Co-production of causal policy simulations of local whole systems approaches to public health

A research proposal seeking Confirmation of Registration to study for Doctor of Philosophy at Leeds Beckett University

John L. Preston

Contents

1 Introduction

1.1 Research context

This section provides a brief overview of the context in which my research is situated. For a more thorough treatment of the relevant literature, broken down by thematic areas, see the literature review.

'Obesity'[1] is a health condition characterised by large deposits of fat in the body. These high levels of fat are risk factors for, and precursors to, diseases such as diabetes, several cancers, cardiovascular disease, hypertension, and osteoarthritis (Meldrum, Morris, and Gambone 2017). Further, 'obesity' has been classified as a 'disease', and as a 'pandemic', with recent studies estimating that as many as 40% of global adults may be 'obese' by 2030 (Meldrum, Morris, and Gambone 2017). In addition to the direct health effects of 'obesity' on the body, people with larger bodies suffer from marginalisation in many areas of their lives, including in access to employment and healthcare (Puhl and Heuer 2009). These effects spread beyond the locus of the individual, and high prevalence of 'obesity' is a driver for increased healthcare expenditure and increased unemployment (Cecchini

[1] Throughout this document, non-neutral terminology about body weight, such as 'obesity', 'overweight', and 'underweight', are placed in 'scare quotes' to highlight their controversial status. For more information, see the critical perspectives on 'obesity' section of the literature review.

and Vuik 2019), both of which have significant knock-on effects for social, economic, and healthcare systems. Due to the increasing prevalence of 'obesity'; the numerous negative consequences at both the individual and social levels; and the presumably preventable nature of the condition, prevention and treatment of 'obesity' has become a global health concern and a priority issue for many public health organisations, including the CDC, WHO, and OHID (formerly PHE). These concerns have been amplified as a result of the COVID-19 pandemic, with research indicating that 'obesity' is a risk factor for contracting COVID-19; developing a more severe COVID-19 illness; being hospitalised; and death (Popkin et al. 2020).

The standard medical model of 'obesity' is problematic for several reasons (Frood et al. 2013). First, the idea that 'obesity' is 'simply' a result of eating too much and not moving enough implies that 'obesity' arises from an individual's personal choices and behaviours, and that individuals are themselves to blame for their conditions. This in turn fuels weight stigma which leads to depression, anxiety, and other mental health problems for people living with 'obesity'. Second, this 'simple' model pushes patients to adopt interventions such as dieting and exercise which may be effective in the short-term, but which often fail in the long-term due to lack of ongoing support. This then lends to weight regain and can result in weight cycling. Third, this perspective discourages considerations of both the underlying or deeper causes of 'obesity', and of the many ways the lives of people living with 'obesity' can be improved through interventions other than weight loss. An individual's ability to eat healthily and live an active live is influenced by their job, income, and built environment, and so changes to these systems can make significant impacts and support more direct interventions. Further, the quality of life of people living with 'obesity' can be improved by making the world around us more accommodating to larger people, particularly in healthcare where patients with 'obesity' have legitimate concerns dismissed due to weight stigma, and where specialist medical equipment is not compatible with patients with larger weights or dimensions (Chrisler and Barney 2017). This is not to say that standard medical approaches to 'obesity' prevention and treatment are never effective, or that we shouldn't be encouraging people to eat healthily and lead an active lifestyle. Rather, the scientific and policy communities are discovering the limits of a purely medical model, and the need for hybrid approaches which include a systems perspective.

The causes of 'obesity' are multifaceted, and occur at many different levels, from the biology and behaviour of individuals; to the social, geographic, and economic factors in which individuals find themselves; to the exposure of individuals to 'obesogenic' chemicals in their environment (Egusquiza and Blumberg 2020; Lobstein and Brownell 2021). These different domains interact to create a complex, nonlinear risk, and this presents significant problems for both researchers and policy-makers, for two reasons. First, it is very difficult to determine the directionality of causal relationships, and to be able to determine which factors are causes *of* 'obesity', and which are caused *by* it. Indeed, in many instances there are feedback loops such that some factors that are a cause at one point in time will be an effect later on, and vice versa. The complexity of this issue is demonstrated by the adverse outcomes faced by people with 'obesity' compared to those of 'normal' weight during the COVID-19 pandemic (Pausé, Parker, and Gray 2021): does 'obesity' induce biological disadvantages which make individuals more susceptible to COVID-19; or is it that people with 'obesity' are more likely to suffer from marginalisations such as poverty and weight stigma which also affect COVID-19 outcomes? Second, even when the causal mechanisms behind 'obesity' are well understood, it is still not always clear what levers are available to policy-makers in our health, economic, and urban systems, and how these causal chains can be weakened or broken. These intersecting and interacting factors, with their uncertain relationships and unclear potential for intervention, are not unique to 'obesity', but lie at the heart of many public health and epidemiological concerns, and as a result, researchers are now turning to complex systems science and causal methods to better understand and prevent public health concerns, including 'obesity' (Galea, Riddle, and Kaplan 2010; Xue et al. 2018).

Systems thinking is an approach to scientific inquiry and problem solving which promotes a non-linear and dynamic perspective. An issue such as 'obesity' is conceived of as a 'system' which can be decomposed into a set of interacting components, and it is through these interactions that the system achieves emergent behaviours such as resilience, which cannot be explained by the individual components in isolation. Systems science, and complex systems science in particular, extends the mentality and qualitative tools afforded by systems thinking with quantitative simulation methods, which allow researchers to explore the dynamics of systems and determine the potential effects of interventions. Causal inference methods are extensions of existing statistical methods to enable researchers to answer causal questions by building statistical and predictive models, which specify the associations or correlations between different variables, on top of process models, which specify

the causal mechanisms which are responsible for creating some observed data. Causal inference and systems science are distinct but complementary approaches, particularly in public health research, where there is a need to understand the causal mechanisms behind complex and dynamic behaviours operating at the population level, with a view to developing effective and adaptable interventions.

In 2015, the Office for Health Improvement and Disparities (formerly Public Health England) commissioned Leeds Beckett University to produce the Whole Systems Obesity (WSO) programme, an approach for developing and implementing local whole systems approaches to 'obesity' prevention and treatment, involving over 50 Local Authorities (Public Health England 2019). The WSO programme represents a breakthrough in the application of systems thinking to service delivery for managing 'obesity', but it does not make use of any dynamic systems modelling or simulation techniques, such as agent-based models (ABMs) or systems dynamics (SD) modelling, nor any causal inference techniques such as structural causal models (SCMs). Thus the current WSO approach relies solely on the stakeholders' qualitative understanding of the system, in order to identify appropriate intervention points and develop effective actions. However, given the complexity of the system and the existence of feedback loops, this is non-trivial, and so the existing purely qualitative nature of the WSO programme can lead to uncertain outcomes, as stakeholders risk overlooking key uncertainties and systems behaviours which would be easily uncovered through quantitative methods. For example, Roberts et al. (2019) develop a qualitative systems model of childhood 'obesity' and then subsequently translate it to a quantitative systems dynamics (SD) model. In their model, participation in organised sports can lead to healthy weight outcomes for children, but the uptake of this is predicated on available parental resources like time and money, and on the child's perceived competence in the sport. When examining the conceptual model, it is easy for stakeholders to reason that increasing the availability of organised sports programmes will reduce incidence of childhood 'obesity', however it is not clear how much this effect will be constrained by parental resources or perceived confidence. If the intervention is to be applied in a population where poverty or child mental health are prevalent issues, this would reduce the impact of the intervention. By supplementing this qualitative analysis of the system with a quantitative simulation, these interactions can be explored more readily by the stakeholders, in order to develop intuition about the system and to inform decision-making. Further, by calibrating the simulation parameters with real-world data from the intervention population and context, it is also possible to get an estimate of the effect size of the intervention, which benefits the appraisal process, as stakeholders become able to look for the most cost-effective interventions, and determine which interventions are feasible given scarce resources.

1.2 Research objectives

My core research is primarily applied and translational in nature, as it involves an extension of the current WSO programme with existing systems simulation and causal inference techniques. My intention is to work directly with Local Authorities to run a series of model building workshops, and then use the outputs of these workshops to produce dynamic simulations of the 'obesity' system, which will then be presented to the Local Authorities as a tool to aid them in understanding the system dynamics, and designing and testing potential interventions and policies. These workshops will be informed by a patient and public involvement and engagement (PPIE) perspective, and stakeholders will include members of the public from a range of backgrounds, including people living with or at risk of 'obesity'. This builds on the strengths of the existing whole systems approach at the heart of the WSO programme, by allowing Local Authorities and researchers to leverage the knowledge of local communities; and promotes the development of socially and epistemically just policies which empower marginalised individuals and communities.

An essential part of this work is the development of a robust extension of the WSO programme. This will require its own programme of research (discussed in the extending local whole systems approaches section of the proposed work). However, at a high level, this approach can be roughly broken down into four phases:

1. Participatory modelling – Stakeholders describe fuzzy qualitative and quantitative aspects of a model in a workshop environment.
2. Simulation implementation – A simulation expert (myself) translates the fuzzy model into a series of numerical computer simulations.
3. Simulation presentation – The simulations and their outputs are presented to the stakeholders.
4. Reflect and repeat – Stakeholders evolve their understanding of the system in response to the simulations, new questions are uncovered, and the simulation modelling process repeats.

This four phase process forms a cycle, wherein once a simulation has been presented to the stakeholders, further workshopping can take place to reflect on what has been learned from the model, and then to modify the model, layering in additional complexity, thus precipitating additional insights. This approach is thus fundamentally iterative and agile, as each cycle of this process informs the implementation of the simulation for the next cycle. An iterative approach such as this affords multiple benefits. First, regular check-ins between simulation experts and other stakeholders such as simulation users and domain experts (including the public and patients) promotes alignment between the mental models of these different groups and will result in simulation models which are more consistent with the group's collective understanding of the system. Second, the simulations developed are able to be adapted to fit the changing needs and questions of the stakeholders over time.

It should be noted that as this stage of this research, the proposed cycle of iterative simulation co-production is subject to change for several reasons. First, there are many additional steps in between the proposed phases, such as model validation and simulation validation, which will extend the production cycle, and the specifics of all of these phases need to be developed and defined. Second, these phases are under-defined and do not account for the changing needs of the stakeholders, as different models or different results may be required at later iterations. Third, these phases only describe a process which is situated across the WSO phases 3 and 4, 'mapping the local system' and 'action', and does not describe how simulations and models can be used as part of the later policy phases of monitoring and evaluation, or how the earlier phases of set-up must be changed to ensure simulation modelling can proceed smoothly and provide maximum benefit. Finally, there may be plenty of opportunities for introducing simulations and causal inference throughout the WSO programme which do not require co-production, and the benefits of a co-production approach must be balanced against the needs and time constraints of the stakeholders for each Local Authority. This affords a great deal of flexibility to my research as I am able to develop a range of methods as I work with different sets of stakeholders. Figure 1 demonstrates the proposed extension along with some of the research questions which need to be addresses as part of the development of this work.

Given this initial outline of the core research, my primary research questions are:

1. How can the local whole systems process be extended to include causal inference and systems simulation components?
2. Do the additions of systems simulation and causal inference methods to the local whole systems process benefit Local Authorities when allocating resources and designing and developing policies and interventions?
3. How can causal inference methods (such as structural causal models) and systems simulation methods (such as agent-based models and systems dynamics models) be combined to deliver valid and empirically-grounded insights about local systems?
4. How can systems simulations and causal inference models be constructed in a co-produced and iterative manner, with stakeholders who are not experts in these methods?

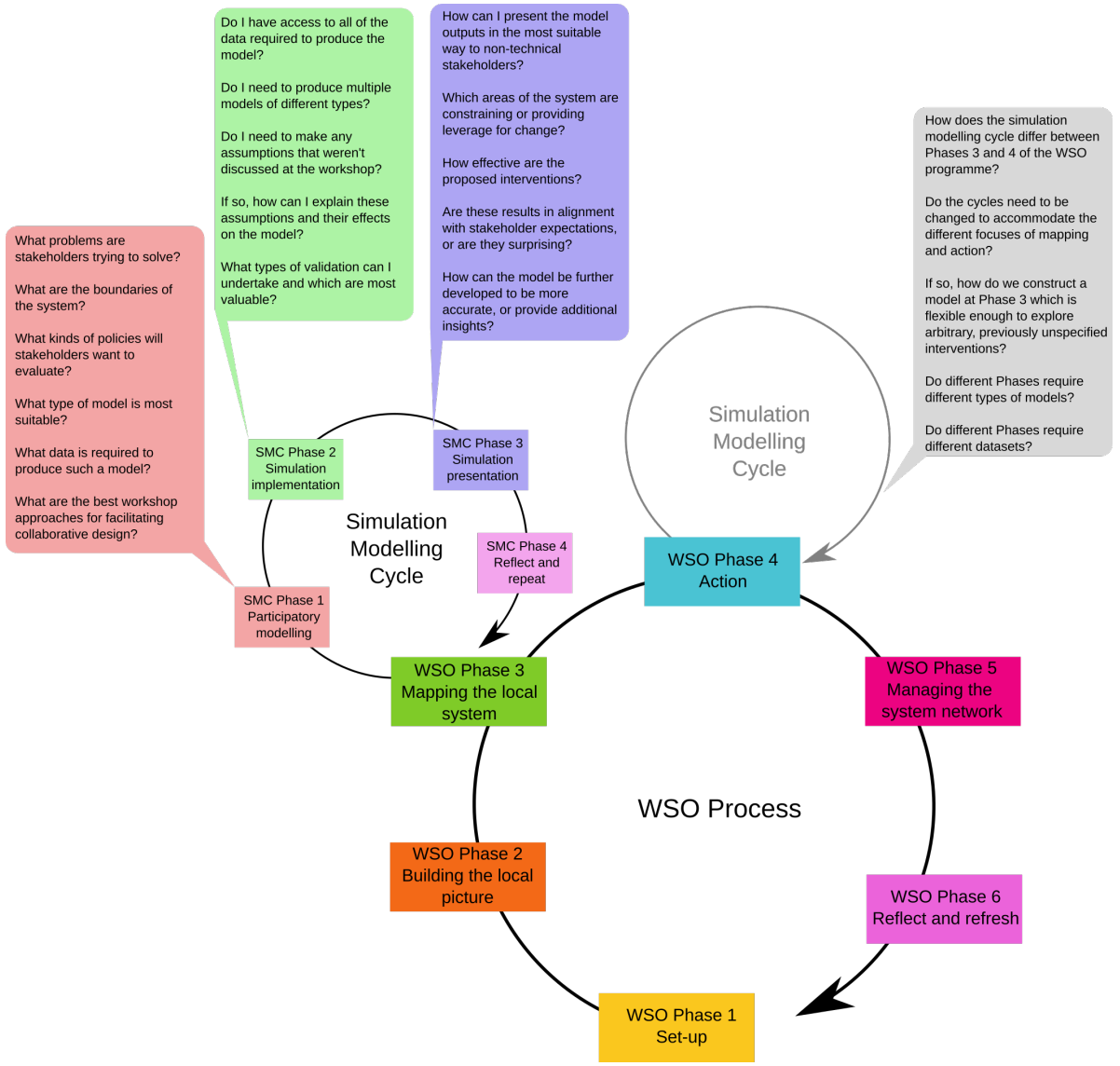

Figure 1: An illustration of the proposed extension to the WSO programme, annotated with open questions at different phases of the process.

