A prototype software framework for ethical implementation of computational economic models and its early application in youth mental health

Matthew P Hamilton1,2,✉, Caroline X Gao1,3,2, Glen Wiesner4, Kate M Filia1,3, Jana M Menssink1,3, Petra Plencnerova5, David Baker1,3, Patrick D McGorry1,3, Alexandra Parker6, Jonathan Karnon7, Sue M Cotton1,3, and Cathrine Mihalopoulos2

**Summary:**  Economists who use modeling techniques to help inform health policy make limited use of the modular and open source approaches that other disciplines use for collaborative development of complex models. We propose a framework for developing a modular open source economic model (MOSCEM) in youth mental health called ready4. The framework includes of set of seven standards for implementing a MOSCEM that is accountable (three standards), reusable (two standards) and updatable (two standards). We provide a rationale for each standard. The framework also includes a modelling toolkit of open access repositories and six R libraries for authoring MOSCEM modules, supplying those modules with data and implementing reproducible modelling analyses. We describe an early application of the framework to implementing a utility mapping study and detail how the MOSCEM components produced by that study meet 18 framework standards. We discuss how the framework will enable us to undertake and synthesise diverse economic modelling studies in youth mental health and highlight some broader implications for undertaking MOSCEMs to explore mental health and other topics.

**Code:**  Visit <https://www.ready4-dev.com> for more information about how to find, install and apply ready4.

1 Orygen, Parkville, Australia  
2 School of Public Health and Preventive Medicine, Monash University, Clayton, 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 Victoria University, Footscray, Australia  
7 Flinders University, Adelaide, Australia

✉ Correspondence: [Matthew P Hamilton <[matthew.hamilton@orygen.org.au](mailto:matthew.hamilton@orygen.org.au)>](mailto:matthew.hamilton@orygen.org.au)

# 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 [WAGSTAFF]. In seeking to solve these problems, health economists typically use models which are simplified and selective representations of systems that are believed to influence human health. These representations can be described in words and pictures (a *conceptual* model), in equations (a *mathematical* model) or in computer code (a *computational* model). The predictions reported in health economic studies are typically generated by the execution of a computer program which applies a computational model to compatible data inputs (e.g., parameter values) and performs a sequence of numeric calculations.

Computational models are now widely used to inform health policy and system design [1,2]. This level of influence has concomitant ethical responsibilities for model developers that are often poorly understood and inadequately fulfilled [<https://doi.org/10.3389/fpubh.2017.00068>, 10.1001/amajethics.2021.599., thompson2022escape, thompson2019escape].

We are developing a model to explore multiple economic questions relating to the mental health of young people aged 12 to 25. When making choices about how to implement this model computationally, we wished to facilitate ethical development and use. However, we are not aware of any software framework for implementing CHEMs that adhere to explicitly stated ethical requirements.

In this paper, we describe:

1. a set of ethical responsibilities for CHEM developers and criteria for assessing responsible CHEM implementations;
2. a prototype software framework for the ethical implementation of CHEMs; and
3. use of the software framework to develop a computational economic model in youth mental health, with an initial focus on outcome valuation.

# Ethical requirements and assessment criteria

We considered prior literature on modelling practice, our own professional experience and the needs of our project, to identify three core ethical responsibilities of CHEM developers, three attributes of CHEMs that enable fulfilment of these responsibilities and six criteria (two for each model attribute) for assessing ethical CHEM implementations.

## Ethical responsibilities of CHEM developers

We believe that health economists have ethical responsibilities relating to the social acceptability, adequacy for purpose and public benefit of how their computational models are developed and used.

Misalignment between the values of model developers and those of the population groups affected by decisions based on their models presents significant ethical risks [thompson2022escape, thompson2019escape. The value judgments of model developers influence the assumptions, selection of model features and standards for evidence that shape the CHEM development process [[https://doi.org/10.1016/j.socscimed.2020.112975](https://doi.org/10.1016/j.socscimed.2020.112975" \t "_blank" \o "Persistent link using digital object identifier)]. These value judgments are rarely made explicit, omissions that may lead to socially unacceptable policy recommendations [9]. For example, to reduce the risk of inequitable policy implementation, it may be important for of a model to predict the benefits of harms of an intervention for different subpopulations [<https://doi.org/10.3389/fpubh.2017.00068>], but model developers may prefer to allocate a project budget to other priorities [thompson2022escape, thompson2019escape].

Health economists have duties both to take sufficient care that a CHEM is adequate for the explicit purpose for which it was developed and to provide potential third party users with the means of assessing its adequacy for their proposed purposes [[10.1186/s12967-020-02540-4](https://doi.org/10.1186/s12967-020-02540-4), https://doi.org/10.1007/s40273-021-01110-w, <https://doi.org/10.3389/fpubh.2017.00068>, thompson2019escape]. Currently, it is common for CHEMs to have poor reproducibility [4–6], insufficient validation [7] and undeclared errors [8], which make achieving these goals more difficult.

