Supplementary Materials C

Title

The Chronic Kidney Disease Care Model: quantifying management strategies for improving patient outcomes in the United States.

Authors

Andrew H Briggs, DPhil,[1]

0000-0002-0777-1997

Ciaran N Kohli-Lynch, PhD,[2]

0000-0001-5516-1862

Adrian R Levy, PhD,[3,4]

0000-0002-8966-0966

Ziyi Lin, MSc,[1]

0000-0002-3656-3601

Satabdi Chatterjee, PhD,[5]

Bonnie MK Donato, PhD, [5]

0000-0001-5099-2456.

Csaba P Kovesdy, MD.[6]

0000-0002-8204-911X

1. Department of Health Services Research & Policy, London School of Hygiene & Tropical Medicine, London, UK.
2. Department of Preventive Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL, US.
3. Department of Community Health and Epidemiology, Dalhousie University, Halifax, NS, Canada.
4. Pelican Research Inc, Halifax, NS, Canada.
5. Boehringer Ingelheim Pharmaceuticals, Ridgefield, CT, US.
6. Division of Nephrology, University of Tennessee Health Science Center, Memphis TN, US.

Corresponding Author information

Andrew H Briggs, DPhil

Professor of Health Economics

Department of Health Services Research & Policy

London School of Hygiene & Tropical Medicine

London, United Kingdom.

Email:andrew.briggs@lshtm.ac.uk

Phone: +44 (0)207 927 2348

AdViSHE
Assessment of the Validation Status of Health-Economic decision models

AdViSHE is divided into five parts, each covering an aspect of validation:

Part A: Validation of the conceptual model (2 questions)
Part B: Input data validation (2 questions)
Part C: Validation of the computerized model (4 questions)
Part D: Operational validation (4 questions)
Part E: Other validation techniques (1 question)

No final validation score is calculated, as the assessment of the answers and the overall validation effort is left to the model users.

Part A: Validation of the conceptual model (2 questions)

Part A discusses techniques for validating the conceptual model.
A conceptual model describes the underlying system (e.g., progression of disease) using a mathematical, logical, verbal, or graphical representation.
Please indicate where the conceptual model and its underlying assumptions are described and justified.

The conceptual model and its underlying assumptions are fully described and justified in the main manuscript (Methods: Model Structure, Model parameters, Sub-model structures, Modelling Interventions amnd presented graphically in Figure 1) and in the Supplementary Materials (Tables S1–S21, Figures S1–S2, and accompanying methodological text). The core model structure, sub-model frameworks, data sources, and assumptions regarding CKD progression, risks, and treatment effects are documented in detail across these sections. The conceptual model was reviewed during internal project meetings and considered appropriate for CKD progression and economic evaluation

A1/ Face validity testing (conceptual model): Have experts been asked to judge the appropriateness of the conceptual model?
If yes, please provide information on the following aspects:
Who are these experts?
What is your justification for considering them experts?
To what extent do they agree that the conceptual model is appropriate?

If no, please indicate why not.

Internal experts reviewed the conceptual structure and confirmed that it appropriately represents the underlying clinical process (CKD progression through KDIGO stages, competing risks for HF, CVD, ESRD, and other-cause death) and is suitable for economic evaluation. The five-sub-model framework and the assumed treatment pathways were judged clinically and methodologically appropriate. These internal experts at the sponsoring company have varying clinical, market access and statistical/economic modellign experience across a number of product lines.

Aspects to judge include: appropriateness to represent the underlying clinical process/disease (disease stages, physiological processes, etc.); and appropriateness for economic evaluation (comparators, perspective, costs covered, etc.).

A2/ Cross validity testing (conceptual model): Has this model been compared to other conceptual models found in the literature or clinical textbooks? If yes, please indicate where this comparison is reported. If no, please indicate why not.

Yes. The conceptual structure is consistent with other published CKD modelling approaches. No major discrepancies requiring justification were identified. It is deliberatley simpler (as a cohort model) than published micro-simulation models and this is addressed in the manuscript.

Part B: Input data validation (2 questions)

Part B discusses techniques to validate the data serving as input in the model.
These techniques are applicable to all types of models commonly used in HE modelling.
Please indicate where the description and justification of the following aspects are given:
search strategy;
data sources, including descriptive statistics;
reasons for inclusion of these data sources;
reasons for exclusion of other available data sources;
assumptions made to assign values to parameters for which no data were available;
distributions and parameters to represent uncertainty;
data adjustments: mathematical transformations (e.g., logarithms, squares); treatment of outliers; treatment of missing data; data synthesis (indirect treatment comparison, network meta-analysis); calibration; etc.

Input data sources (CRIC, Optum, utilities, treatment effects) and all transformations are described in the manuscript and Supplementary Tables S1–S21. Appropriate probability distributions were assigned to all parameters to represent uncertainty (lognormal for relative risks, beta for event probabilities, gamma for costs, and normal distributions for utilities). These represent standard, widely accepted data for CKD progression and economic evaluation.Continuous functions for event probabilities and annual costs were generated through meta-regression-based smoothing of the categorical inputs.