1.3 Personal objectives

This research project will not be the end of my interest in causal policy simulation and co-production. All PhD research has two primary aims: to develop novel research which pushes the boundaries of collective knowledge; and to provide the student with the skills and understanding to undertake additional high-quality and innovative research in the future. In order to ensure I am able to achieve these two goals of research innovation and researcher development to the greatest possible extent, it is necessary to discuss my own motivations, and to examine how this PhD is situated within my longer term ambitions.

I want to reduce and eliminate marginalisation. Marginalisation is both the state wherein, and the process by which, certain individuals and groups are excluded from specific areas of society or access to the same level of material conditions as other members of that society. Examples include inequalities in access to education, healthcare, and finance through practices such as redlining and other forms of ethnic and spatial discrimination; the physical and information barriers faced by disabled people as explained through the social model of disability; and the feedback loops which trap people in cycles of poverty. These different forms of marginalisation are the result of dysfunctional systems in our society, each with their own sets of causes and effects, but also often mutually reinforcing each other. If I wish to work towards eliminating different forms of marginalisation, I therefore require three things. First, I must amass a sufficient understanding of the structures and operations of the systems which create different marginalisations, which itself requires an understanding of systems theory in general, as well as the many systems which constitute our society, and the different forms of marginalisation which are precipitated by these systems. Second, I must collaborate with other people, particularly marginalised people, who are able to provide direct evidence and insight into specific forms of marginalisation through their lived experience, and are able to inform and validate potential interventions which we hope will reduce or eliminate their marginalisation. Finally, I must develop a reputation as a just, effective, and impactful researcher, and develop my connections across academia, government, and the third sector, so that I am able to advocate for de-marginalisation as a priority concern in our society, and in order to undertake the research required to understand how we can eliminate marginalisation, and then follow through with effective action.

I also want to improve the transparency and quality of decision-making in socially powerful organisations and institutions such as governments. By definition, these organisations are able to effect significant change in the systems in which they are embedded, and as such they have both the capacity and duty to work towards de-marginalisation. An essential part of ensuring these organisations are working towards this mission is to ensure that the decision-making processes, and the evidence used as part of these decision-making processes, are open to the public and available for scrutiny by experts. Further, improving the quality of evidence used by organisations and introducing approaches like co-production will make organisations more successful in achieving their missions, particularly in regards to de-marginalisation.

These personal objectives reinforce each other, and are served directly by my PhD research. My project not only provides me with the opportunity to engage directly with the decision-making processes around eliminating health inequalities, but it also affords me with essential skills which I will need in my future pursuit of these objectives, whether that takes place through academia, civil service, or third sector organisations.

2 Proposed work

As previously discussed in the research objectives section, the primary aim of my research is the extension of the existing Whole Systems Obesity programme (Public Health England 2019) with a co-produced dynamic causal policy simulation component. To achieve this will require the delivery of a number of sub-projects, which will be discussed here in approximate chronological order.

2.1 Artificial microdataset methods: a scoping review

Across many disciplines such as health, economics, and transport, there are research questions which cannot feasibly be answered through the use of experimental methods such as randomised controlled trials (RCTs). This may be because the imposition of a particular experimental condition is inherently unethical, such as infecting participants with a known pathogen; because subjects may not precisely adhere to the experimental protocol, such as in diet control studies; or because

a required experimental condition is infeasible or impossible to bring about, such as moving a road to observe the effect on congestion. Because of these limitations, researchers must rely on observational data to conduct their analyses and answer many of their research questions. However, the availability of suitable observational data suffers from its own set of problems. First, there is often a need to preserve the anonymity of individual subjects within a dataset, and so access to microdata (data pertaining to individual subjects) may be restricted, and some data fields may be bucketed or omitted entirely from a microdata release. Second, if researchers wish to avoid sampling a population and gathering their own data, a process which can be very expensive in terms of time and other resources, and instead rely upon existing datasets, they may be limited by the sample size and the selection criteria, and be left with a highly restricted dataset while trying to address a question which is applicable to a much broader population. Both of these problems are usually relevant to researchers who are addressing public policy questions, where there is a desire to look at how different demographic and geographic groups across an entire country are affected by a particular problem, but survey datasets do not include geographic fields of the required resolution, and sampling across demographic and geographic strata may be inconsistent. This presents a problem for my research because in order to enable the dynamic simulation of the health system within a given Local Authority, it is necessary to have access to high-quality microdata which details the health and socioeconomic characteristics of the individuals who live there.

To address these problems in observational studies, researchers from various disciplines have developed many methods over the years for the generation of artificial microdatasets. Many terms have been used to describe these methods, including "population synthesis", "small area estimation" and "static/spatial microsimulation". "Microsimulation" refers to any technique which is intended to simulate data at the micro-level, or the level of individual subjects, and includes both static and dynamic methods. Static microsimulation methods are oriented towards the generation or derivation of microdata at a single point in time, whereas dynamic microsimulation includes methods which allow the properties of subjects to change over time, perhaps in response to policy options or systemic shocks, much like an agent-based model, but without interactions between subjects. Spatial microsimulations are methods which take spatial scale and the geographic placement of subjects into account, and usually "spatial microsimulation" is a shorthand for "spatial static microsimulation", referring to methods which can be used to generate a geographically distributed artificial population. "Small area estimation" refers to techniques which are able to estimate the distribution of demographics within smaller areas than that for which data is readily available, such as census tracts smaller than for which data is provided in a public release of census data. Small area estimation techniques do not necessarily generate microdata, but as they are able to estimate marginal and joint counts of different demographic groups over a range of areas, they can be used to generate an empirical distribution from which microdata can be sampled. "Population synthesis" is a more generic term which does not necessarily include a geographic component, although within health, geography, and transport this is almost always the case. These terms are largely interchangeable and their use seems to be more an artefact of preferred terminology within particular research groups and traditions, rather than a hard delineation of different approaches, and many specific methods can be labelled with all of these terms.

Due to the large number of terms and their use across many distinct disciplines, it is possible that researchers working with one more of these methods are not aware of other developments or new methods in different fields, and so there may be an opportunity to bring together the expertise and body of knowledge around these methods across all disciplines. There is an existing body of review literature on these methods (Hermes and Poulsen 2012; O'Donoghue, Morrissey, and Lennon 2013; Rahman 2017; Smith, Heppenstall, and Campbell 2021; Tanton 2013). However, each of these reviews have been conducted with only a subset of the known relevant terminology. Additionally, only one of these reviews has been systematically conducted and includes a search methodology, but it was specifically restricted to the discipline of health (Smith, Heppenstall, and Campbell 2021). Therefore, a systematically conducted and cross-discipline scoping review would be a useful contribution to the existing literature, and would benefit other ongoing research projects such as SIPHER (Meier et al. 2019) which make extensive use of these methods.

In addition to being the most thorough systematically conducted literature review of these methods to date, this study will also pioneer the use of the scoping review as a tool for discovering methods and techniques across disciplinary boundaries.

2.2 A comparison of artificial microdataset methods

When any research or statistical method is used to inform policy decisions, it is vital that the assumptions embodied within it, and the conditions under which it produces, or fails to produce, accurate results, are well-understood by the researchers using the method, and able to be adequately communicated to consumers of any analysis dependent on this method. This is because if the method, or the results of an application of this method, are used to inform policy decision in a context which is not appropriate given the assumptions and behaviours of the method, the result is likely to be an inaccurate policy recommendation, which may be very expensive and result in harmful outcomes for the public.

Simulation studies are used to understand and evaluate the behaviour and properties of statistical methods, particularly when closed-form analytical solutions are difficult or impossible to obtain. A numerical approach is particularly valuable when investigating a method which is most naturally expressed algorithmically rather than as a series of equations: there is no need to represent the algorithm in algebraic form, instead we can simply apply the algorithm to a known dataset and observe its behaviour directly.

Evaluation studies of statistical methods can be categorised as either simulation studies, empirical studies, or hybrid studies. Simulation studies rely on a known data generating process (DGP) which is used to simulate artificial data, and these simulated artificial data are then fed in to the statistical methods under examination. Empirical studies instead rely on empirical data gathered from the real world. Finally, hybrid studies make use of both simulated and empirical data, often as two separate analyses within a single study, to examine the statistical methods of interest.

Simulation studies have largely the opposite advantages and disadvantages to empirical studies. Because the ground-truth DGP is known in a simulation study, it is possible to make very specific associations between the outputs of the statistical method and the synthetic input dataset, and these associations can be tied back to the underlying DGP itself. Because the DGP is known, this allows us to reason about how the behaviour of the statistical method under examination relates to the DGP, the distributions of different variables, and the relationships between them, and gain additional insight about the behaviour of the statistical method under different scenarios. Additionally, a single DGP can be used to simulate many datasets, which allows us to adopt a Monte Carlo approach and perform a more robust analysis, with confidence that the behaviour of the statistical method is invariant to outliers or other anomalies in the simulated data. Further, because the DGP is known, it can be manipulated to allow investigating the behaviour of the statistical method under different conditions, allowing us to draw additional conclusions about its behaviour across different contexts. Simulation studies are also immune to the data collection issues present in empirical studies, because by definition the data is simulated rather than collected.

To provide a thorough treatment of a given statistical method using a simulation study, and truly 'put it through its paces', requires the development of an artificial DGP of sufficient complexity to generate data containing the nuances we would expect in empirical data. Multilevel, nonnormal, and multidistributional data are common in many disciplines including ecology, economics, health, and the social sciences, and so investigation of the behaviour of statistical methods against these types of data are of interest to large communities of researchers. However, it is very common to see simulation studies conducted using DGPs consisting of a small number of normally distributed variables, although this does vary across different methodological communities. For example, Mahlke, Schultze, and Eid (2019) undertook a simulation study of a novel multilevel structural equation model, which necessitated the simulation of multilevel data, although they also assumed that all variables were normally distributed.

Following on from the previous scoping review, this statistical simulation study will evaluate the performance of a range of artificial microdataset methods. The objective of this study is to test popular methods against a known ground truth and in so doing, to determine how accurately these methods are able to replicate the true joint distribution of health and socioeconomic characteristics at the local and national levels. This work will fill a significant gap in the current knowledge around artificial microdataset methods, by quantifying the limitations of these methods in a reproducible and extensible manner, and bringing scrutiny to a set of advanced statistical techniques which are ubiquitous across many areas of policy-making.

My approach builds on the non-parametric directed acyclic graphs (DAGs) common in the causal inference literature, by representing a DGP as a fully parameterised and executable DAG, where each node in the DAG represents a single variable sampled from an arbitrary but known distribution,

and where incoming arcs to a node represent an arbitrary parameterisation of the distribution for that node subject to the sampled values of the other variables from which these arcs are arriving. This approach provides numerous advantages. First, it provides an exact and intuitive description of the DGP and describes each variable in relation to the variables it is dependent upon. Second, it requires the researcher to consider the specifics of their DGP and encourages them to ground the DGP in a concrete domain, with reference to the real world, which further aids intuitiveness. Third, the approach is extremely flexible and modular, providing the researcher with complete freedom to choose any distribution for each variable and to parameterise it in any way. Fourth, it allows the simulation of multilevel data which arises naturally in the forms of households, schools, and neighbourhoods. These clusters constitute essential sources of ecological heterogeneity which must be replicated in the outputs of artificial microdataset methods.

After development of the DGP, data will be generated and then variables will be split into multiple datasets and masked, to emulate the typical condition where a researcher has access to a census and a survey, with limited geographic granularity and demographic data in each of these sources. Prominent artificial microdataset methods will then be used on these masked datasets to attempt to reconstruct the original ground truth dataset from the DGP, and then statistical evaluations will determine the goodness-of-fit of the recovered dataset to the original.

