

A prototype software framework for transparent, reusable and updatable computational health economic models

Matthew P Hamilton^{1,2,*} Caroline X Gao^{2,3,1} Glen Wiesner⁴ Kate M Filia^{2,3}
Jana M Menssink^{2,3} Petra Plencnerova⁵ David Baker^{2,3}
Patrick D McGorry^{2,3} Alexandra Parker^{6,3} Jonathan Karnon⁷
Sue M Cotton^{2,3} Cathrine Mihalopoulos¹

Abstract

Summary: Most health economic analyses are undertaken with the aid of computers. However, the ethical dimensions of implementing health economic models as software (or computational health economic models (CHEMs)) are poorly understood. We propose that developers and funders of CHEMs share ethical responsibilities to (i) establish socially acceptable user requirements and design specifications; (ii) ensure fitness for purpose; and (iii) support socially beneficial use. We further propose that a transparent (T), reusable (R) and updatable (U) CHEM is suggestive of a project team that has largely fulfilled these responsibilities. We propose six criteria for assessing CHEMs: (T1) software files are open access; (T2) project team contributions and judgments are easily identified; (R1) programming practices promote generalisability and transferability; (R2) licenses restrict only unethical reuse; (U1) maintenance infrastructure is in place; and (U2) new releases are systematically retested and appropriately deprecated. To facilitate CHEMs that meet TRU criteria, we have developed a prototype software framework in the open-source programming language R. The framework comprises six code libraries for authoring CHEMs, supplying CHEMs with data and undertaking analyses with CHEMs. The prototype software framework integrates with services for software development and research data archiving. We determine that an initial set of youth mental health CHEMs we developed with the prototype software framework wholly meet criteria T1-2, R1-2 and U1 and partially meet criterion U2. Our assessment criteria and prototype software framework can help inform and improve ethical implementation of CHEMs. Resource barriers to ethical CHEM practice should be addressed by research funders.

Code: Visit <https://www.ready4-dev.com> for more information about how to find, install and apply the prototype software framework and CHEMs developed with it.

¹ School of Public Health and Preventive Medicine, Monash University, Clayton, Australia

² Orygen, Parkville, Australia

³ Centre for Youth Mental Health; The University of Melbourne, Parkville, Australia

⁴ Heart Foundation, Melbourne, Australia

⁵ headspace National Youth Mental Health Foundation, Melbourne, Australia

⁶ Institute for Health and Sport, Victoria University, Footscray, Australia

⁷ Flinders University, Adelaide, Australia

* Correspondence: Matthew P Hamilton <matthew.hamilton1@monash.edu>

1 Introduction

Health economics is a discipline concerned with problems that arise due to scarce resources, such as how to value health and healthcare, allocate healthcare budgets and configure health services [1]. In seeking to solve these problems, health economists typically use models which are simplified and selective

Table 1: Implications of choosing commercial modelling software or an open-source programming language for the computational implementation of health economic models

| Implications | Commercial modelling software | Open-source programming language |
|------------------|--|--|
| Resources | (+) Will include tools to efficiently accomplish many common modelling tasks. (+) Requires limited or no programming skills. (-) Requires model developers and users to purchase software licenses | (-) Initial development is likely to be time and resource intensive. (+) Over medium term, efficiency savings are possible if: (i) artefacts (code and data) from pre-existing models are leveraged; and / or (ii) maintaining and transferring (e.g. to new decision problems / contexts) model. |
| | Transparency | |
| | (+) Popular tools (e.g. Excel) are readily understood by many. (-) Non seamless integration with scientific manuscript authoring pipelines raises risk of (undetected) transcription errors when reporting results. | (+) Easy to publicly share testable source code while selectively restricting access to confidential data. (+) Easy to make clear who has developed and tested each part. (-) Novel code increases likelihood of (undetected) software errors. |
| Ethical | Reusability | |
| | (+) Files will open and execute correctly for years after project is completed. | (-) Potentially fragile – if not maintained / not bundled with all required and correctly versioned dependencies (+) Facilitates modular implementations that make complex model representations more tractable. (+) Facilitates customised integration with existing data / decision systems. |
| | Updatability | |
| | (-) Selective modification with appropriate attribution can be difficult. | (-) Facilitates collaborative development and maintenance. |

representations of systems that are believed to influence human health. A health economic model should be capable of being described using words and figures (a conceptual model), equations (a mathematical model) and software (a computational model). Health economic scientific manuscripts typically describe a model using its conceptual and mathematical representations but report results that have been generated by its computational representation (i.e., the format that allows computers to perform calculations). The conceptual, mathematical and computational representations of a model are assumed to be isomorphic. However, the ease of independently assessing the validity of this assumption depends in part on how a health economic model has been implemented computationally.

The set of software files implementing a health economic model’s structure, data and algorithms can be called a computational health economic model (CHEM). CHEMs can be developed using commercial modelling software or open-source programming languages such as Python [2] or R [3]. An advantage of commercial modelling software is that users often require limited or no programming skills to develop and apply CHEMs. However, open-source software programming languages may facilitate the development of CHEMs that are more transparent, reusable and updatable [4,5]. The decision about what type of software development platform to use must balance ethical considerations and resource constraints (see Table 1).

Open-source CHEMs remain rare [6,7]. The software used to develop CHEMs was identified as a barrier to open-source model implementations by participants in a 2020 survey of health economists [5]. Other reported barriers relate to difficulties in generalising models, updating models and transferring model data, concerns about the level of public access and lack of interest from decision-makers, and legal, confidentiality and security considerations [5].

Better software frameworks could potentially help address some of the barriers to open-source CHEM development. A software framework is a shared common technology used by developers to collaboratively author software and which is not typically visible to software end-users [8]. A software framework provides a foundation for developing multiple software applications with shared resources (e.g., code and data files), that can be modified to suit specific needs. Advantages of using software frameworks include facilitating code reuse and extension, promoting good programming practice and the capability to provide enhanced functionality and performance without additional effort by developers [9]. Software frameworks have widely been developed and implemented in data science, for example, PyTorch [10] for machine learning using Python.

