

Design and Implementation of a Whole Disease Model of Oral Cancer

This document will describe the design, implementation, and calibration of a WDM of oral cancer (WDMOC) that addresses the limitations of the existing modeling literature with respect to their ability to generate evidence to guide policy-making. Please note that every time I refer to “Chapter X” I am doing so because this document was originally part of my PhD dissertation and I haven’t bothered to edit all this documentation yet. If it really takes off, then I’ll revisit that decision.

1.1 Introduction

The central purpose of the WDM framework is to model the breadth of the entire disease, including preclinical management and detection through treatment to the end of life. Whole disease models should be able to reflect the impact of multiple simultaneous policy and/or technology changes, while adhering to recommended guidelines for model design and implementation.

This document will describe the process through which the WDMOC was created, following methods described by Tappenden[31]. First, the process through which the model was conceived and designed will be described. Next, the model’s characteristics and basic architecture will be summarized. The implementation of the model – i.e., the specific way in which the model was programmed – will be described briefly as well. A description of the statistical techniques employed by the model, and how they were applied to the values and sources of the model’s parameters, will follow. Finally, the model’s baseline output will be described and compared to values observed in the real world to evaluate its validity.

1.2 Model Design and Theoretical Framework

The conceptual framework for the WDMOC was designed drawing on the work by Tappenden[31] which defines a whole disease model as one that:

1. Includes preclinical and post-diagnostic pathways for individuals who may or may not develop a given disease in their lives;
2. Captures different service pathways from system entry to discharge or death;
3. Represents events, costs and outcomes, and structural relationships between these to a level of detail that allows the point at which technologies may change (decision node) to be transferred across the modeled pathway;
4. Allows for the economic evaluation of individual or multiple service changes

WDMs are conceptually defined by three principal attributes:

- **Boundary:** the populations represented within the model – the people who interact with and are affected by changes made within the system the model seeks to represent
- **Breadth:** the phenomena, costs, and consequences included within the model – the types of processes, services, resources, and outcomes that the model will reflect
- **Depth:** the level of detail used to describe and value each phenomenon, cost, and consequence included in the model.

Three principles follow from these attributes:

1. The model boundary and breadth should capture all relevant aspects of the disease and its treatment – from preclinical disease through to death
2. The model should be developed such that the decision node is conceptually transferable across the model
3. The costs and consequences of service elements should be structurally related

A whole disease model of oral cancer must represent the clinical experience of both people with oral cancer (and pre-cancer) and those who do not experience the disease but would be affected by changes in the system (i.e., the general population), in such a way that several potential changes can be evaluated simultaneously. The model must be granular enough to accurately reflect all relevant costs and outcomes, while being broad enough to estimate the impact that upstream changes will have on all downstream events.

1.3 Model Design process

Designing a whole disease model is a five-stage process:

1. Understanding the decision problem: defining who will use the model and what types of economic questions it will be used to answer;

2. Conceptualization and design: building a conceptual representation of the processes that the model will simulate;
3. Implementation modeling: the creation of the model itself, using computer software. Implementation typically requires a time-to-event approach. This stage includes model calibration and uncertainty analysis;
4. Model checking: ensuring that the model entities and processes are behaving as expected. This is an ongoing process during model development;
5. Engaging with the decision: incorporating the results of the model into a policy making process and/or framework.

The development of the WDMOC, through these stages, is described below.

1.3.1 *Stage 1: Understanding the Decision Problem*

There are three key elements of this stage: establishment of a stakeholder group, immersion in relevant evidence, and agreeing what is to be evaluated and why.

Establishment of a Stakeholder Group

A multidisciplinary stakeholder group was established to guide the development of the WDMOC. This group included health care professionals whose areas of expertise are represented within the full scope of the decision problem being modeled. Ten individuals were included based on both their familiarity with oral cancer management and care, and their previous relationship with researchers in the Oral Cancer Prevention Program at the BCCA. The scope of practice within the stakeholder group is described below, and includes five surgeons, three oncologists (surgical, radiation, medical), and seven frontline community practitioners. A description of the stakeholder group can be found in Appendix 4.1.

Each member of the stakeholder team was approached for a one-on-one interview to provide input on the structure of the model. Most interviews were conducted in person, with some conducted via telephone. Members were provided with a draft version of the conceptual model and a document explaining the model's purpose and summarizing its design. During the course of the interview,

members were asked “What structural elements in the model do not match current practice?”; “What could/should be changed about those elements to more closely match current practice?”; and “What are important research questions within your scope of practice that the model could/should be used to address in the future? How might the model have to change to reflect those?” The model structure and inputs were updated iteratively over the course of these interviews and re-presented to members of the committee until broad agreement was reached.