This study will allow me to determine the most appropriate method(s) to use for the following sub-projects, as well as demonstrating the assumptions inherent in these methods and the circumstances under which they become unreliable or fail. This study will be the most comprehensive evaluation of artificial microdataset methods to date, and will also include the development of a novel approach for constructing complex, multilevel datasets for statistical simulation studies.

2.3 Constructing a synthetic population for local public health research

In order to evaluate the prevalence, causes, and consequences of 'obesity' at the Local Authority level, it is necessary to have access to local statistics about the marginal and joint distributions of relevant variables including demographic (age, ethnicity, …), socio-economic (household income, deprivation, …), and anthropometric and health (weight, height, blood pressure, incidence of diseases such as diabetes and cardiovascular disease, …) factors. Further, in order to evaluate the potential effects of health and 'obesity' interventions, this population data must be combined with information on the food and exercise landscape, such as access to nutritious and fast foods, and access to leisure and exercise environments such as parks and gyms. The creation of bespoke datasets to capture all of this information would present a significant cost to Local Authorities in terms of both time and money, and as a result, it is of interest to both researchers and policy-makers to make use of artificial microdataset methods to generate a synthetic local population which accurately models all of these variables.

Using the most appropriate method(s) from the preceding comparison study, data from the UK Census, UK Biobank, Health Survey for England, Scottish Health Survey, and Welsh Health Survey will be combined to generate a synthetic population of Great Britain, which will consist of millions of individuals, each with specific attributes across this wide range of variables. This synthetic population can be used for both static and dynamic analyses. A static analysis would examine the current prevalence of different health determinants and outcomes, such as the prevalence of economic deprivation or limited access to exercise facilities, or the prevalence of cardiovascular disease or diabetes, within the local population. This point-in-time snapshot of the population could then be used as an input to a causal inference (CI) model, to allow researchers and policy-makers to develop an understanding of the causal relationships between determinants such as household income, mediators such as 'obesity', and health outcomes such as hypertension, cardiovascular disease, or mortality. Dynamic analyses would make use of systems simulation methods such as agent-based models (ABMs), dynamic microsimulation models (MSMs), or systems dynamics (SD) models, to project the change in these variables over time, and to investigate the effects of potential interventions on these health trajectories. For example, it may be that development of new gyms in 'exercise deserts', along with financial incentives to enable people on low incomes to access these facilities, could have a substantial effect on the prevalence of 'obesity' and cardiovascular disease within the Local Authority, and this investment would then translate into health savings over the next several years.

This study will be the first time a synthetic population has been generated from the UK Biobank dataset, and the resulting methodology and dataset will be of significant value to the public health and epidemiological communities, due to to the large number of covariates in the UK Biobank.

Additionally, this study presents the option to develop a set of inverse probability weights for the UK Biobank, which will allow researchers who are working exclusively with this dataset to reweight the UK Biobank sample to better fit the UK population, which would make their studies more generalisable.

2.4 Extending local whole systems approaches with participatory modelling methods

As indicated by the literature review on systems thinking and participatory modelling, there is a lot of scope for how participatory action research (PAR) and group model building (GMB) can be applied as participatory methods for the development of systems models. Further, GMB is mostly applied for the development of causal loop diagrams (CLDs) or systems dynamics (SD) models, and Koh, Reno, and Hyder (2018) remark that further work is needed to develop suitable scripts to facilitate the use of GMB for agent-based modelling (ABM). This is also likely to be true for structural causal modelling (SCM). These GMB approaches must also be incorporated into the existing WSO programme and they may extend the delivery of the programme by multiple days, and so it is necessary to consider how GMB workshops can be delivered in the most effective and efficient way. A primary outcome of any participatory modelling process should be the evaporation of at least some of the fog surrounding the complexity of the system: by taking stakeholders through the process of model construction and presenting results using appropriate visualisations and narratives, stakeholders should develop a better understanding of the system structure and behaviours, and gain value from the participatory modelling process itself, as well as from the final outputs of the simulation model and its results.

In addition to the extension of the WSO programme with GMB workshops, additional research is required to inform which types of models are most applicable to different research questions and policy proposals arising throughout the WSO process, and it is of interest to determine the feasibility of translating between different models, deriving one type of model from another, or developing hybrid models that incorporate multiple methods or frameworks. The objectives of this approach would be threefold. First, these methods would save time, by reducing the amount of GMB work required to develop an appropriate ensemble of models. Second, this would allow researchers and stakeholders to leverage all available analytical tools from across the causal inference and systems simulation spectra and to develop greater insights than if a single model had been used. Third, an ensemble of models serves as a form of internal validation, where we can determine if models produce the same outputs given the same set of inputs, and if they respond to policy options in the same way. This would provide all stakeholders with greater confidence that the different perspectives used in understanding the system are consistent with each other.

Successfully application of participatory modelling methods within the WSO programme will require several pieces of qualitative research. First I will need to talk to Local Authorities who have previously utilised, or intend to utilise, the WSO programme, to understand their objectives and needs both broadly around the WSO programme, and more specifically around co-production and dynamic policy simulation. Second, I will need to liaise with the original authors of the WSO programme, to understand how these new methods can be best assimilated into the broader programme and ensure that the quality of the user experience and the programme outputs are enhanced by the addition of participatory modelling methods. Third, I will need to undertake a thorough review of the existing literature around co-production, participatory modelling, and iterative/agile simulation modelling, particularly as they are applied to SCM and ABM models which are not usually constructed in a participatory manner, to develop a specific approach which can be implemented. This approach must be flexible so that as it is implemented with Local Authorities undertaking the WSO programme, it can be adapted over time and to fit the differing needs of different Local Authorities. Finally, development of techniques for translation, derivation, or hybridisation of models will likewise require an appropriate literature review, and may precipitate further method development research.

2.5 Co-production of local whole systems models of 'obesity' and public health

Once the WSO programme has been extended to include a participatory simulation modelling component, it will then be possible to directly deploy this extended WSO programme with various Local Authorities. Along with other researchers, I will undertake at least one full cycle of the WSO programme with at least two Local Authorities, ideally undertaking additional implementations if

there is sufficient time within my PhD programme. As well as resulting in the delivery of the WSO programme, including delivery of its usual outputs such as action plans, this will also serve to allow evaluation of the impact of the participatory simulation modelling component.

2.6 Evaluating the impact of participatory modelling on the local whole systems approach

As my fieldwork concludes, I will synthesise the learnings from the various implementations of the extended WSO programme to evaluate the overall impact of the participatory modelling extension. This work represents the final phase of a larger iterative process around the entirety of my core research, from which it will be possible to develop recommendations for future development of the WSO programme specifically; and dynamic policy simulation and participatory modelling more generally.

3 Literature review

The following sections present a narrative literature review, grouped by primary theme.

3.1 Critical perspectives on 'obesity'

Before I proceed to discuss the literature which is directly relevant to my research, it is first important to provide an overview of critical perspectives on 'obesity' and 'obesity' research. As researchers, the knowledge we wield is a form of power (Gutting 2005) and it is essential that we are able to examine ourselves and our work with an unflinching gaze, lest we (inadvertently) perpetuate oppressive and marginalising structures, process, and attitudes. This is in keeping with the motivations behind, and approach of, Research Justice (Jolivétte 2015).

The terms 'underweight' and 'overweight' are problematic because they present a normative account of body weight and body shape. The terms 'obese' and 'obesity' are also problematic due to their etymology, from the Latin *obedere*, which connects 'obesity' to individual eating habits ("Obese, Adj. And n." n.d.). All of these terms, including 'fat', have been found to have negative connotations in a clinical context, and it is recommended by Puhl (2020) that clinicians should use neutral terminology to describe weight. At the same time, fat scholars and fat activists are reclaiming the term 'fat', in the same way that the LGBTQIA+ community has reclaimed the term 'queer', and this is discussed in multiple fat studies works including Rothblum and Solovay (2009) and Lupton (2018). Throughout this work I have chosen to use the term 'obesity', but to place it in 'scare quotes' as is common in the fat activist literature. This situates my work within the disciplines of (public) health, ('obesity') epidemiology, and bariatrics, and makes use of the standard terminology within these disciplines, while also making explicit the controversial nature of this term; the lack of suitability of this term in a clinical context; points readers towards critical perspectives; and encourages the development of new terminology within these disciplines, which should be inclusive and non-normative.

The classification of 'obesity' as a disease is controversial. Meldrum, Morris, and Gambone (2017) remarks that the "American Medical Association voted in 2013 to recognize obesity as a disease (against the advice of its Public Health and Science Committee)". Boero (2012) traces the history of the American conceptualisation of 'obesity' as a disease through the lens of 'medicalisation', that is, "the processes by which an ever wider range of human experiences comes to be defined, experienced, and treated as a medical condition". Although this process has been undertaken by clinicians and researchers with good intentions, to destigmatise 'obesity' and begin to recognise it not as an individual failing but as an unwanted health condition, the medicalisation of 'obesity' also pathologises fat bodies and treats them as an 'abnormality' which must be 'cured', and this has significant negative effects on people's mental health and social relationships. This is discussed in further detail by Gard and Wright (2005).

Likewise, the classification of trends in 'obesity' as an epidemic or a pandemic is also controversial, and has been analysed through the lens of moral panic by Campos et al. (2006) and Boero (2012). These works document how the framing of 'obesity' as a 'crisis' precipitates further stigma and discrimination against fat people, results in further damage to their mental health, and serves the interests of other actors.

Fat activists point to the vested interests of different actors in the world-system and how this

complicates the science around weight, fat, and 'obesity'. Heyes (2007) discusses how the weight loss industry thrives on the continued existence of individuals who want to lose weight but are unable to do so, which positions the profit motive of these companies in direct opposition to the well-being of fat people and a positive relationship with their own bodies. Likewise, as 'obesity' researchers, our academic statuses are predicated on research and policies which may cause harm to individuals and their relationships with their bodies, and thus it is absolutely essential that we are not only receptive to, but actively seeking, critical perspectives on our own work, and that we are able to recognise and dismantle power structures which also harm ourselves and our capacity for scientific objectivity and impartiality in our work.

Weight stigma is a form of bias and discrimination faced by people with non-typical weights, and it occurs at many different levels beyond that of individual prejudice. Hill et al. (2021) find that weight stigma is pervasive at the policy level, and Chrisler and Barney (2017) show how 'sizeism' limits people's access to timely and appropriate medical care, in addition to the direct effects on mental health and well-being.

Health is not a monolith, which is to say that there is no single variable or scale on which a person's total health can be measured. Rather, health is a qualitative composite of the status of many different domains of the body and the human experience, and it is more accurate to talk about 'cardiovascular health', 'mental health', and 'joint health' as more specific types of or factors of a person's health. In addition to being a more precise way to discuss health, this post-monolithic approach to health can also help to eliminate weight stigma in clinical contexts, as shown by Tylka et al. (2014).

3.2 Causal inference in health and policy-making

Causal inference is an extension of statistical methods which allow researchers to move beyond correlation and to be able to answer causal questions. The phrase "correlation does not imply causation" is well known to researchers: just because the branches are moving when the wind is blowing, does not mean that it is the moving tree branches which are causing the wind! Despite this, it is common for scientific literature to shy away from making causal statements and instead to talk about "associations" and "links". This can result in research consumers, such as policy-makers, the media, and even other researchers, for drawing incorrect or inappropriate conclusions. Hernán (2018) discusses this and points out that the work required to make causal statements, including developing valid causal questions and adjusting for confounding, is neither easy nor simple. Fortunately, we now possess the tools to be able to undertake robust causal analyses.

Both researchers and policy-makers are fundamentally interested in causal questions, which may be of an interventional ("if I do X, will it cause Y?") or counterfactual ("if I *didn't* do X, would Y *still* have happened?") nature. These differ from associational and predictive questions, which instead are focused on the probability of one observation, given another observation ("if I observe X, how likely am I to observe Y?"). Returning to the prior example with the wind and branches, we may observe a correlation between moving branches and wind, and this is sufficient for us to be able to predict that it is windy outside if we can see the branches are moving. However, it does not mean that if we stop the branches from moving, that the wind will subsequently stop blowing! Hernán (2018) provides another example in terms of the relationship between red wine consumption and coronary heart disease:

> Suppose we want to know whether daily drinking of a glass of wine affects the 10-year risk of coronary heart disease. Because there are no randomized trials of long-term alcohol drinking, we analyze observational data by comparing the risk of heart disease across people with different levels of alcohol drinking over 10 years. Say that this analysis yields a risk ratio of heart disease of 0.8 for one glass of red wine per day versus no alcohol drinking. For simplicity, disregard measurement error and random variability – that is, suppose the 0.8 comes from a very large population so that the 95% confidence interval around it is tiny.
>
> The risk ratio of 0.8 is a measure of the association between wine intake and heart disease. Strictly speaking, it means that drinkers of one glass of wine have, on average, a 20% lower risk of heart disease than individuals who do not drink. The risk ratio of 0.8 does not imply that drinking a glass of wine every day lowers the risk of heart disease by 20%. It is possible that the kind of people who drink a glass of wine per day would have a lower risk of heart disease even if they didn't drink wine because, for

example, they have high enough incomes to buy, besides wine, nutritious food and to take time off to exercise, or have better access to preventive health care.

In other words, the risk ratio of 0.8 may be an unbiased measure of the association between wine and heart disease, but a biased (confounded) measure of the causal effect of wine on heart disease. Importantly, we knew this before conducting the study.

