



Model Choice for Quantitative Health Impact Assessment and Modelling: An Expert Consultation and Narrative Literature Review

Natalie Mueller^{1,2,3*}, Rodrigo Anderle⁴, Nicolai Brachowicz¹, Helton Graziadei⁵, Simon J. Lloyd¹, Gabriel de Sampaio Morais⁴, Alberto Pietro Sironi⁴, Karina Gibert⁶, Cathryn Tonne^{1,2,3}, Mark Nieuwenhuijsen^{1,2,3}, Davide Rasella^{1,4,7}

Abstract

Background: Health impact assessment (HIA) is a widely used process that aims to identify the health impacts, positive or negative, of a policy or intervention that is not necessarily placed in the health sector. Most HIAs are done prospectively and aim to forecast expected health impacts under assumed policy implementation. HIAs may quantitatively and/or qualitatively assess health impacts, with this study focusing on the former. A variety of quantitative modelling methods exist that are used for forecasting health impacts, however, they differ in application area, data requirements, assumptions, risk modelling, complexities, limitations, strengths, and comprehensibility. We reviewed relevant models, so as to provide public health researchers with considerations for HIA model choice.

Methods: Based on an HIA expert consultation, combined with a narrative literature review, we identified the most relevant models that can be used for health impact forecasting. We narratively and comparatively reviewed the models, according to their fields of application, their configuration and purposes, counterfactual scenarios, underlying assumptions, health risk modelling, limitations and strengths.

Results: Seven relevant models for health impacts forecasting were identified, consisting of (i) comparative risk assessment (CRA), (ii) time series analysis (TSA), (iii) compartmental models (CMs), (iv) structural models (SMs), (v) agent-based models (ABMs), (vi) microsimulations (MS), and (vii) artificial intelligence (AI)/machine learning (ML). These models represent a variety in approaches and vary in the fields of HIA application, complexity and comprehensibility. We provide a set of criteria for HIA model choice. Researchers must consider that model input assumptions match the available data and parameter structures, the available resources, and that model outputs match the research question, meet expectations and are comprehensible to end-users.

Conclusion: The reviewed models have specific characteristics, related to available data and parameter structures, computational implementation, interpretation and comprehensibility, which the researcher should critically consider before HIA model choice.

Keywords: Health Impact Assessment, Ex-Ante Impact Evaluation, Forecast, Modelling, Policy

Copyright: © 2023 The Author(s); Published by Kerman University of Medical Sciences. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Citation: Mueller N, Anderle R, Brachowicz N, et al. Model choice for quantitative health impact assessment and modelling: an expert consultation and narrative literature review. *Int J Health Policy Manag.* 2023;12:7103. doi:10.34172/ijhpm.2023.7103

Article History:

Received: 19 January 2022
Accepted: 28 January 2023
ePublished: 13 March 2023

*Correspondence to:

Natalie Mueller
Email:
natalie.mueller@isglobal.org

Introduction

Health impact assessment (HIA) is a widely-used process for identifying the health impacts – both, negative and positive of a policy, program, project or other kind of intervention. HIA is a tool that can help foresee how a *counterfactual* may affect population health, by systematically assessing associated health benefits and risks. The counterfactual accounts for how a change in health risk factor exposure distribution under a proposed policy, program or intervention may change a population's health status in comparison to the business as usual (BAU) or other reference situation.¹ The World Health Organization (WHO) defines HIA as a *combination of procedures, methods and tools by which a policy, program, or project may be judged as to its potential effects on the health of a population, and the distribution of its effects within the*

*population,*² and thereby provides a flexible HIA definition.

HIAs can be carried out prospectively (ie, ex-ante), concurrently or retrospectively (ie, ex-post).³ However, as implied by 'assessment' in contrast to 'evaluation,' most HIAs are carried-out prospectively to compare health consequences under contrasting futures.⁴ In the context of this article, we will also understand HIA as a tool for forecasting policy proposals, whereas a comprehensive body of literature exists on health impact evaluation (ie, ex-post assessments).⁵

Societies usually give considerable weight to their self-interest in health,^{6,7} meaning HIAs can generate awareness and have a bearing on the policy decision-making process.⁴ HIA methods can be qualitative or quantitative in nature. Both methods can co-exist and profit from each other.⁸ Qualitative HIA approaches usually draw on evidence that is already

available, hence can be carried-out more rapidly and thereby might be less resource-intensive, but they may be more prone to selective and subjective reporting.^{3,9,10} Quantitative HIA approaches numerically assess the magnitude and direction of the expected health impacts of a proposed policy; this means they may be of special importance for influencing policy, as decision-makers commonly give more weight to outcomes that are measurable.³ Quantitative HIA is, unlike to qualitative HIA, less sensitive to societal opinions and notions, but can usually only consider a selection of health risk factors of interest.

A variety of quantitative HIA models have been used to forecast health impacts. However, each of these models comprises a different HIA application area, data requirements, complexities, assumptions, health risk modelling, limitations, strengths and comprehensibility. Thus, the objective of this study was to review the most relevant models for prospective HIA, in order to provide systematic guidance on model choice for future HIA studies.

Methods

To review the models, we combined consultation of HIA experts with a narrative literature review. The lead and last author of this manuscript (NM and DR) jointly identified established experts working on quantitative HIA in various fields, including environmental and social epidemiology, econometrics, climate change impact modelling, public health, computer and data science, and invited them to contribute to this interdisciplinary, open discussion on health impact modelling. For ease of understanding, it was decided that this review should cover the broad field of quantitative HIA models. Each expert had worked with at least one of the included HIA models in the past, in various fields of application, including urban planning, transport or environmental sciences, fiscal policies, health care systems research, and infectious disease and non-communicable disease (NCD) modelling, covering low-, middle- and high-income policy settings.

The narrative literature review complemented the expert inputs, in order to provide a comprehensive overview and model examples. Thus, we did not intend to provide a systematic review of the literature, nor do we attempt to outline all the conceptual and computational steps of HIA models. Rather we aim to provide public health researchers with insights into the broad field of prospective HIA models, their key features and implications, in a form that may serve as basis for model choice. Below, we review each model individually and relate them to quantitative HIA application. We acknowledge, however, that models are often combined, integrated and hybrid-models exist, and models cannot always be as easily distinguished as we do in this review.

Key Health Impact Modelling Issues

Available data, their respective structures, and the context of the assessment are of high relevance to HIA model choice. From the outset of model development, researchers should work with the end-users (eg, communities affected by the respective policies and their advocates, policy-makers,

practitioners, etc), so as to ensure that the questions asked and analysis performed are feasible, relevant, comprehensible, timely and resource-efficient, while mutual expectations are met.⁸ In this regard, there is value in qualitative methods that complement the quantitative implementation of HIA models, as there is a considerable body of literature that stresses the importance of citizens' participation in planning and decision-making processes,^{11,12} and an increased need for participatory, quantitative HIA implementation has previously been identified.⁸

Model Comparison

To comparatively review the models in terms of application potential for HIA, we developed a data extraction tool based on eight criteria (Table). The criteria and their definitions are:

- 1) Configuration: Core model functioning and unit.
- 2) Counterfactual scenario: Definition of scenario construct and comparative elements (ie, BAU with policy scenario comparison).
- 3) Objectives/purpose: Defines how model outcomes should be treated and interpreted.
- 4) Assumptions: Underlying assumptions regarding policy effects, population and disease characteristics, data structures and parameters.
- 5) Health risk function: Health risk modelling or risk extrapolation to establish associations between exposure level changes due to policy effect and health outcomes.
- 6) Complexity: Model construct and implementation complexity, linear versus dynamic, consideration of heterogeneity versus homogeneity, consideration of covariates, interactions, adaptation, etc.
- 7) Limitations: Model prerequisites and limits.
- 8) Strengths: Model strengths and advantages.

We also developed a flowchart (Figure 1) intended to guide researchers in their HIA model choice, considering the research question and field of HIA application, available data structures and resources, model functioning and specifics, and interpretation of model outputs.

Results

We identified seven HIA models, that have been used or have the potential to forecast health impacts of policy proposals. We comparatively summarized the models in Table. Moreover, Figure 1 provides key considerations for HIA model choice.

The seven reviewed models are: (i) comparative risk assessment (CRA), (ii) time series analysis (TSA), (iii) compartmental model (CM), (iv) structural model (SM), (v) agent-based model (ABM), (vi) microsimulations (MS), and (vii) artificial intelligence (AI)/machine learning (ML).

