Evidence Synthesis:

Navigating an Evolving Landscape

NIHR | National Institute for Health and Care Research

NIHR Complex Reviews Support Unit

University of Glasgow

UNIVERSITY OF LEICESTER

Two example projects

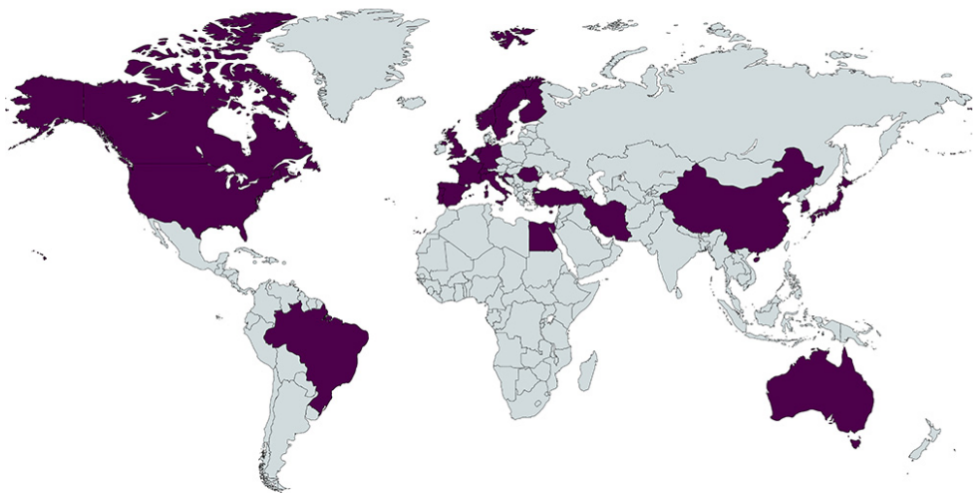
- REcovery and rehabilitation of PeopLE with aphasia after StrokE
  - Marian Brady et al. (Glasgow Caledonian University)
  - Study was supported by the National Institute for Health Research Health Services and Delivery Research (14/04/22); The Tavistock Trust for Aphasia, UK
- Argumentation based evidence synthesis
  - Matt Williams (Imperial), Anthony Hunter (UCL)

RELEASE

REhabilitation and recovery of peopLE with Aphasia after StrokE (RELEASE)

Rationale: The James Lind Alliance partnership between stroke survivors, carers and healthcare professionals listed aphasia twice in the 'top 10' research priorities for life after stroke. A better understanding of what makes aphasia rehabilitation work, would allow treatments to be tailored to specific individuals resulting in more effective and efficient therapy.

Research Activities: We gathered pre-existing data from clinical trials and studies of aphasia treatments after stroke. We pooled these data in a large database and used them to answer new research questions about aphasia. We brought separate databases together to allow us to generate new information about aphasia after stroke and identify future research questions. This informed our understanding of what kind of patients we should be approaching to participate in our study, and when.



Was intended to inform

- The components of aphasia therapy that best inform recovery
- The optimum therapy (timing, intensity, frequency, duration, repetition) and home practice routine
- The usual patterns of recovery (with and without therapy)
- What aspects indicate someone will make a good (or not so good) recovery from aphasia

CRSU Support

- Review and advice during development of analysis plan
  - Including both technical and strategic input
  - Need to be mindful of available analytic resource
- Review and comment on analytic results

APHASIOLOGY
2022, VOL. 36, NO. 4, 513–533
https://doi.org/10.1080/02687038.2021.1897081

Routledge
Taylor & Francis Group

Utilising a systematic review-based approach to create a database of individual participant data for meta- and network meta-analyses: the RELEASE database of aphasia after stroke

Louise R Williams[a], Myzoon Ali[a], Kathryn VandenBerg[a], Linda J Williams[b], Masahiro Abo[c], Frank Becker[d], Audrey Bowen[e], Caitlin Brandenburg[f], Caterina Breitenstein[g], Stefanie Bruehl[h], David A Copland[f], Tamara B Cranfill[i], Marie Di Pietro-Bachmann[j], Pamela Enderby[k], Joanne Fillingham[l], Federica Lucia Galli[m], Marialuisa Gandolfi[n], Bertrand Glize[o,p], Erin Godecke[q], Neil Hawkins[r], Katerina Hilari[s], Jacqueline Hinckley[t], Simon Horton[u], David Howard[v], Petra Jaecks[w], Elizabeth Jefferies[x], Luis MT Jesus[y], Maria Kambanaros[z], Eun Kyoung Kang[aa], Eman M Khedr[bb], Anthony Pak-Hin Kong[cc], Tarja Kukkonen[dd], Marina Laganaro[ee], Matthew A Lambon Ralph[ff], Ann Charlotte Laska[gg], Béatrice Leemann[hh], Alexander P Leff[ii], Roxele Ribeiro Lima[jj], Antje Lorenz[kk], Brian MacWhinney[ll], Rebecca Shisler Marshall[mm], Flavia Mattioli[nn], İlknur Maviş[oo], Marcus Meinzer[pp], Reza Nilipour[qq], Enrique Noé[rr], Nam-Jong Paik[ss], Rebecca Palmer[k], Ilias Papathanasiou[tt], Brigida F Patricio[uu], Isabel Pavão Martins[vv], Cathy Price[ww], Tatjana Prizl Jakovac[xx], Elizabeth Rochon[yy], Miranda L Rose[zz], Charlotte Rosso[aaa], Ilona Rubi-Fessen[bbb], Marina B Ruiter[ccc], Claerwen Snell[ddd], Benjamin Stahl[eee], Jerzy P Szaflarski[fff], Shirley A Thomas[ggg], Mieke Van De Sandt-Koenderman[hhh], Ineke Van Der Meulen[hhh], Evy Visch-Brink[iii], Linda Worrall[f], Heather Harris Wright[jjj], Marian C Brady[a] and The RELEASE Collaborators