Immersion in Relevant Evidence

Chapter 2 describes the relevant evidence consulted in the construction of the WDMOC. The model was also designed according to principles and guidelines set out by International Society for Pharmacoeconomics and Outcomes Research (ISPOR)[30]. The Canadian Association for Drugs and Technologies in Health (CADTH) issue similar guidelines for model-based economic evaluations in Canada[5]. The use of these guidelines ensures consistent quality between modeling exercises, and provides modelers with a set of analytical principles and tools to ensure that the model output is relevant and useful to guide policy making. They also provide recommendations for data sources, statistical analysis, and the form that model inputs should take. The guidelines are written flexibly, allowing modelers some leeway to customize their guidance to fit the particular decision being modeled.

Generally, the CADTH guidelines were consulted to ensure that the model was appropriate for a Canadian policy making context. The characterization of the decision problem, the comparator group, and methodological issues such as discounting and probabilistic analysis were conducted according to the recommendations published by CADTH. The ISPOR guidelines were used to inform technical issues, such as calculating competing risks, parameter estimation, and incorporating parameter uncertainty.

Agreeing What is to be Modeled and Why

Chapter 3 can be thought of as an initial attempt to address the third element of determining what is to be modeled and why, for a single decision context. Through the model design process, and through comparison to existing models, the gaps in the literature were discovered. The committee and stakeholder group provided ongoing input regarding the balance between model complexity and the practicality of deriving model parameter estimates, given these gaps. The stakeholder group provided additional insight into the number of novel technologies that could be evaluated from a more comprehensively designed model. It was apparent from the nature of the gaps in knowledge, and from the complexity and variety of emerging technologies, that several components of the oral cancer management/treatment pathway required modeling in a more granular way than was possible with the approach from Chapter 3. This novel granular approach should be able to credibly evaluate single decisions, but also evaluate multiple decisions in concert given the number of technologies that are likely to require evaluation in the near future. Chapters 5 and 6 describe an approach to these single- and multiple-decision evaluations, respectively.

The question of ‘why’, as described in Chapter 1, relates to the overall goals of HTA: to guide policy making in such a way that produces the greatest possible health outcomes for the population of interest (in this case, British Columbians) from a given level of budgetary constraint.

1.3.2 *Stage 2: Conceptualization and Design*

The structural arrangement of the conceptual model (i.e., how the various elements were organized, and how the relationships between them were described) was informed by a text written by Jaime Caro and colleagues[180]. This text was used primarily as the theoretical background for the discrete event simulation methodology, and formed the basis for much of the programming and the way the model is presented visually.

Elements from previously-published models were also consulted in the design of the WDMOC. The oral cancer screening model published by Speight and colleagues[124] was consulted to establish elements of the model's breadth, particularly with regard to the development, detection, and management of preclinical disease. The model developed in Chapter 3 was developed with the Speight model in mind, while adding some necessary depth to the management of both detected OPLs and detected invasive cancers. Much of this depth, particularly with respect to the management of invasive disease and recurrence, was also taken from clinical guidelines published by the National Comprehensive Cancer Network (NCCN)[47], and by the British Columbia Cancer Agency (BCCA)[37, 46].

In order to translate a conceptual model into an empirical one, data about each step and event within the process is required. However, this kind of data is not always available. Accordingly, model design requires balancing the complexity of the real world and the pragmatic limitations of data availability. Many components of the conceptual model were informed from retrospective cohorts of people who had been treated for oral cancer and pre-cancer within the BCCA. These cohorts are discussed in greater detail in section 4.5.

An initial draft of the model was presented to members of the stakeholder group during the interviews, for their expert feedback (see Appendix 4.2 – Initial Model Structure) Some key findings from these interviews were incorporated into the model's final structure, including:

- Treatment options for invasive cancers – what patient and disease factors influence the type of treatment prescribed. Estimates of treatment duration.
- Relationship between community dentists and oral health specialists with respect to detection, referral, surveillance, and treatment of premalignant lesions (OPLs).
- Role of HPV with respect to cancer incidence and implications for treatment

The stakeholder group also identified a number of ways in which the model's structural assumptions simplify detection, management, and treatment of OPLs. These limitations are discussed in detail in Chapter 7.