The distinction between predictive and causal analyses becomes significant when we consider how we intend to use the results, because interventions aim to create knowable changes in an outcome based on causal knowledge. If the association between red wine consumption and coronary heart disease is purely predictive and not causal, then mandating consumption of red wine will not reduce the incidence of disease. This is not to say that predictive knowledge is not also useful, but rather that public policy and health interventions are implicitly concerned with causal questions, and require causal analyses in order to be effective. Because it is often infeasible or unethical to develop randomised controlled trials (RCTs) in public policy and health research, policy-makers and researchers are dependent on observational data and observational studies, which further complicates causal analysis, because we are unable to directly control factors which may affect our inferences.

3.2.1 The target trial framework

One option available to observational researchers from the causal inference toolbox is the target trial framework, which is explained in Hernán and Robins (2016) and Kutcher et al. (2021). A target trial is a hypothesised randomised controlled trial (RCT) which is used to inform the design of an observational study, and which also exposes the considerable differences between experimental and observational research design. Returning to the wine example, we may ask what an ideal RCT would look like for investigating the effect of red wine consumption on coronary heart disease (CHD).

First, we would need to choose some eligibility criteria for selecting participants into the study, which might be individuals aged 18 to 30 with no prior history of CHD or red wine consumption. This allows us to avoid the fallacy of *'conditioning on the future'*, wherein we develop a selection procedure for our observational study which is dependent on information we could not have known prior to the exposures or outcomes of interest having occurred. In this case, our chosen eligibility criteria prevents us from examining only individuals who eventually develop CHD, which would be equivalent to an RCT trying to recruit people "who will develop CHD": this is impossible because nobody knows what their health outcomes will be in the future!

Second, the target trial framework requires us to specify our treatment strategies, which in the case of our wine-CHD research question would be a series of daily doses of red wine to be consumed, including no wine. Careful consideration of the treatment strategy is important to prevent *healthy user bias*, which occurs when there are structural differences between different treatment groups.

Third, we must define our assignment procedures. In an RCT this would be a truly random assignment, but this is not possible in an observational study. Instead, we seek to identify all potential confounders – variables which may have some effect on both the exposure and outcome – so that we can control for these in our analysis using propensity score methods (discussed later), stratification, regression, or other suitable techniques. In our example, socioeconomic position, household income, and ethnicity are just three of potentially many variables which may affect both red wine consumption and CHD. We may not have data for all of our confounders, but by identifying them we are able to comment on the potential for bias; perform sensitivity analyses to bound our results; and inform future research and surveillance by indicating what data are needed to conduct more robust analyses.

Fourth, we must specify the follow-up period. The beginning of the follow-up period is referred to as 'time zero' and prior to this point all subjects must have met the eligibility criteria. Further, it is only within the follow-up period that subjects must be exposed to the intervention and be able to produce the outcome of interest. For our red wine example, a suitable follow-up specification would be "from randomisation until CHD, death, loss to follow-up, or 10 years after baseline, whichever occurs first". Careful specification of the follow-up period allows us to avoid *immortal time bias*, which occurs when there is a gap between treatment assignment and measurement of outcomes. During this period, it is possible for subjects to be selected for treatment and then die, in which case they would not be included in the final sample, and this will then inflate the survival of the treatment group. There are analytical methods that can be used to eliminate immortal time bias,

but many of these methods require discarding data and all impose additional work during the analysis phase, and so the target trial framework encourages us to choose an appropriate follow-up period as part of the study design.

Fifth, we must define the outcome we are interested in, which for our example would be a diagnosis of CHD within the follow-up window. In an RCT, the treatment status of each participant is kept secret which prevents any external factors from affecting the outcome, but in an observational study this is rarely the case. In our example, the participants of our observational study will know how much wine they are consuming, because they are not in the controlled environment of the hypothetical target trial, where we might be able to give some treatment groups a 'placebo' which looks and tastes like wine but has a different chemical composition. This knowledge may cause participants or people around them to change their behaviour around measurement of the outcome. For example, doctors may be more likely to test for CHD in people who drink no wine, or who drink large amounts of wine, compared to people who drink a small amount of wine. In order to mitigate this, observational studies must carefully consider the provenance of the outcome data and seek an externally validated and well-defined process for monitoring the outcome during the follow-up window.

Sixth, we must specify the causal contrasts of interest, which we are seeking to estimate via our analysis. This is equivalent to defining a robust causal question for our analysis, and this causal effect of interest is sometimes termed the *estimand*, meaning "the thing to be estimated". In the context of a target trial, two common options are the intention-to-treat effect and the per-protocol effect. The intention-to-treat effect is the effect of assignment to a treatment protocol, regardless of how well individuals in that treatment group follow the specified protocol. The per-protocol effect is the effect of correctly adhering to an assigned treatment protocol for the duration of the study. These are often different, especially in studies which occur in uncontrolled non-laboratory environments, because it is not possible for researchers to ensure that subjects are perfectly adhering to their assigned treatment protocol. For our wine example, there is no clearly interpretable 'intention to treat' or treatment assignment, unless participants have been advised by their doctors to drink a set amount of wine! Therefore for our example, only the per-protocol effect would be measured.

Finally, we must specify our analysis plan. We define the methods we will use to determine the causal contrasts of interest, and how we will control for confounders. This will require adjustment for baseline confounders for both intention-to-treat and per-protocol effects, and also for ongoing confounding in the case of per-protocol effect. These adjustments will make use of other methods from the causal inference toolbox such as propensity scores and/or g-methods (discussed later).

The target trial framework has started to see good uptake within the scientific community: at the time of writing Hernán and Robins (2016) has 452 citations within Scopus. However, only 10 of these results are tagged with the keyword "obesity", indicating an opportunity for much greater adoption of this tool by the 'obesity' epidemiology community. Examples of use of the target trial framework within 'obesity' research include Katsoulis et al. (2021), which investigates the effects of weight change on cardiovascular disease (CVD) outcomes; and Ervasti et al. (2019), which investigates the effect of increased physical activity on reduction of work disability in 'overweight' and 'obese' people.

3.2.2 Structural causal models (SCMs)

Pearl (2010) introduces the structural causal model (SCM) approach, which in particular advocates the use of non-parametric directed acyclic graphs (DAGs) as an essential step in the statistical analysis process. Causal DAGs model the underlying data generating process (DGP) which exists in the world, and are designed by the researcher using their existing domain knowledge in order to allow them to answer a specific causal question. Each node in the DAG represents a specific variable or event, and a well drawn DAG orders these events in chronological order, which allows the researcher to avoid *reverse causation* in their analyses. Nodes are connected with arcs to represent that one event may be a cause of, or otherwise have some effect on the outcome or nature of, another event. This means that the *lack* of an arc in a DAG represents a strong assumption that there is *no* causality. Using the 'rules' of SCM, researchers can use their DAG to determine an appropriate adjustment set for a regression analysis which allows them to avoid spurious correlations and biases such as collider bias (Cole et al. 2010).

Figure 2 shows a causal DAG for our red wine example. Let's assume that the total amount of red wine consumed by each subject during follow-up, along with other time-varying measures, has been

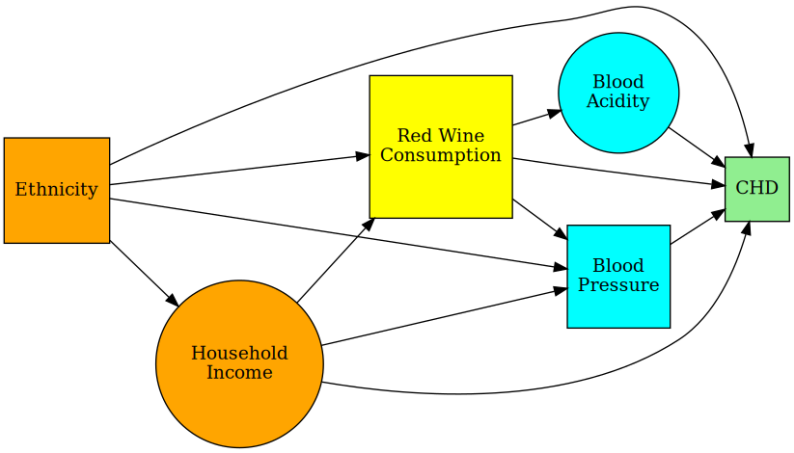

Figure 2: An example of a causal directed acyclic graph.

reduced to a single value, a total or a weekly average or similar, labelled Red Wine Consumption, in yellow, and that this is our *exposure*. CHD, in green, is our *outcome*, and Blood Acidity and Blood Pressure, in blue, are *mediators*, which represent particular causal pathways for the effect of Red Wine Consumption on CHD. Ethnicity and Household Income, in orange, are *confounders*, which have some effect on both the exposure and the outcome, and thus will bias the causal effect of Red Wine Consumption on CHD unless we control for them. Square nodes represent observed covariates whereas circular nodes (Household Income and Blood Acidity) represent unobserved covariates. Based on this DAG, we would be able to estimate the effect of Red Wine Consumption on CHD with a (logistic) regression CHD $\sim$ Red Wine Consumption + Ethnicity, where we control for Ethnicity using stratification or one or more dummy variables.

The DAG has several implications for our analysis. First, our DAG indicates that we should not control for Blood Pressure in our model, because this variable is on a causal path from Red Wine Consumption to CHD. If we were to control for Blood Pressure, this would remove one of the causal pathways from our analysis and would result in an underestimate for the effect of Red Wine Consumption on CHD. Importantly, whether we can control for Blood Pressure or not is dependent on the temporal ordering of the variables. If our Blood Pressure variable represented a single measure taken at baseline, prior to consumption of red wine, then it would not be a mediator but rather a confounder, and so it would be suitable to adjust for. However, in this example, Blood Pressure is measured throughout the follow-up period and will be affected by Red Wine Consumption, and therefore we do not want to exclude the effect of Red Wine Consumption on Blood Pressure as a possible causal factor in development of CHD.

Second, once we have estimated the coefficients or other parameters for our model, our DAG indicates that only one of these coefficients is interpretable in the context of the causal estimand of interest, and thus helps us to avoid a *table 2 fallacy* (Westreich and Greenland 2013). Assume that we had performed the aforementioned regression of CHD $\sim$ Red Wine Consumption + Ethnicity, and we present a table in our study demonstrating the coefficients for Red Wine Consumption and Ethnicity. Although we can interpret the coefficient for Red Wine Consumption as an estimate of the causal effect of this variable on CHD, we cannot interpret the coefficient for Ethnicity in the same way. This is because, if Ethnicity were our exposure, then Red Wine Consumption would become a mediator, and so conditioning on Red Wine Consumption would block some of the causal pathway between Ethnicity and CHD. These tables of coefficients are usually presented as the second table in a study, following on from the descriptive statistics of the different treatment groups, hence why this mistake is known as a table 2 fallacy.

3.2.3 Deterministic variables

Deterministic variables are variables whose value can be computed precisely as a known function of other variables. Examples include total calorie consumption, as a sum of carbohydrates, protein, and fat; BMI, as the ratio of body weight in kilograms to the square of body height in meters; weight class ('normal weight', 'overweight', 'obese', …) as stratifications of BMI; and change scores, as the difference between or ratio of two measures of the same property (weight, height, …) taken at different times. Deterministic variables are problematic in causal analyses because they possess

very strong and tautological associations with their determinants, which can introduce bias into estimates of causal effects while simultaneously giving the impression that a variable has been controlled for.

Tomova et al. (2021) examines the issues around deterministic variables in the context of energy intake in dietary studies. A simulation study is presented which models the effect of sugar consumption on fasting blood plasma glucose (FBPG). Additional dietary components such as saturated and unsaturated fats, alcohol, and protein, are split out as additional dietary exposures, all of which contribute to the total energy intake, a deterministic sum of all dietary components. The authors then investigate six different regression models for estimating the effect of sugar consumption on FBPG. The first model is uncontrolled; the second model adjusts for remaining energy intake, that is calories which come from other sources besides sugars; the third model adjusts for total energy intake, which includes sugars; the fourth model is regressed on the ratio of sugars to total energy, as opposed to being regressed directly on the calories from sugars; the fifth model uses a two-stage approach where first sugars are regressed against total energy and then the residuals are used as the covariate in the FBPG model; and the sixth model independently controls for each dietary component. The first, second and sixth models seek to estimate the total causal effect of sugar on FBPG, which is the effect caused by changing consumption of sugar while keeping intakes from all other sources constant, which in turn leads to a higher total energy consumption. This is different to the average relative causal effect, which is the effect caused by changing consumption of sugar on FBPG while adjusting intakes from other sources so as to keep the total energy consumption constant, and this is the effect sought by models three to five. These difference between these estimands may seem subtle, but this has important consequences for how results of a study can be interpreted and should be translated into clinical or policy guidance or recommendations. The sixth model is able to estimate both total and average relative causal effects, since it is able to adjust for all covariates separately. The simulation study discovered that only the sixth model was completely robust in the presence of confounding and was able to reproduce the true estimate for both effects. In the face of confounding, bias was particularly pronounced for the fourth model, wherein sugars were represented as a ratio of the total energy intake, unless total energy was included as an additional covariate. This demonstrates that taking ratios of variables does not necessarily control for the denominator covariate, and can lead to biased estimates, which is of particular concern to 'obesity' research which relies on BMI as a covariate, instead of estimating the effect of weight while controlling for height as a separate variable.

3.2.4 Time-varying exposures, time-varying confounding, and the g-methods

In our earlier red wine example we made a simplifying assumption that the exposure could be reduced to a single event, but this is unrealistic because people's quantity of wine consumption, and the effects this will have on blood pressure and blood acidity, vary over time. Figure 3 demonstrates a DAG for investigating the effect of exercise on diabetes, in the presence of diet and weight confounders. In this DAG we consider the variables Exercise, Diet, and Weight, at three different points in time.