Comparative Risk Assessment

In CRA the health impact due to an observed health risk factor exposure distribution in a population (ie, BAU) is compared with the health impact under a hypothetical health risk factor exposure distribution (ie, counterfactual), representing the policy proposal.¹³ CRA aim to provide a standardized and comparable estimate of the expected health impact of

Table. Analytical Models for Forecasting of Health Impacts in Health Impact Assessment

Model	Configuration	Counterfactual Scenario	Objective/Purpose	Assumptions	Health Risk Function	Complexity	Limitations	Strengths
CRA	<ul style="list-style-type: none"> • Comparison of BAU vs counterfactual (typically non-healthcare policy), to assess typically NCD impacts • Model unit: population-level 	<ul style="list-style-type: none"> • Comparison of health impacts based on assumed change in health risk factors exposure levels associated with policy 	<ul style="list-style-type: none"> • Standardize and compare health risks and benefits resulting from the change in health risk factors exposure levels • Provide net health impact of policy effect 	<ul style="list-style-type: none"> • Non-temporal model, compares two scenarios with and without change in exposure levels • Frequent assumption of immediate, perfect implementation of policy and build-up of health impacts 	<ul style="list-style-type: none"> • Risk extrapolation from literature, no risk modelling 	<ul style="list-style-type: none"> • Change in exposure level leads to change in health status, static, unidirectional, no feedback • Limited consideration of population heterogeneity and covariates 	<ul style="list-style-type: none"> • Questionable generalizability of risk extrapolation across populations • Exclusive consideration of previously quantified health risk factors • Cannot account for demographic or disease burden changes over time • Limited accounting for competing causes of disease 	<ul style="list-style-type: none"> • Simple but robust model if risk generalizability across populations holds true • High comprehensibility by end-users • Resource-efficient (time, money, computational skills)
TSA	<ul style="list-style-type: none"> • Considers the ordered sequence of values of a variable (health outcome) at equally spaced time intervals and assumes past values of sequence to be good predictor of future values • Model unit: population-level 	<ul style="list-style-type: none"> • Prediction of post-intervention period under assumption that intervention would not have been implemented 	<ul style="list-style-type: none"> • Forecasting future values of sequence based on historical trajectory • Comparison of regression slopes of implemented policy vs BAU (retrospective/ex-post HIA) 	<ul style="list-style-type: none"> • Past values of variable are predictors of future values • Time is independent • Assumptions of autocorrelation (data points have internal structure), seasonality (periodic fluctuations), stationarity (constant mean and variance) 	<ul style="list-style-type: none"> • (Univariate) regression of future values with set of past values • No consideration of covariates, except dynamic regression models • Researcher chooses correlation structure 	<ul style="list-style-type: none"> • Mostly static, forecast of future values based on past values, all other parameters held constant • Limited consideration of population heterogeneity and covariates • Unidirectional, no feedback 	<ul style="list-style-type: none"> • Availability of historic data • Reliability depends on correctly identifying and accounting for time trends • Health outcomes are rarely stationary • Structural breaks lead to instability • Retrospective/ex-post HIA only as policy is implemented 	<ul style="list-style-type: none"> • Robust model if detailed time-series data is available and past is indeed good predictor of future • Resource-efficient (time, money, computational skills)
CM	<ul style="list-style-type: none"> • Individuals are assigned to health-state compartments, they progress along • Used to estimate health-state transitions for infectious/NCD developments and how policies may change these transitions • Model unit: population-level 	<ul style="list-style-type: none"> • Comparison of BAU vs policy scenario and changed health-state transition probabilities 	<ul style="list-style-type: none"> • Forecast infectious/NCD evolution, considering disease and population specific health-state transition probabilities 	<ul style="list-style-type: none"> • Compartmental applications (aggregation of individuals into health states) • Individuals in each compartment have same transition probability 	<ul style="list-style-type: none"> • System of ordinary differential equations • Mostly risk extrapolation from literature, but also risk modelling (researcher decides on model parameters) 	<ul style="list-style-type: none"> • Models range from simple to complex, depending on number and relationships between compartments, consideration of heterogenous population and disease characteristics 	<ul style="list-style-type: none"> • Same health state transition probability for all individuals (population proportions) in the same compartment • Different models for the same infectious/NCD, depending on which population and disease-specific characteristics are considered 	<ul style="list-style-type: none"> • Robust model for compartmental disease and population characteristic transitions • Basic models are resource-efficient (time, money, computational skills)

Table. Continued

Model	Configuration	Counterfactual Scenario	Objective/Purpose	Assumptions	Health Risk Function	Complexity	Limitations	Strengths
SM	<ul style="list-style-type: none"> Forecasts health impacts using regression analysis, considering interdependencies of endogenous and exogenous factors on causal pathway Model unit: population-level 	<ul style="list-style-type: none"> Comparison BAU vs policy scenario and changing endogenous and exogenous parameters to represent policy effects under varying conditions 	<ul style="list-style-type: none"> Forecasts policy impacts under different conditions and environments, considering interdependencies of endogenous and exogenous factors on causal pathway 	<ul style="list-style-type: none"> Policy effect differs according to different conditions and environments studied Causal interdependencies of varying endogenous and exogenous factors lead to differential policy effects 	<ul style="list-style-type: none"> Multivariate regression analysis used for prediction Change of parameters to represent policy effect Risk modelling (researcher decides on model parameters) 	<ul style="list-style-type: none"> Functional forms of equations affect model estimation Need for good data- and theory-driven insight into interdependencies of endogenous and exogenous factors on causal pathway 	<ul style="list-style-type: none"> Researcher decides on model parameters, which influences results, as it is impossible to capture every endogenous and exogenous factor of reality (limited system) Fully-specified models needed to reduce uncertainty Resource-intensive (time, money, computational skills) 	<ul style="list-style-type: none"> Considers varying effects under varying conditions and environments (interdependencies of endogenous and exogenous factors) Goes beyond conclusions of conventional empirical studies that provide reduced-form causal relationships
ABM	<ul style="list-style-type: none"> A set of agents with defined attributes on the micro-scale interact with and adapt to each other and their environment over space and time, and produce social phenomenon on macro-scale Model unit: individual-level 	<ul style="list-style-type: none"> Comparison of BAU vs policy-induced changes in the environment and/or agent's properties, rules actions and interactions 	<ul style="list-style-type: none"> Bridging individual-level assumptions and population-level dynamics and thereby exemplifying a complex social phenomenon Uncover adaptive system responses that can help understand the long-term policy impact 	<ul style="list-style-type: none"> Researchers chooses agent's properties, rules, actions, time and environment and thereby causal mechanisms Simplification of the real-world, best if data and theory-driven 	<ul style="list-style-type: none"> Causation via a social phenomenon Computational model Risk extrapolation from literature, or risk modelling (researcher decides on model parameters) 	<ul style="list-style-type: none"> Highly complex (modeling agent interaction, adaptation and feedbacks) Detailed individual-level data required 	<ul style="list-style-type: none"> Sensitivity to modeling decisions made by researcher Potential lack of transparency on causal pathways and uncertainties involved Lack of real-world data for model calibration and validation Potentially limited comprehensibility by end-user Resource-intensive (time, money, computational skills) 	<ul style="list-style-type: none"> Complex system insight into causal mechanisms and adaptive system responses Consideration of agent heterogeneity (properties, rules, actions), interaction and adaptation Useful to study social phenomenon that arises from changing dynamics at micro-level
MS	<ul style="list-style-type: none"> Individuals' trajectories are simulated, with transitions between states and conditions over time, but with no interaction and adaptation Model unit: individual-level 	<ul style="list-style-type: none"> Comparison BAU vs policy-induced changes of individuals' trajectories 	<ul style="list-style-type: none"> Provide detailed policy impact for individuals or subgroups and thereby gain insight into equity implications 	<ul style="list-style-type: none"> All unit-level parameters need to be known or are reasonably imputable 	<ul style="list-style-type: none"> Multivariate regression analysis used for prediction Risk extrapolation from literature, or risk modelling (researcher decides on model parameters) 	<ul style="list-style-type: none"> Highly complex (modeling individuals' trajectories) Detailed individual-level data required 	<ul style="list-style-type: none"> Difficult to explain in detail Potentially limited comprehensibility by end-user Resource-intensive (time, money, computational skills) 	<ul style="list-style-type: none"> Individuals' trajectory insight Consideration of population heterogeneity (policy effects and health outcomes)
AI/ML	<ul style="list-style-type: none"> Computer algorithms recognize patterns in complex and changing data and learn from them Model unit: individual-level 	<ul style="list-style-type: none"> Supervised learning predicts a particular social phenomenon by discovering relevant patterns in data 	<ul style="list-style-type: none"> Algorithm is trained to detect pattern in data and can detect them under changing data structures 	<ul style="list-style-type: none"> A set of rules is applied to the algorithm that leads to the recognition of the desired pattern and the algorithm evolves by learning from it 	<ul style="list-style-type: none"> Computational model Risk modelling (algorithm decides on simulation parameters and functional forms) 	<ul style="list-style-type: none"> Highly complex (algorithm development) 	<ul style="list-style-type: none"> Quickly-evolving field with no standards defined Difficult to explain in detail Overfitting of data limits forecasting performance Limited comprehensibility by end-user Resource-intensive (time, money, computational skills) 	<ul style="list-style-type: none"> Insight into complex, real-world non-linear associations under changing data and conditions

Abbreviations: ABM, agent-based model; AI, artificial intelligence; HIA, health impact assessment; BAU, business as usual; CM, compartmental model; CRA, comparative risk assessment; ML, machine learning; MS, microsimulations; SM, structural model; TSA, time series analysis; NCD, non-communicable disease.

