mean ratings: 3.1 (regulatory); 2.4(HTA)). The majority of HTA respondents consider that the clinical evidence requested for antineoplastic agents by their respective HTA body is akin to that of other HTA bodies (weighted mean rating of 3.8). Conversely, HTA experts were of the opinion that their agencies' evidentiary requirements are not on par to those being requested by the EMA (weighted mean rating of 2.4). This difference was found to be statistically significant with a p-value of 0.006. Decision-makers also had conflicting views on the quality of evidence generated for antineoplastic agents in the pre- and post-authorisation phases (p-value for pre-authorisation phase: 0.01; p-value for post-authorisation phase: 0.04). A mean of 73% versus 21% for the regulatory and HTA opinions respectively rated the quality of evidence as good, very good or excellent.

Regulatory experts expressed a firmer standpoint (weighted mean rating of 3.8) than HTA respondents (weighted mean rating of 3.5) that divergences in clinical evidence requirements negatively impacts patient access to novel cancer treatments. From a list of 6 stakeholder groups, patients were ranked by both decision-makers as the top stakeholder to benefit from enhanced collaboration between regulators and HTA throughout the medicinal product life cycle.

CONCLUSIONS

Scientific expert opinions indicate that clinical evidence needs for antineoplastic agents are not optimally aligned between regulatory and HTA bodies. Decision-makers perceive the quality of evidence generated for medicines indicated in malignancy differently. Regulatory and HTA experts recognise patients as the main stakeholders to gain from greater collaborative initiatives. Findings from this study are intended to stimulate calls for more effective alignment between the two facets, potentially driving faster patient access to novel cancer treatments.

REFERENCES

1 Cressman S, Browman GP, Hoch JS, Kovacic L, Peacock SJ. A time-trend economic analysis of cancer drug trials. Oncologist. 2015;20(2):729-736.

2 Leyens L, Brand A. Early patient access to medicines: Health technology assessment bodies need to catch up with new marketing authorization methods. Public Health Genomics. 2016;19:187-191.

3 Martinalbo J, Bowen D, Camarero J, Chapelin M, Demolis P, Foggi P, et al. Early market access of cancer drugs in the EU. Ann Oncol. 2016;27:96-105.

4 Wang T, McAuslane N, Liberti L, Leufkens H, Hovels A. Building synergy between regulatory and HTA agencies beyond processes and procedures – Can we effectively align evidentiary requirements? A survey of stakeholder perceptions. Value Health. 2018;21(6):707-714.

5 Tafuri G, Pagnini M, Moseley J, Massari M, Petavy F, Behring A et al. How aligned are the perspectives of EU regulators and HTA bodies? A comparative analysis of regulatory-HTA parallel scientific advice. Br J Pharmacol. 2016;82:965-973.

6 Lynn MR. Determination and quantification of content validity. Nurs Res. 1986;35(6):382-385.

Abstract accepted as a poster presentation at the 79th FIP World Congress of Pharmacy and Pharmaceutical Sciences held in Abu Dhabi, United Arab Emirates, 22-26 September 2019:

Outcomes studied in leukaemia clinical trials: A need for harmonisation?

Dylan Said, John Joseph Borg, Maresca Attard Pizzuto, Anthony Serracino-Inglott

Background: Significant inconsistency in outcomes reported in cancer clinical trials (CTs) is documented, which presents as a barrier to evidence synthesis, increasing the complexity of regulatory, policy and healthcare decision-making.

Purpose: To analyse the variability of efficacy outcomes reported in leukaemia CTs.

Method: Interventional Phase II to Phase IV CTs investigating the main leukaemia subtypes acute lymphoblastic leukaemia (ALL), acute myeloid leukaemia (AML), chronic lymphocytic leukaemia (CLL) and chronic myeloid leukaemia (CML) were identified from the EU Clinical Trials Register over an 11-year date range (2007-2017). Therapeutic CTs reporting efficacy data in the English language for investigational medicinal products of chemical, biological and biotechnological origin were included in the study. Unique primary and secondary efficacy endpoints were extracted from the selected CT protocols. The level of heterogeneity in reported outcomes was compared between the leukaemia subtypes.

Results: The register search generated 666 CTs, with 378 meeting inclusion criteria. Thirty-six unique efficacy measures were identified and grouped into the endpoint categories of survival (n=5), response rates and biomarkers (n=16), time-to-event (n=6) and other (n=9). The mean number of outcomes reported per CT for each main subtype was 4 (ALL), 4.7 (AML), 4.9 (CLL) and 4.1 (CML). Efficacy studies conducted in CLL demonstrated the greatest variability with 65% (n=66) and 4% (n=41) of the trials studying 4-10 outcomes and >10 outcomes respectively.