Now that we have unrolled the Exercise variable as a time-varying exposure, our non-baseline Diet and Weight variables serve two roles: when Diet at time $t = 1$ is affected by Exercise at time $t = 0$, Diet is a mediator on a causal pathway towards Diabetes; but when Diet at time $t = 1$ affects Exercise at time $t = 1$, Diet is a confounder which is biasing the causal effect of Exercise! We colour these nodes with a blue-orange gradient to indicate that their status as both a mediator and a confounder varies, depending on which time step we are examining. In order to eliminate this time-dependent confounding, we will need to use a method from the family of 'g-methods', short for 'generalised methods'. An introduction to the g-methods are given by Naimi, Cole, and Kennedy (2017), but I will discuss the application of one method to a question involving 'obesity'.

Shakiba et al. (2018) investigate the effect of different types of 'obesity' on coronary heart disease (CHD) in the presence of time-varying confounding. This study used data from the Atherosclerosis Risk in Communities (ARIC) Study, a prospective cohort study of 15,792 subjects of 45-64 year of age across 4 communities in the US. Participants were examined 4 times. The exposures of interest for Shakiba et al. (2018) were general 'obesity' (GOB), defined as BMI ≥ 30 kg/m^2; and abdominal 'obesity' (AOB), defined as waist circumference ≥ 102 cm in men and ≥ 88 cm in women, and waist-to-hip ratio ≥ 0.9 in men and ≥ 0.85 in women. Confounders in the model include typical socio-economic factors such as age, sex, race, and education level; lifestyle factors such as calorie intake, physical activity level, and smoking and drinking statuses; and presence of health conditions

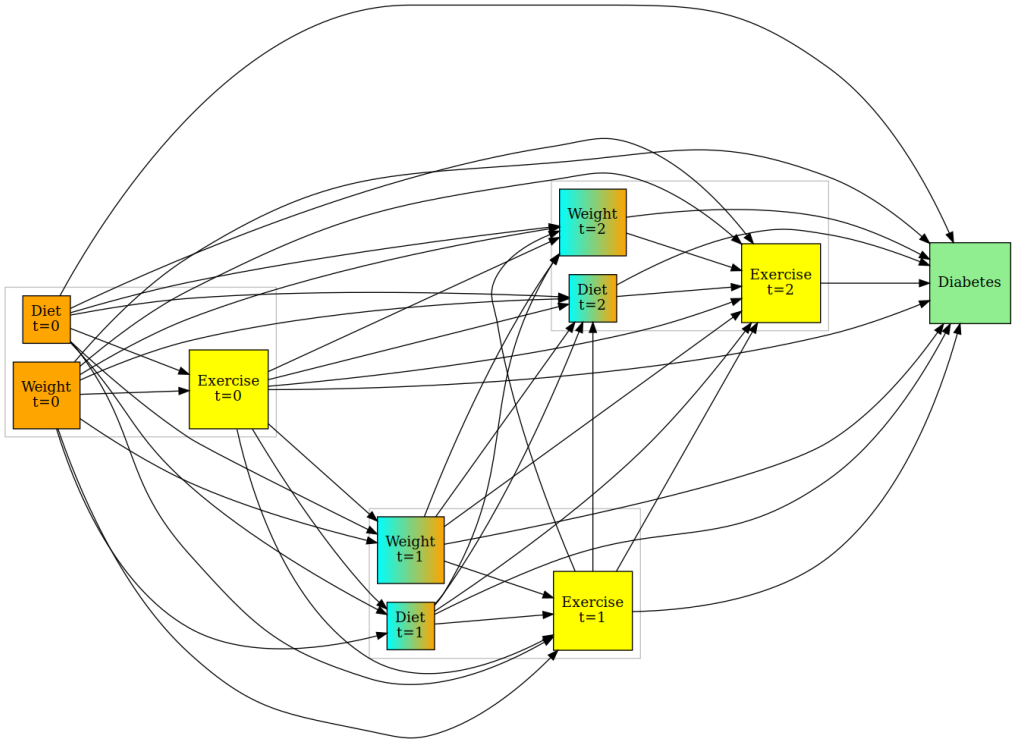

Figure 3: An causal directed acyclic graph demonstrating a time-varying exposure. In this graph, we observe Diet, Weight and Exercise at three different points in time, and we are interested in the effect of Exercise on Diabetes. Exercise at one point in time will affect Weight and Diet at the next point in time, and so these nodes serve as mediators of the causal effect of Exercise on Diabetes. However, Diet and Weight at an earlier point in time act as confounders which change the level of Exercise. This dual confounder-mediator relationship is known as time-varying confounding, and these nodes are shaded with a gradient to represent their nature depending on if they are upstream or downstream from the exposure at each point in time.

such as hypertension or diabetes. It was assumed that at each point in time, in addition to affecting CHD status at that time, 'obesity' status also affects other time-varying covariates, as well as the 'obesity' and CHD statuses at all later points in time. The model additionally included unobserved time-varying confounders, to account for covariates which were not measured. In order to estimate this model, the authors used a g-method called 'g-estimation' in tandem with a structural nested accelerated failure-time (SNAFT) model. The SNAFT is essentially a survival model where each follow-up visit constitutes the start of a new observational trial, 'nested inside' of the surrounding trials from the previous visits. g-estimation then proceeds by embedding the equations for this model inside of another equation which represents that the conditional dependence of the exposure based on the outcomes are zero. The authors compare their estimates from g-estimation with three standard regression approaches: a first model which does not include any metabolic mediators; a second model which does adjust for metabolic mediators; and a third model which is time-adjusted. For each measure of 'obesity' (BMI, waist circumference, and wait-to-hip ratio), the metabolic mediator model and time-adjusted model both reported small effect sizes with confidence intervals covering the null effect line (where the hazard ratio is 1), whereas the no metabolic mediator model and the g-estimation model both consistently report larger negative effect sizes, with only the g-estimation of BMI presenting a confidence interval which crosses the null effect line. The study demonstrates how controlling for mediators removes causal effects and results in underestimates of effect sizes, indicating the need in longitudinal studies for techniques like the g-methods which are able to account for this. Further, the study also demonstrates the importance of non-BMI measures for determining the potential health outcomes caused by 'obesity.

3.2.5 The 'obesity' paradox, selection and collider biases, and reverse causation

The 'obesity paradox' is the observation that in many studies relating measures of body weight or adiposity to mortality or other negative health outcomes, individuals who are 'overweight' may have better outcomes than those who are 'normal' weight. This shows as a dip in the typical J-shaped

risk ratio curve of BMI vs mortality. The 'paradox' is the cause of a great deal of speculation within the research community, and it is not clear to what degree it may be caused by statistical biases, or if the primary driver is an underlying protective benefit of certain distributions of fat against certain health risks.

Additionally, BMI-based definitions of 'obesity' are insufficient to capture the specific health effects of high levels of fat, because these effects are dependent on the distribution of fat. Piché et al. (2018) discuss the significantly different effects on health between subcutaneous 'obesity' and visceral 'obesity', and Burkhauser and Cawley (2008) discuss how different measures of adiposity affect the distribution of 'obesity' across different demographics. Together, these pieces of literature indicate that much more nuanced approaches are needed for measuring 'obesity' at the population level, and for diagnosing 'obesity' at the individual level, in order for public health professionals and clinicians to be able to target interventions at those people who are most susceptible to weight-related diseases and negative health outcomes.

BMI-based definitions of 'obesity' are particularly unsuitable for understanding the emerging phenomena of 'normal weight obesity' (NWO) and 'metabolically healthy obesity' (MHO). Blüher (2020) reviews the existing literature on MHO and concludes that "up to a third of people with obesity do not exhibit overt cardiometabolic abnormalities" and that both MHO and metabolically unhealthy 'obesity' (MUO) can be transient and that individuals can move between these two phenotypes over time. Bosello, Donataccio, and Cuzzolaro (2016) further discusses the NWO phenotype and indicates that these individuals, despite having a BMI $< 25\text{kg/m}^2$, are at significantly higher risk of cardiovascular disease and have higher mortality rates. These studies further reinforce the preceding conclusion, that more nuanced approaches to adiposity-related health consequences are essential. NWO is of particular concern from a public health perspective as it indicates that there may be a significant portion of individuals at risk of cardiovascular disease who are not being accounted for by national and international surveillance of 'obesity'.

Westreich (2012) discusses an observation by Berkson in 1946 that studies of exposure-outcome relationships in hospital in-patients are fundamentally biased. Consider an investigation in to the effect of the presence of a diabetes diagnosis on the development of a severe urinary tract infection (UTI) in hospital patients. Diabetes and UTI are competing reasons for visiting hospital: most people will have one of these problems, rather than both. Therefore, even if there is no causal relationship between diabetes and UTI, by selecting only individuals who are in hospital, a spurious correlation is introduced between diabetes and UTI, and it would appear that diabetes has a protective effect! Although Berkson's original paradox arose in a clinical cohort and within an interventional context, there are still lessons to be learned which are important to population-level observational analyses. This is because Berkson's paradox is a special case of two important biases: selection bias, and collider bias.

Selection bias refers to any condition where the selection criteria for the sample study biases the estimated effect. In addition to Berkson's paradox, where it arises as a result of stratifying on a variable caused by both the exposure and the outcome, selection bias can also arise simply by choosing a non-representative subset of the population, if we are trying to derive an estimate which is generalisable to that population.

Collider bias is caused by conditioning on a variable which would have multiple incoming arcs if drawn as a DAG, so called because the causal arrows 'collide' at this node. Berkson's paradox arises because both diabetes and UTI are causes of a hospital visit, making hospital visit a collider, which is then stratified on as part of the selection procedure. Hence this issue is also referred to as collider stratification bias.

Reverse causation refers to assignment of an incorrect causal direction between a cause and an outcome, such that the outcome precedes the cause. An example would be: does a sedentary lifestyle lead to weight gain, or does weight gain lead to a sedentary lifestyle? Often there are feedback loops in a system which cause both such statements to be true, but in order to undertake a causal inference analysis we must decide which of these effects we want to estimate, and structure our DAG appropriately. In cases where the feedback loop is believed to play a significant role in the causal chain, a time-varying analysis should be conducted as described in the prior section time-varying exposures, time-varying confounding, and the g-methods. Importantly, although cause and effect can oscillate in a time-varying analysis, we must still determine what the causal direction of the relationship is between these variables at each point in time: it is only through a longitudinal approach that we are able to allow a mediator at the present time to become a confounder at a future time. In the case of sedentary behaviour and weight gain, Ekelund et al. (2008) examined

this but did not make use of appropriate techniques such as the g-methods, which indicates a need for a greater awareness of these techniques with the 'obesity' epidemiology community.

A slightly different meaning of 'reverse causation' is sometimes used within the 'obesity' literature, referring to the effect of illness on the loss of body weight. Rather than being a true instance of reverse causation, where it is not clear to the researcher what the causal relationship may be between 'obesity' and disease, this is more accurately described as confounding, because these illnesses are developing prior to selection into study (Flegal et al. 2011). Such effects would serve as drivers for the 'obesity paradox', as the inclusion of ill individuals with diseases which cause significant loss of body weight, such as cancer, will result in a relatively better risk ratio for 'overweight' and 'obese' individuals.

It is likely that confounding due to disease (in the sense of weight loss caused by illness) and collider stratification bias (due to smoking also being a cause of cardiovascular disease, diabetes, cancers, etc. in addition to 'obesity') are major factors in the cause of the 'obesity paradox'. Stokes and Preston (2015) investigate these issues using the US National Health and Nutrition Examination Survey (NHANES) dataset. After restricting their sample to individuals with sufficient data for all covariates, Cox proportional hazards models for mortality were fitted to investigate the effect of smoking and changes in weight within a subset of the sample who were diagnosed with cardiovascular disease (CVD) prior to baseline. The hazard ratio for subjects from the CVD subsample with 'overweight' or 'obesity' at survey time is below 1 when compared with those with 'normal' weight at survey time, which is a clear instance of the 'obesity paradox'. However, this seeming protective effect is removed when the analysis is restricted to never-smokers, and also when individuals who have been 'overweight' or 'obese' prior to the survey are separated out from the 'normal' weight group who had never been 'overweight' or 'obese'. These results demonstrate the significant effect that reverse causation and collider bias can have on effect estimates.

Sperrin et al. (2016) present a mathematical treatment to show that collider bias is insufficient to fully explain the effect size of the 'paradox'. They derive a logistic regression from an idealised DAG which represents the effect of a binary exposure such as 'obesity' on a binary outcome such as mortality, in the presence of (potentially unmeasured) confounding. They conclude that the effect of collider bias must be large, compared to the true causal effect under investigation, in order for this to be the sole cause of the 'paradox'. However, they only investigate binary variables, which will result in information loss and lead to larger residual confounding. Additionally, they do not work with any simulated or empirical data, and so further research is needed to determine how this effect maps on to continuous models of weight vs mortality.