A high-level software framework for implementing open-source CHEMs in R has previously been developed with the primary aim of improving model transparency [11]. However, to gain widespread adoption, a software framework for CHEMs may need to provide tools to help model developers meet a broader range of ethical practice requirements. There are multiple challenges to overcome. Firstly, the ethical responsibilities of public health computational modellers remain poorly understood [12] and the computational modelling landscape is changing fast [13]. Secondly, software frameworks can be challenging and time consuming to create and may become excessively complex over time [9]. Thirdly, software frameworks can be difficult to learn, often requiring model developers to undergo specialist training [8].

In this paper, we:

- (i) propose a set of ethical responsibilities for CHEM project teams and criteria for assessing ethical CHEM implementations;
- (ii) describe a prototype software framework we have developed to support the ethical implementation of CHEMs; and
- (iii) assess an initial set of youth mental health CHEMs developed with our software framework against our proposed ethical CHEM implementation criteria.

2 Ethical issues

We considered literature on modelling practice and reflected on our prior experience with developing and using CHEMs to identify: (i) some core ethical responsibilities of CHEM project teams; (ii) attributes of CHEMs that can suggest fulfillment of these responsibilities; and (iii) criteria against which these attributes can be assessed.

2.1 Ethical responsibilities of CHEM project teams

Boden and McKendrick [12] propose a framework for ethical public health modelling based on the criteria of independence (concerning how modeller subjectivity shapes model design), beneficence / non-maleficence (concerning model quality and utility), transparency (concerning the need for policymakers to reliably evaluate model strengths and weaknesses) and justice (concerning the social obligations of modellers to consider and communicate ethical issues about model use). However, these criteria are not specific to the computational model, but are designed to apply to a modelling project as a whole, including a model’s conceptual and mathematical representations. The authors propose 13 questions to evaluate ethical risk across the four criteria, but of these, only “is the model code open source or available on request” (for transparency) and “has the model been verified, i.e., does it do what the modeller wants it to do?” (for beneficence / non-maleficence) are specific to model’s representation as software.

To identify ethical issues that are specific to CHEMs, we first considered the distinct phases in a software development project. Software development lifecycle (SDLC) models take different approaches to the inclusion, naming, definition, sequencing and iteration of project phases [14]. However, SDLC models typically have components that can principally map to the concepts of *planning* (e.g., identification of system / user requirements and design specifications), *implementation* (e.g., development and testing) and *release* (e.g., deployment, system integration, maintenance and support). For simplicity, we propose one overarching ethical responsibility for CHEM project teams for each phase. These responsibilities are to:

- (i) **establish socially acceptable user requirements and design specifications** (during CHEM planning);
- (ii) **ensure fitness for purpose** (of CHEM implementation); and
- (iii) **support socially beneficial use** (on CHEM release).

Transparency underpins all three responsibilities (Table 2). Boden and McKendrick’s other ethical criteria can be selectively mapped to each phase’s responsibility: justice and independence to the planning responsibility; independence to the implementation responsibility; and beneficence / non-maleficence and justice to the release responsibility. All three ethical responsibilities have implications for both professional practice and project resourcing. As it is unreasonable to place expectations of modellers that they are not resourced to fulfil, these responsibilities should apply jointly to health economic model developers and funders.

Misalignment between the values of computational model developers and those of the population groups affected by decisions based on their models presents significant ethical risks [15,16]. The value judgments of health economic modellers are rarely adequately specified, omissions that may lead to socially unacceptable policy recommendations [16]. These value judgments influence the assumptions, selection of model features and standards for evidence that shape health economic model projects [17]. For example, as illustrated in Table 1, project resource constraints may require model developers to make trade-offs when agreeing CHEM **user requirements and design specifications**. These trade-offs in turn determine how straightforward it is for a CHEM to be subsequently amended to reflect alternative value judgments relating to model conceptualisation and mathematical formalisation (Table 2).

Modellers have duties both to ensure a computational model is **fit for purpose** and to provide potential third-party users with the means of assessing its adequacy for their proposed uses [6,12,18,19]. However, it is common for health economic models to have serious methodological flaws [20,21]; insufficient validation [22–24], poor reproducibility [25–27]; and undeclared errors [28]. Appropriate computational implementation choices can help address many of these shortcomings, for example by automating quality assurance checks and facilitating manual reviews by third parties.

A CHEM will have limited **social benefit** if it is rarely used, if mis-used or when its acceptability and adequacy rapidly decay. Reuse of CHEMs as components of other models can potentially make model development more efficient [29,30]. However, health economic models face challenges related to transferability across jurisdictions [30] that create barriers to reuse. Without ongoing maintenance, a CHEM risks becoming less reliable with time [30] and is at risk of being deployed for purposes for which it is poorly suited [13]. Currently, health economic models are rarely implemented computationally in a manner that facilitates routine updates [31], thus limiting the temporal window within which a CHEM can be validly applied.

2.2 CHEM attributes associated with ethical modelling practice

The responsibilities of model project teams over a CHEM lifecycle are easier to state than to measure. However, aspects of ethical modelling practice may be inferred from measurable attributes of CHEMs. As described in Table 2, we believe that the creators of transparent, reusable and updatable (TRU) CHEMs are likely to have fulfilled a number of their ethical obligations. We therefore selected these model attributes to use as the basis for deriving ethical assessment criteria.

Table 2: Transparent, reusable and updatable (TRU) computational health economic models (CHEMs) are suggestive of ethical modelling practice

| | Social acceptability | Fitness for purpose | Beneficial use |
|--------------------|--|--|---|
| Transparent | Modeller judgments, model features and verification checks can be reviewed by third parties. | | Decision-makers can understand the strengths and weakness of models before applying them. |
| Reusable | Models can be modified to reflect alternative value judgments. | Third party use increases likelihood of errors being identified. | Can inform more decision problems, in more contexts with less duplicative effort. |
| Updatable | | Models that are maintained can be validly used for longer. | Models can be validly used for longer. |

Taking steps to make computational models **transparent** is an important part of ethical public health modelling practice [12]. Guidance on transparency in health economic modelling recommend that model code and data should be clearly documented [32]. Notably, the same guidelines, published over ten years ago, do not recommend sharing model code and data. However, more recent healthcare modelling guidance does recommend public dissemination of such artefacts [18]. Online repository services such as Zenodo [33] and Dataverse [34] provide persistent storage solutions that generate a Digital Object Identifier (DOI) for each code and data collection. Ensuring that calculations are correct and consistent with model specifications is an essential part of CHEM quality assurance [35]. The extensiveness of such verification checks can be reported using the concept of code coverage [36] - the proportion of model code that has been explicitly tested. Tests should ideally combine both unit tests (to verify that small, isolated sections of code produce the correct output when run independently) and acceptance tests (to verify that the correct output is produced when multiple code components are run together to perform tasks that meet core user-requirements [37]). The nature and extent of individual model authorship contributions can become unclear when models are implemented over longer time-frames with a large and changing group of collaborators [15]. This issue can be addressed by use of online repository services such as GitHub [38], that provide citation tools and can transparently record all individual code contributions to a modelling project over its lifecycle.