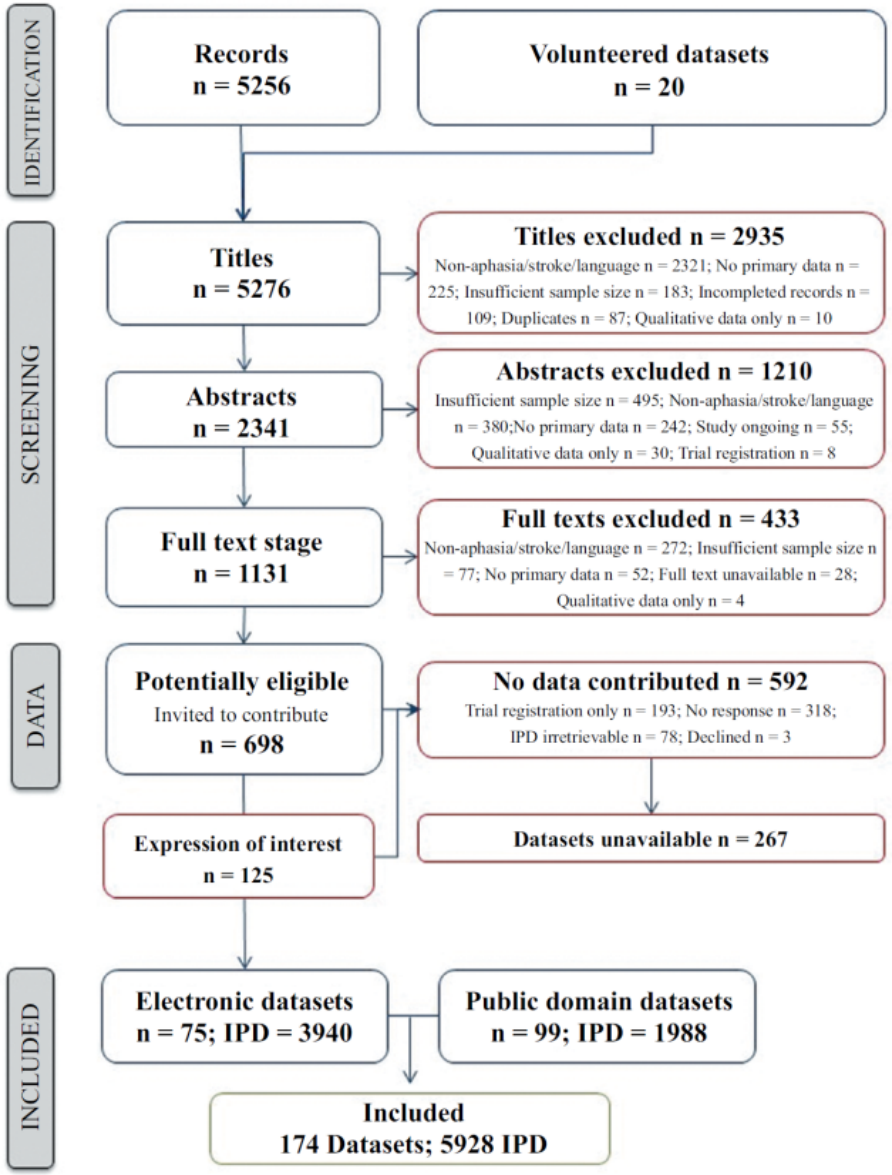


Figure 1. PRISMA flow diagram: Eligible database identification and contribution.

Interventions

Table 8. Speech and language therapy (SLT) Interventions in the RELEASE dataset.

Speech and language therapy intervention descriptor		Datasets N = 67 (%)	IPD n = 2330 (%)
SLT method of delivery	Face-to-face	60 (89.6)	1957 (84.0)
	Computer	15 (22.4)	315 (13.5)
	Telephone	1 (1.5)	15 (0.6)
	Constraint induced aphasia therapy	7 (10.4)	113 (4.8)
	Self-managed	4 (6.0)	106 (4.5)
	One-to-one	47 (70.1)	1613 (69.2)
	Group	8 (11.9)	148 (6.4)
	Mixed	9 (13.4)	207 (8.9)
Theoretical approach	Semantic	2 (1.5)	34 (1.5)
	Phonological	9 (13.4)	124 (5.3)
	Semantic and phonological	15 (22.4)	260 (11.2)
	Functional and pragmatic	8 (11.9)	246 (10.6)
	Constraint induced aphasia therapy	7 (10.4)	113 (4.8)
	Melodic intonation therapy	4 (6.0)	61 (2.6)
	Conversational partner training	2	55 (2.4)
Target of Impairment	Spoken language	41	734 (31.5)
	Auditory comprehension	4	68 (2.9)
	Auditory comprehension & spoken language	24	651 (27.9)
	Reading	1	10 (0.4)
	Writing	0	0

Note: Categories were not mutually exclusive; an intervention may span categories or appear more than once.

Outcomes

Table 2. Data availability for language outcomes.

Language outcome	*Datasets N = 174 (%)*	*IPD n = 5928 (%)*
Overall language ability	80 (46.0)	2699 (45.5)
Naming	75 (43)	2886 (48.7)
Other spoken language	9 (5.2)	380 (6.4)
Auditory comprehension	71 (40.8)	2750 (46.4)
Reading comprehension	12 (6.9)	770 (13.0)
Writing	13 (7.5)	724 (12.2)
Function communication – observer rated	29 (16.7)	1591 (26.8)
Functional communication – self rated	3 (1.7)	68 (1.1)

Key-IPD Individual Participant Data; % percentage; N = total datasets; n = total IPD

Table 3. Overall language ability assessment tools (at baseline) included in RELEASE – Datasets and IPD where measure is reported, available, missing from report, or unavailable.

	Datasets	
Overall Language Ability Assessment	Reported (IPD available, missing)	Assessed but unavailable (IPD)
Aachen Aphasia Test (AAT) overall Severity Score	1 (12,0)	15 (537)
Afazi Dil Değerlendirme Testi (ADD)	1 (30,23)	0 (0)
Aphasia Handicap Scale (AHS)	2 (39,19)	0 (0)
Aphasia Severity Rating Scale (ASRS)	14 (441,6)	0 (0)
Boston Assessment of Severe Aphasia (BASA)	1 (15,0)	0 (0)
Comprehensive Aphasia Test (CAT)	3 (433,37)	5 (180)
Norsk Grunntest for Afaxi (NGA)	3 (62,0)	0 (0)
Porch Index of Communicative Ability (PICA)	8 (171,1)	0 (0)
Short Norsk Grunntest for Afasi (Short NGA)	2 (241,0)	0 (0)
Sprachsystemtisches Aphasie Screening (SAPS)	1 (133,9)	0 (0)
Standard Language Test of Aphasia (SLTA)	2 (24,0)	1 (36)
Western Aphasia Battery - Aphasia Quotient*	35 (733,0)	0 (0)
Western Aphasia Battery-Revised Aphasia Quotient	6 (69,0)	1 (18)
Western Aphasia Battery-Cantonese	1 (105,0)	0 (0)
Western Aphasia Battery-Japanese	1 (24,0)	0 (0)
Western Aphasia Battery- Korean	3 (125,3)	0 (0)
Western Aphasia Battery-Persian	2 (86,0)	0 (0)

Key *Anchor Measure; IPD Individual Participant Data.