B1 / Face validity testing (input data):
Have experts been asked to judge the appropriateness of the input data?
If yes, please provide information on the following aspects:
Who are these experts?
What is your justification for considering them experts?
To what extent do they agree that appropriate data have been used?

If no, please indicate why not.

Yes.
The inputs were reviewed by the same multidisciplinary onternal team and judged appropriate, representative, and consistent with current CKD evidence. Assumptions for missing SEs and smoothing methods were documented.
Experts confirmed that:
• All input relationships behave as expected clinically: risks increase with worsening CKD severity (lower eGFR, higher uACR), and costs/utilities follow expected gradients across KDIGO categories.
• All input values fall within realistic clinical and economic ranges, matching the patterns seen in CRIC data, Optum costs, and published evidence.
• No structural or numerical anomalies were detected: no negative probabilities, unrealistic transitions, or inconsistent parameter values.

Aspects to judge may include but are not limited to: potential for bias; generalizability to the target population; availability of alternative data sources; any adjustments made to the data.

B2 / Model fit testing:
When input parameters are based on regression models, have statistical tests been performed?
If yes, please indicate where the description, justification, and outcomes of these tests are reported.

If no, please indicate why not.

Yes. Regression-based inputs were tested for statistical fit.
Alternative model specifications were compared using R^2, AIC, and BIC, and the best-fitting functional forms were selected.
Model fits were also assessed using:
• visual inspection of fitted curves and residuals,
• comparison against original data patterns from CRIC and Optum,
• checks of numerical stability and convergence under repeated runs.
No inconsistencies or implausible patterns were identified.

Simulation-based numerical stability: a 1,000-iteration probabilistic sensitivity analysis was performed using the regression-based inputs. Outputs showed no systematic drift or numerical instability, indicating that the regression-derived functions behave consistently across the full parameter space.

Examples of regression models include: disease progression based on survival curves; risk profiles using regression analysis on a cohort; local cost estimates based on multi-level models; meta-regression; quality-of-life weights estimated using discrete choice analysis; mapping of disease-specific quality-of-life weights to utility values.
Examples of tests include: comparing model fit parameters (R^2, AIC, BIC); comparing alternative model specifications (covariates, distributional assumptions); comparing alternative distributions for survival curves (Weibull, lognormal, logit); testing numerical stability and convergence of the model; visually testing model fit and/or regression residuals.

Part C: Validation of the computerized model (4 questions)

Part C discusses various techniques for validating the model as it is implemented in a software program.
If there are any differences between the conceptual model (Part A) and the final computerized model, please indicate where these differences are reported and justified.

The computerized model in Excel reflects the conceptual structure. An independent replication in R was completed for double programming and consistency checking. All validation activities described below confirm correct translation of the conceptual model, numerical stability, and absence of coding errors.

C1 / External review:
Has the computerized model been examined by modelling experts?
If yes, please provide information on the following aspects:
Who are these experts?
What is your justification for considering them experts?
Can these experts be qualified as independent?
Please indicate where the results of this review are reported, including discussion of any unresolved issues.

If no, please indicate why not.

Yes. The computerized model was reviewed by an independent economic modeller involved not involved in its development. They checked overall logic, implementation, formulas, and structure, and confirmed that the model matches the conceptual design.
This reviewer was not responsible for the direct coding of the Excel version and therefore qualifies as independent reviewer for the computerized implementation.

What was examined?
absence of coding bugs;
correct flow of cohort between model components;
meta-regression outputs;
logical code structure optimized for computational speed;
alignment between Excel and R versions.

Aspects to judge may include: absence of apparent bugs; logical code structure optimized for speed and accuracy; appropriate translation of the conceptual model.

C2 / Extreme value testing:
Has the model been run for specific, extreme sets of parameter values in order to detect any coding errors?

If yes, please indicate where these tests and their outcomes are reported.
If no, please indicate why not.

Yes. The model was run with extreme baseline inputs and wide range of baseline inputs, treatment combinations, and TSA Low/High values, deterministic/probabilistic runs and TSA.
Outcome: All runs were stable, produced logical results, and showed no numerical or coding issues.

Examples include: zero and extremely high mortality; extremely beneficial, detrimental, or no treatment effect; zero or extremely high treatment or healthcare costs.

C3 / Testing of traces:
Have patients been tracked through the model to determine whether its logic is correct?

If yes, please indicate where these tests and their outcomes are reported.
If no, please indicate why not.

Yes. Traces for selected profiles were inspected and matched expected movement across states.No missing transitions; all probabilities summed to ≤1 at each cycle.Post-event sub-models triggered correctly and accumulated costs/QALYs as intended. Excel and R results were consistent. Survival curves from the Excel implementation were compared to R-generated curves.

In cohort models, this involves listing the number of patients in each disease stage at one or several time points (e.g., Markov traces). In individual patient simulation models, this involves following several patients throughout their disease progression.

C4 / Unit testing:
Have individual sub-modules of the computerized model been tested?