Through this process, the boundary, breadth, and depth of the WDMOC were determined for the conceptual model:

Model boundary

Based on feedback from the committee and the stakeholder group, The WDMOC was designed to simulate a population of adult British Columbians (age 20-60) who would be at risk of developing oral cancer within their lifetimes. Because anyone could potentially develop oral disease, this includes all adult members of the population, excluding those who currently have oral cancer. Because changes to the availability of dental care would likely impact the rate at which preclinical oral disease can be detected, people who do not have access to a dentist were also included. The model's structure is summarized in Section 4.3, and described in Figures 4.1 through 4.6.

It should be noted that the model does not adequately reflect the extent to which regional factors influence availability and type of treatment, especially in the context of premalignancy. While the scope of the model is provincial, it is worth noting that many aspects of the model structure are reflective of the Lower Mainland of British Columbia (i.e., Vancouver and the surrounding area) rather than being truly representative of the whole province. This regional bias and its implications will be discussed in greater depth in Chapter 7, but briefly it was agreed that this represented a reasonable 'starting point' upon which a more comprehensive model that reflects the complexity of practice outside the Lower Mainland could be built in the future.

Model breadth

The WDMOC is designed to reflect the full treatment/management pathway of oral cancer from premalignant disease to death. The model is divided into five principal 'components', each reflecting a related group of health care services used to address clinically meaningful stages of disease progression. The structure of each component will be described in greater detail in section 4.5.

Model depth

The level of detail used within each model component required a balance between the complexity necessary to adequately address the technologies undergoing assessment and the availability of data to inform parameters. Many cancer screening and treatment processes are highly individualized, and it is impractical (if not impossible) to build a model that is capable of reflecting all possible options for all possible people. Several simplifying assumptions were made, each of which will be discussed in Chapter 7.

Because the ultimate purpose of the WDMOC is to evaluate the impact of changes in health policy and technologies, members of the stakeholder group contributed suggestions of potentially impactful and/or emerging technologies (including policies, programs, and services) for future evaluation beyond the initial implementation of the model as described in this dissertation. These suggestions, and the steps necessary to implement them within the WDMOC, will be described in Chapter 7.

1.4 Conceptual Model of the WDMOC

The WDMOC simulates the trajectory of hypothetical people ('entities') through the oral cancer pathway from preclinical disease through the development of invasive cancer to death from terminal illness. The following section will describe the structure of the WDMOC, and the path that entities can follow from creation to termination.

The WDMOC involves the creation of entities (simulated people), whose disease status is informed by a “Natural History” model, and whose health care interactions are informed by a “Clinical Trajectory” model. The Natural History model describes the development of *de novo* OPL and its progression to invasive squamous cell carcinoma (SCC) of increasing severity. The Clinical Trajectory model describes the health care system processes through which premalignancies and invasive SCC are detected and managed, and is divided into five interacting ‘components’.

The structure of the model, by each component, is presented in the following section. The structure and entity path assumptions were determined iteratively through review of the literature, input from the stakeholder team, and the available data to inform parameter estimates (which will be described in greater detail in section 4.5).

Entity Creation

The process by which an entity is created precedes both the Natural History and Clinical Trajectory models. Entities are assigned a set of personal characteristics that informs the way they will move through the model:

- Age at start of model
- Sex (binary M/F)
- Smoking status (ever/never, by sex)
- Alcohol use (heavy/non-heavy, by sex)
- Access to a dentist
- Date of death from causes other than oral cancer

This list of characteristics were drawn from the literature reviewed in Chapter 2, with input from members of the stakeholder group who agreed that they were reasonably comprehensive for this initial model. The entity may also start the model with an undetected OPL, based on their age and sex. This prevalence-based approach was used in a previously published decision model[124], and its limitations

will be discussed in Chapter 7. The OPL is assigned a risk profile of 'high', 'medium' or 'low', based on its level of loss of heterozygosity (LOH), informed by data published in the literature (see section 4.5).