CLINICAL AND POPULATION SCIENCES

Dosage, Intensity, and Frequency of Language Therapy for Aphasia: A Systematic Review–Based, Individual Participant Data Network Meta-Analysis

The REhabilitation and recovery of peopLE with Aphasia after StrokE (RELEASE) Collaborators*

BACKGROUND AND PURPOSE: Optimizing speech and language therapy (SLT) regimens for maximal aphasia recovery is a clinical research priority. We examined associations between SLT intensity (hours/week), dosage (total hours), frequency (days/week), duration (weeks), delivery (face to face, computer supported, individual tailoring, and home practice), content, and language outcomes for people with aphasia.

METHODS: Databases including MEDLINE and Embase were searched (inception to September 2015). Published, unpublished, and emerging trials including SLT and ≥10 individual participant data on aphasia, language outcomes, and time post-onset were selected. Patient-level data on stroke, language, SLT, and trial risk of bias were independently extracted. Outcome measurement scores were standardized. A statistical inferencing, one-stage, random effects, network meta-analysis approach filtered individual participant data into an optimal model examining SLT regimen for overall language, auditory comprehension, naming, and functional communication pre-post intervention gains, adjusting for a priori–defined covariates (age, sex, time poststroke, and baseline aphasia severity), reporting estimates of mean change scores (95% CI).

RESULTS: Data from 959 individual participant data (25 trials) were included. Greatest gains in overall language and comprehension were associated with >20 to 50 hours SLT dosage (18.37 [10.58–26.16] Western Aphasia Battery–Aphasia Quotient; 5.23 [1.51–8.95] Aachen Aphasia Test–Token Test). Greatest clinical overall language, functional communication, and comprehension gains were associated with 2 to 4 and 9+ SLT hours/week. Greatest clinical gains were associated with frequent SLT for overall language, functional communication (3–5+ days/week), and comprehension (4–5 days/week). Evidence of comprehension gains was absent for SLT ≤20 hours, <3 hours/week, and ≤3 days/week. Mixed receptive-expressive therapy, functionally tailored, with prescribed home practice was associated with the greatest overall gains. Relative variance was <30%. Risk of trial bias was low to moderate; low for meta-biases.

CONCLUSIONS: Greatest language recovery was associated with frequent, functionally tailored, receptive-expressive SLT, with prescribed home practice at a greater intensity and duration than reports of usual clinical services internationally. These exploratory findings suggest critical therapeutic ranges, informing hypothesis-testing trials and tailoring of clinical services.

REGISTRATION: URL: https://www.crd.york.ac.uk/PROSPERO/; Unique identifier: CRD42018110947.

GRAPHIC ABSTRACT: A graphic abstract is available for this article.

Key Words: aphasia ■ big data ■ comprehension ■ language therapy ■ meta-analysis ■ stroke

Stroke. 2022;53:956–967

Dosage (total therapy hours)

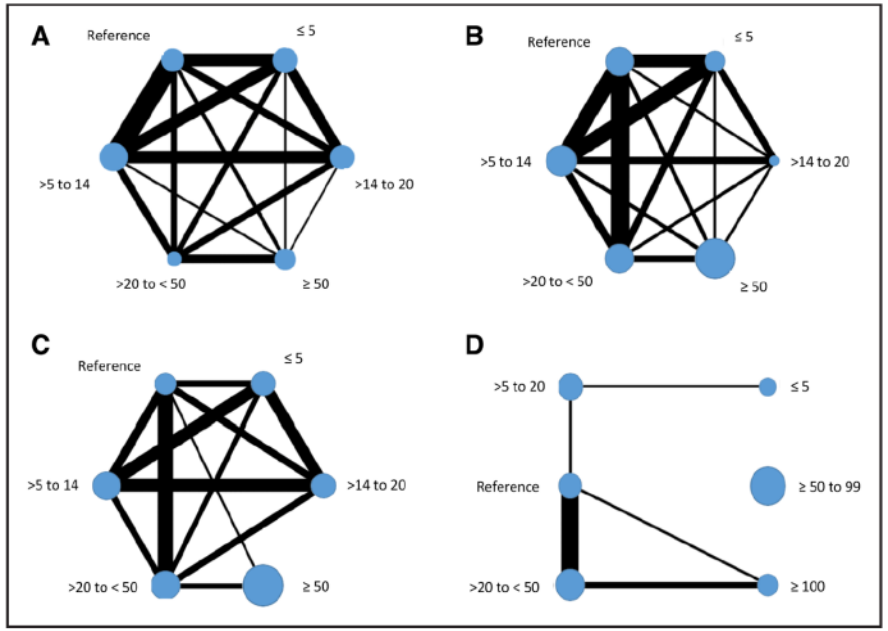


Figure 2. Dosage (total speech and language therapy hours) by language outcome.
Overall language ability (**A**), functional communication (**B**), auditory comprehension (**C**), and naming (**D**).