Even an acceptable and adequate CHEM will have limited public benefit if it not much used or when its acceptability and adequacy rapidly decay. Model reuse advances the scientific goals of generalisability (application without adaptation) and transferability (selective reuse and/or modification of model components) [https://link.springer.com/article/10.2165/11313670-000000000-00000 and can facilitate more efficient model implementations [14]. Common barriers to model re-use include commercial and legal considerations [13], as well as challenges related to model transferability across jurisdictions [GARCIA-MOCHON – Check still valid for rephrase]. The temporal window for valid application of CHEMs is often limited by implementation choices that rarely facilitate routine updates [18]. Without ongoing maintenance, a model may become less reliable with time, deterioration that model users may be unaware of, and has a growing risk of being deployed for purposes for which it is poorly suited [[10.1098/rsos.172096](https://doi.org/10.1098%2Frsos.172096" \t "_blank)].

## Criteria for assessing ethical CHEM implementation

The acceptability, adequacy and public benefit responsibilities for CHEM developers are easier to state than to measure. It may therefore be pragmatic for assessment criteria for ethical modelling practice to instead be based on measurable attributes of the models themselves.

As described in Table 1 [AUTOMATE\_REF], we believe that implementing CHEMs that are transparent, reusable and updatable (TRU) can enable modellers to meet their ethical obligations. We therefore selected these model attributes to use as the basis for deriving assessment criteria.

Transparency has been recommended as a core criterion for assessing ethical public health modelling practice [<https://doi.org/10.3389/fpubh.2017.00068>]. Guidance on transparency in health economic modelling recommended that model code and data should be clearly documented, potentially with different versions for technical and non-technical users [12]. Notably, the same guidelines, published over ten years ago, did not include recommendations on sharing model code and data. However, more recent guidance recommends publicly dissemination of healthcare model artefacts using online repository services [2]. Repositories such as Zenodo [37] and Dataverse [38] provide persistent storage solutions that generate a Digital Object Identifier (DOI) for each code and data collection. An essential component of quality assuring health economic models is verification - ensuring that calculations are correct and consistent with model specifications [44]. The extensiveness of verification checks can be reported using the concept of code coverage [45] - 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 configured amd run to perform tasks that meet core user-requirements [Martin\_2003). 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 [10]. This issue can be addressed by use of online repository services such as GitHub [36], that provide citation tools and can transparently record all individual code contributions to a modelling project over its lifecycle.

Assessment criteria for a transparent CHEM:

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

Making a CHEM’s code, data and documentation publicly available is helpful but insufficient for promoting model re-use. The choices that CHEM developers make about model implementation and licensing will also shape who can use a model and for what purposes. Using open-source development platforms and licenses can aid both generalizability and transferability. Compared to using commercial modelling software, authoring CHEMs in an open-source language like R [@RCORE2022] makes it easier to store model algorithms and data in distinct files and locations (as opposed to hard coding - embedding data such as parameter values into source code) which facilitates selective modification of model components. This benefit can be further enhanced if model developers adopt a modular approach, in which a model is constructed from multiple reusable and replaceable sub-models (modules) [19]. To grant permissions to others to use, test and adapt models and their components, health economists can avail of two broad categories of open source licensing options. Some guidance strongly recommends the use of permissive licensing [39] that provides users with great flexibility as to the purposes (including commercial) for which content can be re-used. An alternative approach is to use copyleft licenses [46] that can require content users to distribute any derivative works they create under similar open-source arrangements.

Assessment criteria for a reusable CHEM:

* R1: Model code and data are implemented to facilitate both generalizability and transferability.
* R2: Terms of use allow anyone to reuse model code and non-confidential data, in whole or in part, without charge, and for purposes that include the creation of derivative works.

To remain valid for longer, models should be continually updated and refined as new evidence emerges and healthcare systems evolve [17]. Ensuring that a model is regular reviewed to identify and implement required improvements is a receommended defence against model validity decay [[10.1098/rsos.172096](https://doi.org/10.1098%2Frsos.172096" \t "_blank)]. Key enablers of sustainable maintenance of open source research software are committed, adequately resourced core development team and active user community [https://www.jmir.org/2021/12/e20028/pdf]. Currently, the core development team for a CHEM will be typically be funded to produce a project end-point deliverable whose specifications are well defined early in the project. For more complex and multi-purpose CHEMs, particularly those designed to be incorporated into decision support systems, it may be better for development teams to adopt Agile Software Development, an approach that has been recommended for complex public health software projects [<https://doi.org/10.3389/fpubh.2022.899874>]. An Agile model will be less clearly specified in the initial project plan, but will instead continually develop in response to the requirements and feedback of users, who are provided with an initial, simplified working version of the model at the earliest feasible opportunity. 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 [REF] provide collaborative code development tools [[https://doi.org/10.1016/j.giq.2015.09.004](https://doi.org/10.1016/j.giq.2015.09.004" \t "_blank" \o "Persistent link using digital object identifier)] that help elicit, integrate and reconcile contributions from multiple contributors and to ensure each update is uniquely identifiable and retrievable. It is important that verification checks are rerun with each model update, a task that can be automated using the software development practice of Continuous Integration [54]. The risk of model revisions having unintended consequences for third party users can be mitigated through the use of deprecation conventions [REF] that take an informative and staged approach to retiring outdated model code and data.

Assessment criteria for an updatable CHEM:

* U1: Resources and infrastructure are in place to support sustained development, testing, maintenance and version control of a model in collaboration with model users.
* U2: Each new release of a model is retested, with changes implemented to minimize disruptions for existing model users.