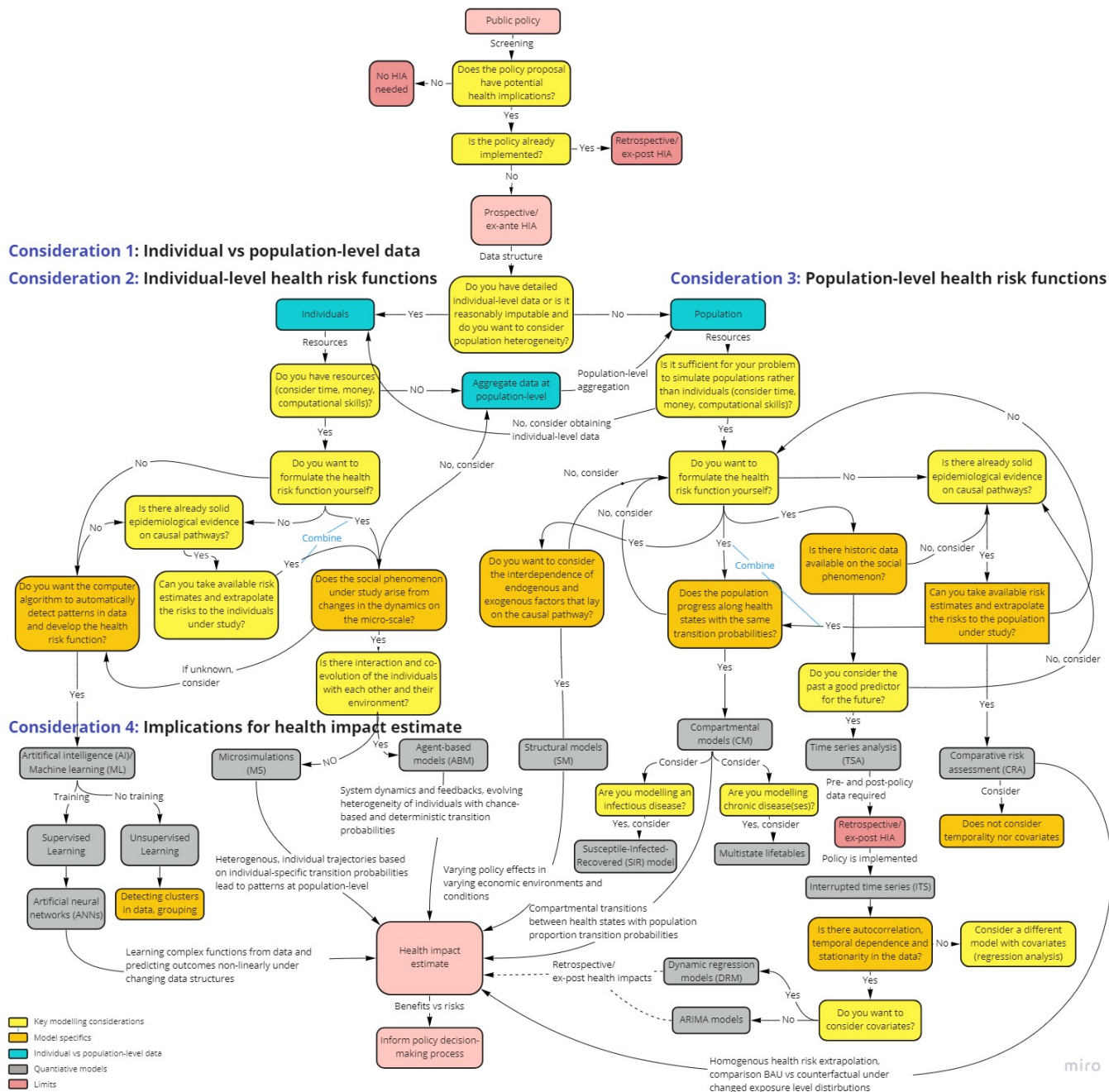


Figure 1. Flowchart for Health Impact Assessment Model Choice Considerations. Abbreviations: HIA, health impact assessment; BAU, business as usual.

shifting from the BAU to the counterfactual scenario, and the distribution thereof among the population.^{13,14} CRA, in comparison to the other models, should be understood as a comparison of BAU with an alternative “future” scenario, and therefore, CRAs are more static and can only restrictedly consider temporal elements such as an aging population. CRAs typically deal with long-term exposure and NCD outcomes (eg, changes in premature deaths, cases of NCDs),¹³ can consider multiple health risk factors and aim to provide a net health impact (ie, positive or negative) of the policy proposal.^{13,15} Steps in CRA are commonly (a) definition of counterfactual (ie, policy proposal); (b) identification of health risk factors of interest; (c) estimation of population health risk factor exposure distribution; (d) selection of health outcomes of interest; (e) selection of exposure-response

functions (ERFs) from the scientific literature to quantify the association between the health risk factor exposure level and the health outcome; (f) combination of health risk factor exposure data and ERFs for each population subgroup under consideration; (g) quantification of the magnitude of the expected distribution of health impact in the population, typically via potential impact fractions; (h) quantification of uncertainty and provision of the range of the potential effect.¹⁶

Application in Health Impact Assessment

CRA is frequently used in HIA studies, with the counterfactual scenario representing the policy proposal (often non-healthcare) under study, and therefore an alternative “future.” Common assumptions are immediate and perfect implementation of the policy, as well as homogenous uptake

and unfolding of policy effects among the population. CRA typically draws on health risk functions (ie, ERFs) from previous studies on the assumption that they are generalizable to the study population. Typically, simple CRAs are static and cross-sectional, meaning they simply compare BAU with an alternative health risk factor exposure distribution. They hold all population parameters constant, do not consider potential varying health risks for subgroups, and assume unidirectional relations from cause to effect (ie, no feedback).

Common limitations include a selective consideration of previously quantified health risk factors and disregard of potentially important factors with weak evidence base; limited consideration of interactive, synergetic effects of multiple health risk factors, leading to potential double-counting of health impacts¹⁷; lack of temporal considerations (eg, changing population structure over time) and inability to account for competing causes of disease¹⁸; frequent assumption of no time-lags in health impacts (ie, immediate impacts); the lack of consideration of required individuals' behavior change¹⁹; limited consideration of benefit-risk trade-off distortion when comparing long-term benefits with immediate risks (eg, long-term health benefits of physical activity compared to the immediate risk of suffering a traffic accident, related to a policy that promotes cycling for transport¹⁹); questionable generalizability of health risk functions (ie, ERFs) as they are extrapolated from another population²⁰ and capture the population rather than individual-level.

Despite its uncertainties and limitations, strengths of CRAs are that they can be carried out relatively rapidly, are less data and resource intensive than other HIA models (eg, models that estimate individuals' trajectories, such as ABM or MS), and provide an evidence-based, health net impact that can be used to inform the policy decision-making process with a typically high comprehensibility for end-users (eg, communities affected by the policy and their advocates, policy-makers, practitioners, etc). CRAs have been widely used, for instance, to model comparatively a wide range of health risk factors within the Global Burden of Disease study,^{21,22} or more specifically, to model city-specific the health impacts of urban and transport planning interventions.²³⁻²⁵

Time Series Analysis

Time series data refers to a time sequence of observations of a certain variable, usually at equal time intervals and for a single unit of analysis, eg, an individual or a geographic region, etc. TSA typically draws on the assumption that the past is a good predictor of the future. Several TSA models are available to forecast health outcomes. Considering the nature of the outcome variable, temporal dependence, and covariates, guide in selecting the most suitable model.

The autoregressive (AR) model assumes the value of a variable at a given moment is dependent on its value at previous time points (ie, autocorrelation), and regresses the current value with a set of past observations of the sequence. The autoregressive moving average (ARMA) model additionally introduces temporal dependence of the random errors, meaning the assumption that the next observation is the mean of all past observations. The AR and ARMA models both

rely on the assumption of stationarity, which implies that the mean and variance are constant over time (ie, non-seasonal, and when the time-series fluctuates, it does so uniformly around a particular time). Since the stationarity assumption is commonly unrealistic in health outcomes, as they usually have a seasonal component, the ARIMA (autoregressive integrated moving average) model is preferred, which accommodates non-stationary behaviors in time-series.²⁶ ARIMA models achieve stationarity by taking a series of differences, ie, measuring how many non-seasonal differences are needed to achieve stationarity. As usual in statistical modelling, the ARIMA model consists of identification, estimation and application. During the identification step, autocorrelation and partial autocorrelation functions check seasonality, and other trends in the series. These steps allow to find the appropriate linear model for which parameters are estimated through the maximum likelihood estimation. Finally, the model is applied to make forecasts for future points in time.