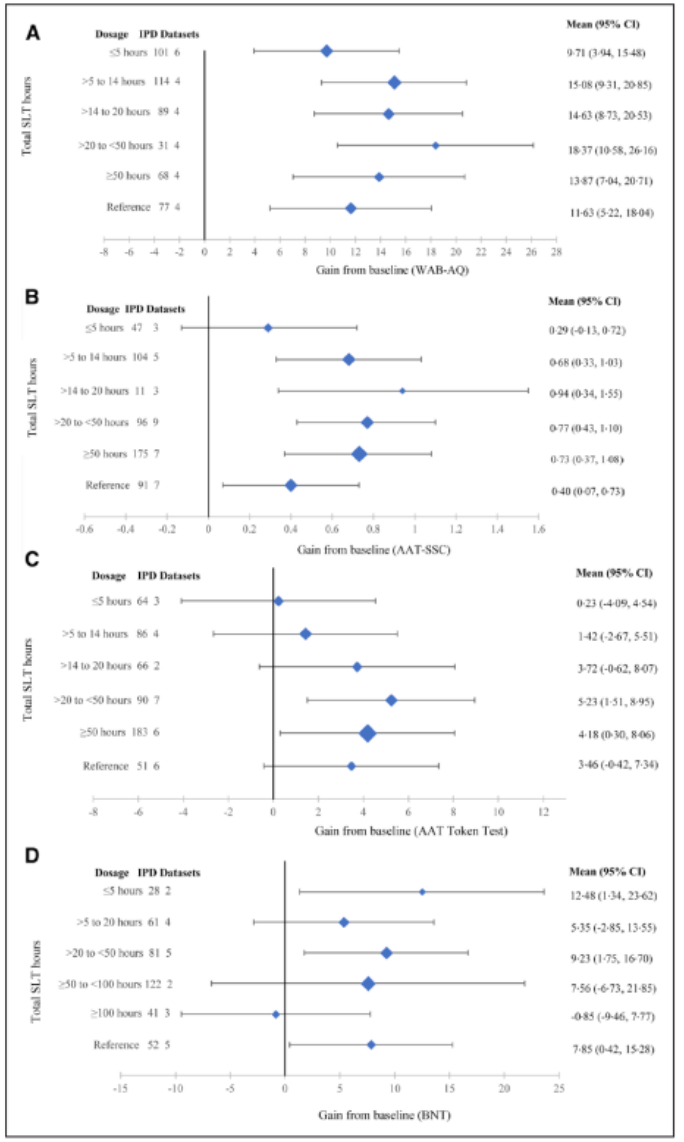


Figure 4. Dosage (total speech and language therapy [SLT] hours) and associated gains from baseline (mean; 95% CI).
Overall language (**A**): Western Aphasia Battery–Aphasia Quotient (0–100); 480 individual participant data (IPD; 11 randomized controlled trials [RCTs]); functional communication (**B**): Aachen Aphasia Test–Spontaneous Speech Communication (AAT-SSC; 0–5); 524 IPD (14 RCTs); auditory comprehension (**C**): Aachen Aphasia Test (AAT) Token Test (0–50); 540 IPD (16 RCTs); naming (**D**): Boston Naming Test (BNT; 0–60); 385 IPD (13 RCTs).

CLINICAL AND POPULATION SCIENCES

Predictors of Poststroke Aphasia Recovery

A Systematic Review-Informed Individual Participant Data Meta-Analysis

The REhabilitation and recovery of peopLE with Aphasia after StrokE (RELEASE) Collaborators*

BACKGROUND AND PURPOSE: The factors associated with recovery of language domains after stroke remain uncertain. We described recovery of overall-language-ability, auditory comprehension, naming, and functional-communication across participants' age, sex, and aphasia chronicity in a large, multilingual, international aphasia dataset.

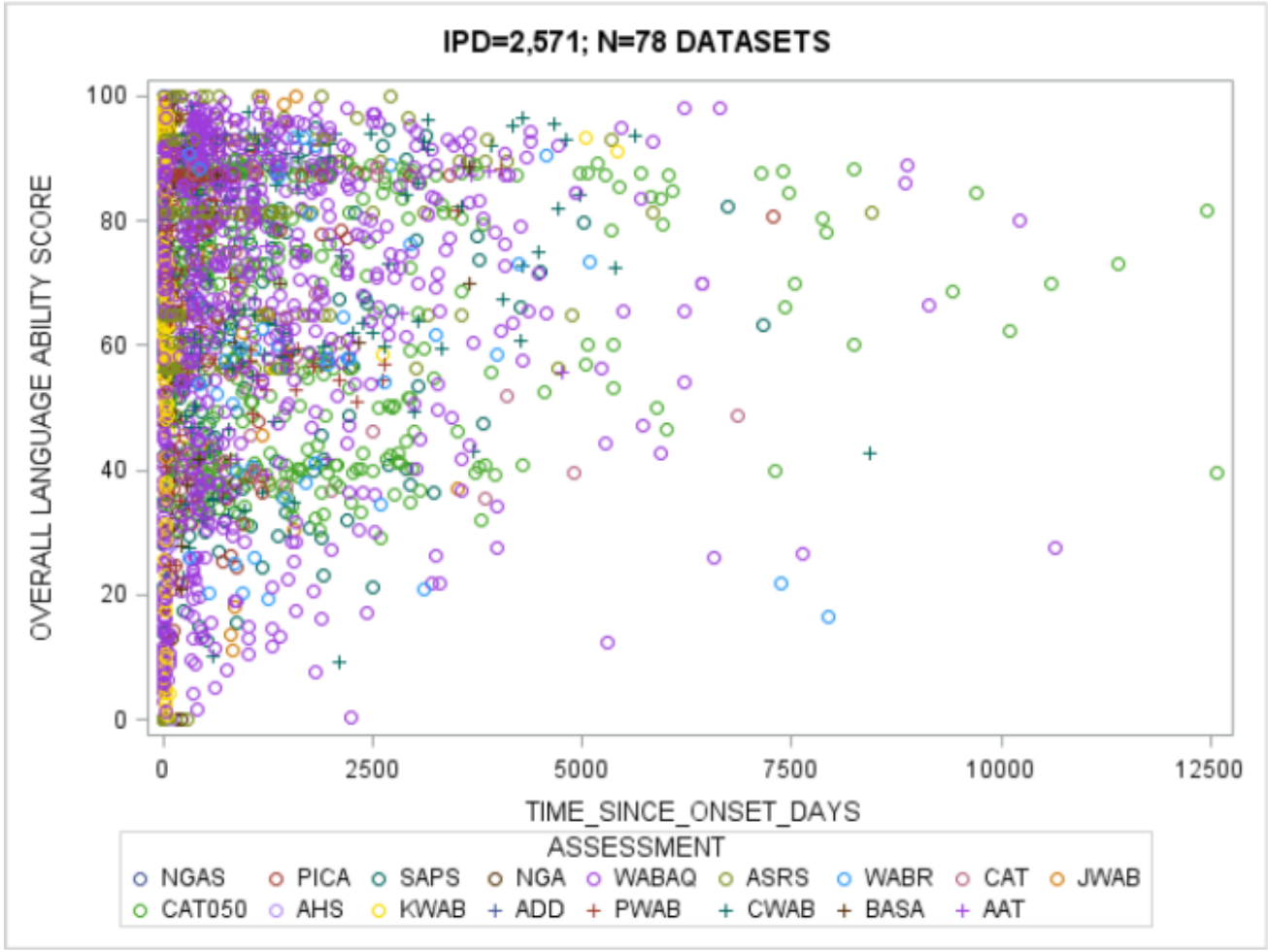
METHODS: Individual participant data meta-analysis of systematically sourced aphasia datasets described overall-language ability using the Western Aphasia Battery Aphasia-Quotient; auditory comprehension by Aachen Aphasia Test (AAT) Token Test; naming by Boston Naming Test and functional-communication by AAT Spontaneous-Speech Communication subscale. Multivariable analyses regressed absolute score-changes from baseline across language domains onto covariates identified a priori in randomized controlled trials and all study types. Change-from-baseline scores were presented as estimates of means and 95% CIs. Heterogeneity was described using relative variance. Risk of bias was considered at dataset and meta-analysis level.

