IPECAD Modeling Workshop 2023 Cross Comparison Challenge on Cost-Effectiveness Models in Alzheimer's Disease

Ron Handels[1,2], William L Herring[3], Farzam Kamgar[3], Sandar Aye[2], Ashley Tate[2], Colin Green[2,4], Anders Gustavsson[5], Anders Wimo[2], Bengt Winblad[2], Anders Sköldunger[2], Lars Lau Raket[6], Chelsea Bedrejo Stellick[7], Jakub Hlavka[8,9], Yifan Wei[8], Javier Mar[10], Myriam Soto-Gordoa[11], Inge de Kok[12], Chiara Brück[12], Robert Anderson[13], Peter Pemberton-Ross[14], Michael Urbich[14], Eldon Spackman[15], Bryan Tysinger[8], Raphael Wittenberg[13], Linus Jönsson[2]

[1] Alzheimer Centrum Limburg, Maastricht University, Netherlands; [2] Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Stockholm, Sweden; [3]RTI Health Solutions; [4] Biogen UK; [5] Quantify Research; [6] Lund University; [7] University of Alberta; [8] University of Southern California; [9] Masaryk University; [10] Osakidetza Basque Health Service; [11] Mondragon Unibertsitatea; [12] Erasmus Medical Center; [13] London School of Economics; [14] Biogen; [15] University of Calgary

Aim

We cross-compared AD decision-analytic models for a hypothetical disease-modifying treatment starting in mild cognitive impairment (MCI) due to AD to improve understanding, transparency and credibility of health-economic model results.

Background

Using clinical trial treatment effect outcomes in decision-analytic models presents several key challenges:

1. Models use **different endpoints and scales** compared to clinical trials to inform decision-makers and clinical practice on relevant outcomes.
2. The choice of which treatment **effect domain(s)** to consider (e.g., cognition, function and/or behavior; single or multiple) is often driven by model design and complicated by limited evidence on causality and correlation among domains (carrying the risk of spurious correlations between treatment effects and health-economic outcomes).
3. Trial statistical analysis differ from **health-economic modeling methods** (events, transitions, post-trial analyses).
4. Trials are designed for observing treatment effect rather than estimating the size of the health **benefit in a routine care population**.

Methods

We organized the workshop in alignment with guidelines for multi-model comparisons [den Boon, 2019; Eddy, 2012].

1. Draft benchmark scenario (previous workshop recommendations).
2. Identify models:
   - Search for eligible models and modeling groups (systematic literature review, ad-hoc networking, and open ISPOR call; n=46).
   - Exclude models and modeling groups not able to adhere to the benchmark scenario (n=8), not responding the invitation (n=12), or with insufficient resources to apply the benchmark scenario (n=16).
   - Participating models and modeling groups (n=10).
3. Invite for workshop and for finetuning **benchmark scenario**.
4. Summarize outcomes and discuss model differences during workshop.
5. Disseminate findings after review by all participants.

Benchmark scenario – Standardized patient

- Demographics: person mean age of 70 years, men/women separately
- Diagnosis: amyloid confirmed AD-type MCI
- Setting: memory clinic in United States (US)
- Intervention: disease-modifying treatment in addition to standard care
- Comparator: standard of care
- Treatment specification:
  - Discontinuation: 10% (e.g., due to amyloid-related imaging abnormalities)
  - Treated up to and including mild dementia
  - Waning: 5% per year
  - No treatment costs included (i.e., treatment costs set to $0)
- Background mortality: U.S. life table for 2019
- Time horizon: 25 years
- Discount rate: 3.5% per year

Details on the IPECAD workshop

www.ipecad.org/workshop

Acknowledgment

Contact: ron.handels@maastrichtuniversity.nl

We thank the great effort by the modeling groups listed in table 3 for their participation in the workshop. We acknowledge ADNI for synthetic data generation

Benchmark scenario – Hypothetical trial estimates

Based on synthetic correlated data based on ADNI data (QR-code; ISPOR poster ID 132729; https://github.com/ronhandels/synthetic-correlated-data)