A CHEM that is **reusable** also signals ethical modelling practice. Although making a CHEM’s code, data and documentation publicly available is increasingly considered good practice, it is insufficient to enable model re-use. Key concerns for health economists when considering whether to reuse a model are generalisability (application without adaptation) and transferability (selective reuse and/or modification of model components) [39]. Writing model algorithms as collections of functions (short, self-contained and reusable software routines that each perform a discrete task) is good scientific computing practice [40] and promotes selective reuse. Computational implementations that store model code and data in distinct files and locations (as opposed to embedding data such as parameter values into source code) are easier to selectively modify. Modular implementations enable models to be constructed from multiple independently reusable and replaceable sub-models (modules) [41]. The programming concept inheritance [42], deployed in object-oriented programming approaches, can facilitate the duplication and selective modification of models. Granting permissions to others to use, test and adapt models and their components, can be facilitated by two broad categories of open source licenses. Some guidance strongly recommends the use of permissive licensing [40] that provides users with great flexibility as to the purposes (including commercial) for which content can be re-used. Alternatively, copyleft licenses [43] can be used to require that derivative works created by content users remain open-source.

Models should be **updatable** so that they remain valid for longer, evolving as new evidence emerges and the systems being modelled change [30,44]. Ensuring that a model is regularly reviewed to identify and implement required improvements is a recommended defence against model validity decay [13]. Sustainable maintenance of open source research software requires both a core development team and an active user community [45]. Online communities can be an efficient means of engaging model users in testing each version of a model, identifying issues and suggesting improvements. Services such as GitHub [38] provide collaborative code development tools [46] that help elicit, integrate and reconcile contributions from multiple contributors and to ensure each update is uniquely identifiable and retrievable. It is also important that verification checks are rerun with each model update, a task that can be automated using the software development practice of continuous integration [47]. The risk of software revisions having unintended consequences for third party users can be mitigated through the use of deprecation conventions [48] that take an informative and staged approach to retiring outdated CHEMs.

2.3 Assessment criteria

For each CHEM attribute that suggests ethical CHEM modelling practice, we identified assessment criteria.

How **transparent** a CHEM is can be assessed against the criteria:

- **T1:** Software files (executable code, testing procedures and outcomes and non-confidential data) are available in open access repositories.
- **T2:** It is easy to see who funded, developed and tested each part of the CHEM and to identify the modelling team’s assumptions, judgments and theories about CHEM development and use.

How **reusable** a CHEM is can be assessed against the criteria:

- **R1:** Programming practices facilitate both generalisability and transferability.
- **R2:** Licenses restrict only unethical use and allow anyone to ethically reuse model code and non-confidential data, in whole or in part, without charge, and for purposes that include the creation of derivative works.

How **updatable** a CHEM is can be assessed against the criteria:

- **U1:** Maintenance infrastructure is in place to support version control and collaboration with model users.
- **U2:** Each new release of a CHEM is systematically retested, with changes implemented to minimise disruptions for existing model users.

3 Prototype software framework

To support the development of CHEMs that meet TRU assessment criteria, we have created a prototype software framework called ready4. The software framework aims to provide a toolkit for: (i) enabling modular implementation of CHEMs; (ii) authoring and documenting CHEM modules; (iii) managing the labelling and transfer of CHEM input and output data; and (iv) authoring reproducible analyses that apply CHEMs to compatible data.

To achieve these goals, we have implemented the software framework as R [3] code libraries that integrate with a number of online services and which are supported by a documentation website.

3.1 R libraries

A library in the R language will typically depend on multiple other R libraries, all of which potentially have different authors. As the number of third-party dependencies of an R library grows so does the fragility of that library (e.g., the library may cease to work as intended due to changes in one of its dependency libraries). To reduce the fragility of our framework we implemented it as multiple R libraries rather than

one R library. In total we authored six novel R libraries to implement the ready4 framework, all of which have distinct purposes and dependencies (Table 3).

One framework library provides a **foundation for modular CHEM implementations**. In modular model implementations, modules need to be able to share inputs and outputs with each other and to be run as independent models [49]. To achieve this, the foundation framework library defines a template CHEM module (using R’s S4 class system), which can be used to create other CHEM modules with a common set of inherited properties. One of these inherited properties is a novel syntax of 15 core commands that enable CHEM module algorithms to be consistently named. The foundation library also contains tools for retrieving web-based information on CHEM modules, datasets and analysis programs authored with the framework and tools for partially automating updates to a project documentation website.

Three framework libraries are designed to help streamline and standardise workflows for **authoring CHEM modules** from the template module. The R language supports functional and object-oriented programming paradigms [50]. Authoring with each of these paradigms is facilitated by a dedicated CHEM module authoring library. One module authoring library contains tools for writing functions in a consistent house style and automatically generating basic documentation for each function. A second module authoring library contains tools to help streamline and standardise the authoring and documenting of novel CHEM modules. A third module authoring library provides tools for disseminating themed bundles of CHEM modules as R libraries that are:

- documented (with a website and PDF manuals);
- licensed (using the copyleft GNU GPL-3 [66] by default);
- easily citable (citation information can be retrieved within an R session or from hosting repositories); and
- quality assured (each update triggers continuous integration workflows, including any acceptance and unit tests created by module library authors).

A library for **managing CHEM data** contains tools for supplying CHEM modules with input data ingested from local (i.e., a user’s computer) or remote (online repositories) locations, labelling CHEM module datasets and exporting CHEM module data to online repositories. A library for **authoring reproducible analyses** contains tools to help write programs that apply CHEM modules to compatible datasets for the purpose of undertaking health economic analyses. These analysis programs can be either self-documenting (code is integrated with plain English explanations of what it does) or trigger the creation of explanatory documents (e.g., a scientific manuscript).