RESULTS: Assessments at baseline (median=43.6 weeks poststroke; interquartile range [4–165.1]) and first-follow-up (median=10 weeks from baseline; interquartile range [3–26]) were available for n=943 on overall-language ability, n=1056 on auditory comprehension, n=791 on naming and n=974 on functional-communication. Younger age (<55 years, +15.4 Western Aphasia Battery Aphasia-Quotient points [CI, 10.0–20.9], +6.1 correct on AAT Token Test [CI, 3.2–8.9]; +9.3 Boston Naming Test points [CI, 4.7–13.9]; +0.8 AAT Spontaneous-Speech Communication subscale points [CI, 0.5–1.0]) and enrollment <1 month post-onset (+19.1 Western Aphasia Battery Aphasia-Quotient points [CI, 13.9–24.4]; +5.3 correct on AAT Token Test [CI, 1.7–8.8]; +11.1 Boston Naming Test points [CI, 5.7–16.5]; and +1.1 AAT Spontaneous-Speech Communication subscale point [CI, 0.7–1.4]) conferred the greatest absolute change-from-baseline across each language domain. Improvements in language scores from baseline diminished with increasing age and aphasia chronicity. Data exhibited no significant statistical heterogeneity. Risk-of-bias was low to moderate-low.

CONCLUSIONS: Earlier intervention for poststroke aphasia was crucial to maximize language recovery across a range of language domains, although recovery continued to be observed to a lesser extent beyond 6 months poststroke.

Stroke. 2021;52:1778–1787

Time from stroke onset to intervention



Conclusions

- Greatest improvement for enrollment within 1-month poststroke across all language domains.
- Improvements in mean absolute scores from baseline diminished with increasing time since stroke
- Yet still exceeded established group-level benchmarks of significant change for overall-language-ability

Argumentation based synthesis

Aggregation of Clinical Evidence using Argumentation: A Tutorial Introduction

Anthony Hunter* and Matthew Williams†

March 13, 2014

Abstract

In this tutorial, we describe a new framework for representing and synthesizing know
clinical trials involving multiple outcome indicators. The framework offers a formal a
aggregating clinical evidence. Based on the available evidence, arguments are generatec
ing that one treatment is superior, or equivalent, to another. Evidence comes from r
clinical trials, systematic reviews, meta-analyses, network analyses, etc. Preference c
arguments are used that are based on the outcome indicators, and the magnitude of thos
indicators, in the evidence. Meta-arguments attack (i.e they are counterarguments to)
that are based on weaker evidence. An evaluation criterion is used to determine wh
winning arguments, and thereby the recommendations for which treatments are superic
proach has an advantage over meta analyses and network analyses in that they aggregate evidence according to a single outcome indicator, whereas our approach combines evidence according to multiple outcome indicators.

Artificial Intelligence in Medicine 56 (2012) 173–190

Contents lists available at SciVerse ScienceDirect

Artificial Intelligence in Medicine

journal homepage: www.elsevier.com/locate/aiim

Aggregating evidence about the positive and negative effects of treatments

Anthony Hunter[a,*], Matthew Williams[b]

[a] Department of Computer Science, University College London, London WC1E 6BT, UK
[b] Department of Clinical Oncology, University College Hospital, London NW1 2PG, UK

ARTICLE INFO

Article history:
Received 12 September 2011
Received in revised form
20 September 2012
Accepted 22 September 2012

Keywords:
Computational models of argument
Argument systems
Knowledge aggregation
Evidence aggregation
Evidence-based medicine
Clinical recommendations

ABSTRACT

Objectives: Evidence-based decision making is becoming increasingly important in healthcare. Much valuable evidence is in the form of the results from clinical trials that compare the relative merits of treatments. In this paper, we present a new framework for representing and synthesizing knowledge from clinical trials involving multiple outcome indicators.
Method: The framework generates and evaluates arguments for claiming that one treatment is superior, or equivalent, to another based on the available evidence. Evidence comes from randomized clinical trials, systematic reviews, meta-analyses, network analyses, etc. Preference criteria over arguments are used that are based on the outcome indicators, and the magnitude of those outcome indicators, in the evidence. Meta-arguments attacks arguments that are based on weaker evidence.
Results: We evaluated the framework with respect to the aggregation of evidence undertaken in three published clinical guidelines that involve 56 items of evidence and 16 treatments. For each of the three guidelines, the treatment we identified as being superior using our method is a recommended treatment in the corresponding guideline.
Conclusions: The framework offers a formal approach to aggregating clinical evidence, taking into account subjective criteria such as preferences over outcome indicators. In the evaluation, the aggregations obtained showed a good correspondence with published clinical guidelines. Furthermore, preliminary computational studies indicate that the approach is viable for the size of evidence tables normally encountered in practice.

Individual Arguments for glaucoma treatment

ID	Left	Right	Outcome indicator	Outcome value	Net outcome	Sig	Type
e_{01}	BB	NT	visual field prog	0.77	superior	no	MA
e_{02}	BB	NT	change in IOP	-2.88	superior	yes	MA
e_{03}	BB	NT	respiratory prob	3.06	inferior	no	MA
e_{04}	BB	NT	cardio prob	9.17	inferior	no	MA
e_{05}	PG	BB	change in IOP	-1.32	superior	yes	MA
e_{06}	PG	BB	acceptable IOP	1.54	superior	yes	MA
e_{07}	PG	BB	respiratory prob	0.59	superior	yes	MA
e_{08}	PG	BB	cardio prob	0.87	superior	no	MA
e_{09}	PG	BB	allergy prob	1.25	inferior	no	MA
e_{10}	PG	BB	hyperaemia	3.59	inferior	yes	MA
e_{11}	PG	SY	change in IOP	-2.21	superior	yes	MA
e_{12}	PG	SY	allergic prob	0.03	superior	yes	MA
e_{13}	PG	SY	hyperaemia	1.01	inferior	no	MA
e_{14}	CA	NT	convert to COAG	0.77	superior	no	MA
e_{15}	CA	NT	visual field prog	0.69	superior	no	MA
e_{16}	CA	NT	IOP $>$ 35mmHg	0.08	superior	yes	MA
e_{17}	CA	BB	hyperaemia	6.42	inferior	no	MA
e_{18}	SY	BB	visual field prog	0.92	superior	no	MA
e_{19}	SY	BB	change in IOP	-0.25	superior	no	MA
e_{20}	SY	BB	allergic prob	41.00	inferior	yes	MA
e_{21}	SY	BB	drowsiness	1.21	inferior	no	MA