# Software framework

A software framework is a shared common technology used by software developers to collaboratively author software applications and which is not typically visible to application end-users [<https://jserd.springeropen.com/articles/10.1186/s40411-018-0050-8>]. Advantages of using software frameworks include facilitating code reuse and extension, promoting good programming practice and the capability to provide enhanced functionality and performace without additional effort by developers [[DOI:10.4236/jsea.2014.78061](http://www.scirp.org/journal/PaperInformation.aspx?PaperID=47999)]. However, software frameworks can be challenging and time consuming to create [[DOI:10.4236/jsea.2014.78061](http://www.scirp.org/journal/PaperInformation.aspx?PaperID=47999)] and then difficult for others to learn, often requiring developers to undergo specialist training [<https://jserd.springeropen.com/articles/10.1186/s40411-018-0050-8>]. There is also a risk that a software framework may become excessively complex over time [[DOI:10.4236/jsea.2014.78061](http://www.scirp.org/journal/PaperInformation.aspx?PaperID=47999)].

To support the collaborative development of CHEMs that meet TRU assessment criteria, we have created a prototype software framework called ready4. We have designed this framework as a lightweight extension to existing programming infrastructure with the goal of supporting the development of CHEMs that are both open-source (for transparency and reusability) and modular (for reusability and updatability).

## 3.1 Framework libraries

To work within the popular open-source programming environment R, the ready4 software framework is implemented as six development version R code libraries. The R libraries collectively provide model developers with tools for authoring CHEMs, supplying those CHEMs with data and using CHEMs to implement reproducible modelling analyses. The six novel libraries and the preexisting third-party R libraries they depend on are summarised in Table **[2](#cpkgs)** [AUTOMATIC\_REF].

Standardization and interoperability are core requirements of implementing a modular approach. Model modules need to be able to share inputs and outputs with each other and to be run as independent models [20]. To achieve this goal, we adopt an object oriented paradigm in which each CHEM module includes both a data structure (specifying the required properties of data that can validly be supplied to a module) and a set of algorithms (specifying the operations that can be performed on data contained in a module instance).

The foundational framework,,data structure CHEM data structures be created and that enable module to be consistently namedThe ready4 library also contains tools for retrieving web based information on model modules, datasets and analysis programs authored with the framework and for partially automating updates to a project documentation website.

Three R libraries are designed to help standardize workflows for authoring, documenting, testing and disseminating new model modules. The ready4pack library is designed to integrate with GitHub and provides tools for authoring model modules and desseminating them as themed bundles in R libraries libraries that are:

* documented (with a website, a PDF manual itemising selected contents and a PDF manual itemising all contents);
* 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 unit tests created by module library authors).

The ready4pack library depends on two other module authoring libraries. Writing model algorithms as collections of functions (short, self-contained and reusable software routines that each perform a discrete task), has been recommended as good practice for scientific computing [39]. The ready4fun library contains tools for authoring functions in a consistent house style that automatically generates basic documentation for each function. Functions to implement model algorithms can be associated with a module via a special type of function called a method. Tools from the ready4class [68] library can help streamline and standardise the authoring of module data structures and their associated methods and to automatically generate basic documentation for each module.

The ready4use library [69] contains tools for supplying model modules with data stored in online repositories (hosted on a Dataverse installation or on GitHub), labelling these datasets and then sharing them via online repositories. The ready4show library [70] contains tools to help author R Markdown programs that combine model modules and datasets to undertake analyses. These programs are either self-documenting (code is easy to understand and integrated with plain English explanations of what it does) or trigger the creation of separate documents (e.g. a scientific manuscript).

## Framework integration with online services

To be used in the development of CHEMs, our software framework requires the establishment and configuration of a number of online services (see Availability of data and materials).

We created a GitHub organisation (a collection of code repositories) where all code (libraries, programs and sub-routines) that we author with the framework is stored and version controlled. We configured individual repositories in our GitHub organisation that are used for code library projects to use GitHub actions 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) [56]. To track our code coverage, we linked our GitHub organisation to an account we established at codecov [57]. To facilitate the creation and hosting of documentation websites, we enabled GitHub Pages in each repository we 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.

# Application

Our main motivation for developing the software framework was to enable us to implement TRU computational models that address multiple economic topics in youth mental health and which we can combine as modules of an overarching CHEM (also called “ready4”) that progressively extends in scope and functionality.

## Economic topics

Currently, wusing thesoftware framework develop, apply and share youth mental health computational models in

## Case study: Utility mapping

Our initial application of the ready4 software framework was to undertake a previously described study [75] to develop utility mapping models appropriate for use in samples of young people presenting to primary mental health services. The ready4 software framework was used 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 [76], scoring health utility [77], specifying utility mapping models [78] and implementing reproducible utility mapping studies [79];
* a development version library of functions for finding and using utility mapping models developed with these tools [80];
* data collections of synthetic populations for testing model modules [81] and study input and results data [82];
* programs for replicating all steps from data ingest to manuscript reporting [83], applying utility mapping models to new data [84] and generating a synthetic representation of the study dataset [85];
* subroutines for creating a catalogue of utility mapping models [86] and generating a draft scientific manuscript [87] for studies implemented with these modules.