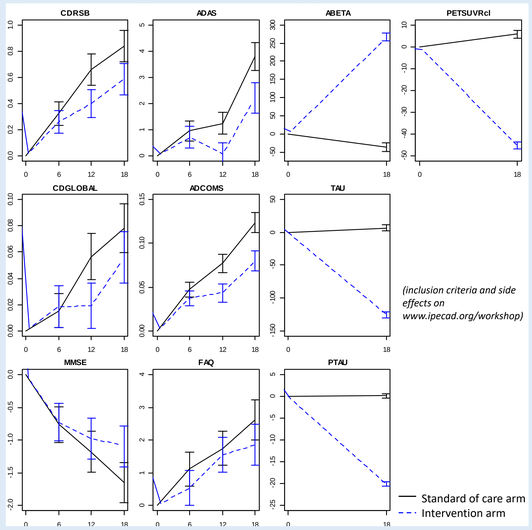
Table 1: benchmark baseline characteristics.

	Control (SOC + placebo) n=654	Intervention (SOC + DMT) n=654	p-value
Age (years), mean (SD)	73 (6.7)	73 (6.8)	0.420
Female, %	42%	39%	0.400
Education (years), mean (SD)	16.5 (3.5)	16.1 (3.6)	0.079
ApoE e4/e4, %	18%	19%	0.620
CDR SOB, mean (SD)	1.71 (1.01)	1.70 (0.99)	0.870
MMSE, mean (SD)	27.4 (2.4)	27.2 (2.4)	0.350
ADAS-Cog13, mean (SD)	18.0 (7.4)	18.6 (7.3)	0.110
ADCOMS, mean (SD)	0.23 (0.12)	0.23 (0.12)	0.580
FAQ, mean (SD)	4.30 (6.72)	4.43 (6.73)	0.730
CSF ABeta24 (pg/mL), mean (SD)	718 (242)	705 (232)	0.340
CSF total tau (pg/mL), mean (SD)	311 (133)	325 (146)	0.067
CSF phosphorylated tau (pg/mL), mean (SD)	31 (15.0)	33 (16.5)	0.023
Amyloid PET SUVr centiloid, mean (SD)	68 (40)	72 (42)	0.041

Table 2: benchmark effectiveness results over 18 months.

	Control (SOC + placebo) n=654	Intervention (SOC + DMT) n=654	p-value
CDR global, n (%)			0.463
-0.5 (from 0.5 to 0)	34 (5%)	54 (8%)	
+0 (from 0.5 to 0.5)	484 (74%)	473 (72%)	
+0.5 (from 0.5 to 1)	136 (21%)	127 (19%)	
+1.5 (from 0.5 to 2)	0	0	
+2.5 (from 0.5 to 3)	0	0	
CDR SOB, mean (SE)	0.84 (0.06)	0.59 (0.06)	0.004
MMSE, mean (SE)	-1.7 (0.2)	-1.1 (0.2)	0.021
ADAS-Cog13, mean (SE)	3.8 (0.3)	2.2 (0.3)	0.000
ADCOMS, mean (SE)	0.12 (0.01)	0.08 (0.01)	0.000
FAQ, mean (SE)	2.6 (0.3)	1.9 (0.3)	0.058
CSF ABeta24 (pg/mL), mean (SE)	-36 (6)	267 (6)	0.000
CSF total tau (pg/mL), mean (SE)	6.8 (2.3)	-125.7 (2.3)	0.000
CSF phosphorylated tau (pg/mL), mean (SE)	0.2 (0.3)	-20.1 (0.3)	0.000
Amyloid PET SUVr centiloid, mean (SE)	5.9 (0.9)	-45.3 (0.9)	0.000

Figure 1: benchmark effectiveness results: change from baseline over 18 months.



Participating models

Table 3: Selected details on benchmark scenario implementation for participating models.