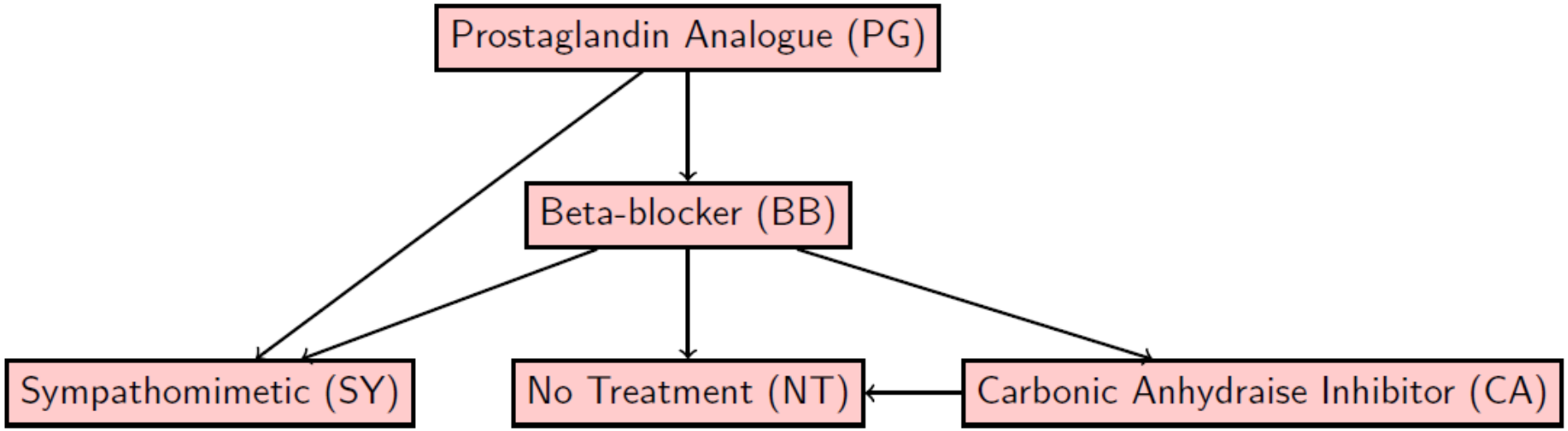


Figure 1: Example of a superiority graph. This concerns treatments for glaucoma and it has been generated by our approach using the evidence table given in Table 1. There is an arc for each pair of treatments that we compared in one or more trials. If a pair of treatments were not compared in any trial, then there is no arc between them. When there is an arrow from treatment τ_1 to τ_2, then it means that our study found τ_1 to be superior to τ_2.

CRSU Support

- Development of case study and interactive web based app (ongoing)
- A prototype interactive tool to allow patients to explore available data and make decisions based on their individual preferences

Typical NMA

Cancer Management and Research

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

First-line treatment strategies for newly diagnosed chronic myeloid leukemia: a network meta-analysis

This article was published in the following Dove Press journal:
Cancer Management and Research

Kang-Kang Chen[1]
Tai-Feng Du[1]
Ku-Sheng Wu[2]
Wei Yang[3]

[1]Department of Preventive Medicine and MPH Education Center, Shantou University Medical College, Shantou, Guangdong Province, China; [2]Department of Preventive Medicine, Shantou University Medical College, Shantou, Guangdong Province, China; [3]Department of Thoracic Surgery, Administrative Office, Shantou University Medical College Cancer Hospital, Shantou, Guangdong Province, China

Objectives: With bosutinib proven to be available for frontline treatment, there are currently four frontline treatments as well as an additional strategy with high-dose imatinib for newly diagnosed chronic myeloid leukemia (CML). Due to the lack of direct comparison of high-dose imatinib, dasatinib, nilotinib, and bosutinib, we summarized the evidence to indirectly compare the efficacy among these treatment options.

Methods: In total, 14 randomized clinical trials including 5,630 patients were analyzed by direct and mixed-treatment comparisons. Outcomes assessed were the following: complete cytogenetic response at 12 months; major molecular response at 12, 24, and 36 months; deep molecular response at 12, 24, 36, and 60 months; early molecular response at 3 months; progression-free survival (PFS); overall survival (OS); and Grade 3 or 4 adverse events (AEs).

Results: The Bayesian network meta-analysis demonstrated that high-dose imatinib was less effective than all new-generation tyrosine kinase inhibitors and had a higher probability of Grade 3 or 4 AEs. For molecular response, 300 mg of nilotinib was likely to be the preferred frontline treatment, as demonstrated by higher response rates and faster, deeper, and longer molecular response. For PFS and OS, there were high likelihoods (79% and 74%, respectively) that 400 mg of nilotinib was the preferred option. For AEs, standard-dose imatinib has the highest probability (65%) of being the most favorable toxicity profile.

Conclusion: Considering the efficacy and toxicity profile, it is not recommended to use high-dose imatinib for treatment. This analysis also showed that nilotinib has the highest probability to become the preferred frontline agents for treating CML.

Keywords: CML, tyrosine kinase inhibitor, imatinib, bosutinib, dasatinib, nilotinib

Typical NMA Results

Response Endpoint

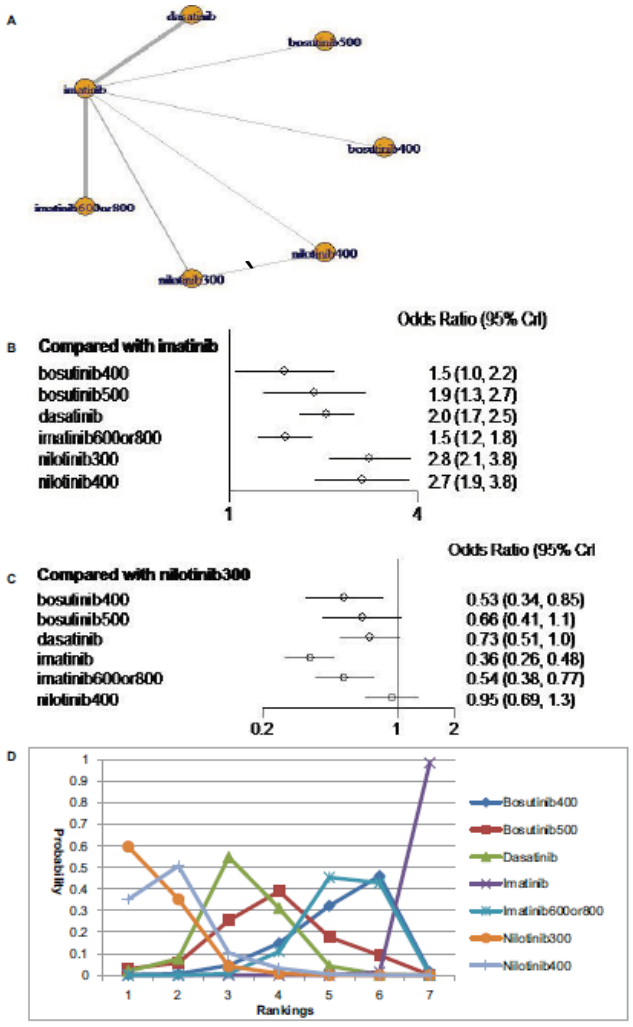


Figure 3 Analysis of major molecular response at 12 months: (A) network diagram; (B) forest plot, with imatinib as the comparator; (C) forest plot, with nilotinib 300 mg as the comparator; and (D) SUCRA plot.
Notes: Imatinib = standard-dose imatinib; bosutinib400 = bosutinib 400 mg daily; bosutinib500 = bosutinib 500 mg daily; nilotinib300 = nilotinib 300 mg daily; nilotinib400 = nilotinib 400 mg daily; imatinib600_800 = high-dose imatinib.
Abbreviations: CrI, credible interval; SUCRA, surface under the cumulative ranking.