As a result of these issues — failure to adjust for relevant confounders; collider bias; BMI being a deterministic variable; the presence of time-varying exposures; reverse causation — consensus has yet to emerge on the relationship between BMI and mortality, and these studies indicate is an urgent need to bring causal inference methods further into 'obesity' research. As an example of the ongoing debate in this area, Angelantonio et al. (2016) conducted a large meta-analysis to reveal J-shaped risk-ratio curves of BMI vs mortality, but the methodology used was criticised by Flegal, Ioannidis, and Doehner (2019) on the grounds that the majority of the data available was excluded in a non-systematic manner. Although there are additional issues related to the original study, it seems clear that more transparent use of causal methods could have informed the design of the original study and addressed the issue relating to the exclusion of ever-smokers from the dataset. While the science remains unresolved, Flegal and Ioannidis (2018) suggests that the label of 'paradox' should be abandoned, as this terminology unscientifically pre-supposes that 'obesity' is *always* a risk factor, when in some exposure-outcome relationships there may be a valid causal mechanism for an apparent protective effect, such as that of subcutaneous fat on hip fractures in the event of a fall.

3.3 Systems thinking and participatory methods in health and policy-making

Systems are any 'thing' which we can view as ensembles of separate components, which interact with each other, and through these interactions realise emergent complex or higher-level behaviours or outcomes. For example, consider a bicycle: it has wheels, a frame, pedals, a chain, and a rider. Each of these components can behave in a certain way on their own, but when we put them together we achieve new behaviours, as the components push and pull on each other, constrain each other, and depend on each other. Systems thinking is both a mindset, and a set of tools, to analyse 'things', and problems, as systems. Systems thinking promotes a nonlinear, dynamic, and holistic

perspective and encourages us to think beyond a single component or problem and to be deliberate about where the boundaries of our analyses lie. In doing so, systems thinking prompts us to move from an attitude of examining a single intervention-problem relationship, and towards looking at the effects our interventions might have on the system as a whole. Systems thinking can be contrasted with reductionist approaches, which look at individual components in isolation and independently of their relationships with other components.

One advantage of systems thinking is that it enables a collaborative approach to comprehending systems and solving problems within them. Many of the methods in the systems science toolbox are applied in a participatory context, where a systems practitioner works with one or more domain experts to develop a model of the system under investigation. Systems science and participatory methods are separate worlds, but they have a large degree of overlap. In order to provide an overview of these different methods for both systems thinking and participatory modelling, I will discuss five specific research projects which make use of some combination of systems and participatory methods, and I will expand on the techniques used in the context of each of these studies.

3.3.1 The LIKE programme: participatory action research (PAR), causal loop diagrams (CLD), and group model building (GMB)

Waterlander et al. (2020) discusses the protocol for the LIKE programme, an approach to supporting healthy living and healthy weight in 10 to 14 year olds living in Amsterdam East, Netherlands, using qualitative systems methods and participatory approaches.

Baum, MacDougall, and Smith (2006) provides a brief overview of participatory action research (PAR), and it can be seen to have its foundations in two concepts. The first is that action should be a core part of the research process, because PAR intends to create change in the world, which requires action; and because action is a mechanism by which research can be evaluated and from which further investigation can take place. This research-action loop is particularly valuable for systems research, where the effects of actions may be hard to predict, and where direct intervention in a system offers a way to develop a greater understanding about its structure and behaviour. The second foundational component of PAR is that the subjects of research and different stakeholders who are embedded in the system under investigation should be brought into the research and action processes as collaborators. The motivation for participation is threefold: individuals within the system possess innate domain knowledge through their lived experience, which can precipitate greater understandings of the system under investigation; individuals within the system are likewise in a stronger position to act on and change that system in an effective way, with a view to the potential effects of that action; and power relations between researchers and research subjects should be dissolved, because individuals and communities are key stakeholders in the creation of their own reality, and this must be respected in order to achieve outcomes of social justice and research justice.

The LIKE programme implements PAR by bringing adolescents and their families directly into the research process as co-creators and co-researchers. Within the school context, the LIKE programme initiates groups of adolescents and researchers to develop a collective understanding of the drivers of different health outcomes, and actions that can be taken, within a peer context:

> The PAR groups include 6-8 adolescents per group and 1-2 facilitating academic researchers. PAR groups are set up for adolescents in a school setting and specifically for adolescents with obesity in the health care setting. PAR groups meet weekly throughout a predefined period, and in these meetings, they conduct research among their peers on their needs regarding healthy behaviours and potential actions toward stimulating healthy behaviours. Through capacity-building workshops, adolescents learn basic research principles and methods that enable them to conduct research. Furthermore, adolescents and families are encouraged to develop an individual and collective view on their current lifestyle using photo voice. This is facilitated in several sessions where participants take photos, collect and select images, discuss and reflect on the images and bring about action. Each PAR group summarizes their research results in a CLD [causal loop diagram].

Causal loop diagrams (CLDs) are one of the core methods used by systems researchers, and Schaffernicht (2010) provides an overview of the technique and several issues with it. At its most basic, a CLD is a graph of the structure of a system, where the nodes represent components or variables, and the arcs represent causal links or flows of information or other resources. These arcs

are typically annotated with 'polarities' of positive (+) or negative (-), indicating the effect that a change in one variable will have on another. For example, in a CLD to investigate population growth, an arrow from net births to population would have a positive polarity because an increase in the net births would lead to an increase in the population. Once all arcs have been annotated with polarities, it is possible to discover loops of variables that lead to positive or negative feedback. For example, as net births increase population, so does population increase net births, which leads to a self-reinforcing or positive feedback loop, which exerts exponential upwards pressure on the population.

Within the LIKE programme, CLDs take the place of one of the primary research outputs. PAR groups summarise their research in the form of CLDs, and these CLDs are then used to validate potential actions, to determine if they will be effective and target feedback loops within the system. These CLDs are then again updated throughout the evaluation process, as participants discover more about the system and as it changes due to their actions, and through this process the CLDs serve as knowledge artefacts which allow the programme to move between research and action in an iterative process.

In addition to the PAR workshops, the LIKE programme also makes use of an approach known as group model building (GMB). Vennix (1999) outlines the evolution of GMB and it can be summarised as an approach to the development of systems models (such as CLDs, although other models can also be produced) in a group workshop context. Typically, a long GMB session is broken down into a series of smaller group processes which are facilitated by 'scripts', which are narrative descriptions of the flow of each process, and which are used to aid in process delivery. Andersen and Richardson (1997) describes some scripts that may be used for problem definition, elicitation of system structures, and parameterisation. Within the LIKE programme, GMB is implemented with societal stakeholders at the local government and community levels across Amsterdam East, in order to build capacity for systems thinking within the policy system.

3.3.2 PHIMS for HCI: qualitative methods, soft systems methodology (SSM), and social network analysis (SNA)

Conte et al. (2017) describes a university-policymaker-practice research partnership to investigate the impact of the Population Health Information Management System (PHIMS) on the delivery of the Healthy Children Initiative (HCI) across New South Wales, Australia, making heavy use of both traditional qualitative and systems qualitative methods.

Traditional qualitative methods used by these researchers include ethnography, semi-structured interviews, and workshops. Ethnography is an approach which involves direct observation of subjects in their day-to-day activities, in order to allow the researcher to gather a deep qualitative understanding of the experiences and practices of the subjects. In the context of this study, researchers observed staff at different local health districts (LHDs), and at the state co-ordination level, and their interactions with the PHIMS system. Semi-structured interviews are a one-to-one qualitative data gathering method where a researcher asks a research subject a set of pre-defined questions, but also allowing for deviation from these questions and the capacity for the researcher to 'drill in' to the subjects answers and develop a deeper understanding. For this work, researchers carried out semi-structured interviews with policy-makers who are consumers of information from PHIMS to understand the effects of the system on the development and promotion of health policies. Workshops are any method taking place in a many-to-one context of participants to researchers, and for this study the researchers used a participatory action research (PAR, see previous section on the LIKE programme) to present initial findings and recalibrate their approach before undertaking a final analysis. Tracy (2019) provides an overview of all of these methods.

Social network analysis (SNA) is an approach whereby a set of social entities, such as individuals or organisations, and their interactions, are represented in a (possibly directed) graph. Nodes in the graph represent individual entities and arcs are added to represent the presence of the particular social relationship of interest. One example of an SNA graph would be a citation graph, showing a network of authors who are connected if one has cited the other. Social networks are fundamentally systems, whose components are indviduals, and so SNA (and network analysis more generally) falls within the systems science toolkit. SNA allows social science researchers to combine qualitative and quantitative data about social structures and relations into a single model and to utilise the tools of graph theory to understand these relations. For example, a centrality index can be used to discover 'central' entities with the social network, which might correspond to their level of connectedness, or their propensity for being included in certain information flows. Butts

(2008) provides a thorough overview of SNA approaches, and Hunter et al. (2019) demonstrates the effectiveness of social network interventions for health outcomes. For 'obesity' in particular, this evidence shows that 'obesity' can 'spread' through social networks, and so SNA is a very promising tool for developing interventions. Within the PHIMS for HCI research, the researchers use SNA to develop an understanding of which actors collaborate with whom; which actors are commonly called on for support and advice; which actors share similar ideas and attitudes; and the degree of contact across different LHDs.

Conte and Davidson (2020) expands on the research methodology to discuss how researchers have used the rich picture method to crystallise their findings and present these back to policy-makers via a workshop. Unlike a causal loop diagram (CLD), there is no formal definition for how a rich picture should be constructed, nor what its contents can be. Rather, they are a free-form and open-ended technique to allow individuals to collaborate on and communicate a 'map' of a system, based on whatever structural or behavioural aspects of the system are deemed most important by the creators. Rich pictures originate in the soft systems methodology (SSM), which is an approach to understanding and changing systems which are messy, complicated, and which lack clear problem statements or component or system boundaries, such as social systems (Checkland and Poulter 2020). Rich pictures can be applied in a participatory context where participants collaborate with each other and a researcher to develop a system map, but the PHIMS for HCI researchers apply SSM and rich pictures in a different way, where the researchers themselves build a rich picture as a representation of their learnings from the analysis of their qualitative data. The researchers spent a day workshopping together to produce a rich picture summary of their findings, and this was then presented at another workshop to a group of policy-makers.

3.3.3 Childhood 'obesity' in Australia: group model building (GMB) of a systems dynamics (SD) model

Roberts et al. (2019) describes a systems dynamics (SD) model for evaluating the potential effects of various interventions on the prevalence of childhood 'obesity' in New South Wales, Australia. Their methodology leans heavily on the group model building (GMB) approach (see earlier section on the LIKE programme) and is documented in more detail in Freebairn et al. (2016).

Systems dynamics (SD) modelling is a quantitative realisation of a stock and flow diagram, which itself can be seen as a more formal evolution of the causal loop diagram (CLD). Whereas nodes in a CLD may represent either stocks or flows, a stock and flow diagram makes the distinction between these explicit. Meadows (2009) uses the stock and flow model as the foundation for demonstrating systems thinking and systems modelling. A stock is a store of some quantity, such as a bath tub, which stores water; a bank account, which stores money; or a human body, which stores energy (among other things). Flows are portals which connect stocks and allow quantities to flow between stocks, and in and out of a system. Examples of flows include a tap or a plughole, which allow water to flow in and out of a bath tub; transactions which allow money to flow in and out of a bank account; and food consumption and exercise, which allow energy to flow in and out of a human body. Stocks serve as a memory of the history of flows in to and out from them, and the stock and flow approach allows for the modelling of dynamic equilibria (where a stock maintains its value over time due to an equal in-flow and out-flow); the modelling of feedback loops (which may be reinforcing or balancing, depending on the presence of other loops affecting the same stocks and flows); and the elicitation of systems archetypes, which are common structures of stocks and flows which can be applied to understand a broad set of systems as part of a systems thinking approach. SD models arise from the precise quantification of a stock and flow model, allowing the flows to be simulated, thus permitting the values of stocks and flows to be traced over time and to allow researchers to develop an understanding of the behaviours and equilibria of the system under investigation.

To develop their SD model of childhood 'obesity' in New South Wales, the researchers began by assembling a stakeholder panel consisting of policy experts, academics, health economists, public health experts, a biostatistician, and experts in dynamic simulation modelling. Three workshops were organised to undertake development and make use of the model, including calibration and validation. Unfortunately Roberts et al. (2019) does not provide further detail on the participatory modelling process but instead refers back to an earlier project, Freebairn et al. (2016), where a similar approach was used. Although it is not possible to comment on the structure of each of the three workshops used as part of this childhood 'obesity' research, the implication of their prior documentation is that they have adopted a GMB approach, and that steps undertaken across all

workshops include defining the problem to be addressed and the boundaries of the system to be modelled; development of the conceptual map which will correspond to the stock and flow network underlying the SD model when it is coded; identifying data sources for model parameterisation; determining potential intervention leverage points within the system and potential interventions to be tested; validation of the model by expert assessment of assumptions and behaviours, reproduction of expected behaviours under specific conditions, reasonableness of parameters, sensitivity analyses, and statistical tests; and exploration of potential interventions, including development of an action plan detailing which interventions will be deployed in the real-world and when. The SD model was parameterised using data from the Australian Census and data about the status of existing 'obesity' interventions. Stakeholders prioritised nine interventions of interest to be tested in the model, including increased provision of healthy food choices in government settings; bans on advertising of unhealthy foods; tax on sugar-sweetened beverages; and programmes to support increased physical activity. Four scenarios were tested: a 'business as usual' baseline; enhancement of existing programmes; enhancement of existing programmes plus expansion of sport and built environment and active transport infrastructures; and enhancement of existing programmes plus expansion of the complete set of interventions. The New South Wales government have a target to reduce prevalence of childhood 'overweight' and 'obesity' by 5% by 2025 relative to a 2015 baseline, and so prevalence by age strata are the primary output of the model which stakeholders are interested in and by which each scenario will be evaluated. The results of the SD model indicate that only scenarios three and four, both requiring the delivery of several new programmes, are consistent with achieving the government's 5% target, and the researchers conclude that significant cross-portfolio policy action is required.