If yes, please provide information on the following aspects:
Was a protocol that describes the tests, criteria, and acceptance norms defined beforehand?
Please indicate where these tests and their outcomes are reported.

If no, please indicate why not.

Yes. Sub-modules were tested individually, key formulas and named ranges were checked, and macros were updated and verified. Redundant elements from earlier versions (e.g., 5-year incidence blocks) were removed.

Examples include: turning sub-modules on and off; altering global parameters; testing messages (e.g., warning against illegal inputs), drop-down menus, named areas, switches, labelling, formulas and macros; removing redundant elements.

Part D: Operational validation (4 questions)

Part D discusses techniques used to validate the model outcomes.

Validation activities included expert review, comparison with previous model versions, testing with alternative inputs, and consistency checks against patterns observed in the CRIC population.

D1 / Face validity testing (model outcomes):
Have experts been asked to judge the appropriateness of the model outcomes?

If yes, please provide information on the following aspects:
Who are these experts?
What is your justification for considering them experts?
To what extent did they conclude that the model outcomes are reasonable?

If no, please indicate why not.

Yes. Model outcomes (event rates, ESRD incidence, QALYs, lifetime costs) were reviewed by experts involved in the model development and review. These experts have extensive experience in CKD epidemiology and health-economic modelling. They concluded that the outcomes produced by the continuous model were reasonable and consistent with expectations for CKD populations.

Outcomes may include but are not limited to: (QALYs), deaths, hospitalizations, total costs.

D2 / Cross validation testing (model outcomes):
Have the model outcomes been compared to the outcomes of other models that address similar problems?
If yes, please provide information on the following aspects:
Are these comparisons based on published outcomes only, or did you have access to the alternative model?
Can differences in outcomes between your model and other models be explained?
Please indicate where this comparison is reported, including discussion of comparability with your model.

If no, please indicate why not.

Informal comparsion is made in the manuscript with similar models in the literature based on published outcomes. The model presented is deliberatley simpler than published models but produces similar final outcomes in terms of cost and QALYs. Formal cross validation with other models was not performed as we did not have access to the published models.

Other models may include models that describe the same disease, the same intervention, and/or the same population.

D3 / Validation against outcomes using alternative input data:
Have the model outcomes been compared to the outcomes obtained when using alternative input data?

If yes, please indicate where these tests and their outcomes are reported.
If no, please indicate why not.

No alternative data was available for inputs.

Alternative input data can be obtained by using different literature sources or datasets, or by splitting the original dataset into two parts (one for calibration, one for validation).

D4 / Validation against empirical data:
Have the model outcomes been compared to empirical data?

If yes, please provide information on the following aspects:
Are these comparisons based on summary statistics or patient-level datasets?
Have you been able to explain any difference between the model outcomes and empirical data?
Please indicate where this comparison is reported.
If no, please indicate why not.

Yes. Model outcomes were compared against empirical patterns from the CRIC dataset (dependent validation). Trends in ESRD incidence, CVD events, costs, and utilities followed the expected patterns across KDIGO categories. Independent external datasets were not available, but no inconsistencies with published CKD epidemiology were identified. The model projects to lifetimes which is beyond the availability of empirical data.

D4.A Comparison against data sources on which the model is based (dependent validation).

Yes. Model trends match empirical CRIC patterns. No independent external dataset was available for additional validation.

D4.B Comparison against a data source not used to build the model (independent validation).

No. The model was not validated against an external dataset not used in its development, as no suitable independent CKD data source was available for this analysis.

Part E: Other validation techniques (1 question)

E1 / Other validation techniques:
Have any other validation techniques been performed?

If yes, indicate where the application and outcomes are reported, or provide a short summary here.

Yes. Several additional validation techniques were performed:
1) Structured walk-throughs: Full Markov traces, state transitions, and payoffs for each Traces were checked manually. This process identified and corrected several small issues (e.g., misaligned cost formulas, incorrect ESRD probability reference, duplicated "All-fatal ESRD" entry, displaced cell references).
2) Naïve benchmarking: simple manual checks ("back-of-the-envelope") verified that event risks, costs, utilities, and treatment effects behaved as expected. Simplified manual calculations were used to verify reasonableness of event rates, treatment risk reductions, accumulated costs, utilities, and time-to-event patterns across KDIGO categories.
3) Double programming / full replication: the entire model was independently reproduced in Excel and R.
This included complete reproduction of:
• Dashboard SA, Dashboard PSA
• Tornado diagrams
• Incremental cost-effectiveness outputs
• Event-incidence distributions
• First-event composition
• Full lifetime Markov traces across all time horizons
4) Subgroup and scenario checks: model performance was confirmed across all KDIGO categories and treatment combinations.
5) No formal heterogeneity tests were performed.

Examples of other validation techniques: structured "walk-throughs" (guiding others through the conceptual model or program step-by-step); naïve benchmarking ("back-of-the-envelope" calculations); heterogeneity tests; double programming (two developers program components independently or in different software packages to check consistency).