We created a checklist (Table 3) 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 the outputs from our utility mapping study may be assessable as having satisfactorily met four of the six criteria (T1, T2, R2 and U1) and to have partially met two criteria (R1 and U2). The main shortcomings that we identified when applying the assessment criteria were the need for additional development before the CHEM modules would be sufficiently generalizable for valid application in datasets that measure health utility with different instruments and the general lack of unit testing.

## Model documentation

We developed a versioned model documentation website ([www.ready4-dev.com](http://www.ready4-dev.com)) that provides guidance to model developers on how to use and contribute improvements to the ready4 software framework and model. The documentation website was developed using the Hugo framework [59], Docsy theme [60] and Algolia DocSearch [https://docsearch.algolia.com] and is hosted using the Netlify [61] service. We used functions from the ready4 R 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 the source code in its GitHub repository was edited.

# Discussion

Ethical practice is a core expectation of health researchers and computational methods underpin most quantitive research, yet an understanding of what constitutes ethical computational modelling practice in health is underdeveloped [<https://doi.org/10.3389/fpubh.2017.00068>]. The modeller responsibilities, enabling model attributes and model implementation assessment criteria that we propose can help address this gap. We have created a prototype software framework to demonstrate how ethical CHEM implementations can be operationalized. We have used this framework to develop initial, preliminary modules of a TRU economic model in youth mental health.

The ethical responsibilities and enabling model attributes we describe have both commonalities and distinctive features compared to a previous ethical framework for computational modelling in public health [<https://doi.org/10.3389/fpubh.2017.00068>]. The authors of that framework propose 13 questions to evaluate ethical risk across the four criteria of independence, transparency, benficience and justice. Their descriptions of the four criteria at least partially map to either our proposed modeler responsibilities (“justice” to “social acceptability”, “independence” to “adequacy for purpose” and “beneficience” to “public benefit”) or enabling model attributes (“transparency”). However, while our six assessment criteria are specific to three attributes (TRU) of the computational implementation of the model, the prior ethical framework includes questions relevant to the the conceptual and mathematical models and the potential impacts of model use. Examples of these more general evaluation questions include (for the justice criterion) “is any lack of knowledge about important parameters attributable to uncertatiny or variability?” and (for the beneficience criterion) “if a policy is based on the model evidence, is it more likely to be effective and beneficial than a decision made in the absence of the model?”. The less numerous and more focused assessment criteria we propose may potentially be more tractable to implement in reviews of models authored by third parties and as the basis for designing software frameworks to support ethical computational model implementation.

We believe that our software framework is likely to be adequate for the narrow purpose of implementing the ready4 model, but too fragile to be anything other than a prototype for a framework that could satisfactorily meet the needs of other modelling teams and projects. A major reason for this distinction is that our software framework was developed with the needs of only one group of developers in mind – ourselves. We began developing a prior version of our software framework in December 2018 to try to automate and standardize a number of tasks we were repeadedly performing when authoring the code to implement our youth mental health model. However as the amount of code we wrote grew, our initial framework proved too cumbersome to use. We implemented a major overhaul, making early versions of the current framework source code public in August 2020 and continued to extend and refine the framework as we identified issues from applying it to the development and use of ready4 model modules. As we illustrated by Table 3, our framework is now capable of being used by us to largely satisfy our TRU criteria.

Commiting to a software framework, particularly a project with modular implementation, many contributors and multi-annual planning horizon, is a major decision. A software framework to support ethical CHEM implementations would only merit adoption if it is both sufficiently robust (it has been developed and tested for purposes closely aligned with project goals) and sustainable (supported by an adequately staffed core team and active user community). We believe that projects to develop and promote adoption of such software frameworks, both in R and other open source languages such as python, would be a significant step forward. A software framework for ethical CHEMs would ideally incorporate a base set of features useful to developers of health economics across all health conditions, with the capability for community-led extensions that are tailored to the needs of modellers focused on specific health-conditions.

Our prototype framework has a number of features that subsequent work may find useful to incorporate. Firstly, developing a software framework to work within an existing and widely used open source programming language, can keep framework scope relatively narrow (making it more tractable to develop, maintain and learn) while readily leveraging and coherently integrate other modelling tools written in that language (e.g. the dependency libraries we list in Table 2). Secondly, implementation that combines both object oriented and functional programming paradigms can avail of the modular and syntactical simplicity benefits of the former, while reducing the limitation that object oriented approaches bundle artefacts (most famously described as: **“you wanted a banana but what you got was a gorilla holding the banana”). Thirdly agile and documentation…**

Such a project would be highly aligned with the goals of some research funders, who have have been encouraged to support methodological innovation to improve model transferability [100] and for who reducing waste in research is a core responsibility [92].

Research into health economic and public health modelling practice suggest that many if not most CHEMs will fall short of meeting the transparency [4–8], reusability [13,25,26] and updatability [Sampson\_2017, <https://doi.org/10.3389/fpubh.2022.899874>]] criteria we propose.

The strong in-principle support from health economists for open-source health economic models suggest that these shortfalls in ethical practice are not caused by lack of interest

reproducibility [4–6], and uncertain validity [7,8], limited reusability [13,25,26] of health economic models

Addressing this issue will require technical innovation, skills development and changes to the way research proposals and research careers are assessed.