3.3.4 FAAMC: group model building (GMB) of an agent-based model (ABM)

Koh, Reno, and Hyder (2019) discusses the development and evaluation of the Food Accessibility Agent-based Model in Central Columbus, Ohio (FAAMC). The FAAMC is an agent-based model (ABM) which has been constructed using a group model building (GMB) approach.

Agent-based models (ABMs) provide a ground-up approach to systems modelling and simulation. Whereas systems dynamics (SD) models operate from the top-down, explicitly describing the relationships between components and flows between different stocks (see the earlier section on childhood 'obesity' in Australia), ABMs instead define the operation of individual components (the agents) and then allow the system's dynamics to emerge from these low-level behaviours and interactions. In the context of social science and epidemiology, each agents usually corresponds to an individual person, and each agent has its own state, a set of variables of interest such as location, energy expenditure, money, and so on. Agents may have different roles or types, and as an ABM simulation proceeds, each agent will become heterogeneous with the others, as they each experience different aspects of the world and make different choices. Macal and North (2005) provides an introduction to the ABM modelling approach.

For FAAMC, the researchers have applied the group model building (GMB) methodology, and this is documented in Koh, Reno, and Hyder (2018). Typically GMB is used in the development of causal loop diagrams (CLDs) or systems dynamics (SD) models (see earlier section on the LIKE programme), which makes this research particularly innovative. To develop the ABM, the researchers assembled a panel of participants from academia, public health, education, and NGOs, to take part in two workshop sessions. The first workshop served to define the problem (food availability and accessibility across households); determine an appropriate boundary for the model; define the different types of agents and their state variables and interactions; and discuss potential policy options to be evaluated with the model. Part of this first workshop involved the development of conceptual models or mind maps which outlined the factors affecting food accessibility. After the first workshop, the ABM was implemented to a near-final state, so that it could be used in the second workshop. The modellers presented the ABM back to the participants in the second workshop, and sought to validate the model by querying the participants about the realism of the behaviours of the agents. New conceptual models were developed and all feedback was incorporated into the final model, which was then re-distributed to all participants via a Web platform. This research shows that GMB can be applied to the development of more complex or bottom-up models such as ABMs, although the researchers remark that further work is needed in this area to develop appropriate GMB scripts, given that GMB is traditionally applied for the development of top-down models.

4 Timeline

I intend to undertake this PhD on a part-time basis, at least for the first 1 to 2 years. This will provide me with a period of five to eight years to complete my research. This long time frame is ideal for the proposed work because systems solutions to 'obesity' are inherently continuous and time-lagged, requiring iteration and ongoing development, and with a delay of potentially multiple years before results can be observed. Additionally, because the proposed work is highly collaborative in nature, requiring the support of one or more Local Authorities undertaking one or more cycles of the WSO process, this longer time frame means I am less burdened by the velocity of my collaborators and their respective organisations.

Figure 4 presents a Gantt chart of an expected timeline for my research, under the assumption that I will participate in two iterations of the extended WSO process, and that these can be run somewhat in parallel. These timespans are only an estimate and are intended to be generous, to provide me with ample manoeuvrability in the event of unexpected delays or complexity. These timelines are also high-level and do not include tasks such as applying for access to the UK Biobank or other datasets.

4.1 Current progress

I have already started to make good progress on my research. The protocol for the scoping review of artificial microdataset methods is complete and ready for registration, at which point I can begin to search for literature and screen results. I have also begun work on the setup for the simulation study in the comparison paper which will follow this review, and this works should itself be sufficient for publishing as a tutorial paper, separate to the comparison project. Finally, I am in the process of applying to access to the UK Biobank dataset, a process which may take several months for approval to be granted, and so it is important that I am able to apply for access as soon as possible.

4.2 Future work

There are also many opportunities for extensions to the core work to further aid in intervention design, which would be suitable as future research following on from this PhD. One possibility is to apply cutting-edge research in causal Bayesian optimisation (Aglietti et al. 2020) in order to automate the discovery of potentially fruitful interventions in local 'obesity' systems. With this approach, once a suitable causal model of the system had been constructed, it could be annotated with additional information about the feasibility and costs of intervening in different areas of the system, and then an optimisation problem could be constructed. This optimisation problem would take the form of a mathematical function which can be numerically 'minimised' to discover a set of optimal intervention points in the system: sets of actions which result in maximum impact for minimal cost.

Another possibility for future research is to investigate the 'transportability' (Pearl and Bareinboim 2014) of causal insights across different local systems, to determine if there are common factors in the 'obesity' systems of different Local Authorities. A transportable model is one which is structurally invariant in different local contexts, such that the model can be build once, and then it only needs to be re-parameterised in order to be applicable within a different environment. It is likely that some structural differences would exist between different contexts, although this would provide researchers and Local Authorities with a common starting point for modelling of the 'obesity' system, and facilitate rapid intervention and policy design across many Local Authorities, allowing the WSO programme to be implemented more efficiently and with more targeted use of resources.

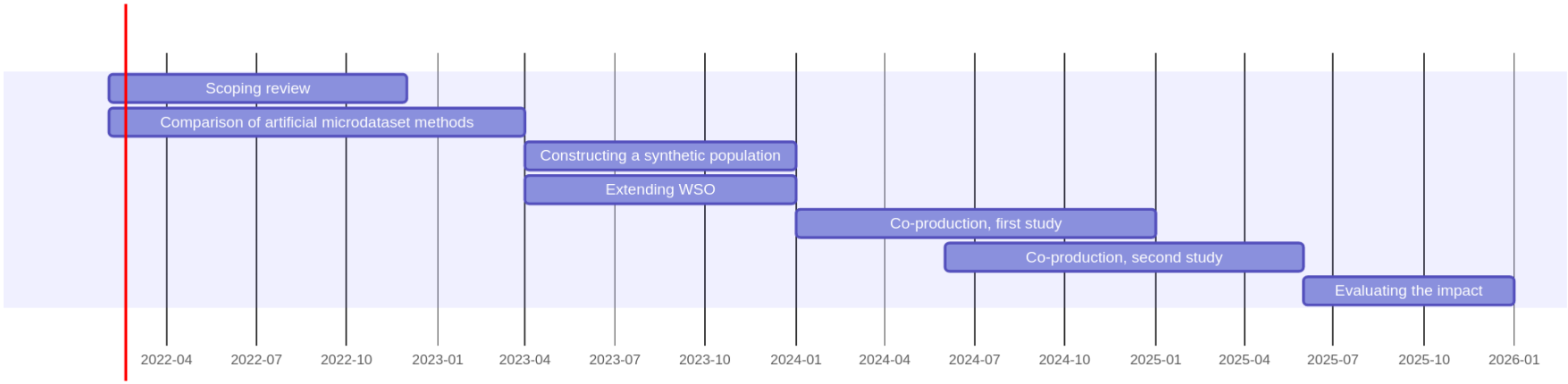

Figure 4: An estimated Gantt chart for my PhD research.

5 References

Aglietti, Virginia, Xiaoyu Lu, Andrei Paleyes, and Javier González. 2020. "Causal Bayesian Optimization." In *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, 3155–64. PMLR. https://proceedings.mlr.press/v108/aglietti20a.html.

Andersen, David F., and George P. Richardson. 1997. "Scripts for Group Model Building." *System Dynamics Review* 13 (2): 107–29. https://doi.org/10.1002/(SICI)1099-1727(199722)13:2%3C107::AID-SDR120%3E3.0.CO;2-7.

Angelantonio, Emanuele Di, Shilpa N. Bhupathiraju, David Wormser, Pei Gao, Stephen Kaptoge, Amy Berrington de Gonzalez, Benjamin J. Cairns, et al. 2016. "Body-Mass Index and All-Cause Mortality: Individual-Participant-Data Meta-Analysis of 239 Prospective Studies in Four Continents." *The Lancet* 388 (10046): 776–86. https://doi.org/10.1016/S0140-6736(16)30175-1.

Baum, Fran, Colin MacDougall, and Danielle Smith. 2006. "Participatory Action Research." *Journal of Epidemiology and Community Health* 60 (10): 854–57. https://doi.org/10.1136/jech.2004.028662.

Blüher, Matthias. 2020. "Metabolically Healthy Obesity." *Endocrine Reviews* 41 (3): 405–20. https://doi.org/10.1210/endrev/bnaa004.

Boero, Natalie. 2012. *Killer Fat: Media, Medicine, and Morals in the American "Obesity Epidemic": Media, Medicine, and Morals in the American "Obesity Epidemic"*. 1st edition. New Brunswick, N.J: Rutgers University Press.

Bosello, Ottavio, Maria Pia Donataccio, and Massimo Cuzzolaro. 2016. "Obesity or Obesities? Controversies on the Association Between Body Mass Index and Premature Mortality." *Eating and Weight Disorders - Studies on Anorexia, Bulimia and Obesity* 21 (2): 165–74. https://doi.org/10.1007/s40519-016-0278-4.

Burkhauser, Richard V., and John Cawley. 2008. "Beyond BMI: The Value of More Accurate Measures of Fatness and Obesity in Social Science Research." *Journal of Health Economics* 27 (2): 519–29. https://doi.org/10.1016/j.jhealeco.2007.05.005.

Butts, Carter T. 2008. "Social Network Analysis: A Methodological Introduction." *Asian Journal of Social Psychology* 11 (1): 13–41. https://doi.org/10.1111/j.1467-839X.2007.00241.x.

Campos, Paul, Abigail Saguy, Paul Ernsberger, Eric Oliver, and Glenn Gaesser. 2006. "The Epidemiology of Overweight and Obesity: Public Health Crisis or Moral Panic?" *International Journal of Epidemiology* 35 (1): 55–60. https://doi.org/10.1093/ije/dyi254.

Cecchini, Michele, and Sabine Vuik. 2019. "The Heavy Burden of Obesity." Paris: OECD. https://doi.org/10.1787/3c6ec454-en.

Checkland, Peter, and John Poulter. 2020. "Soft Systems Methodology." In *Systems Approaches to Making Change: A Practical Guide*, edited by Martin Reynolds and Sue Holwell (Retired), 201–53. London: Springer. https://doi.org/10.1007/978-1-4471-7472-1_5.

Chrisler, Joan C., and Angela Barney. 2017. "Sizeism Is a Health Hazard." *Fat Studies* 6 (1): 38–53. https://doi.org/10.1080/21604851.2016.1213066.

Cole, Stephen R, Robert W Platt, Enrique F Schisterman, Haitao Chu, Daniel Westreich, David Richardson, and Charles Poole. 2010. "Illustrating Bias Due to Conditioning on a Collider." *International Journal of Epidemiology* 39 (2): 417–20. https://doi.org/10.1093/ije/dyp334.

Conte, Kathleen P., and Seanna Davidson. 2020. "Using a 'Rich Picture' to Facilitate Systems Thinking in Research Coproduction." *Health Research Policy and Systems* 18 (1): 14. https://doi.org/10.1186/s12961-019-0514-2.

Conte, Kathleen P., Sisse Groen, Victoria Loblay, Amanda Green, Andrew Milat, Lina Persson, Christine Innes-Hughes, et al. 2017. "Dynamics Behind the Scale up of Evidence-Based Obesity Prevention: Protocol for a Multi-Site Case Study of an Electronic Implementation Monitoring System in Health Promotion Practice." *Implementation Science* 12 (1): 146. https://doi.org/10.1186/s13012-017-0686-5.

Egusquiza, Riann Jenay, and Bruce Blumberg. 2020. "Environmental Obesogens and Their Impact on Susceptibility to Obesity: New Mechanisms and Chemicals." *Endocrinology* 161 (3). https://doi.org/10.1210/endocr/bqaa024.

Ekelund, Ulf, Søren Brage, Herve Besson, Stephen Sharp, and Nicholas J Wareham. 2008. "Time Spent Being Sedentary and Weight Gain in Healthy Adults: Reverse or Bidirectional Causality?" *The American Journal of Clinical Nutrition* 88 (3): 612–17. https://doi.org/10.1093/ajcn/88.3.612.

Ervasti, Jenni, Jaakko Airaksinen, Jaana Pentti, Jussi Vahtera, Sakari Suominen, Marianna Virtanen, and Mika Kivimäki. 2019. "Does Increasing Physical Activity Reduce the Excess Risk of Work Disability Among Overweight Individuals?" *Scandinavian Journal of Work,*

Environment & Health 45 (4): 376–85. https://www.jstor.org/stable/26746264.

Flegal, Katherine M., Barry I. Graubard, David F. Williamson, and Richard S. Cooper. 2011. "Reverse Causation and Illness-related Weight Loss in Observational Studies of Body Weight and Mortality." *American Journal of Epidemiology* 173 (1): 1–9. https://doi.org/10.1093/aje/kwq341.

Flegal, Katherine M., and John P. A. Ioannidis. 2018. "The Obesity Paradox: A Misleading Term That Should Be Abandoned." *Obesity* 26 (4): 629–30. https://doi.org/10.1002/oby.22140.

Flegal, Katherine M., John P. A. Ioannidis, and Wolfram Doehner. 2019. "Flawed Methods and Inappropriate Conclusions for Health Policy on Overweight and Obesity: The Global BMI Mortality Collaboration Meta-Analysis." *Journal of Cachexia, Sarcopenia and Muscle* 10 (1): 9–13. https://doi.org/10.1002/jcsm.12378.