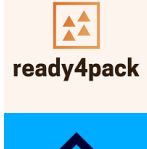
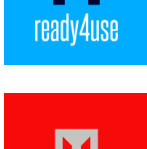

3.2 Online services

Framework libraries are designed to be used in conjunction with a number of third-party online services that we established and configured accounts with.

We created a GitHub organisation (a collection of code repositories) where source code that we author is stored and version controlled. We configured individual repositories in our GitHub organisation to implement continuous integration. By default, code libraries authored with our framework will use continuous integration to assess compliance with policies specified by the Comprehensive R Archive Network (CRAN) [51]. To track our code coverage, we linked our GitHub organisation to an account we established at codecov [52]. To facilitate the creation and hosting of documentation websites, we enabled GitHub Pages in each repository used for code library development.

We also created a Zenodo community - a collection of permanent, uniquely identified repositories. We then linked our Zenodo community and GitHub organisation so that every time we specify a version of code in one of our GitHub repositories as a “release”, a copy of that code is automatically created on Zenodo with a DOI. Finally, to manage model datasets, we created a dedicated collection within the Harvard Dataverse installation.

Table 3: Software framework R libraries

| Package | Purpose | R library dependencies |
|---|---|--|
|  | Provide a template and novel syntax for modular CHEM implementations and tools for finding interoperable CHEM modules, datasets and reproducible analysis programs. | assertthat bib2df dataverse dplyr fs Hmisc kableExtra knitr lifecycle magrittr methods natmanager piggyback purrr readr readxl rlang rmarkdown rvest stats stringi stringr testit testthat tibble tidyRSS tools utils zen4R |
|  | Streamline and standardise the authoring and documenting of functions that support transferable and generalisable model algorithms. | desc devtools dplyr gert Hmisc knitr lifecycle lubridate magrittr methods piggyback pkgdown purrr readxl ready4 ready4show ready4use rlang sinew stats stringi stringr testit testthat tibble tidyr tools usethis utils xfun |
|  | Streamline and standardise the authoring and documenting of new interoperable CHEM modules. | devtools dplyr fs gtools Hmisc knitr lifecycle magrittr methods purrr ready4 ready4fun ready4show rlang stats stringi stringr testit testthat tibble tidyr usethis utils |
|  | Help bundle and disseminate newly created CHEM modules as R libraries that are documented, licensed and quality assured. | dataverse dplyr knitr lifecycle magrittr methods purrr ready4 ready4class ready4fun rlang stringr testthat tibble tidyr utils |
|  | Help manage the labelling and transfer of data between CHEM modules and local and remote data repositories. | data.table dataverse dplyr fs Hmisc knitr lifecycle magrittr methods piggyback purrr readxl ready4 ready4show rlang stats stringi stringr testit testthat tibble tidyr utils |
|  | Facilitate the use of CHEM modules in programs to implement integrated and reproducible data ingest, analysis and reporting pipelines. | dataverse DescTools dplyr flextable grDevices here Hmisc kableExtra knitr knitrBootstrap lifecycle magrittr methods officer purrr ready4 rlang rmarkdown stringi stringr testthat tibble tidyr utils xtable |

3.3 Documentation website

We developed a framework documentation website (<https://www.ready4-dev.com>) that provides guidance to model developers on how to use and contribute improvements to the ready4 software framework and CHEM modules developed with it. The documentation website is versioned, which means documentation relating to prior versions of framework software can be archived, retrieved and viewed.

The documentation website was developed using the Hugo framework [53], Docsy theme [54] and Algolia DocSearch [55] and is hosted using the Netlify [56] service. We used functions from our foundation framework library to partially automate website updates relating to available CHEM modules, datasets and analysis programs. We linked our Netlify account to our GitHub organisation so that the project website would automatically update whenever its source code (publicly available in a GitHub repository) was edited.

4 Application in youth mental health

We are using the ready4 software framework to implement multiple transparent, reusable and updatable CHEMs in youth mental health (see Figure 1). Some initial outputs from this work are publicly available as development version releases.

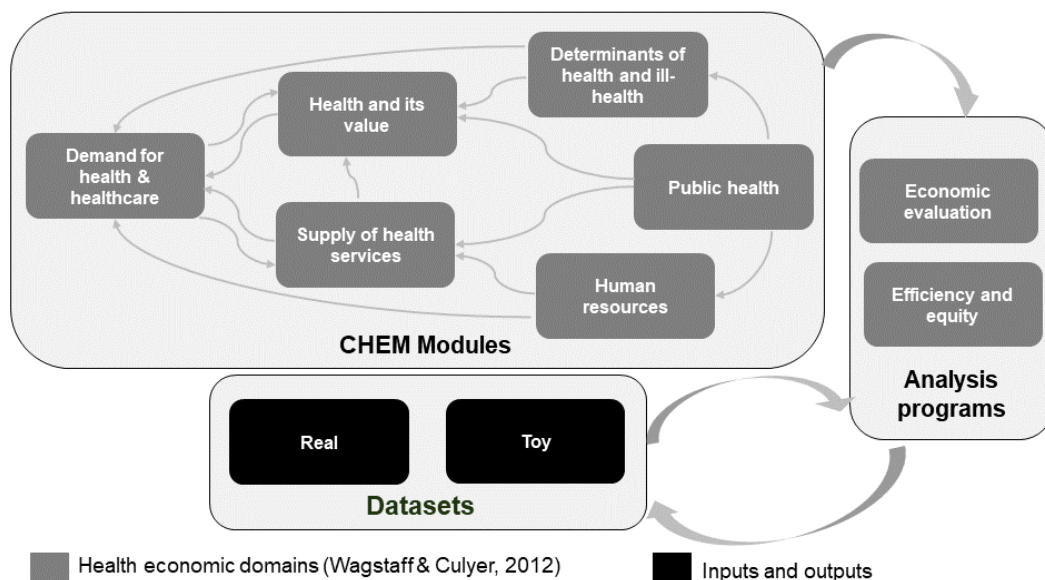


Figure 1: High level summary of planned implementation of youth mental health economic model

4.1 Economic topics

Currently, we are using the ready4 software framework to develop, apply, and share youth mental health CHEMs in four of the twelve domains of health economics identified by Wagstaff and Culyer [1]:

- health and its value (our projects: models to map psychological and functional measures to health utility);
- determinants of health and ill-health (our projects: models for creating synthetic household populations with key risk and protective factors for mental disorders);
- demand for health and health care (our projects: spatial epidemiology and help-seeking choice models); and

- supply of health services (our projects: a model of primary mental health care services for young people).