Application in Health Impact Assessment

TSA are substantially different from other HIA forecasting models, because they can only predict, based on historic data, what would have happened if the counterfactual (ie, policy) would not have been implemented. Extended TSA models can overcome this limitation: The interrupted time series (which may be AR, ARMA, or ARIMA), is a quasi-experimental method used to estimate the impact of a policy implemented at given moment over a single time-series. Interrupted time series, however, do not allow the inclusion of covariates; that is, traditional ARIMA models assume an outcome depends only on its past levels. In contrast, regression models can consider the influence of covariates but not the temporal dependence of the time-series. Dynamic regression models, which may be considered extended ARIMA models, overcome these limitations and can simultaneously include covariates and account for temporal dependence,^{27,28} to give a flexible forecast of the post-intervention period.

Nevertheless, the above approaches require multiple observations that are taken pre- and post-policy to robustly assess the impact, and are therefore mostly applied in retrospective HIA. The effect of the policy is evaluated by comparing differences in the regression slopes – that is, the difference in the trajectory of the health outcome of interest under the counterfactual (ie, policy implemented) compared to the BAU. However, TSA may be used for forecasting if the historic trajectory is known and the researcher has an idea of how a policy may change this trajectory (ie, the change in regression slopes). Examples of retrospective HIA using TSA include modeling the impact of an Australian health policy restricting the conditions under which an antipsychotic medicine could be subsidized, predicting the number of monthly drug dispensings,²⁹ or assessing mortality impacts of air quality changes due to a Hong Kong shipping emission policy.³⁰

Compartmental Models

CMs are typically used for modelling infectious diseases but may also be used to model NCDs. The main CM assumption

is that the individuals composing the study population can be aggregated into (mutually-exclusive) compartments, according to their health status, and proportionally move from one compartment to another, as their health status changes. CMs are usually run with a system of ordinary differential equations, and may be deterministic or stochastic. One of the simplest models in infectious disease research is the Susceptible-Infected-Recovered model. In the model, all individuals are susceptible to the disease at the begin of the epidemic and there are one or more initially infected cases. The progression from S to I, at each time step, is usually defined by the product of the number of susceptible individuals, the number of infected individuals, and the effective contact rate (ie, the proportion of contacts that transmit the disease). The recovery rate – the percentage of infected individuals that recover at each time step – is another fundamental parameter in a Susceptible-Infected-Recovered model. More sophisticated CM models are typically used, being customized to the infectious disease and population of interest, considering demographics or other heterogenous population characteristics. As a result, the same disease can be modeled through different CMs, depending on which parameters are included in the model.³¹

For NCDs, various CMs exist. For instance, the RIVM Chronic Disease Model represents 28 NCDs with compartmental transitions occurring due to changes in risk factor classes, incidence, remission, disease progress, and mortality.³² Also, multistate lifetables can generally be understood as CMs as well, as they follow a population experiencing state transitions over time as they age, given changing transition rates and probabilities,³³ and (often) the simplifying assumption that diseases occur independently of one another.¹⁸ The proportional multistate lifetable allows to model for multiple diseases as proportions of the population can co-exist in more than one compartment, which relaxes the assumption that compartments are mutually-exclusive.³⁴

Application in Health Impact Assessment

CMs have been widely used to assess infectious disease interventions, including both, vaccination strategies,³⁵ and non-pharmaceutical interventions, such as wearing face masks for preventing the spread of COVID-19.³⁶ NCD intervention strategies have also been assessed, where the counterfactual, ie, the proposed policy, generally represents changed transition probabilities relative to BAU. For instance, multistate lifetables have been used to model health outcomes of policies targeting modifiable health risk factors, such as physical activity or obesity.¹⁸

Structural Models

SMs consider how a specific (health) outcome relates to relevant exogenous factors in the specific environment studied, ie, the potential interactions and interdependencies of a policy with other existing policies (macro-level), and endogenous factors, ie, individual characteristics and behaviors (micro-level).³⁷ Therefore, SMs provide a framework for understanding how changes in health outcome might relate to distinct environments and conditions.

The researcher needs to define the variable the model shall compute (ie, health outcome), and which endogenous and exogenous variables to include based on the structural connections between them. There may be multiple plausible models that fit the data equally well. Therefore, the researcher needs to look at the satisfying description of the behavior of the included variables, by checking the available data against theory. Hence, the researcher needs to define a set of formulations with unknown parameters, compute for each formulation the values which give the best explanation of past evolutions, using as criteria statistical testing and compliance to theory. The model estimation can call for the introduction of new variables, or changes in their definitions. Once the model is satisfactorily estimated it can be used for forecasting. SMs can rely on linear methods, but more typically SMs are composed of non-linear estimations and dynamic stochastic models.³⁸

Application in Health Impact Assessment

SMs can be used to forecast policy impacts, by drawing on the BAU estimated regression models and changing their parameters to represent the policy, and its impacts on the interdependencies of the endogenous and exogenous factors in the model.^{5,39} Unlike reduced-form estimations, which consider a one-way relationship between an intervention and an outcome, SMs explicitly specify interdependencies between endogenous and exogenous variables, and therefore, provide insight into indirect effects, mechanisms and factors that lay on the causal pathway.⁴⁰ SMs are widely used in econometrics, however, are increasingly being used in health research. For instance, SMs have been applied to forecast the impacts of fiscal policies (eg, restrictions or expansions of fund transfers from government to municipalities) on different health indicators in Brazil (eg, infant mortality, antenatal care visits, access to primary care, premature mortality).⁴¹

Agent-Based Models

ABMs are simulation models that may predict the appearance of a macro-scale social phenomenon, that arises from changing the dynamics on the micro-scale (ie, agents' behaviors, actions and interactions in a system).⁴² In ABMs, agents are autonomous individuals/units that behave according to given decision-making rules, within a specified institutional environment (eg, under certain policies), and in defined-interactions, including feedback-interactions. In contrast to other models, ABMs allow agents to adapt (either biologically or behavioral via learning) to each other and their environment over time.⁴³ By explicitly modelling every individual, ABMs have the advantage that no population-level aggregation is needed, allowing for representation of real-world heterogeneity of individuals (eg, biological, behavioral, demographic diversity, etc). Moreover, ABMs can take different spatial structures across space and time, which are difficult to represent with other models that rely on mean-field approximations and aggregation.⁴³

The different elements in ABMs can be organized according to the "Properties, Actions, Rules, Time, Environment" framework.⁴³ Properties, actions and rules define the agents,

while time and environment define the context. Properties are individual characteristics of the agents, actions define the specific behavior of the agents and can lead to a change in the agent's own properties or rules, the properties of other agents, and the environment. For every action included, the researcher must define rules under which the action is triggered, and each action must have consequences. Rules define how agents choose actions, change properties and interact with each other and their environment. Time is the unit in which change in rules, actions, properties or environment are defined. The environment provides the boundaries and context for the agents and their interactions and can change over time as a result of agent action or as a result of an intervention policy.

In ABMs, events occur either by chance, according to probability distributions defined by the researcher, or deterministic behavioral rules not subject to chance.⁴⁴ Hence, in ABMs some elements remain beyond empirical observation, while other elements are statistically estimated,⁴⁴ for instance, by drawing and extrapolating health risk functions from the literature to the agents based on their individual exposure and risk.

ABMs are complex and flexible to work with, which is part of their power and at the same time limits them. Assumptions on properties, actions, rules, time and environment need to be well grounded and are best if data-driven or theory-based. In contrast to other models, ABMs cannot be compactly described by a set of mathematical equations, which makes it difficult to critically assess them. To overcome this, it has become practice to accompany ABMs with "Overview, Design Concepts and Details plus Decision-Making" criteria, which aim to standardize the model description.⁴⁵ Translating an ABM design into computational code is challenging, specific choices of functional form or algorithms can affect the results of an ABM, and one can get lost in translation during the conceptual to computational transition.⁴³

Application in Health Impact Assessment

With regards to HIA, ABMs are useful to explore the potential impacts of policies in dynamic social and physical environments, therefore, considering non-linear relations and being performed under more realistic conditions. ABMs can unveil adaptive and potentially diverse responses, and thus, provide insight into differential policy uptake and impacts. ABMs have been applied in the field of social epidemiology, eg, modelling the associations between gender, socio-economic status and smoking behavior,⁴⁶ and income inequalities, residential segregation and dietary intake.⁴⁷ Recently, ABMs have been applied to demonstrate the effect of social distancing (micro-scale) on reducing the spread of COVID-19 (macro-scale).⁴⁸

Microsimulations

MS are increasingly being used in public health research because of their ability to evaluate intervention effects considering population heterogeneities.⁴⁹ MS are able to simulate the differential effects of policies across specific subgroups, and therefore also provide insight into health equity implications.^{50,51} The MS modelling process is based

on the simulation of each unit (ie, individuals, municipalities, regions, etc) that compose the population under study, according to mutually-exclusive and collectively-exhaustive health states.⁵² In contrast to ABMs, in MS, individual units do not take decision-based actions nor interact.