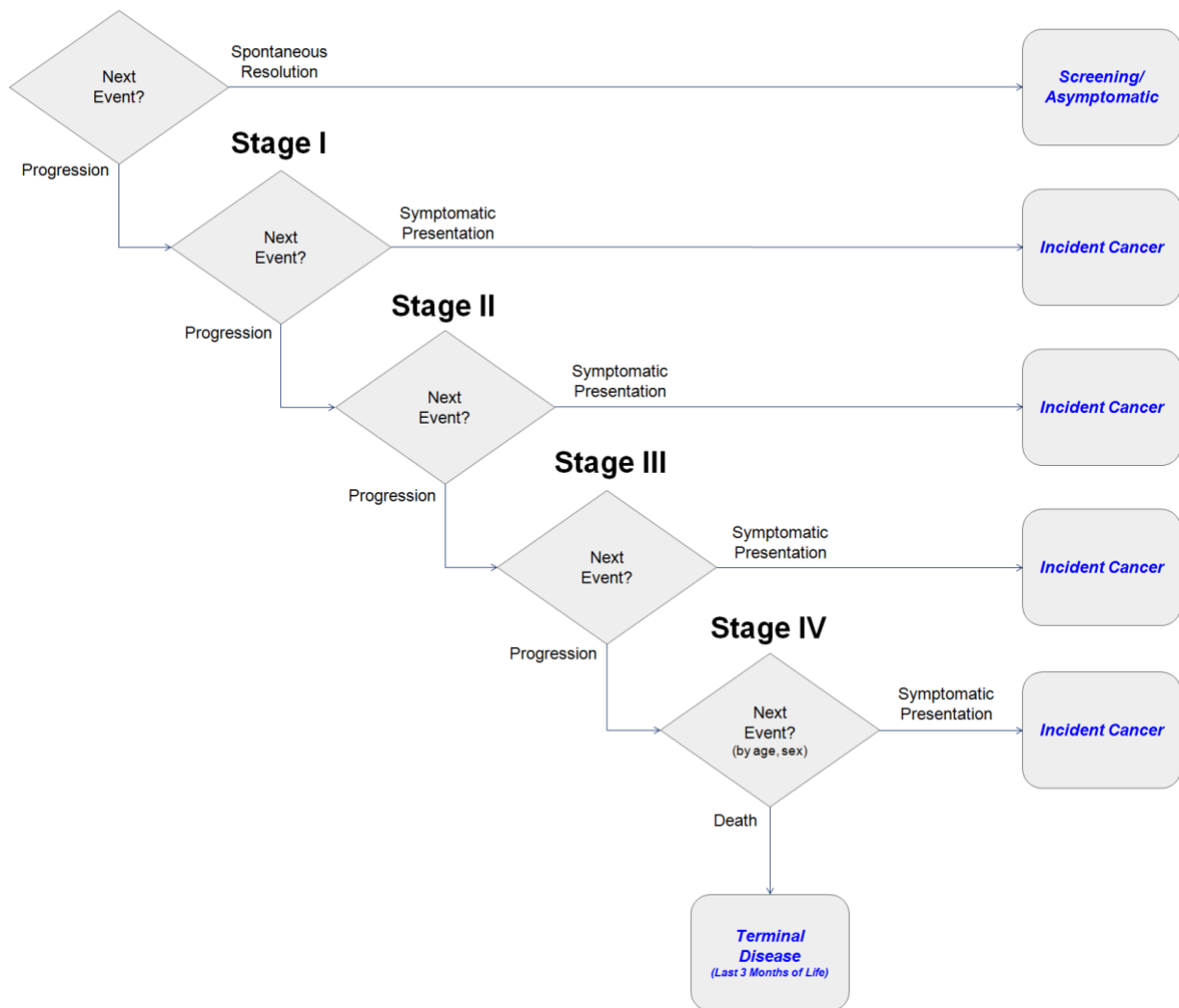
1.4.1 *Natural History Model*

Based on demographic characteristics (age, sex), newly-created entities may have an OPL that can be detected through screening. An OPL may progress to SCC, or it may spontaneously resolve (i.e., return to normal, non-diseased epithelial tissue), based on the entity's age, sex, smoking status, and LOH risk profile.

SCCs start at stage I and may progress to a higher stage (i.e., $I \rightarrow II \rightarrow III \rightarrow IV$), or may present symptoms that drive a person to seek medical care outside routine care, at which point the disease is detected and diagnosed – this part of the process occurs in the 'Incident Cancer' component of the Clinical Trajectory model. Based on expert input, the WDMOC assumes that terminal undetected stage IV cancers (i.e., an entity can die from an undetected stage IV cancer) are detected symptomatically three months before death or less – this part of the process occurs in the 'Terminal Disease' component of the Clinical Trajectory model. The process is illustrated in Figure 4.1.

Figure 1.1 – The Natural History model component

Undetected OPL



It is important to note that this modeling approach uses *prevalent* OPL cases rather than *incident* ones. This approach was adapted from a previously-published oral cancer screening model[124], but places meaningful limits on the WDMOC's function. The reasons for this choice and the implications of the resulting limitations will be discussed in Chapter 7.

1.4.2 *Clinical Trajectory Model*

The Clinical Trajectory model is divided into five components, each representing a set of health care system processes for management of oral cancer at various stages. Entities move through the components according to their disease status (i.e., their progression within the Natural History model) and their clinical history (i.e., the events that have happened previously in the Clinical Trajectory model).