Grade 3 or 4 Adverse Event Endpoint

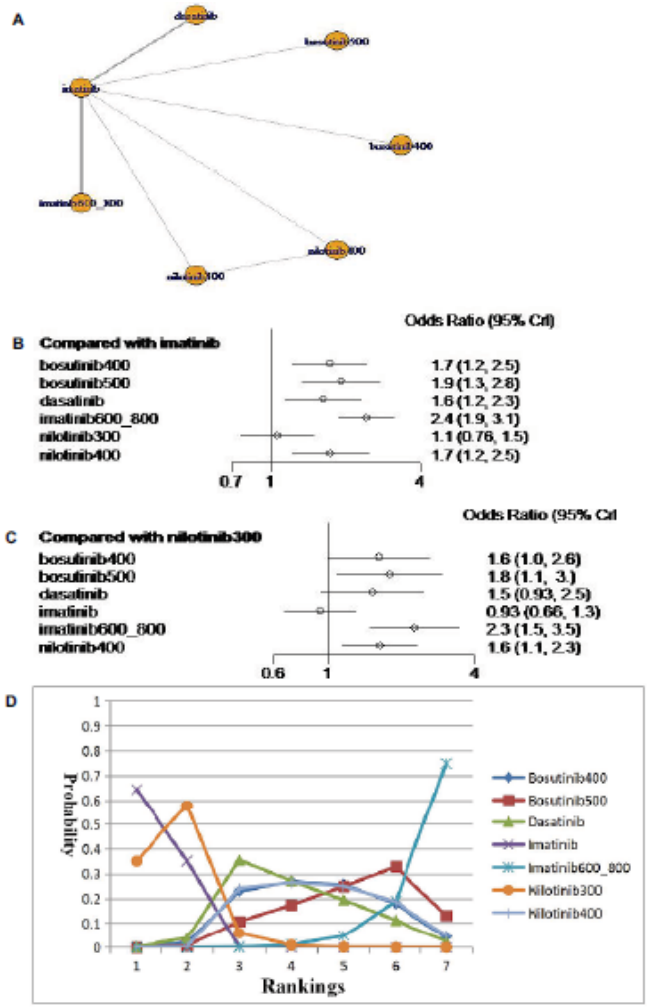


Figure 5 Analysis of Grade 3 or 4 AEs: (A) network diagram; (B) forest plot, with imatinib as the comparator; (C) forest plot, with nilotinib 300 mg as the comparator; and (D) SUCRA plot.
Notes: Imatinib = standard-dose imatinib; bosutinib400 = bosutinib 400 mg daily; bosutinib500 = bosutinib 500 mg daily; nilotinib300 = nilotinib 300 mg daily; nilotinib400 = nilotinib 400 mg daily; imatinib600_800 = high-dose imatinib.
Abbreviations: CrI, credible interval; SUCRA, surface under the cumulative ranking.

Arguments based on individual studies

Argumentation Based Synthesis

Take a screenshot

Make arguments based on:

- (•) Individual Studies
- ( ) Pairwise meta-analyses
- ( ) Indirect treatment comparisons

Choose Treatments

- [x] Bosutinib 400
- [x] Dasatinib 100
- [x] Nilotinib 300
- [x] Imatinib 400

Choose Endpoints

- [ ] Diarrhea
- [ ] Fatigue
- [x] MMR (12 wks)
- [ ] Nausea
- [ ] Neutropenia
- [ ] Rash
- [ ] Thrombocytopenia

Statistically Significant Results Only (p<0.05)

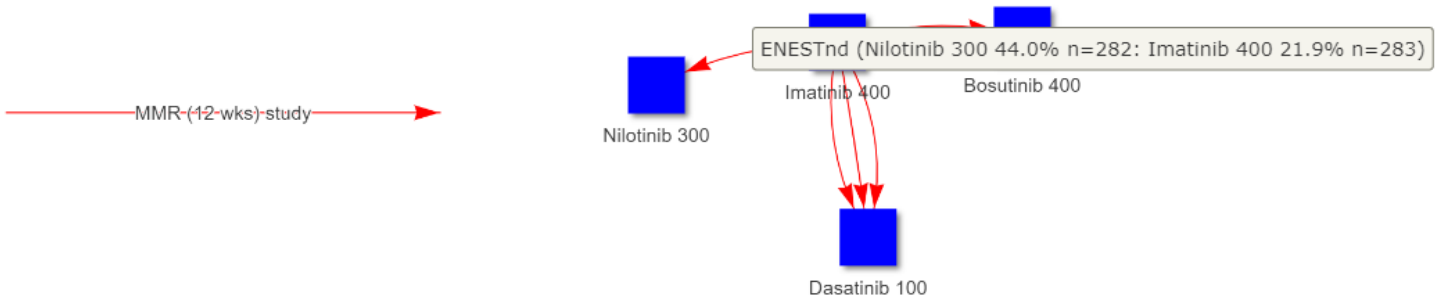
- ( ) Yes
- (•) No

Summarise Arguments

- ( ) Yes
- (•) No

Arrows point to better treatment

Argument Graph | Argument Details



Arguments based on individual studies

Argumentation Based Synthesis

Take a screenshot

Make arguments based on:

- (•) Individual Studies
- ( ) Pairwise meta-analyses
- ( ) Indirect treatment comparisons

Choose Treatments

- [x] Bosutinib 400
- [x] Dasatinib 100
- [x] Nilotinib 300
- [x] Imatinib 400

Choose Endpoints

- [ ] Diarrhea
- [ ] Fatigue
- [x] MMR (12 wks)
- [ ] Nausea
- [ ] Neutropenia
- [ ] Rash
- [ ] Thrombocytopenia

Statistically Significant Results Only (p<0.05)