MS represent a bottom-up approach to obtain population-level results: MS simulate the trajectories of all units and summarize individual unit results to obtain aggregated population-level insight. In MS, each unit changes its state based on unit-specific transition probabilities, which makes MS models different from CMs where the same transition probability applies to all units in a given compartment. MS transition probabilities can follow the Markov assumption assuming that the transition probability depends only on the current state, can consider the states or values of other demographic and socioeconomic variables of the unit, or other more complex transition rules.⁵³ MS models follow a cohort of individuals/units and either estimate health risk functions (ie, ERFs) for this cohort directly, or more commonly take and extrapolate health risk functions from the literature. A useful distinction of MS is between static or dynamic: the static models seek to reproduce the impact of a policy on the units at a given time point, while the dynamic models track unit trajectories and transitions over time.⁵⁴

Application in Health Impact Assessment

MS represent a powerful tool to simulate the impact of policies with a high level of precision not achievable with other models. MS could be considered among the most flexible models to evaluate policies on different population and health outcomes and MS can consider interaction and synergistic effects of multiple health risk factors. However, MS require large amounts of data to holistically represent individuals' heterogeneity. MS usually require the construction of complex codes that use parallel computing and are time and computationally demanding.⁵² To overcome these barriers, in the context of public health research, some user-friendly platforms have been created to allow policy-makers and practitioners to develop customized HIAs and to facilitate and speed-up the execution of MS,⁵⁵ such as the UK Health Forum microhealth simulation model⁵⁶ and the Organization for Economic Cooperation and Development Strategic Public Health Planning model.⁵⁷ Further examples of MS application in the context of public health research include: estimations of the impact of sugar-sweetened beverage taxation on obesity in India,⁵⁸ the effects of fiscal policies on premature mortality in Scotland,⁵⁹ and recently, the cost-effectiveness of COVID-19 control strategies in South Africa.⁶⁰

Artificial Intelligence/Machine Learning

AI is a broad scientific discipline, aiming to understand and disentangle complex non-linear systems. Despite being a wide discipline, the most popular areas of AI is ML, which refers to the study of computer algorithms that develop automatically by recognizing patterns in data and learning from them.

ML is divided in two main areas, supervised and non-supervised learning. Supervised learning tackles predictive problems, with the aim to gain insight into a particular

phenomenon, emerging from labeled data observations the algorithm is trained with.⁶¹ Unsupervised learning solves grouping problems or characterization of interactions between variables, and helps to find unknown patterns in data, with the algorithm trying to learn from underlying data structures.⁶¹ Two common unsupervised learning tasks are clustering and dimensional reduction. In clustering, the algorithm attempts to group data observations into meaningful clusters (eg, grouping patients with similar disease evolution, etc). In dimensional reduction, the algorithm reduces the number of variables by identifying redundant or noisy attributes and grouping similar or correlated attributes for better model training.

Among ML models, artificial neural networks (ANNs) have been developed as a computational analogy of brain networks of neurons to learn complex functions from data in order to predict (health) outcomes non-linearly. ANNs can solve problems that would be too complex for classical statistical methods as ANNs have self-learning capabilities that enable them to adapt and produce results as data is changing or more data becomes available. ANNs are built like neuron nodes of the human brain, called processing units, interconnecting a web. The processing units are made up of input and output units. The input units receive various forms of data and based on an internal weighing system and activation functions, the ANN attempts to learn and produce an output report. An ANN initially goes through a training phase where it learns to recognize patterns in data. During this supervision, ANN compares the actual output produced with what it was meant to provide. The difference between both outputs is adjusted by correcting the weights of the connections until the difference between actual and desired outcomes has the lowest possible error.

In comparison to more traditional statistical methods, ML tends to be free of distributional assumptions that often are violated in real-world applications and they can account for non-linear relationships in a longitudinal manner. Instead of the researcher defining the model parameters, it is the computer that decides which parameters to include to best represent the data and reveal complex patterns. ML algorithms can examine larger data sets and disentangle patterns much quicker than other traditional methods. While freedom of distributional assumptions could be considered a strength, ML methods are limited by the characteristics of the training database by fitting it extremely well (ie, overfitting problem), ignoring potentially important real-world aspects that are not included in the training database.^{62,63}

Application in Health Impact Assessment

ML is increasingly being used in public health/HIA research and has been used to predict healthcare outcomes either numerically or qualitatively,⁶¹ eg, describing diagnostics, by drawing on supervised learning, or to identify units that can be grouped together, eg, patients with similar disease evolution, hospitals with similar structure, by drawing on unsupervised learning. A 2019 example used ANNs to identify individuals at risk of in-hospital mortality.⁶⁴ The model not only allowed researchers to model the probability of death longitudinally,

but also was able to document how various timestamps contributed to the prediction. Therefore, the researchers were able to provide targeted interventions, as well as forecast changes in the in-hospital mortality probability over time.

Discussion

We reviewed seven models that can be used in HIA modelling, to gain insight into the health impacts of policy proposals. The reviewed models have different configurations, fields of application, assumptions, health risk functions, complexities, limitations and strengths.

Before choosing a model, the researcher needs to consider that model input assumptions match the available parameter and data structures, while the model outputs must match the research question and need to be comprehensible and acceptable to the HIA end-users (eg, communities affected and their advocates, policy-makers, practitioners, etc, while available resources (ie, time, costs, computational skills), are overarching and decisive (Figure 2).⁶¹

The research question needs to be defined clearly, best if in consultation with the end-users,⁸ from the outset of the project. The inventory of existing literature and explanatory theory, available data and its structure, and needed team expertise need to be clear. Agreement on research question, timeliness, resources available, and expected use of the outputs between the model conception and design teams and potential end-users are important for model choice and set the boundaries of the work.⁶¹

Comparing the seven reviewed models in their structural configurations, the first difference is the modelling unit: while CRA, TSA, CM, and SM model outcomes on the population level, ABM, MS, and ML model individuals and their trajectories, from which eventually population-level patterns derive, ie, social phenomenon under study. The modelling unit, ie, individual vs population-level, together with the complexity of the models, usually determine the amount and variety of input data required, with CRA and TSA usually requiring less data, and ABM and MS the most. However, there is considerable variability and there can be very advanced CRA or simple MS. The structural complexity of the model also determines the requirement of resources and computational skills. ABM, MS, and ML typically demand more advanced mathematical and programming competencies and might require parallel computing, while CRA, TSA, CM, and SM are typically easier implementable. However, with increasing computing power and access to cloud computing, this distinction is less pronounced. If researchers decide to do multiple uncertainty simulations (eg, Monte Carlo) to better quantify the range of potential effect, these can considerably increase model run times.

All models have the potential for prospective HIA application, except TSA only partially, as TSA models typically require post-intervention observations (unless the change in regression slope is known). CRA outputs should be interpreted as a comparison between BAU and some (future) alternative, as CRA are inherently non-temporal and cannot account for changing population structures, while the other models are simulation approaches and can account

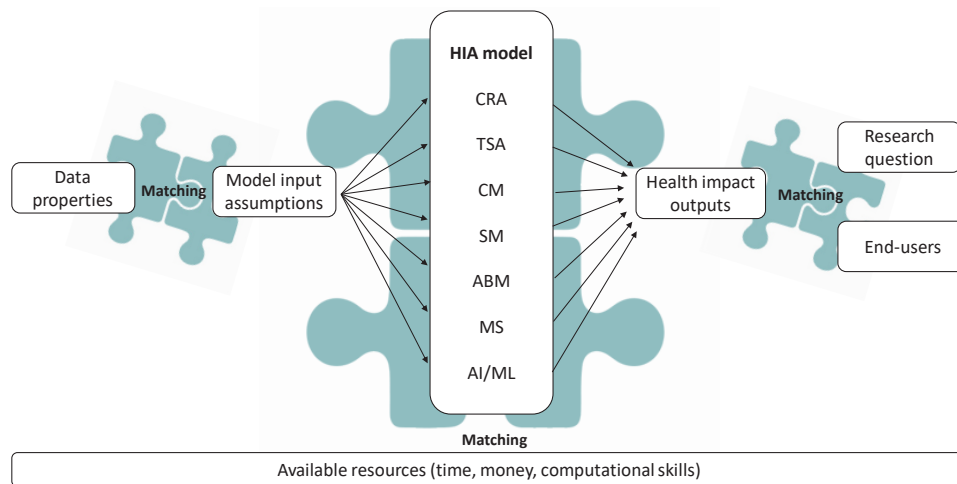


Figure 2. Interlinkages of Key Considerations for Health Impact Assessment Model Choice. The model input assumptions must match the data structure and parameter properties that will be used for model input, while health impact outputs must match the research question to be answered and must be comprehensible to the end-user the HIA is conducted for. Available resources in terms of time, money and computational skills are overarching and decisive. The figure is adapted from Gibert et al.⁶¹ Abbreviations: HIA, health impact assessment; CRA, comparative risk assessment; TSA, time series analysis; CM, compartmental model; SM, structural model; ABM, agent-based model; MS, microsimulations; AI/ML, artificial intelligence/machine learning.