Currently, it takes “an extraordinary amount of idealism” to dedicate the substantial time and resources required to author, test, document and maintain even fragile prototype research software that could instead be used to write scientific publications [https://doi.org/10.12688/f1000research.23224.2 ].

The use of shared software frameworks may help reduce the burden associated with implementing some aspects of ethical modelling practice. Collaboration between teams of health economists can also make some complex modelling projects more feasible [14]. Developing networks of modellers working on common health conditions has been recommended as a strategy for improving model validity [28] and some of us are part of a nascent initiate of this type in mental health [94].

However, we think the current low rates of adoption by health economists of open source approaches [13,25,26] that facilitate collaboration will only change slowly unless there is significant and strategic investments made by research funders. Currently, incentive structures for health economists do not promote the dedication of large quantities of time to enable peers to reuse their work.

Ideally health economists would explore these complex topics in partnership with modellers from other disciplines (in particular epidemiology and health services research) and a wide range of stakeholders such as other researchers, policymakers, service planners and community members. Modular and open source approaches would facilitate such investigations by breaking down ambitious and long term goals into manageable time-bound discrete projects, each progressed by different teams. To facilitate such an approach, a common framework of standards and tools would be needed. To be suitable for such a task our framework would need additional development, with its overall architecture reviewed for scalability and suitability and to provide better integration with and use of other open source languages (particularly python) and repositories. Whatever MOSCEM infrastructure is developed, its resilience would depend on a community of open source contributors sufficiently large and active to ensure that all core modules are maintained even after their original authors cease their involvement.

However, having features that facilitate accountability, reuse and updating is not the same as being accountable, reused and updated. If diverse groups of stakeholders do not review model components, suggest improvements and develop alternatives, then little progress is made towards enhancing model legitimacy. Similarly, making code and data publicly available does not guarantee that others will know of the existence of these tools, trust their validity and find them easy to use. Without reuse, errors in model artefacts are more likely to remain undetected. Even when errors are detected, they still need to be fixed, but maintaining code and data requires ongoing resourcing through a combination of centralised infrastructure and an active open source community.

To progress from a technical capability to behavioural outcomes, both our framework and MOSCEM need further work. Currently all the framework and model module libraries we have developed are available only as “development” releases. An early priority for us is to undertake the additional development, testing and documenting of these libraries so that we can submit production versions of each library to CRAN [56]. Making an R library available on CRAN is normally a prerequisite for a high level of use.

Some of the issues we have discussed in the context of the development of our model or health economic modelling in mental health are potentially relevant to health economists and funders of health economic research more generally.

Proactive measures by funders to encourage more accountable, reusable and updatable health economic models is not a need confined to mental health.

However, funders also need credible proposals to support and this is an area for health economists interested in MOSCEMs to prioritize. Health economists could use existing and new special interest groups to identify opportunities and enablers of more collaborative approaches to model development, potentially as the basis for future funding proposals.

Adopting MOSCEMs will expand the type of skillset typically engaged in health economic modelling projects, with a much greater role for data-scientists, software engineers and online community builders. The requirement for these roles should be incorporated into project proposals. Not all efforts by health economists to promote MOSCEMs need to depend on the decisions of research funders. Releasing selected subsets of unmaintained model artefacts in open source repositories is still better than not providing access to any code and data and can typically be accomplished within existing project budgets. Developing knowledge and skills of MOSCEMS can be advanced by making small contributions (e.g. improvements to documentation, code contributions) to open source projects. Our project website [58] includes details of multiple ways to contribute to ready4.

and the significant deficits in our understanding of the systems in which mental disorders emerge and are treated [95] suggest that there are a number of candidate topics in mental health that might benefit from pooling of efforts. The weak theoretical underpinnings for understanding complex mental health systems [96] may be a place to start. It remains unclear why increased investments in mental health care have yet to discernibly reduce the prevalence and burden of mental disorders[97]. The literature, and evidence base, regarding how the requirements, characteristics and performance of mental health services are shaped by spatiotemporal context needs to be further developed [98]. There is also a need for better evidence to identify the social determinants of mental disorders most amenable to preventative interventions, and for which population sub-groups such interventions would be most effective [99].

## Implications for implementing ready4

The most direct implication of the development of the ready4 framework is that it makes it feasible for us to implement a MOSCEM in youth mental health. The standards specified by the framework have enabled us to partially automate workflows for developing and applying ready4 through use of the framework’s modelling toolkit. We have demonstrated the practical utility of the modelling toolkit by applying it to authoring, documenting and disseminating ready4 module libraries [76–79], datasets [81,82]; analyses [83–85], reporting templates [86,87] and prediction tools [80]] used in a utility mapping study [75]. The standardised and partially automated workflows used in creating and sharing these artefacts has the potential to generate significant efficiencies as we apply the ready4 framework to undertaking new economic studies.

We have also been able to demonstrate the interoperability of the initial ready4 modules developed with the modelling toolkit. The program used to implement the utility mapping analysis [83] combines modules from four module libraries ([76–79]) and two framework libraries [69,70]. Example literate programs published on the ready4 documentation website [58] use toy data [81] to illustrate the potential for ready4 modules to facilitate study replication and transferability. As demonstrated by the checklist we developed (Table 3), our framework’s standards also provide a mechanism for to assess the extent to which the ready4 MOSCEM meets explicit objectives.