- ( ) Yes
- (•) No

Summarise Arguments

- ( ) Yes
- (•) No

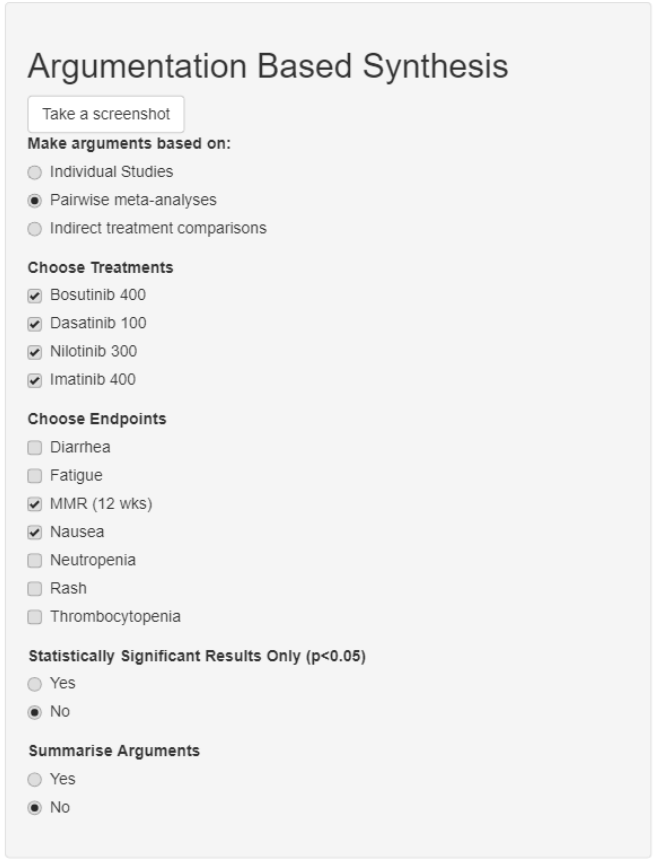
Arrows point to better treatment

Argument Graph | Argument Details

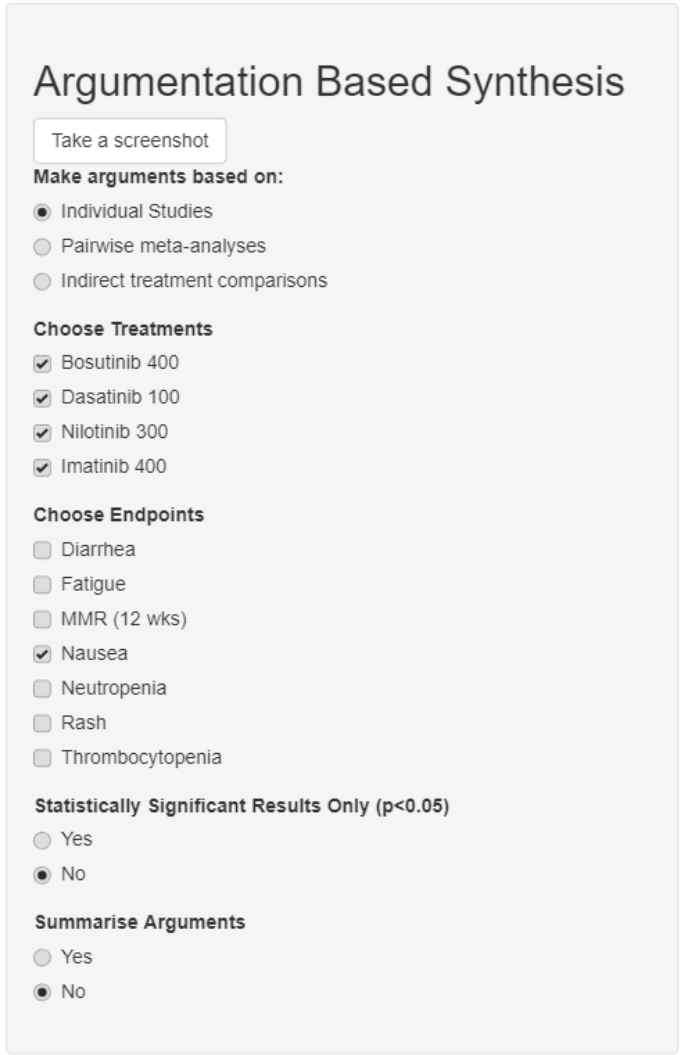
	Endpoint	Treatment	Compared With	Study	Effect (95% CI)	Scale	Treatment is	P<0.05
1	MMR (12 wks)	Bosutinib 400	Imatinib 400	BFORE (Bosutinib 400 47.2% n=246: Imatinib 400 36.9% n=241)	0.1(0.02 to 0.19)	Risk Difference	superior	sig
2	MMR (12 wks)	Dasatinib 100	Imatinib 400	DASISION (Dasatinib 100 45.9% n=259: Imatinib 400 28.1% n=260)	0.18(0.1 to 0.26)	Risk Difference	superior	sig
3	MMR (12 wks)	Dasatinib 100	Imatinib 400	NordCML006 (Dasatinib 100 81.8% n=22: Imatinib 400 58.3% n=24)	0.23(-0.02 to 0.49)	Risk Difference	superior	not sig
4	MMR (12 wks)	Dasatinib 100	Imatinib 400	S0325 (Dasatinib 100 58.6% n=99: Imatinib 400 41.2% n=136)	0.17(0.05 to 0.3)	Risk Difference	superior	sig
5	MMR (12 wks)	Nilotinib 300	Imatinib 400	ENESTnd (Nilotinib 300 44.0% n=282: Imatinib 400 21.9% n=283)	0.22(0.15 to 0.3)	Risk Difference	superior	sig

Arguments based on network meta-analysis

Nausea, effectiveness, and ‘treatment burden’ arguments conflict



Nausea Data (severe nausea is rare)



Arrows point to better treatment

Argument Graph | Argument Details

	Endpoint	Treatment	Compared With	Study	Effect (95% CI)	Scale	Treatment is	P<0.05
1	Nausea	Bosutinib 400	Imatinib 400	BFORE (Bosutinib 400 0.4% n=270: Imatinib 400 0.4% n=267)	0(-0.01 to 0.01)	Risk Difference	superior	not sig
2	Nausea	Dasatinib 100	Imatinib 400	DASISION (Dasatinib 100 0.4% n=260: Imatinib 400 0.4% n=260)	0(-0.01 to 0.01)	Risk Difference	indeterminate	not sig
3	Nausea	Dasatinib 100	Imatinib 400	S0325 (Dasatinib 100 0.8% n=124: Imatinib 400 1.5% n=195)	-0.01(-0.03 to 0.02)	Risk Difference	superior	not sig
4	Nausea	Nilotinib 300	Imatinib 400	ENESTnd (Nilotinib 300 0.4% n=279: Imatinib 400 0.4% n=282)	0(-0.01 to 0.01)	Risk Difference	inferior	not sig

Personal learning

- It is good to be ambitious
- There is always more to learn
- It is important to be pragmatic