Potential future directions are to supplement this work with CHEMs in two additional Wagstaff and Culyer domains of: (i) public health (to model the impact of selected fiscal policy and regulation options on young people’s mental health); and (ii) human resources (to model the supply and behaviours of the youth mental health workforce). Our ultimate aim is to flexibly combining all our CHEMs in analyses that help answer questions in two additional Wagstaff and Culyer domains:

- efficiency and equity (our goal: assess the distributional impacts and identify the optimal targeting of youth mental health interventions); and
- economic evaluation (our goal: assess the cost-utility of competing policy options for improving the mental health of young people).

Although we are principally interested in using CHEMs to answer policy questions relating to the mental health of young people in Australia, we want to facilitate CHEM transferability to other jurisdictions. Our CHEMs are being derived from and applied to real data (which can be empirical, simulated or assumption, so long as it is appropriate for use in analysis intended to inform decision-making) from Australia. To help demonstrate the potential use of CHEMs in other decision contexts, we also create toy datasets. Data created for illustration purposes is prominently labelled as not for use in decision-analysis.

4.2 Case study: health and its value

We have previously described a study [57] to develop utility mapping models for use in samples of young people presenting to primary mental health services. The ready4 software framework was used in that study to develop CHEM modules, supply those modules with data and implement modelling analyses, creating the following artefacts:

- development version module libraries for describing and validating youth mental health human record datasets [58], scoring health utility [59], specifying utility mapping models [60] and implementing reproducible utility mapping studies [61];
- a development version library of functions for finding and using utility mapping models developed with these tools [62];
- collections of real data (study input and results [63]) and toy data (synthetic populations for testing model modules [64]);
- programs for replicating all steps from data ingest to manuscript production [65], applying utility mapping models to new data [66] and generating a synthetic representation of the study dataset [67]; and
- subroutines for creating a catalogue of utility mapping models [68] and generating a draft scientific manuscript [69] for studies implemented with these modules.

We created a checklist (Table 4) that we used to subjectively assess these study outputs against TRU criteria. For each criterion, we provided a global assessment of whether it was met using the responses “yes”, “no” or “partial”. We believe our utility mapping study CHEMs have satisfactorily met five of the six criteria (T1, T2, R1, R2 and U1) and have partially met one criterion (U2). The main shortcomings that we identified when applying the assessment criteria was that we have yet to adequately implement unit testing of the R libraries authored as part of this study.

Table 4: Assessment of a CHEM implementation in youth mental health using Transparent, Reusable and Updatable (TRU) criteria

| Criterion | Met? | Detail |
|--|---------|---|
| T1 Software files are open access | Yes | All source code and testing procedures are available in public GitHub repositories, with each code release persistently available in a Zenodo repository. As the study dataset contains confidential patient records, it was not published. Instead, a synthetic representation of the study dataset is persistently available in a repository in the Harvard Dataverse. Data files to support out of sample application of models are published at the same location. |
| T2 Project team contributions and judgments are easily identified | Yes | All module libraries, programs, and datasets are distributed with citation information. Public GitHub repositories detail author code contributions over project development history. Model catalogues persistently available on the Harvard Dataverse describe the predictive performance of models under multiple usage regimes. Each code library is documented with worked examples of how to apply CHEM modules. Analysis and reporting programs are self-documenting. Sub-routines for generating reports are documented with README files. |
| R1 Programming practices promote generalisability and transferability | Yes | CHEMs are written using both functional and object-oriented paradigms. Code library documentation websites include hypothetical examples of generalisability (applying study algorithm to estimate mapping models from new data with the same predictor and outcome variables) and transferability (adapting study algorithm to develop mapping models from datasets with different predictor variables and outcomes measured with a different utility instrument). |
| R2 Licenses restrict only unethical reuse | Yes | All code is distributed using GPL-3 licenses. Datasets use an amended version of template terms provided by Harvard Dataverse, allowing reuse of data subject to some ethical restrictions (e.g., use in efforts to reidentify study participants is prohibited). |
| U1 Maintenance infrastructure is in place | Yes | All code is version controlled using git and GitHub, with semantic versioning. Each code library has a specified maintainer and guidance for potential code contributors is available on the project website. |
| U2 New releases are systematically retested and deprecated | Partial | Continuous integration is used for all code libraries, primarily for acceptance testing. Only limited use is made of unit testing. Retired library code is deprecated using tools from the lifecycle R library. Library documentation articles and datasets are also deprecated. |

5 Discussion

Ethical practice is a core expectation of health researchers and computational methods underpin most quantitative research, yet an understanding of what constitutes ethical public health modelling practice is underdeveloped [12]. We have addressed this gap with our proposed responsibilities for model developers and funders, enabling CHEM attributes and CHEM implementation assessment criteria.

Currently, many if not most existing CHEMs are insufficiently transparent [22,25–27], not reusable [6,7], nor updatable [31,70]. Addressing these issues will require changes to the tools, practices and funding underpinning health economic model development.

As illustrated by Table 4, we have developed a prototype software framework that can be used to implement youth mental health CHEMs that largely satisfy our TRU criteria. However, in its current form our software framework is too fragile to be recommended as anything more than a prototype for supporting the development needs of other modelling teams and projects. We developed our prototype software framework primarily to address the user requirements of one group – ourselves. Factors such as user enjoyment, usability, active user-community and supporting resources influence adoption of software frameworks [8]. We currently lack the resources required to address these issues for our software framework.