	Markov-type					Micro-simulation-type			
Name	Biogen	CPEC	IPECAD	Bedrejo et al.	SveDem	BASQDEM	FEM	RTI-HS	MISCAN-Dementia
Abbreviation	(company name)	Care Policy Evaluation Centre	International Pharmaco-economic Collaboration on Alzheimer's Disease	(author name)	Swedish Dementia Registry	Basque dementia model	Future Elderly Model	(company name)	Micro-simulation SCreening ANalysis (Dementia)
Reference	[Herring, 2021]	[Anderson, 2018]	[Green, 2019]	[Replication of Lin, 2021]	[Wimo, 2020]	[Mar, 2020]	[Goldman, 2018]	[Herring, 2017]	[Brück, 2023]
Developer(s) (bold indicates present during the workshop)	PP-R, MU (Biogen)	RA, RW (London School of Economics)	RH (Maastricht University), CG (Biogen), AG (Quantify Research)	CBS (University of Alberta), ES (University of Calgary)	AW, RH (Karolinska Institute)	JM (Osakidetza Basque Health Service), MS-G (Mondragon Unibertsitatea)	BT (USC), YW (USC), JH (USC, Masaryk University)	WLH, FK (RTI Health Solutions)	CB, IdK (Erasmus Medical Center)
Choice of treatment effect outcome(s)	CDR-SB, change from baseline	MMSE, change from baseline	CDR-SB, time to dementia	CDR-SB, time to dementia	CDR global, change from baseline	CDR-SB, change from baseline	CDR-SB, change from baseline	MMSE and FAQ, changes from baseline	CDR global, change from baseline
Rationale	Health states in the model defined using CDR-SB ranges	Health states in the model defined using MMSE ranges	Measure captures multiple domains and is sensitive in early AD	Health states in the model defined using CDR-SB ranges	Transition to dementia is clinically relevant; aligns with use of synthetic control arm for MCI natural history	Time to dementia in the model relies on equation for CDR-SB progression over time	Cognitive states in the model are defined using CDR-SB, which also contributes to the staging of dementia	Aligned with the cognitive (MMSE) and functional (DAD) domains in the model (with mapping from FAQ to DAD)	Appropriate for the model's time to dementia approach; aligns with use of synthetic control arm for MCI natural history
Implementation method	Applied as a HR to transition probabilities from MCI to mild AD and from mild AD to moderate AD	Applied as a RR to transition probabilities from MCI to mild AD and from mild AD to moderate AD	HR applied directly to transition probability from MCI to mild AD; calibration required to model treatment effect in mild AD	Applied as a HR to transition probabilities from MCI to mild AD and from mild AD to moderate AD	Applied as a RR to transition probabilities from MCI to mild AD and from mild AD to moderate AD	Relative difference in CDR-SB change from baseline applied as treatment term coefficient in mixed model equation for CDR-SB progression	Applied as a RR to transition probabilities from MCI to mild AD and from mild AD to moderate AD	Relative reduction in MMSE and FAQ change from baseline applied to MMSE and DAD annual rates of decline, assuming a linear mapping between FAQ and DAD	Multiplicative factor for time in MCI calibrated to the CDR global transitions at 18 months for intervention arm; calibrated factor also applied to time in mild AD
MCI natural history	Synthetic trial control arm	[Vos, 2015]	[Vos, 2015]	[Potashman, 2021]	Synthetic trial control arm	Synthetic trial control arm; [van Oudenhoven, 019; Soininen, 2017]	[Wei, 2022]	Synthetic trial control arm; [Sapkota, 2021]	Synthetic trial control arm; [Vermunt, 2019]

Results

9 models (5 Markov and 4 microsimulation) implemented the benchmark scenario (see table 3). 1 other participating model's results lacked face validity and are thus not presented.

Treatment implementation methods (see table 3)

Choice of treatment effect outcome & scale used:

- Cognition domain, using MMSE
- Composite cognition and function domain, using CDR-SB, CDR-global
- Function domain, using FAQ

Pre-model analysis:

- Ratio of change from baseline intervention/control (dichotomous CDR-g)
- Ratio of change from baseline intervention/control (continuous CDR-SB, MMSE)
- Survival model to obtain relative risk (CDR-SB ≥4.5)

Between-model variability

Figure 2: time in state in control (left) and difference between control and intervention (right).