for temporal changes and eg, aging populations. While CRA exclusively draw on health risk extrapolation, and therefore assume generalizability of risks to the population under study, all other models can estimate health risk functions, or draw on a combination of risk estimation and risk extrapolation, such as CM, ABM, and MS. While in TSA, CM, SM, ABM, and MS it is the researcher who decides on parameter inclusion for risk estimation, which should be robustly data-driven and theoretically-grounded, in ML it is the algorithm that decides automatically on the model form and function. Hence, ML models are very flexible as they tend to be free of distributional assumptions, eg, variable's interdependences, higher order interactions, linearity, that the other models depend on, but which are often violated in real-world applications. Nevertheless, ML models are prone to over-fitting, meaning that the model fits extremely well to the training data it was constructed on, but might poorly forecast future values under changed conditions, which might limit ML in HIA application.⁶³ This raises generally the question whether extrapolation of available health risk functions from one population to another (as in CRA, and possibly CM, ABM, and MS) are of higher or lower order of validity than obtaining health risk functions from a single dataset with a set of observations in a given population (all other models).

A key consideration for HIA is timeliness, namely providing evidence of the policy impact *before* its implementation, to influence the decision-making process. Often, a narrow time window is available, after which outcomes are no longer relevant, because a decision on policy implementation has been made. The researcher, according to research question, data structures, end-users and available resources (ie, time, money, computational skills), needs to decide whether to choose a more complex or simpler model and how much added value a more complex model has over a simpler model that achieves similar results. Modelling should follow the principles of parsimony, which refers to preferring simpler models over complex ones, if the simpler model produces

robust results. Hence, a simple, comprehensive model, eg, in terms of number of model parameters, distributional assumptions, model form, etc, is preferable over a complex model where the number of parameters is exhaustive and distributional assumptions and model form cannot be understood anymore. However, the robustness of very simple models might be difficult to assess, especially if they try to model a complex social phenomenon. On the other hand, very complex models, without proper description, could be considered as “black boxes” by end-users, running the risk of not being understood and outputs not being used (eg, in the policy decision-making process). Concerns may arise from the choice of formal tools, eg, computer language and modelling packages, because often softwares and algorithms are not standardized, open-source and are continually evolving.

During the HIA modelling, a multidisciplinary team of experts is desirable, including experts in theory and mechanisms, experts in mathematical modelling and computational coding that help translating the model from concept to code, and experts that can translate the outputs into policy and societal impact. Generally, there is a call for more participatory, quantitative approaches in HIA.⁸ Participation allows planners and policy-makers to gain a more detailed insight into stakeholder behaviors and preferences and is valuable to backup assessment procedures of policy proposals. Moreover, participation allows an increase in public acceptability of decisions, build stronger consensus and reduce conflicts.⁸

The WHO HIA definition also calls for assessing “*the distribution of [policy] effects within the population,*”²² and thereby emphasizes that health impacts of policies can vary for different subgroups, according to varying genetic, lifestyle, or socio-economic, etc, susceptibilities. Studying these differential impacts will help identify most vulnerable subgroups and will pinpoint where intervention is most urgently needed for equitable and just health outcomes. In this regard, individual-level models that allow for identification

of subgroup impacts (ie, ABMs, MS, ML) are particularly valuable to understand which individuals/subgroups are disproportionately affected, which is key in defining equitable intervention strategies.

Several common challenges in health impact forecasting have previously been identified⁶⁵ and need consideration in the health impact modelling discussion. Challenges relate to (a) modelling the time course of the counterfactual, ie, immediate policy implementation versus phase-in and the resulting time-varying effects of health drivers on health outcomes; (b) time-lags in health benefit build-up and considerations of discounting future impacts; (c) uncertainty in addressing structural and parametric uncertainties; (d) general uncertainty in projecting the future, related to health drivers, demographic and socioeconomic trends, and health outcomes; and (e) how to deal with structural breaks and unforeseen events (eg, pandemics, world economic crisis, etc), leading to instability of model parameters.

In the context of Georg Box's reflection of *all models being wrong, but some being useful*, it is important to highlight that all models are limited systems and can only reflect a small fraction of the real-world complex systems. Models are very much tied to the datasets they were constructed on and might miss aspects that are not captured in the data but are important in the real-world. Therefore, having a good theoretical understanding of the social construct under study, including unobserved entities or unmeasured variables, is a prerequisite for model choice and model functioning (eg, structural and parametric robustness). Also, statistical modelling, meaning the fitting of the model to the underlying observational dataset (ie, empirical analysis), should be done complementary to the simulation, meaning the theoretical representation of the real world that goes beyond the observational data (ie, theoretical analysis).⁴⁴ While statistical analysis is probably closer to the data, it is not necessarily closer to reality and empirical models can potentially lack firm theoretical grounding, not accounting for underlying system dynamics and unobserved but influential entities. On the other hand, theoretical models, are potentially speculative, untested and not fundamentally empirical. Therefore, statistical models need to move closer to theoretical models and vice versa. SM might serve as an example of a statistical model moving closer to a theoretical model. SM can unravel complex processes, temporal ordering, causal linkages and direct and indirect effects, despite being fairly statistical and empirical by only including variables that have been measured.⁴⁴ ABMs might serve as an example of a theoretical model having moved closer to a statistical model, intermingling chance with determinism. In ABMs events occur either by chance, according to probability distributions or are governed by strong behavioral rules which are determinative and not subjective to chance. Nevertheless, in ABMs many elements remain beyond empirical observation and one has a theoretical model that is partly estimated statistically.⁴⁴

In line with Epstein's thoughts that even the best models are *fruitfully wrong* and *illuminating abstractions*,⁶⁶ one needs to recognize that models are limited systems, while they have the power to form the conceptual foundations of their respective

fields and are *headlights in dark unexplored territory*. Models can surprise, make us curious, lead to new questions, allow us to doubt and thereby, help us to base our beliefs on evidence and not on authority.⁶⁶

As common in HIA studies, where evidence on causal inferences is lacking, due to a lack of supportive data or theory, researchers feed in assumptions on model parameter characteristics that carry uncertainties. Uncertainties in the model may be structural (ie, associated with model configuration) and parametric (ie, associated with model parameters, eg, slope of ERFs, threshold levels, etc). Uncertainties need to be transparently defined and quantified by giving ranges of confidence on plausible values. Uncertainty must be effectively communicated with the model design and results, and expectations about accuracy in forecasting of policy impacts must be managed.

A final consideration is the model validation. For the majority of simpler models, a face-validity of the model (first order validation, ie, determination by experts that the model reflects the understanding of the available evidence and science) and an internal validation (second order validation, ie, verification of the adequacy of the codes and algorithms used for the modelling process) could be considered sufficient.⁶⁷ While for some models, especially the ones with complex dynamics and feedbacks, a validation of third or higher order is recommended, which would be an external validation (third order validation, ie, model output are compared with empirical observation not considered in the model development), predictive validation (fourth order validation, ie, assessing model's ability to predict empirical results that were not available during the model development) or a cross-model validation (ie, comparison of results among different models for the same or sufficiently similar analyses).⁶⁷

Strengths and Limitations

To our knowledge, this is the first narrative review comparing different mathematical models that have been or can be used for health impact forecasting of public policy proposals, and we provide key consideration for model choice for researchers who wish to engage in HIA modelling. The expert consultation and supportive narrative literature review add robustness to our findings. However, this is not a systematic review of HIA models nor relevant literature; therefore, we acknowledge that partially selective reporting, also of model examples, might have occurred. Moreover, as stated above, we recognize that models are often more integrated and cannot be distinguished as simply as we represent them here. We intend to describe basic model setup and functioning and acknowledge that working definitions in reality are much broader, hybrid-models exist, and distinguishing between model types as we do in this manuscript is often not possible.

Conclusions

We reviewed seven different models that can be used to forecast the health impacts of policy proposals, and therefore are relevant in the field of HIA. Each of these models has specific characteristics, related to available parameter and data structures, computational implementation, and

comprehensibility, and the researcher should consider them critically before choosing a model. Model input assumptions need to match the available data structure, while the outputs of the modelling must match the research question to be answered and need to be comprehensible to the end-users the HIA is conducted for, while available resources (ie, time, costs, computational skills) are overarching and decisive. Moreover, good communication and transparency during the entire HIA process, from research question definition, model conception and implementation, to interpretation of modelling outputs, is needed to make HIA outcomes relevant to inform the policy-decision making process and generate impact for society.