However, our prototype has a number of features that future work on CHEM software frameworks may find useful to incorporate. Firstly, developing a software framework to work within an existing and widely used open source programming language such as R, can keep framework scope relatively narrow. This makes it more tractable to develop, maintain and learn, and facilitates integration with other modelling tools written in that language (e.g., the dependency libraries we list in Table 3). Secondly, implementation that combines both object-oriented and functional programming paradigms can avail of the modular and syntactical simplicity benefits of the former, while limiting needless bundling of code artefacts. Thirdly, a sensible trade-off needs to be found between transparent code implementation (which requires clear and sufficiently detailed documentation) and Agile Software Development (for which a foundational principle is prioritising the development of working code over writing documentation [71]). Our software framework makes this trade off by enforcing the use of consistent code style conventions and file organisation which in turn enables automated generation of simple documentation at every CHEM module update. All model data-structures and algorithms are therefore always documented (at least minimally, with machine authored content), meaning model developers have a requirement to write customised documentation less frequently. Many of the features of our prototype software framework have the potential to be significantly enhanced through integration with evolving code and documentation authoring capabilities of large language models.

Our software framework provides the outline of a technical solution to some of the challenges of open source CHEM development. However, better tools alone are insufficient and policy and institutional changes are also required. Reducing waste in research is a core responsibility of research funders [72] and funding the development of CHEMs that are not adequately understood, reused or updated is wasteful. Existing incentive structures for health economists generally do not promote facilitating peers to reuse their work and funding is rarely adequate to maintain and routinely update a model. Currently, it can take “an extraordinary amount of idealism” to commit to authoring and maintaining research software [73]. Previously recommended strategies for more beneficial health economic research investments include support for harmonised ethical standards for model development [12], methodological innovation to improve model transferability [74], networks of modellers working on common health conditions [75], and centralized infrastructure such as open source model repositories [5] and a standard platform for model implementations [22]. Development of software frameworks to support ethical CHEM implementations could enable and enhance each of these strategies.

A future software framework for ethical CHEMs would ideally incorporate a base set of features useful to developers of computational models across all domains of public health and health economics, with the capability for community-led extensions that are tailored to the needs of modellers focused on specific health-conditions. Such approaches will likely be important to ensure a sufficient potential user-base to make CHEM software frameworks sustainable.

6 Conclusion

We have identified criteria that can be used to systematically assess the extent to which the computational implementation of health economic models adheres to transparency, reusability and updatability criteria. We have developed a prototype software framework that can support the implementation of CHEMs that largely meet these criteria. Our framework can be used as a prototype for developing future software frameworks to support ethical implementation of CHEMs. Barriers to ethical CHEM practice arising from resource constraints should be addressed by research funders.

Acknowledgement

The authors would like to acknowledge the contribution of John Gillam who provided advisory input to this research.

Availability of data and materials

The most up to date and comprehensive source of documentation on our framework and model is available at <https://www.ready4-dev.com> . Development versions of all code repositories referenced in this article are available in <https://github.com/ready4-dev/> . Archived code releases are available in <https://zenodo.org/communities/ready4> . All data repositories referenced in this article are available in <https://dataverse.harvard.edu/dataverse/ready4> .

Ethics approval

Software framework development did not involve human subject research and was not ethically reviewed. The utility mapping worked example is a previously reported study that was reviewed and granted approval by the University of Melbourne’s Human Research Ethics Committee, and the local Human Ethics and Advisory Group (1645367.1).

Funding

Software framework development was funded by Orygen, VicHealth, Victoria University and an Australian Government Research Training Program (RTP) Scholarship. The utility mapping study used as a worked example was funded by the National Health and Medical Research Council (NHMRC, APP1076940), Orygen and headspace.

Conflict of Interest

None declared.

References

1. Wagstaff A, Culyer AJ. Four decades of health economics through a bibliometric lens. *Journal of health economics*. Elsevier; 2012;31: 406–439.
2. Van Rossum G, Drake FL. *Python 3 reference manual*. Scotts Valley, CA: CreateSpace; 2009.
3. R Core Team. *R: A language and environment for statistical computing* [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2022. Available: <https://www.R-project.org/>
4. Incerti D, Thom H, Baio G, Jansen JP. R you still using excel? The advantages of modern software tools for health technology assessment. *Value in Health*. Elsevier; 2019;22: 575–579.
5. Pouwels X, Sampson CJ, Arnold RJG. Opportunities and barriers to the development and use of open source health economic models: A survey. *Value Health*. 2022;25: 473–479. doi:10.1016/j.jval.2021.10.001
6. Feenstra T, Corro-Ramos I, Hamerlijnck D, Voorn G van, Ghabri S. Four aspects affecting health economic decision models and their validation. *PharmacoEconomics*. 2022;40: 241–248. doi:10.1007/s40273-021-01110-w
7. Emerson J, Bacon R, Kent A, Neumann PJ, Cohen JT. Publication of decision model source code: Attitudes of health economics authors. *PharmacoEconomics*. 2019;37: 1409–1410. doi:10.1007/s40273-019-00796-3
8. Myllärniemi V, Kujala S, Raatikainen M, Sevoón P. Development as a journey: Factors supporting the adoption and use of software frameworks. *Journal of software engineering research and development*. SpringerOpen; 2018;6: 1–22.
9. Edwin NM. *Software frameworks, architectural and design patterns*. Journal of Software Engineering and Applications. Scientific Research Publishing; 2014;2014.
10. Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, et al. PyTorch: An imperative style, high-performance deep learning library. *Proceedings of the 33rd international conference on neural information processing systems*. Red Hook, NY, USA: Curran Associates Inc.; 2019.
11. Alarid-Escudero F, Krijkamp EM, Pechlivanoglou P, Jalal H, Kao S-YZ, Yang A, et al. A need for change! A coding framework for improving transparency in decision modeling. *Pharmacoeconomics*. Springer; 2019;37: 1329–1339.
12. Boden LA, McKendrick IJ. Model-based policymaking: A framework to promote ethical “good practice” in mathematical modeling for public health policymaking. *Frontiers in Public Health*. 2017;5. doi:10.3389/fpubh.2017.00068
13. Calder M, Craig C, Culley D, De Cani R, Donnelly CA, Douglas R, et al. Computational modelling for decision-making: Where, why, what, who and how. *Royal Society open science*. The Royal Society Publishing; 2018;5: 172096.
14. Ruparelia NB. *Software development lifecycle models*. ACM SIGSOFT Software Engineering Notes. ACM New York, NY, USA; 2010;35: 8–13.
15. Thompson E. *Escape from model land: How mathematical models can lead us astray and what we can do about it*. New Yourk: Basic Books; 2022.
16. Duckett S. *A journey towards a theology of health economics and healthcare funding*. Theology. SAGE Publications Sage UK: London, England; 2022;125: 326–334.
17. Harvard S, Werker GR, Silva DS. Social, ethical, and other value judgments in health economics modelling. *Social Science & Medicine*. 2020;253: 112975. doi:<https://doi.org/10.1016/j.socscimed.2020.112975>
18. Erdemir A, Mulugeta L, Ku JP, Drach A, Horner M, Morrison TM, et al. Credible practice of modeling and simulation in healthcare: Ten rules from a multidisciplinary perspective. *Journal of translational medicine*. 2020;18: 369. doi:10.1186/s12967-020-02540-4
19. Thompson EL, Smith LA. *Escape from model-land*. Economics. De Gruyter Open Access; 2019;13.