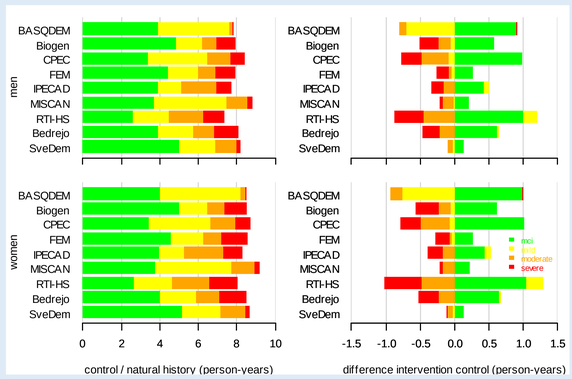
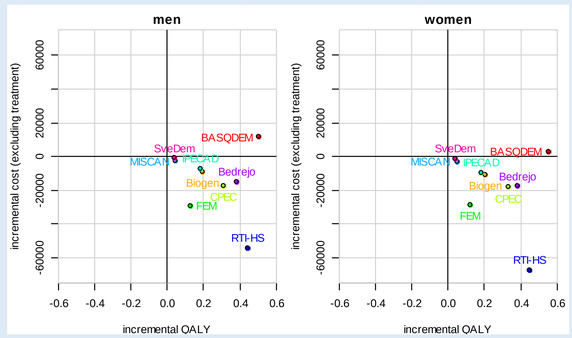


Figure 3: Incremental QALYs and costs (no treatment costs included).



Underlying causes for variation

- Chosen trial treatment effect outcomes varied largely in ratio in change from baseline in the hypothetical trial intervention/control arm (relative risk MMSE=0.65, FAQ=0.73, CDR-SB=0.70 when using reported estimate and 0.80 when using Cox regression on individual-level data, CDR-global=0.93) and seemed strongly related to the model's predicted gained time in MCI (correlation coefficient 0.84).
- Gained time in MCI was related to time in MCI (i.e., natural disease history) (correlation coefficient 0.63).
- QALY gain and cost difference could be explained by the gained time in MCI (correlation coefficient 0.89 and 0.50 respectively) (excluding treatment costs).

Discussion

The following cascade of factors may explain model differences:

1. model design determined choice of treatment effect outcome (e.g., CDR-SB, MMSE, etc.)
2. which was associated to the relative effect size (7% to 35%)
3. which together with natural history (faster progression creating a larger window for benefit) and assumptions on waning and discontinuation determined the time in states (particularly MCI)
4. resulting in differences in
   1. incremental QALYs (both more time in less severe states and increased life expectancy) and
   2. incremental costs (although less straightforward, as cost-savings due to more time in less severe states are offset by additional costs in life-years gained)

Comparison to previous studies

Variation in natural history, difference in time in MCI [Handels, 2023] and QALY gain [Handels, 2019] was relatively similar or had explainable differences.

Recommendations for modelers and reimbursement agencies

- Place greater focus on the MCI stage of a model (rather than to detail the dementia stage) as this stage drives variation in model outcomes.
- Address choice of treatment effect outcome and implementation into model in sensitivity analyses.
- Address uncertainties around predicted long-term model outcomes through prospective registries.
- Report model outcomes in standardized reporting table 4:

Suggested standardized template reporting table 4: *proportion persons in state over time in control and intervention strategy over 25-year time horizon (sex-specific, undiscounted, no half-cycle correction).*

	Control					Intervention				
Year	MCI	Mild dementia	Moderate dementia	Severe dementia	Death	MCI	Mild dementia	Moderate dementia	Severe dementia	Death
Control										
0										
1										
2										
3										
…										
25										

Limitations

- Of the invited modeling groups, 12 did not respond and 16 could not participate due to limited resources.
- Voluntary participation limited the availability of resources to generate evidence not specified in the benchmark scenario (e.g., implementing (surrogate biomarker) outcomes).

Invitation

We invite modelers internationally to submit their model results of this benchmark scenario or real-world interventions in a standardized format to our IPECAD open continuous cross-comparison. See https://osf.io/jv85a/ for more information.

Abbreviations: ABETA, amyloid beta; AD, Alzheimer's Disease; ADAS, Alzheimer's Disease Assessment Scale; ADCOMS, Alzheimer's Disease Composite Score; ADNI, Alzheimer's Disease Neuroimaging Initiative; CDR-SB, Clinical Dementia Rating sum of boxes; CSF, cerebrospinal fluid; DMT, disease-modifying treatment; FAQ, Functional Activities Questionnaire; MCI, mild cognitive impairment; MMSE, Mini-Mental State Examination; PET SUVr, Positron Emission Tomography Comparison of Standardized Uptake Value Ratio; QALY, quality-adjusted life year; SD, standard deviation; SOC, standard of care; US, United States

Maastricht University

Karolinska Institutet