The transferability claims we make for our existing modules are to date supported only by example programs using toy data. Our future work aims to address this with real world studies that apply modules to different concepts and contexts. Our current work program also aims to create new ready4 modules for modelling help-seeking choice, spatial epidemiology, household populations and primary mental health services that we hope will provide others with more reasons to use ready4 and contribute to its development. To facilitate code contributions by third parties, our libraries for authoring modules [67] require some additional development to make them easier to use by third parties without knowledge of the naming and directory structure conventions we use in authoring code.

## Implications for economic modelling in mental health

Open source approaches have been recommended to help develop the mental health modelling field [88] but only one mental health related model (in Alcohol Use Disorder [89]) is currently indexed in the Open Source Models Clearinghouse [25,90]. We are aware of just one other open source mental health model - a reference model in Major Depressive Disorder - that is currently in development [91].

Approximately 4,000 mental health focused economic evaluation reports were produced between 2000 to 2019 [34]. The intellectual asset represented by this literature could be enhanced if many of the models described in these reports could be brought and kept up to date and made available in formats that maximised transferability to diverse decision contexts. We believe that modular and open source approaches would be well suited to accomplishing this goal and that the framework we have developed could act as an early prototype for solving some of the technical challenges of this task. Ideally such a program of research would be resourced to be sustained over the medium to long term and to engage a diverse network of investigators, contributors and advisers from high, middle and low income countries.

In addition to extracting more value from the existing health economic knowledge base in mental health, there is an opportunity for research funders to shape how future health economic models in mental health are undertaken. Mental health topics accounted for 268 of the 2829 (10%) peer reviewed economic evaluations undertaken during 26 month period in 2012-2014 identified by a review [93]. The ongoing annual output of economic research in mental health that focuses on the other 11 domains identified by Wagstaff and Culyer [35] is probably also substantial. Funders should provide support for the projects and infrastructure to promote greater collaboration, interoperability, transferability and maintenance of future mental health modelling projects.

## General issues for health economists and health research funders

Of the known barriers to adoption of open source models by health economists (including issues like intellectual property and confidentiality [15,27]), our experience suggests that the biggest challenges may be the enormous effort required to first prepare model code and data for public release in formats that facilitate appropriate reuse by third parties and to then maintain and continually improve potentially large numbers of digital artefacts.

# Conclusion

We have developed a framework for undertaking a MOSCEM in youth mental health and demonstrated its use by applying it to undertake a utility mapping study. We intend using this framework to undertake and synthesise multiple types of economic research. With further development the framework could be applied to extracting greater value from existing economic models in mental health and to facilitate collaborations between health economists and other stakeholders to address complex mental health modelling challenges. Although MOSCEMs provide a promising opportunity to advance the health economcis field, action from funders is required to realise this potential.

Software f

Software f

**Table 1: How transparent, reusable and updatable models can promote ethical practice.**

|  |  |  |  |
| --- | --- | --- | --- |
| **Model attribute** | **Promotes:** | | |
|  | **Acceptability** | **Adequacy for purpose** | **Public benefit** |
| **Transparent** | Enables assessment. | | Reduces risk of inappropriate use. |
| **Reusable** | Allows adaptation by users with different value judgments. | Facilitates independent testing by third parties. | Enables application to different decision contexts and reduces duplication of modeler effort. |
| **Updatable** | Maintains validity. | Extends lifetime. |

**Table 2** **:** ready4 framework R libraries

| Package | Focus | Depends on these R libraries |
| --- | --- | --- |
| ready4 | Foundation | 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 |
| ready4fun | Module algorithms | desc devtools dplyr generics 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 |
| ready4class | Module structures | devtools dplyr fs gtools Hmisc knitr lifecycle magrittr methods purrr ready4 ready4fun ready4show rlang stats stringi stringr testit testthat tibble tidyr usethis utils |
| ready4pack | Module libraries | dataverse dplyr knitr lifecycle magrittr methods purrr ready4 ready4class ready4fun rlang stringr testthat tibble tidyr utils |
| ready4use | Datasets | data.table dataverse dplyr fs Hmisc knitr lifecycle magrittr methods piggyback purrr readxl ready4 ready4show rlang stats stringi stringr testit testthat tibble tidyr utils |
| ready4show | Analyses | 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 |

**Table 3: Assessment of utility mapping ready4 CHEM modules against TRU criteria.**

|  |  |  |
| --- | --- | --- |
| Criteria | Met? | Detail |
| T1 Open access | Yes | All source code and testing procedures are available in public GitHub repositories, with each code release persistently available on Zenodo.  Study dataset is not publicly available as it contains confidential patient health data. A synthetic representation of the study dataset and the data files to apply models to out of sample data are persistently available on the Harvard Dataverse. |
| T2 Authors & beliefs | Yes | All code libraries, programs, sub-routines and datasets are distributed with tools for appropriate citation.  Code development on GitHub means author contributions over time are visible.  Model catalogues are persistently available on the Harvard Dataverse describe how predictive performance of models varies under multiple usage regimes.  Each code library is documented with worked examples of how to apply modules.  Analysis and reporting programs are self-documenting.  Sub-routines for generating reports are documented with README files. |
| R1 Generalisability & transferability | Partial | Model code is written using both functional and object-oriented paradigms.  Code library websites include hypotethical examples that apply modules for a replication study (same predictors and outcomes, but with different variable names) and generalization to develop models using both different predictors and outcomes measured with a different utility instrument. The models in the latter case perform poorly and a combination of better synthetic data for the generalized examples and additional development to facilitate generalized applications of study algorithms is probably required. |
| R2 Terms | Yes | All code is distributed using GPL-3 licenses.  Datasets use amended version of template provided by Harvard Dataverse, allowing reuse of data subject to some ethical restrictions (e.g., use in efforts to re-dentify study participants is prohibited) |
| U1 Infrastructure | 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 Retesting & Deprecation | Partial | Continuous integration 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. |