20. Carletto A, Zanzuzzi M, Sammarco A, Russo P. Quality of health economic evaluations submitted to the Italian Medicines Agency: Current state and future actions. *International Journal of Technology Assessment in Health Care*. Cambridge University Press; 2020;36: 560–568. doi:10.1017/S0266462320000641
21. Wonder M, Dunlop S. Assessment of the quality of the clinical evidence in submissions to the Australian Pharmaceutical Benefits Advisory Committee: Fit for purpose? *Value in Health*. 2015;18: 467–476. doi:https://doi.org/10.1016/j.jval.2015.02.011
22. Ghabri S, Stevenson M, Möller J, Caro JJ. Trusting the results of model-based economic analyses: Is there a pragmatic validation solution? *Pharmacoeconomics*. 2019;37: 1–6. doi:10.1007/s40273-018-0711-9
23. Kolovos S, Bosmans JE, Riper H, Chevreul K, Coupé VM, Tulder MW van. Model-based economic evaluation of treatments for depression: A systematic literature review. *Pharmacoeconomics-open*. Springer; 2017;1: 149–165.
24. Haji Ali Afzali H, Gray J, Karnon J. Model performance evaluation (validation and calibration) in model-based studies of therapeutic interventions for cardiovascular diseases: A review and suggested reporting framework. *Applied Health Economics and Health Policy*. Springer; 2013;11: 85–93.
25. Jalali MS, DiGennaro C, Guitart A, Lew K, Rahmandad H. Evolution and reproducibility of simulation modeling in epidemiology and health policy over half a century. *Epidemiologic Reviews*. 2021;43: 166–175. doi:10.1093/epirev/mxab006
26. McManus E, Turner D, Sach T. Can you repeat that? Exploring the definition of a successful model replication in health economics. *Pharmacoeconomics*. 2019;37: 1371–1381. doi:10.1007/s40273-019-00836-y
27. Bermejo I, Tappenden P, Youn J-H. Replicating health economic models: Firm foundations or a house of cards? *Pharmacoeconomics*. 2017;35: 1113–1121. doi:10.1007/s40273-017-0553-x
28. Radeva D, Hopkin G, Mossialos E, Borrill J, Osipenko L, Naci H. Assessment of technical errors and validation processes in economic models submitted by the company for NICE technology appraisals. *International Journal of Technology Assessment in Health Care*. 2020;36: 311–316. doi:10.1017/S0266462320000422
29. Arnold RJG, Ekins S. Time for cooperation in health economics among the modelling community. *Pharmacoeconomics*. 2010;28: 609–613. doi:10.2165/11537580-000000000-00000
30. Garcia-Mochon L, Rovira Forns J, Espin J. Cost transferability problems in economic evaluation as a framework for an European health care and social costs database. *Cost Effectiveness and Resource Allocation*. Springer; 2021;19: 1–12.
31. Sampson CJ, Wrightson T. Model registration: A call to action. *Pharmacoeconomics - Open*. 2017;1: 73–77. doi:10.1007/s41669-017-0019-2
32. Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, Wong JB. Model transparency and validation: A report of the ISPOR-SMDM modeling good research practices task force-7. *Medical Decision Making*. 2012;32: 733–743. doi:10.1177/0272989X12454579
33. European Organization For Nuclear Research, OpenAIRE. Zenodo [Internet]. CERN; 2013. doi:10.25495/7GXK-RD71
34. Quantitative Social Science I for. Dataverse [Internet]. Harvard University; 2007. Available: <https://dataverse.org>
35. Büyükkaramikli NC, Rutten-van Mölken MPMH, Severens JL, Al M. TECH-VER: A verification checklist to reduce errors in models and improve their credibility. *Pharmacoeconomics*. 2019;37: 1391–1408. doi:10.1007/s40273-019-00844-y
36. Eric Wong W, Debroy V, Choi B. A family of code coverage-based heuristics for effective fault localization. *Journal of Systems and Software*. 2010;83: 188–208. doi:https://doi.org/10.1016/j.jss.2009.09.037
37. Martin RC. Agile software development: Principles, patterns, and practices. Prentice Hall PTR; 2003.