The five components are organized as follows:

1. **Screening/Asymptomatic**
2. **Oral Premalignant Lesion (OPL)**
3. **Incident Cancer**
4. **Follow-up**
5. **Terminal Disease**

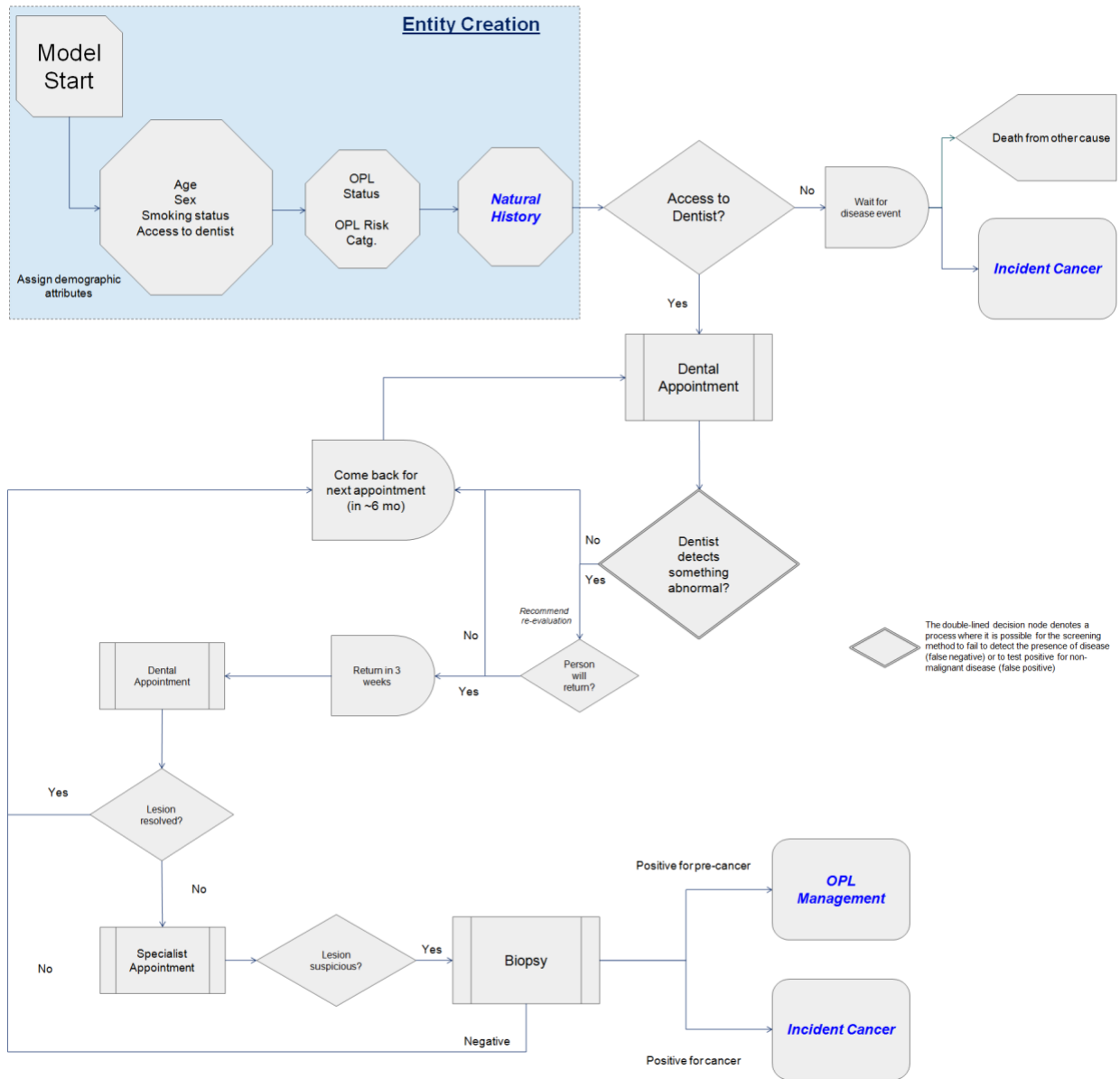
Entities pass through these components from the start of the model run (i.e., the creation of an entity) until they reach a terminal state, which simulates either death from oral cancer or from another cause.

The paths that entities can take through each component are described in detail below.