Acknowledgements

We thank James Woodcock and Belen Zapata-Diomedes for their valuable comments on this manuscript. ISGlobal authors acknowledge support from the Spanish Ministry of Science and Innovation and State Research Agency through the “Centro de Excelencia Severo Ochoa 2019-2023” Program (CEX2018-000806-S), and support from the Generalitat de Catalunya through the CERCA Program.

Ethical issues

Not applicable.

Competing interests

Authors declare that they have no competing interests.

Authors' contributions

NM and DR conceptualized the work, reviewed the literature, analyzed the data, drafted and revised the manuscript. RA, NB, HG, SJL, GdSM, APS, KG, CT, and MN reviewed the literature, provided expert input on the respective models and revised the manuscript. All authors approved the final manuscript.

Authors' affiliations

¹ISGlobal, Barcelona, Spain. ²Universitat Pompeu Fabra (UPF), Barcelona, Spain. ³CIBER Epidemiología y Salud Pública (CIBERESP), Madrid, Spain. ⁴Institute of Collective Health (ISC), Federal University of Bahia (UFBA), Salvador, Brazil. ⁵School of Applied Mathematics, Getulio Vargas Foundation, Rio de Janeiro, Brazil. ⁶Intelligent Data Science and Artificial Intelligence Research Center, Universitat Politècnica de Catalunya (IDEAI-UPC), Barcelona, Spain. ⁷Hospital Clínic—Universitat de Barcelona, Barcelona, Spain.

References

- Glymour MM, Spiegelman D. Evaluating public health interventions: 5. Causal inference in public health research—do sex, race, and biological factors cause health outcomes? *Am J Public Health*. 2017;107(1):81-85. doi:10.2105/ajph.2016.303539
- World Health Organization (WHO). Health Impact Assessment: Main Concepts and Suggested Approach. Gothenburg Consensus Paper. Gothenburg: WHO; 1999.
- Joffe M, Mindell J. Health impact assessment. *Occup Environ Med*. 2005; 62(12):907-912. doi:10.1136/oem.2004.014969
- Wismar M, Blau J, Ernst K, Figueras J. The Effectiveness of Health Impact Assessment: Scope and Limitations of Supporting Decision-Making in Europe. Copenhagen: WHO Regional Office for Europe; 2007. http://www.euro.who.int/__data/assets/pdf_file/0003/98283/E90794.pdf. Accessed October 2, 2016.
- Khandker SR, Koolwal GB, Samad HA. *Handbook on Impact Evaluation: Quantitative Methods and Practices*. Vol 1. Washington, DC: World Bank Publications; 2010.
- Jacobs LR, Shapiro RY. Questioning the conventional wisdom on public opinion toward health reform. *PS Polit Sci Polit*. 1994;27(2):208-14. doi:10.2307/420272
- Burstein P. Bringing the public back in: should sociologists consider the impact of public opinion on public policy? *Soc Forces*. 1998;77(1):27-62. doi:10.1093/sf/77.1.27
- Nieuwenhuijsen MJ, Khreis H, Verlinghieri E, Mueller N, Rojas-Rueda D. Participatory quantitative health impact assessment of urban and transport planning in cities: a review and research needs. *Environ Int*. 2017;103:61-72. doi:10.1016/j.envint.2017.03.022
- Mindell J, Ison E, Joffe M. A glossary for health impact assessment. *J Epidemiol Community Health*. 2003;57(9):647-651. doi:10.1136/jech.57.9.647
- Nieuwenhuijsen MJ, Ristovska G, Davdand P. WHO environmental noise guidelines for the European region: a systematic review on environmental noise and adverse birth outcomes. *Int J Environ Res Public Health*. 2017;14(10):1252. doi:10.3390/ijerph14101252
- Mumpower JL. Selecting and evaluating tools and methods for public participation. *Int J Technol Policy Manag*. 2001;1(1):66-77. doi:10.1504/ijtpm.2001.001745
- van de Kerkhof M. Making a difference: on the constraints of consensus building and the relevance of deliberation in stakeholder dialogues. *Policy Sci*. 2006;39(3):279-299. doi:10.1007/s11077-006-9024-5
- Murray C, Ezzati M, Lopez A, Rodgers A, Vander Hoorn S. Comparative quantification of health risks: conceptual framework and methodological issues. In: Ezzati M, Lopez A, Rodgers A, Murray C, eds. *Comparative Quantification of Health Risks: Global and Regional Burden of Disease Attributable to Selected Major Risk Factors*. Geneva: World Health Organization; 2004:1-38.
- Vander Hoorn S, Ezzati M, Rodgers A, Lopez AD, Murray CJ. Estimating attributable burden of disease from exposure and hazard data. In: Ezzati M, Lopez A, Rodgers A, Murray C, eds. *Comparative Quantification of Health Risks: Global and Regional Burden of Disease Attributable to Selected Major Risk Factors*. Geneva: World Health Organization; 2004:2129-2140.
- Veerman JL, Barendregt JJ, Mackenbach JP. Quantitative health impact assessment: current practice and future directions. *J Epidemiol Community Health*. 2005;59(5):361-370. doi:10.1136/jech.2004.026039
- Mueller N, Nieuwenhuijsen MJ, Rojas-Rueda D. Quantitative health impact and burden of disease assessment of traffic-related air pollution. In: Khreis H, Nieuwenhuijsen M, Zietsman J, Ramani T, eds. *Traffic-Related Air Pollution*. Elsevier; 2020:339-359. doi:10.1016/b978-0-12-818122-5.00013-2
- Mueller N, Rojas-Rueda D, Basagaña X, et al. Urban and transport planning related exposures and mortality: a health impact assessment for cities. *Environ Health Perspect*. 2017;125(1):89-96. doi:10.1289/ehp220
- Mytton OT, Tainio M, Ogilvie D, Panter J, Cobiac L, Woodcock J. The modelled impact of increases in physical activity: the effect of both increased survival and reduced incidence of disease. *Eur J Epidemiol*. 2017;32(3):235-250. doi:10.1007/s10654-017-0235-1
- Mueller N, Rojas-Rueda D, Cole-Hunter T, et al. Health impact assessment of active transportation: a systematic review. *Prev Med*. 2015;76:103-114. doi:10.1016/j.ypmed.2015.04.010
- Murray C, Ezzati M, Lopez A, Rodgers A, Vander Hoorn S. *Comparative Quantification of Health Risks: Conceptual Framework and Methodological Issues*. Vol 1. Geneva: World Health Organization; 2004.
- Lozano R, Naghavi M, Foreman K, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*. 2012; 380(9859):2095-2128. doi:10.1016/s0140-6736(12)61728-0
- Forouzanfar MH, Alexander L, Anderson HR, et al. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet*. 2015;386(10010):2287-2323. doi:10.1016/s0140-6736(15)00128-2
- Mueller N, Rojas-Rueda D, Khreis H, et al. Changing the urban design of cities for health: the superblock model. *Environ Int*. 2020;134:105132. doi:10.1016/j.envint.2019.105132
- Mueller N, Rojas-Rueda D, Salmon M, et al. Health impact assessment of cycling network expansions in European cities. *Prev Med*. 2018;109:62-70. doi:10.1016/j.ypmed.2017.12.011
- Woodcock J, Tainio M, Cheshire J, O'Brien O, Goodman A. Health effects of the London bicycle sharing system: health impact modelling study. *BMJ*. 2014;348:g425. doi:10.1136/bmj.g425
- Hamilton JD. *Time Series Analysis*. Princeton, New Jersey: Princeton University Press; 1994.