38. github. GitHub [Internet]. 2007. Available: <https://github.com/>
39. Augustovski F, Iglesias C, Manca A, Drummond M, Rubinstein A, Marti SG. Barriers to generalizability of health economic evaluations in latin america and the caribbean region. *Pharmacoeconomics*. 2009;27: 919–29. doi:10.2165/11313670-000000000-00000
40. Wilson JAC, Greg AND Bryan. Good enough practices in scientific computing. *PLOS Computational Biology*. Public Library of Science; 2017;13: 1–20. doi:10.1371/journal.pcbi.1005510
41. Pan M, Gawthrop PJ, Cursons J, Crampin EJ. Modular assembly of dynamic models in systems biology. *PLoS computational biology*. Public Library of Science San Francisco, CA USA; 2021;17: e1009513.
42. Hourani H, Wasmi H, Alrawashdeh T. A code complexity model of object oriented programming (OOP). 2019 IEEE jordan international joint conference on electrical engineering and information technology (JEEIT). 2019. pp. 560–564. doi:10.1109/JEEIT.2019.8717448
43. Foundation TFS. What is copyleft? [Internet]. Available: <https://www.gnu.org/copyleft/>
44. Jenkins DA, Martin GP, Sperrin M, Riley RD, Debray TPA, Collins GS, et al. Continual updating and monitoring of clinical prediction models: Time for dynamic prediction systems? *Diagnostic and Prognostic Research*. 2021;5: 1. doi:10.1186/s41512-020-00090-3
45. Ye Y, Barapatre S, Davis MK, Elliston KO, Davatzikos C, Fedorov A, et al. Open-source software sustainability models: Initial white paper from the informatics technology for cancer research sustainability and industry partnership working group. *J Med Internet Res*. 2021;23: e20028. doi:10.2196/20028
46. Mergel I. Open collaboration in the public sector: The case of social coding on GitHub. *Government Information Quarterly*. 2015;32: 464–472. doi:<https://doi.org/10.1016/j.giq.2015.09.004>
47. Shahin M, Ali Babar M, Zhu L. Continuous integration, delivery and deployment: A systematic review on approaches, tools, challenges and practices. *IEEE Access*. 2017;5: 3909–3943. doi:10.1109/ACCESS.2017.2685629
48. Zhou J, Walker RJ. API deprecation: A retrospective analysis and detection method for code examples on the web. *Proceedings of the 2016 24th ACM SIGSOFT international symposium on foundations of software engineering*. 2016. pp. 266–277.
49. Barros C, Luo Y, Chubaty AM, Eddy IM, Micheletti T, Boisvenue C, et al. Empowering ecological modellers with a PERFICT workflow: Seamlessly linking data, parameterisation, prediction, validation and visualisation. *Methods in Ecology and Evolution*. Wiley Online Library; 2023;
50. Chambers JM. Object-Oriented Programming, Functional Programming and R. *Statistical Science*. Institute of Mathematical Statistics; 2014;29: 167–180. doi:10.1214/13-STS452
51. Statistical Computing RF for. The comprehensive r archive network [Internet]. 2022. Available: <https://cran.r-project.org>
52. Codecov [Internet]. Available: <https://about.codecov.io/>
53. Hugo: The world’s fastest framework for building websites [Internet]. Available: <https://gohugo.io>
54. Docsy [Internet]. Available: <https://www.docsy.dev>
55. Docsearch [Internet]. Available: <https://docsearch.algolia.com>
56. Netlify [Internet]. Available: <https://www.netlify.com>

57. Hamilton MP, Gao CX, Filia KM, Menssink JM, Sharmin S, Telford N, et al. Mapping psychological distress, depression and anxiety measures to AQoL-6D utility using data from a sample of young people presenting to primary mental health services. medRxiv. Cold Spring Harbor Laboratory Press; 2022; doi:10.1101/2021.07.07.21260129
58. Hamilton M, Gao C. youthvars: Describe and Validate Youth Mental Health Datasets [Internet]. Zenodo; 2022. doi:10.5281/zenodo.6084467
59. Hamilton M, Gao C. Scorz: Score questionnaire item responses [Internet]. Zenodo; 2022. doi:10.5281/zenodo.6084824
60. Hamilton M, Gao C. specific: Specify Candidate Models for Representing Mental Health Systems [Internet]. Zenodo; 2022. doi:10.5281/zenodo.6116701
61. Gao C, Hamilton M. TTU: Implement Transfer to Utility Mapping Algorithms [Internet]. Zenodo; 2022. doi:10.5281/zenodo.6130155
62. Hamilton MP, Gao CX. Youthu: Transform youth outcomes to health utility predictions [Internet]. Zenodo; 2022. doi:10.5281/zenodo.6210978
63. Hamilton MP, Gao CX, Filia KM, Menssink JM, Sharmin S, Telford N, et al. Transfer to AQoL-6D Utility Mapping Algorithms [Internet]. Harvard Dataverse; 2021. doi:10.7910/DVN/DKDIB0
64. Hamilton MP. Synthetic (fake) youth mental health datasets and data dictionaries [Internet]. Harvard Dataverse; 2021. doi:10.7910/DVN/HJXYKQ
65. Hamilton M, Gao C. Complete study program to reproduce all steps from data ingest through to results dissemination for a study to map mental health measures to AQoL-6D health utility [Internet]. Zenodo; 2022. doi:10.5281/zenodo.6212704
66. Hamilton M, Gao C. aqol6dmap_use: Apply AQoL-6D Utility Mapping Models To New Data [Internet]. Zenodo; 2022. doi:10.5281/zenodo.6416330
67. Hamilton MP. aqol6dmap_fakes: Generate fake input data for an AQoL-6D mapping study [Internet]. Zenodo; 2022. doi:10.5281/zenodo.6321821
68. Hamilton M. ttu_md1_ctlg: Generate a template utility mapping (transfer to utility) model catalogue [Internet]. Zenodo; 2022. doi:10.5281/zenodo.6116385
69. Hamilton MP. ready4-dev/ttu_lng_ss: Create a Draft Scientific Manuscript For A Utility Mapping Study [Internet]. Zenodo; 2022. doi:10.5281/zenodo.5976988
70. Kokol P, Blažun Vošner H, Kokol M, Završnik J. Role of agile in digital public health transformation. *Frontiers in Public Health*. 2022;10. doi:10.3389/fpubh.2022.899874
71. Beck K, Beedle M, Van Bennekum A, Cockburn A, Cunningham W, Fowler M, et al. Manifesto for agile software development. Snowbird, UT; 2001;
72. Chalmers I, Bracken MB, Djulbegovic B, Garattini S, Grant J, Gülmezoglu AM, et al. How to increase value and reduce waste when research priorities are set. *The Lancet*. Elsevier; 2014;383: 156–165.
73. Anzt H, Bach F, Druskat S, Löffler F, Loewe A, Renard BY, et al. An environment for sustainable research software in germany and beyond: Current state, open challenges, and call for action. *F1000Research*. Faculty of 1000 Ltd; 2020;9.
74. Craig P, Di Ruggiero E, Frolich KL, Mykhalovskiy E, White M, Campbell R, et al. Taking account of context in population health intervention research: Guidance for producers, users and funders of research. National Institute for Health Research; 2018;
75. Sampson CJ, Arnold R, Bryan S, Clarke P, Ekins S, Hatswell A, et al. Transparency in decision modelling: What, why, who and how? *PharmacoEconomics*. 2019;37: 1355–1369. doi:10.1007/s40273-019-00819-z