1.4.3 *Screening/Asymptomatic Component*

Entities with access to a dentist will be seen at regular intervals for a dental checkup. If the entity has developed a premalignant lesion, it may be detected in a routine exam. If no lesion or other abnormality is detected, the entity will return for their next appointment after a period of time. The component is described graphically in Figure 4.2.

Figure 1.2 – Screening/Asymptomatic model component



Based on input from the community practitioners within the stakeholder group, entities with a detected lesion are asked to return in three weeks. If the lesion persists beyond at three-week follow-up, the entity is referred to an oral health specialist (periodontist, oral medicine specialist, oral surgeon) for additional scrutiny. The specialist will perform a biopsy of any lesion that is deemed suspicious for premalignancy.

Premalignant lesions are detected in this way, and referred for pre-malignant management (OPL component). Invasive cancers may also be detected in the course of routine dental care and are referred for curative treatment (Incident Cancer component). Entities with non-malignant lesions and/or lesions that resolve within the three-week period return for routine dental checkups after a period of time.

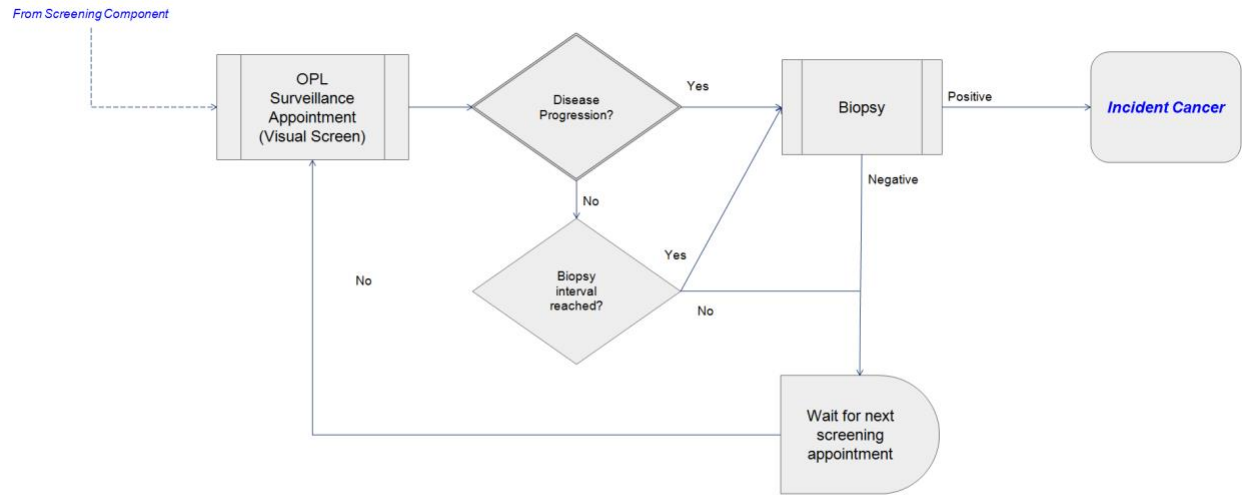
It is possible for the screening procedure to return a false negative (i.e., the entity *has* premalignant or malignant disease, but a negative test), in which case they will not be re-screened until the next screening appointment. Their disease may progress during this time.

Entities with no dental access will not have premalignant lesions detected, and any invasive cancer that may result can only be detected symptomatically. If they do not develop oral cancer, they will eventually die of another cause. The rate of non-oral cancer related death is described in section 4.5.

1.4.4 *Oral Premalignant Lesion Component*

Entities with a detected premalignant lesion will undergo regular evaluations by a specialist for evidence of progression to malignant disease. If progression is suspected, the entity will undergo a diagnostic biopsy. The entity's OPL may be biopsied at regular intervals as well, after a period of time (based on stakeholder group input). If no progression is suspected or detected, the entity will return for another evaluation after a period of time. Detected invasive cancers are referred for treatment (Incident Cancer component). The component is described graphically in Figure 4.3.