Freebairn, L., J. Atkinson, P. Kelly, G. McDonnell, and L. Rychetnik. 2016. "Simulation Modelling as a Tool for Knowledge Mobilisation in Health Policy Settings: A Case Study Protocol." *Health Research Policy and Systems* 14 (1): 71. https://doi.org/10.1186/s12961-016-0143-y.

Frood, Sarah, Lee M. Johnston, Carrie L. Matteson, and Diane T. Finegood. 2013. "Obesity, Complexity, and the Role of the Health System." *Current Obesity Reports* 2 (4): 320–26. https://doi.org/10.1007/s13679-013-0072-9.

Galea, Sandro, Matthew Riddle, and George A Kaplan. 2010. "Causal Thinking and Complex System Approaches in Epidemiology." *International Journal of Epidemiology* 39 (1): 97–106. https://doi.org/10.1093/ije/dyp296.

Gard, Michael, and Jan Wright. 2005. *The Obesity Epidemic: Science, Morality and Ideology*. London: Routledge. https://doi.org/10.4324/9780203619308.

Gutting, Gary. 2005. *The Cambridge Companion to Foucault*. Cambridge University Press.

Hermes, Kerstin, and Michael Poulsen. 2012. "A Review of Current Methods to Generate Synthetic Spatial Microdata Using Reweighting and Future Directions." *Computers, Environment and Urban Systems* 36 (4): 281–90. https://doi.org/10.1016/j.compenvurbsys.2012.03.005.

Hernán, Miguel A. 2018. "The C-Word: Scientific Euphemisms Do Not Improve Causal Inference From Observational Data." *American Journal of Public Health* 108 (5): 616–19. https://doi.org/10.2105/AJPH.2018.304337.

Hernán, Miguel A., and James M. Robins. 2016. "Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available." *American Journal of Epidemiology* 183 (8): 758–64. https://doi.org/10.1093/aje/kwv254.

Heyes, Cressida J. 2007. *Self-Transformations: Foucault, Ethics, and Normalized Bodies*. Oxford ; New York: Oxford University Press.

Hill, Briony, Heidi Bergmeier, Angela C. Incollingo Rodriguez, Fiona Kate Barlow, Alexandra Chung, Divya Ramachandran, Melissa Savaglio, and Helen Skouteris. 2021. "Weight Stigma and Obesity-Related Policies: A Systematic Review of the State of the Literature." *Obesity Reviews* 22 (11): e13333. https://doi.org/10.1111/obr.13333.

Hunter, Ruth F., Kayla de la Haye, Jennifer M. Murray, Jennifer Badham, Thomas W. Valente, Mike Clarke, and Frank Kee. 2019. "Social Network Interventions for Health Behaviours and Outcomes: A Systematic Review and Meta-Analysis." *PLOS Medicine* 16 (9): e1002890. https://doi.org/10.1371/journal.pmed.1002890.

Jolivétte, Andrew. 2015. *Research Justice: Methodologies for Social Change*. Policy Press.

Katsoulis, Michail, Bianca D. Stavola, Karla Diaz-Ordaz, Manuel Gomes, Alvina Lai, Pagona Lagiou, Goya Wannamethee, et al. 2021. "Weight Change and the Onset of Cardiovascular Diseases: Emulating Trials Using Electronic Health Records." *Epidemiology (Cambridge, Mass.)* 32 (5): 744–55. https://doi.org/10.1097/EDE.0000000000001393.

Koh, Keumseok, Rebecca Reno, and Ayaz Hyder. 2018. "Designing an Agent-Based Model Using Group Model Building: Application to Food Insecurity Patterns in a U.S. Midwestern Metropolitan City." *Journal of Urban Health* 95 (2): 278–89. https://doi.org/10.1007/s11524-018-0230-1.

———. 2019. "Examining Disparities in Food Accessibility Among Households in Columbus, Ohio: An Agent-Based Model." *Food Security* 11 (2): 317–31. https://doi.org/10.1007/s12571-019-00900-7.

Kutcher, Stephen A., James M. Brophy, Hailey R. Banack, Jay S. Kaufman, and Michelle Samuel. 2021. "Emulating a Randomised Controlled Trial With Observational Data: An Introduction to the Target Trial Framework." *Canadian Journal of Cardiology* 37 (9): 1365–77. https://doi.org/10.1016/j.cjca.2021.05.012.

Lobstein, Tim, and Kelly D. Brownell. 2021. "Endocrine-Disrupting Chemicals and Obesity Risk: A Review of Recommendations for Obesity Prevention Policies." *Obesity Reviews* 22 (11): e13332. https://doi.org/10.1111/obr.13332.

Lupton, Deborah. 2018. *Fat*. Routledge.
Macal, C. M., and M. J. North. 2005. "Tutorial on Agent-Based Modeling and Simulation." In *Proceedings of the Winter Simulation Conference, 2005.*, 14 pp.–. https://doi.org/10.1109/WSC.2005.1574234.
Mahlke, Jana, Martin Schultze, and Michael Eid. 2019. "Analysing Multisource Feedback with Multilevel Structural Equation Models: Pitfalls and Recommendations from a Simulation Study." *British Journal of Mathematical and Statistical Psychology* 72 (2): 294–315. https://doi.org/10.1111/bmsp.12149.
Meadows, Donella H. 2009. *Thinking in Systems: A Primer*. Edited by Diana Wright. First edition. London: Routledge.
Meier, Petra, Robin Purshouse, Marion Bain, Clare Bambra, Richard Bentall, Mark Birkin, John Brazier, et al. 2019. "The SIPHER Consortium: Introducing the New UK Hub for Systems Science in Public Health and Health Economic Research." *Wellcome Open Research* 4 (November): 174. https://doi.org/10.12688/wellcomeopenres.15534.1.
Meldrum, David R., Marge A. Morris, and Joseph C. Gambone. 2017. "Obesity Pandemic: Causes, Consequences, and Solutions—but Do We Have the Will?" *Fertility and Sterility* 107 (4): 833–39. https://doi.org/10.1016/j.fertnstert.2017.02.104.
Naimi, Ashley I, Stephen R Cole, and Edward H Kennedy. 2017. "An Introduction to g Methods." *International Journal of Epidemiology* 46 (2): 756–62. https://doi.org/10.1093/ije/dyw323.
O'Donoghue, Cathal, Karyn Morrissey, and John Lennon. 2013. "Spatial Microsimulation Modelling: A Review of Applications and Methodological Choices." *International Journal of Microsimulation* 7 (1): 26–75. https://doi.org/10.34196/ijm.00093.
"Obese, Adj. And n." n.d. *OED Online*. Oxford University Press. Accessed February 2, 2022. https://www.oed.com/view/Entry/129577.
Pausé, Cat, George Parker, and Lesley Gray. 2021. "Resisting the Problematisation of Fatness in COVID-19: In Pursuit of Health Justice." *International Journal of Disaster Risk Reduction* 54 (February): 102021. https://doi.org/10.1016/j.ijdrr.2020.102021.
Pearl, Judea. 2010. "Causal Inference." In *Proceedings of Workshop on Causality: Objectives and Assessment at NIPS 2008*, 39–58. PMLR. https://proceedings.mlr.press/v6/pearl10a.html.
Pearl, Judea, and Elias Bareinboim. 2014. "External Validity: From Do-Calculus to Transportability Across Populations." *Statistical Science* 29 (4): 579–95. https://doi.org/10.1214/14-STS486.
Piché, Marie-Eve, Paul Poirier, Isabelle Lemieux, and Jean-Pierre Després. 2018. "Overview of Epidemiology and Contribution of Obesity and Body Fat Distribution to Cardiovascular Disease: An Update." *Progress in Cardiovascular Diseases*, Update on Obesity and Cardiovascular Diseases, 61 (2): 103–13. https://doi.org/10.1016/j.pcad.2018.06.004.
Popkin, Barry M., Shufa Du, William D. Green, Melinda A. Beck, Taghred Algaith, Christopher H. Herbst, Reem F. Alsukait, Mohammed Alluhidan, Nahar Alazemi, and Meera Shekar. 2020. "Individuals with Obesity and COVID-19: A Global Perspective on the Epidemiology and Biological Relationships." *Obesity Reviews* 21 (11): e13128. https://doi.org/10.1111/obr.13128.
Public Health England. 2019. "Whole Systems Approach to Obesity: A Guide to Support Local Approaches," 104.
Puhl, Rebecca M. 2020. "What Words Should We Use to Talk about Weight? A Systematic Review of Quantitative and Qualitative Studies Examining Preferences for Weight-Related Terminology." *Obesity Reviews* 21 (6): e13008. https://doi.org/10.1111/obr.13008.
Puhl, Rebecca M., and Chelsea A. Heuer. 2009. "The Stigma of Obesity: A Review and Update." *Obesity* 17 (5): 941–64. https://doi.org/10.1038/oby.2008.636.
Rahman, Azizur. 2017. "Estimating Small Area Health-Related Characteristics of Populations: A Methodological Review." *Geospatial Health* 12 (1). https://doi.org/10.4081/gh.2017.495.
Roberts, Nick, Vincy Li, Jo-An Atkinson, Mark Heffernan, Geoff McDonnell, Ante Prodan, Louise Freebairn, et al. 2019. "Can the Target Set for Reducing Childhood Overweight and Obesity Be Met? A System Dynamics Modelling Study in New South Wales, Australia." *Systems Research and Behavioral Science* 36 (1): 36–52. https://doi.org/10.1002/sres.2555.
Rothblum, Esther, and Sondra Solovay. 2009. *The Fat Studies Reader*. NYU Press.
Schaffernicht, Martin. 2010. "Causal Loop Diagrams Between Structure and Behaviour: A Critical Analysis of the Relationship Between Polarity, Behaviour and Events." *Systems Research and Behavioral Science* 27 (6): 653–66. https://doi.org/10.1002/sres.1018.
Shakiba, Maryam, Mohammad Ali Mansournia, Arsalan Salari, Hamid Soori, Nasrin Mansournia, and Jay S Kaufman. 2018. "Accounting for Time-Varying Confounding in the Relationship Between Obesity and Coronary Heart Disease: Analysis With G-Estimation: The ARIC Study." *American Journal of Epidemiology* 187 (6): 1319–26. https://doi.org/10.1093/aje/kwx360.
Smith, Dianna M., Alison Heppenstall, and Monique Campbell. 2021. "Estimating Health over

Space and Time: A Review of Spatial Microsimulation Applied to Public Health." *J* 4 (2): 182–92. https://doi.org/10.3390/j4020015.

Sperrin, Matthew, Jane Candlish, Ellena Badrick, Andrew Renehan, and Iain Buchan. 2016. "Collider Bias Is Only a Partial Explanation for the Obesity Paradox." *Epidemiology (Cambridge, Mass.)* 27 (4): 525–30. https://doi.org/10.1097/EDE.0000000000000493.

Stokes, Andrew, and Samuel H. Preston. 2015. "Smoking and Reverse Causation Create an Obesity Paradox in Cardiovascular Disease." *Obesity* 23 (12): 2485–90. https://doi.org/10.1002/oby.21239.

Tanton, Robert. 2013. "A Review of Spatial Microsimulation Methods." *International Journal of Microsimulation* 7 (1): 4–25. https://doi.org/10.34196/ijm.00092.

Tomova, Georgia D, Kellyn F Arnold, Mark S Gilthorpe, and Peter W G Tennant. 2021. "Adjustment for Energy Intake in Nutritional Research: A Causal Inference Perspective." *The American Journal of Clinical Nutrition*, no. nqab266 (July). https://doi.org/10.1093/ajcn/nqab266.

Tracy, Sarah J. 2019. *Qualitative Research Methods: Collecting Evidence, Crafting Analysis, Communicating Impact*. John Wiley & Sons.

Tylka, Tracy L., Rachel A. Annunziato, Deb Burgard, Sigrún Daníelsdóttir, Ellen Shuman, Chad Davis, and Rachel M. Calogero. 2014. "The Weight-Inclusive Versus Weight-Normative Approach to Health: Evaluating the Evidence for Prioritizing Well-Being over Weight Loss." *Journal of Obesity* 2014 (July): e983495. https://doi.org/10.1155/2014/983495.

Vennix, Jac A. M. 1999. "Group Model-Building: Tackling Messy Problems." *System Dynamics Review* 15 (4): 379–401. https://doi.org/10.1002/(SICI)1099-1727(199924)15:4%3C379::AID-SDR179%3E3.0.CO;2-E.

Waterlander, Wilma E., Angie Luna Pinzon, Arnoud Verhoeff, Karen den Hertog, Teatske Altenburg, Coosje Dijkstra, Jutka Halberstadt, et al. 2020. "A System Dynamics and Participatory Action Research Approach to Promote Healthy Living and a Healthy Weight Among 10–14-Year-Old Adolescents in Amsterdam: The LIKE Programme." *International Journal of Environmental Research and Public Health* 17 (14): 4928. https://doi.org/10.3390/ijerph17144928.

Westreich, Daniel. 2012. "Berkson's Bias, Selection Bias, and Missing Data." *Epidemiology (Cambridge, Mass.)* 23 (1): 159–64. https://doi.org/10.1097/EDE.0b013e31823b6296.

Westreich, Daniel, and Sander Greenland. 2013. "The Table 2 Fallacy: Presenting and Interpreting Confounder and Modifier Coefficients." *American Journal of Epidemiology* 177 (4): 292–98. https://doi.org/10.1093/aje/kws412.

Xue, H., L. Slivka, T. Igusa, T. T. Huang, and Y. Wang. 2018. "Applications of Systems Modelling in Obesity Research." *Obesity Reviews* 19 (9): 1293–1308. https://doi.org/10.1111/obr.12695.