27. Brodersen KH, Gallusser F, Koehler J, Remy N, Scott SL. Inferring causal impact using Bayesian structural time-series models. *Ann Appl Stat.* 2015;9(1):247-274. doi:10.1214/14-aos788
28. Hyndman R, Athanasopoulos G. Dynamic regression models. In: *Forecasting: Principles and Practice*. 2nd ed. Monash University; 2018. <https://otexts.com/fpp2/dynamic.html>.
29. Schaffer AL, Dobbins TA, Pearson SA. Interrupted time series analysis using autoregressive integrated moving average (ARIMA) models: a guide for evaluating large-scale health interventions. *BMC Med Res Methodol.* 2021;21(1):58. doi:10.1186/s12874-021-01235-8
30. Mason TG, Chan KP, Schooling CM, et al. Air quality changes after Hong Kong shipping emission policy: an accountability study. *Chemosphere.* 2019;226:616-624. doi:10.1016/j.chemosphere.2019.03.173
31. Siettos CI, Russo L. Mathematical modeling of infectious disease dynamics. *Virulence.* 2013;4(4):295-306. doi:10.4161/viru.24041
32. RIVM. A Conceptual Framework for Budget Allocation in the RIVM Chronic Disease Model. A Case Study of Diabetes Mellitus. 2005. <https://www.rivm.nl/bibliotheek/rapporten/260706001.pdf>.
33. Al Mamun A. Multistate life tables in public health. In: *Life History of Cardiovascular Disease and Its Risk Factors: Multistate Life Table Approach and Application to the Framingham Heart Study (Population Studies)*. Purdue University Press; 2003.
34. Briggs AD, Wolstenholme J, Blakely T, Scarborough P. Choosing an epidemiological model structure for the economic evaluation of non-communicable disease public health interventions. *Popul Health Metr.* 2016;14:17. doi:10.1186/s12963-016-0085-1
35. Ojal J, Griffiths U, Hammit LL, et al. Sustaining pneumococcal vaccination after transitioning from Gavi support: a modelling and cost-effectiveness study in Kenya. *Lancet Glob Health.* 2019;7(5):e644-e654. doi:10.1016/s2214-109x(18)30562-x
36. Xiang Y, Jia Y, Chen L, Guo L, Shu B, Long E. COVID-19 epidemic prediction and the impact of public health interventions: a review of COVID-19 epidemic models. *Infect Dis Model.* 2021;6:324-342. doi:10.1016/j.idm.2021.01.001
37. Williams MJ. External validity and policy adaptation: from impact evaluation to policy design. *World Bank Res Obs.* 2019;35(2):158-191. doi:10.1093/wbro/lky010
38. Carroll CD. The method of endogenous gridpoints for solving dynamic stochastic optimization problems. *Econ Lett.* 2006;91(3):312-320. doi:10.1016/j.econlet.2005.09.013
39. McCarthy M, Biddulph JP, Utley M, Ferguson J, Gallivan S. A health impact assessment model for environmental changes attributable to development projects. *J Epidemiol Community Health.* 2002;56(8):611-616. doi:10.1136/jech.56.8.611
40. Low H, Meghir C. The use of structural models in econometrics. *J Econ Perspect.* 2017;31(2):33-58. doi:10.1257/jep.31.2.33
41. Castro MC, Massuda A, Almeida G, et al. Brazil's unified health system: the first 30 years and prospects for the future. *Lancet.* 2019;394(10195):345-356. doi:10.1016/s0140-6736(19)31243-7
42. Hedstrom P. *Dissecting the Social: On the Principles of Analytical Sociology*. Cambridge: Cambridge University Press; 2005. doi:10.1017/cbo9780511488801
43. Hammond RA. Considerations and best practices in agent-based modeling to inform policy. In: Wallace R, Geller A, Ogawa VA, eds. *Assessing the Use of Agent-Based Models for Tobacco Regulation*. Washington, DC: The National Academy of Sciences; 2015:161-193. <https://www.ncbi.nlm.nih.gov/books/NBK305917/>.
44. Burch TK. Computer simulation and statistical modeling: rivals or complements? In: Burch TK, ed. *Model-Based Demography*. Cham: Springer; 2018:67-77. doi:10.1007/978-3-319-65433-1_4
45. Müller B, Bohn F, Dreßler G, et al. Describing human decisions in agent-based models—ODD + D, an extension of the ODD protocol. *Environ Model Softw.* 2013;48:37-48. doi:10.1016/j.envsoft.2013.06.003
46. Chao D, Hashimoto H, Kondo N. Dynamic impact of social stratification and social influence on smoking prevalence by gender: an agent-based model. *Soc Sci Med.* 2015;147:280-287. doi:10.1016/j.socscimed.2015.08.041
47. Auchincloss AH, Riolo RL, Brown DG, Cook J, Diez Roux AV. An agent-based model of income inequalities in diet in the context of residential segregation. *Am J Prev Med.* 2011;40(3):303-311. doi:10.1016/j.amepre.2010.10.033
48. Stevens H. Why Outbreaks Like Coronavirus Spread Exponentially, And How to "Flatten the Curve." *The Washington Post*; 2020. <https://www.washingtonpost.com/graphics/2020/world/corona-simulator/>. Published April 14, 2020.
49. Abraham JM. Using microsimulation models to inform U.S. health policy making. *Health Serv Res.* 2013;48(2 Pt 2):686-695. doi:10.1111/1475-6773.12052
50. Rasella D, Basu S, Hone T, Paes-Sousa R, Ocké-Reis CO, Millett C. Child morbidity and mortality associated with alternative policy responses to the economic crisis in Brazil: a nationwide microsimulation study. *PLoS Med.* 2018;15(5):e1002570. doi:10.1371/journal.pmed.1002570
51. Rasella D, Hone T, de Souza LE, Tasca R, Basu S, Millett C. Mortality associated with alternative primary healthcare policies: a nationwide microsimulation modelling study in Brazil. *BMC Med.* 2019;17(1):82. doi:10.1186/s12916-019-1316-7
52. Krijkamp EM, Alarid-Escudero F, Enns EA, Jalal HJ, Hunink MGM, Pechlivanoglou P. Microsimulation modeling for health decision sciences using R: a tutorial. *Med Decis Making.* 2018;38(3):400-422. doi:10.1177/0272989x18754513
53. Roberts M, Russell LB, Paltiel AD, Chambers M, McEwan P, Krahn M. Conceptualizing a model: a report of the ISPOR-SMDM modeling good research practices task force-2. *Med Decis Making.* 2012;32(5):678-689. doi:10.1177/0272989x12454941
54. Rutter CM, Zaslavsky AM, Feuer EJ. Dynamic microsimulation models for health outcomes: a review. *Med Decis Making.* 2011;31(1):10-18. doi:10.1177/0272989x10369005
55. Richiardi MG, Richardson RE. JAS-mine: a new platform for microsimulation and agent-based modelling. *Int J Microsimul.* 2017; 10(1):106-134. doi:10.34196/ijm.00151
56. UK Health Forum. UK Health Forum microhealth simulation model. <https://ukhealthforum.org.uk/our-work/>. Published 2022. Accessed July 18, 2022.
57. Organization for Economic Cooperation and Development (OECD). OECD's SPHeP Models. A Tool to Inform Strategic Planning in Public Health. <https://www.oecd.org/health/OECD-SPHEP-Models-Brochure-2020.pdf>. Published 2020. Accessed July 18, 2022.
58. Basu S, Vellakkal S, Agrawal S, Stuckler D, Popkin B, Ebrahim S. Averting obesity and type 2 diabetes in India through sugar-sweetened beverage taxation: an economic-epidemiologic modeling study. *PLoS Med.* 2014; 11(1):e1001582. doi:10.1371/journal.pmed.1001582
59. Richardson E, Fenton L, Parkinson J, et al. The effect of income-based policies on mortality inequalities in Scotland: a modelling study. *Lancet Public Health.* 2020;5(3):e150-e156. doi:10.1016/s2468-2667(20)30011-6
60. Reddy KP, Shebl FM, Foote JHA, et al. Cost-effectiveness of public health strategies for COVID-19 epidemic control in South Africa: a microsimulation modelling study. *Lancet Glob Health.* 2021;9(2):e120-e129. doi:10.1016/s2214-109x(20)30452-6
61. Gibert K, Izquierdo J, Sánchez-Marré M, Hamilton SH, Rodríguez-Roda I, Holmes G. Which method to use? An assessment of data mining methods in Environmental Data Science. *Environ Model Softw.* 2018;110:3-27. doi:10.1016/j.envsoft.2018.09.021
62. Lachowycz K, Jones AP. Towards a better understanding of the relationship between greenspace and health: development of a theoretical framework. *Landsc Urban Plan.* 2013;118:62-69. doi:10.1016/j.landurbplan.2012.10.012
63. Peng Y, Nagata MH. An empirical overview of nonlinearity and overfitting in machine learning using COVID-19 data. *Chaos Solitons Fractals.* 2020; 139:110055. doi:10.1016/j.chaos.2020.110055
64. Shickel B, Loftus TJ, Adhikari L, Ozrazgat-Baslanti T, Bihorac A, Rashidi P. DeepSOFA: a continuous acuity score for critically ill patients using clinically interpretable deep learning. *Sci Rep.* 2019;9(1):1879. doi:10.1038/s41598-019-38491-0
65. Remais JV, Hess JJ, Ebi KL, et al. Estimating the health effects of greenhouse gas mitigation strategies: addressing parametric, model, and valuation challenges. *Environ Health Perspect.* 2014;122(5):447-455. doi:10.1289/ehp.1306744
66. Epstein JM. Why model? *J Artif Soc Soc Simul.* 2008;11(4):12.
67. Dahabreh IJ, Chan JA, Earley A, et al. A review of validation and calibration methods for health care modeling and simulation. In: *Modeling and Simulation in the Context of Health Technology Assessment: Review of Existing Guidance, Future Research Needs, and Validity Assessment*. Rockville, MD: Agency for Healthcare Research and Quality; 2017.