Figure 1.3 – Oral Premalignant Lesion management model component

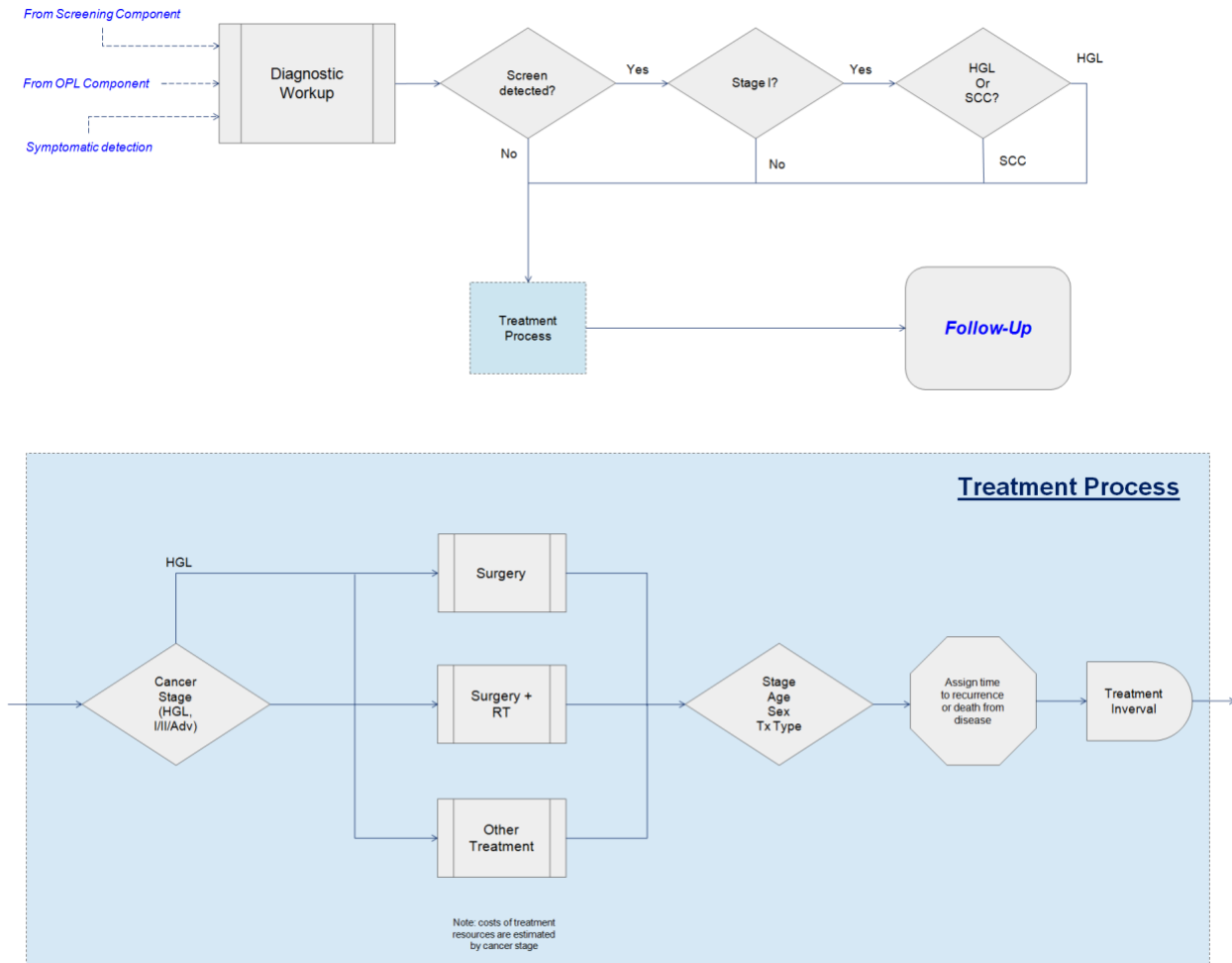


The double-lined decision node denotes a process where it is possible for the screening method to fail to detect the presence of disease (false negative) or to test positive for non-malignant disease (false positive)

1.4.5 *Incident Cancer Component*

Entities with a detected invasive cancer will undergo a diagnostic workup to determine the disease stage. If a stage I disease was detected through screening (either in the Screening/Asymptomatic component or the OPL component), it may be classified as either a high-grade lesion (HGL) or a squamous cell carcinoma (SCC), reflecting the possibility that some lesions may be referred for treatment before they develop invasive malignant characteristics. Stage I cancers detected symptomatically are assumed to be SCC. The model treats all cancers of stage II or higher as SCC – a simplifying assumption made based on feedback from the stakeholder committee. The component is described graphically in Figure 4.4.