In this article we described a framework that we developed to help us implement ready4 - a MOSCEM in youth mental health. We outlined framework standards for an accountable, reusable and updatable MOSCEM and described the modelling toolkit we created for applying those standards to the development and use of the ready4 MOSCEM. We also provided an overview of an initial set of MOSCEM modules developed with the framework to implement a utility mapping study. We reviewed the modules, datasets and analyses generated by that study against framework standards. The work we have described has potential implications for the development of the ready4 MOSCEM and for health economic modelling in mental health. A number of issues have more general relevance to health economic modellers and funders of health economic research.

Consistent use of meaningful naming conventions when authoring code is recommended [39,40]. Code can be made easier to follow by using the practices of abstraction [41], where only simple, high level commands are routinely exposed to reviewers, and polymorphism [42], where the same command (e.g. “simulate”) can be reused to implement different algorithms of the same type. Programs to implement model analyses can be made comprehensible to even non-technical users through the use of literate programming techniques and tools like RMarkdown [43] that integrate computer code with plain English descriptions.

Transcription errors - mistakes introduced when transferring data between sources, models and reports - are very common in health economic models [8]. The risk of these errors might be lower if there was full transparency across all steps in a study workflow. Scientific computing tools now make it relatively straightforward to author programs that reproducibly execute all steps in data ingest, processing and reporting [39].

Code and data should be distributed with tools that make it easy for potential users to appropriately cite each model artefact.

To make model code and data widely re-usable by others, it is important to provide users with appropriate and explicit permissions. For code, it may be appropriate to adopt the prevailing open source licensing practice within the programming language being used. For data, it may not be sufficient to simply choose between a permissive license like the Public Domain Dedication (CC0) [47] or a copyleft option such as the Attribution-Share Alike (CC-BY-SA) [48]. In addition to ensuring that data is ethically appropriate for disseminate in open access repositories, responsible custodianship of some de-identified or aggregated data may involve using or adapting template terms of use [49] which have a number of ethical clauses (for example, prohibiting efforts to re-identify research participants).

Clear distinctions should be made between model modules (code that defines abstract data structures and the algorithms that can be applied to data described by these structures), model datasets (digital information such as parameter values, unit records, etc) and model analyses (code that links model datasets to model modules and specifies the algorithms to apply to data associated with each module).

The software development practice of encapsulation [41] can be used to help ensure that model modules continue to work as intended when they are combined [50]. In some cases, combining modules may mean new versions of modules have to be created to better account for interaction effects. The concept of inheritance [41] can be used to write code that efficiently achieves this objective as well as to facilitate selective editing of modules when transferring models to different decision contexts [50]. Model modules of a similar type or purpose can be efficiently distributed and documented by bundling them as code libraries. It is good practice to make available test or toy data to demonstrate the use of model algorithms [39].

Statistical models are a common output of health economic evaluations, but they are often not reported in a format that enables others to confidently and reliably re-use them for out of sample prediction [51]. Open source approaches can help address this by disseminating code artefacts that enable easy and appropriate use of a statistical model to make predictions with new data. However, great care must be exercised when publicly releasing model artefacts derived from data on human subjects as they may by default embed a copy of the source dataset. Sensitive dataset copies must therefore be replaced (for example, with synthetic data) and the amended artefact’s predictive performance then retested before any public release. Another way to make MOSCEMs easier to use is to develop simple user-interfaces for non-technical users.

Issues such as privacy and confidentiality can limit public release of some sensitive health economic model artefacts [15].

selectively restrict access to data that are confidential, while disseminating all other model artefacts. Separating code and data will also make it easier

Model authors may wish to facilitate reuse in both contexts to which their model can be

Such flexibility is useful when transferring a health economic model developed for one jurisdiction for application in another, as this task typically involves retaining some model features and updating others [16].

When used in conjunction with toolkit repositories, the six R libraries provide support for implementing 17 out of 20 framework standards (Table **[1](#timelygls)**). Standards relating to safe dissemination of statistical models (R8), user-interface development (R9) and deprecation conventions (U4) are better met through using existing third party R libraries. Preparing statistical models for dissemination can be accomplished with standard R data management tools like the dplyr [71] and purrr [71] libraries. User-interfaces are typically developed with the shiny [72] library, for which a tutorial aimed at health economists is available [73]. The library lifecycle [74] provides tools for R developers to consistently deprecate their code.

## Paradigm

Barriers to health economists adopting open source approaches include concerns about intellectual property, confidentiality, model misuse and the resources required to support open source implementations [15,27].