Conclusion: Leukaemia CTs, particularly in CLL, were observed to report outcomes that are highly diverse, demanding future work on the harmonisation of efficacy parameters.

Abstract accepted as a poster presentation at the European Society for Medical Oncology (ESMO) 2019 congress held in Barcelona, Spain, 27 September – 1 October 2019:

Efficacy endpoints studied in clinical trials for early-onset leukaemia

Dylan Said, John Joseph Borg, Maresca Attard Pizzuto, Anthony Serracino-Inglott

Background: Acute lymphoblastic leukaemia (ALL) is associated with high survival rates in paediatrics and young adults and less favourable outcomes in older adults. Clinical trials (CTs) studying survival-based endpoints in ALL lead to significantly prolonged trial durations, resulting in delays to medicines access. Biomarkers in ALL have been correlated with longer-term outcomes, potentially accelerating the clinical development of novel treatments. The aim of this study was to characterise efficacy measures investigated in ALL CTs conducted in the European Union (EU)/Economic Area (EEA).

Methods: Interventional Phase II to Phase IV ALL CTs registered in the EU Clinical Trials Register over an 11-year period (2007-2017) were identified. Therapeutic CTs reporting efficacy data in the English language for investigational medicinal products of chemical, biological and biotechnological origin were included in the study. A protocol design or age filter was not applied to avoid limiting the scope of outcomes identified. Primary and secondary efficacy endpoints were extracted from the selected CTs and categorised according to type of measurement. A data mining process was performed to detect trends in outcomes studied.

Results: The data set comprised 68 CTs representing about 20,000 patients. The majority of trials (69%, n=47) recruited patients from the adult population (18-64 years) and conducted only Phase II studies (62%, n=42). Twenty-three unique efficacy endpoints were identified and stratified into the clusters of survival (n=4), time-to-event (n=3), response rates and biomarkers (n=11) and other (n=5). Fifty-three per cent (n=36) of the trials reported 4-10 outcomes, with a mean of 4 outcomes per CT (range 1-10). The principal endpoints evaluated in CTs consisted of overall survival (CTs: 63%, n=43), minimal residual disease (CTs: 50%, n=34), event-free survival (CTs: 40%, n=27) and disease-free survival (CTs: 40%, n=27).

Conclusions: The high uptake of minimal residual disease as an efficacy parameter in ALL CTs is in line with reported findings confirming the prognostic value of this marker on clinical outcomes. Heterogeneity in the selection of efficacy endpoints was observed which warrants future work on the standardisation of efficacy outcomes in ALL CTs.

Abstract accepted at The Professional Society for Health Economics and Outcomes Research (ISPOR) Europe 2019 conference held in Copenhagen, Denmark, 2-6 November 2019:

Trends in efficacy parameters studied in clinical trials for acute myeloid leukaemia

Dylan Said, John Joseph Borg, Maresca Attard Pizzuto, Anthony Serracino-Inglott

Objectives: Biomarker-based endpoints have been validated as surrogate measures to predict overall survival (OS) in acute myeloid leukaemia (AML), potentially shortening clinical trial (CT) durations and expediting patient access. The aim was to characterise efficacy endpoints over time for CTs in AML conducted in the European Union (EU)/Economic Area (EEA).

Methods: Interventional Phase II to Phase IV AML CTs registered in the EU Clinical Trials Register over an 11-year period (2007-2017) were identified. Therapeutic CTs reporting efficacy data in the English language for investigational medicinal products of chemical, biological and biotechnological origin were included in the study. Primary and secondary efficacy endpoints were extracted from the selected CTs and categorised according to type of measurement. Descriptive and inferential statistics were operated to detect trends in reported outcomes.

Results: The data set comprised 161 CTs representing about 47,000 patients, with the majority of trials (67%, n=108) recruiting only patients from the adult (18-64 years) and elderly (>=65 years) populations. Twenty-nine unique efficacy measures were identified and stratified into the endpoint clusters of survival (n=4), time-to-event (n=5), response rates and biomarkers (n=13) and other (n=7) parameters. Sixty-eight per cent (n=110) of the trials reported 4-10 outcomes, with a mean of 5 per CT. The principal outcomes examined included OS (CTs: 78%, n=126), complete response rate (CR) (CTs: 52%, n=84) and event-free survival (EFS) (CTs: 44%, n=71). The endpoints EFS, OS and complete response rate with incomplete blood recovery (CRi) registered the highest frequency change of selection pre- and post-2012 (EFS: 30%, $p<0.001$; OS: 27%, $p<0.001$; CRi: 18%, $p=0.01$).

Conclusions: The increase in uptake of EFS and CRi as efficacy parameters in AML CTs is in line with previous findings correlating these measures to OS. Heterogeneity in the selection of endpoints was observed which warrants future work on the development of a core set of outcomes.