Our interest in modular and open source approaches developed when we began seeking an appropriate framework for undertaking and validly synthesising diverse types of economic research in mental health. Mental disorders impose high health, social and economic burdens worldwide [29,30]. Much of this burden is potentially avertable [31], but poorly financed and organised mental health systems are ill-equipped for this challenge [32,33]. A substantial economic literature already exists to assess the affordability and value for money of mental health interventions [34]. This economic evaluation work is an essential prerequisite for improving allocative efficiency in mental health, but could be of greater value to systems planners if integrated with a broader program of economic research.

However, there now appears to be strong in principle support from many health economists for greater use of open-source CHEMs [15]. However, open-source CHEMS remain relatively rare [13,25,26] and better guidance for how to implement CHEMs as open-source projects is needed [28].

Funding for health economic modelling projects rarely extend to provision of medium term support for model updates and improvements. The career trajectories of health economists can also mitigate against adequate maintenance of a model. For example, it is relatively common for model authors to have moved on from the team that owns the model and / or from working on the health condition for which the model was developed.

Computational health economic models (CHEMs) are widely used, influential, increasingly complex and subject to potentially under-appreciated limitations.

There is significant scope for the acceptability, adequacy for purpose and social benefit of computational health economic models (CHEMs) to be enhanced.

## 2.2 Desireable CHEM attribues

These issues have the potential to compound as CHEMs become more complex – with concomitant accountability obligations for model authors [12,13].

An existing ethical framework for public health modelling [<https://doi.org/10.3389/fpubh.2017.00068>] suggests four criteria for considering the suitability of models to inform policymaking – independence, transparency, beneficence and justice.

The growth in the volume and breadth of published health economic analysis [REF] suggests that substantial public funds are now invested in developing CHEMs. The social returns from this investment could be enhanced if CHEMs could be more readily and appropriately used by all who could benefit from them and if the lifetime for their valid application could be extended.

Not in – representativeness & engagement [Conceptual model development, partially addressed under TRU]

Adherence to good practice guidance is an essential requirement for healthcare modelling [2]. Peer reviewed articles on health economic modelling practice typically consider issues specific to computational implementation in conjunction with other issues such as development of the conceptual model, the identification and selection of input data and study reporting.

Individual guidelines from this literature that are specific to the computational implementation of models address issues of both development and use. Guidelines for the development of CHEMs include recommendations on code organisation, data file management, version control [[10.1186/s12967-020-02540-4](https://doi.org/10.1186/s12967-020-02540-4)], verification [[10.1186/s12967-020-02540-4](https://doi.org/10.1186/s12967-020-02540-4), <https://doi.org/10.1093/epirev/mxab006>],XXXXXXXXXXX) and validation [[10.1186/s12967-020-02540-4](https://doi.org/10.1186/s12967-020-02540-4), <https://doi.org/10.1007/s40273-021-01110-w])>]. Guidelines on how model authors can support appropriate use of CHEMs address issues such as availability [[10.1186/s12967-020-02540-4](https://doi.org/10.1186/s12967-020-02540-4), <https://doi.org/10.1093/epirev/mxab006>, https://doi.org/10.1007/s40273-021-01110-w], developer documentation [[10.1186/s12967-020-02540-4](https://doi.org/10.1186/s12967-020-02540-4)], fitness for purpose [[10.1186/s12967-020-02540-4](https://doi.org/10.1186/s12967-020-02540-4), https://doi.org/10.1007/s40273-021-01110-w], reproducibility [[10.1186/s12967-020-02540-4](https://doi.org/10.1186/s12967-020-02540-4), <https://doi.org/10.1371/journal.pcbi.1010856>, <https://doi.org/10.1093/epirev/mxab006>], re-use [https://doi.org/10.1007/s40273-021-01110-w], terminology [**<https://doi.org/10.1002/jrsm.1333>**], user-documentation [[10.1186/s12967-020-02540-4](https://doi.org/10.1186/s12967-020-02540-4), <https://doi.org/10.1007/s40273-021-01110-w>] and user-interfaces [https://doi.org/10.1007/s40273-021-01110-w].

We are developing a computational model to explore multiple economic topics relating to the mental health of young people aged 12 to 25 called ready4 ([www.ready4-dev.com](http://www.ready4-dev.com)).

We have developed a framework that:

* specifies a set of guidelines for implementing a transparent, reusable and updatable CHEM;
* provides a toolkit of online services and novel software for implementing a youth mental health model that meet these standards.

## Modelling toolkit

We developed a toolkit to help us develop ready4 as a CHEM that meets all six TRU standards. The toolkit is comprised of accounts that we have established and configured using existing online services and novel software that we have written as R libraries (for details, see Availability of Data and Materials).

Advantages of modular models include feasibility (large projects are broken into smaller tasks, with each component independently developed and tested) and flexibility (making it easier to selectively replace or update specific parts of a model and to scale up or down the level of granularity) [19]. Modular approaches are currently being used to facilitate the development of complex computational models in disciplines such as biology [19], ecology [20] and neuroscience [21]. In health economics the related and enabling concept of reference models has been recommended [22], but peer reviewed studies describing modular health economic models remain relatively rare, though examples exist in infectious disease [23] and cardiology [24].

Modular models also provide an opportunity for multiple modelling teams to contribute to, test and reuse models. To enhance this capacity, modular models may be implemented as open source projects that give others liberal permissions to access and use model source code and data [19–21].

####