Figure 1.4 – Incident cancer model component



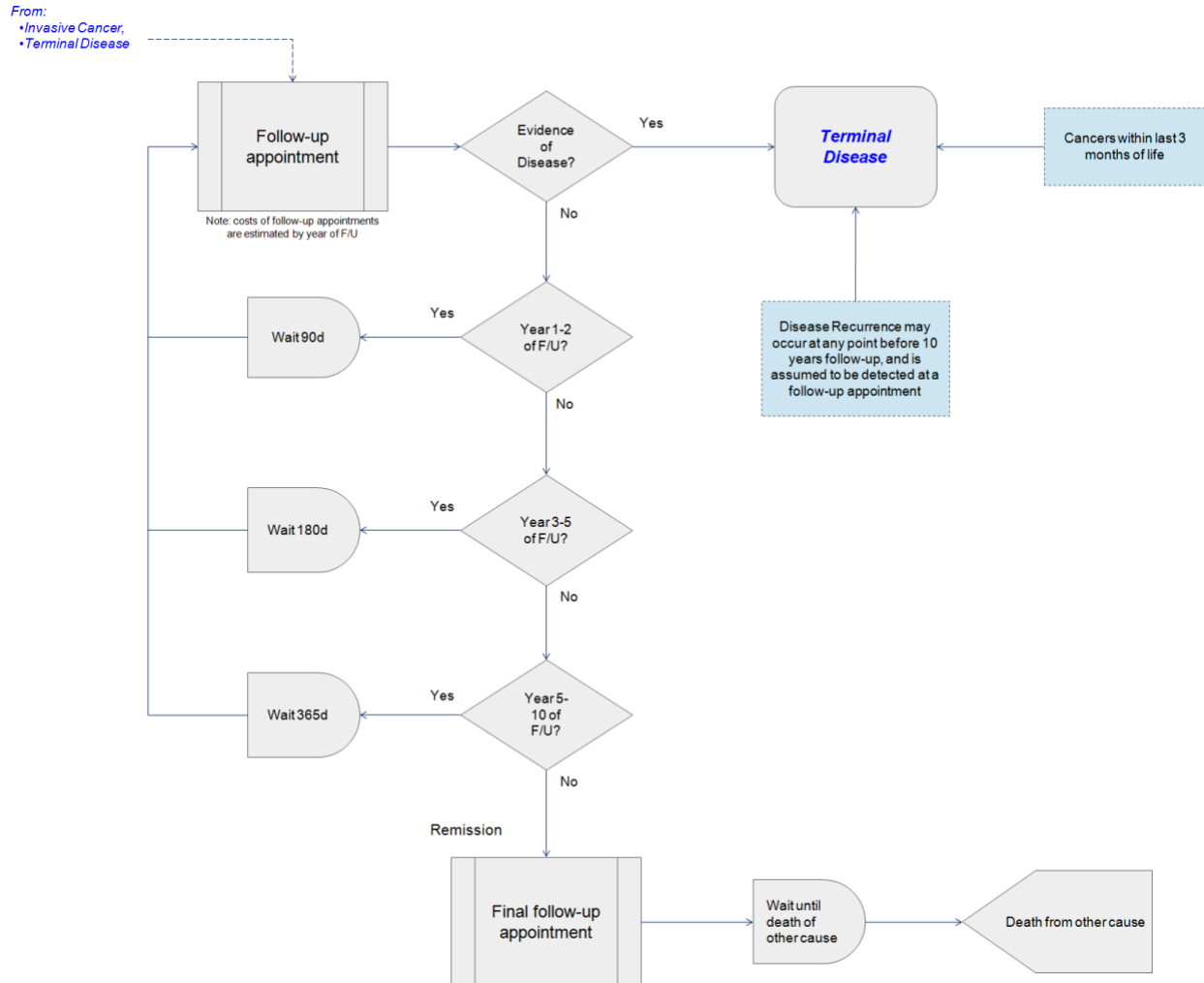
Cancers at all stages undergo a Treatment Process, based on stage at diagnosis. Treatment falls into three categories: surgery alone, surgery with adjuvant external beam radiotherapy (RT), or another treatment that may include any combination of surgery, chemotherapy, and RT. All HGL are managed with surgery alone. These assumptions were informed by feedback from the stakeholder group and the availability of data as described in section 4.5.

Based on demographic and clinical characteristics (stage, age at detection, sex, and treatment type), an entity is assigned a time at which they either experience disease recurrence or death from disease. After a period of time, entities begin attending post-treatment follow-up (Follow-up component).

1.4.6 *Follow-up Component*

After their treatment is finished, entities return at regular intervals to see their oncologist(s) to evaluate their tumour site for evidence of disease progression. The frequency of evaluations is determined by the length of time since the entity was treated, becoming less frequent as the entity approaches ten years of follow-up care, reflecting clinical practice guidelines and input from the stakeholder group. If the entity is alive and disease-free after ten years, it is assumed to be in full remission and will die of a cause other than oral cancer at the time determined in the Natural History model. The component is described graphically in Figure 4.5.

Figure 1.5 – Cancer follow-up model component

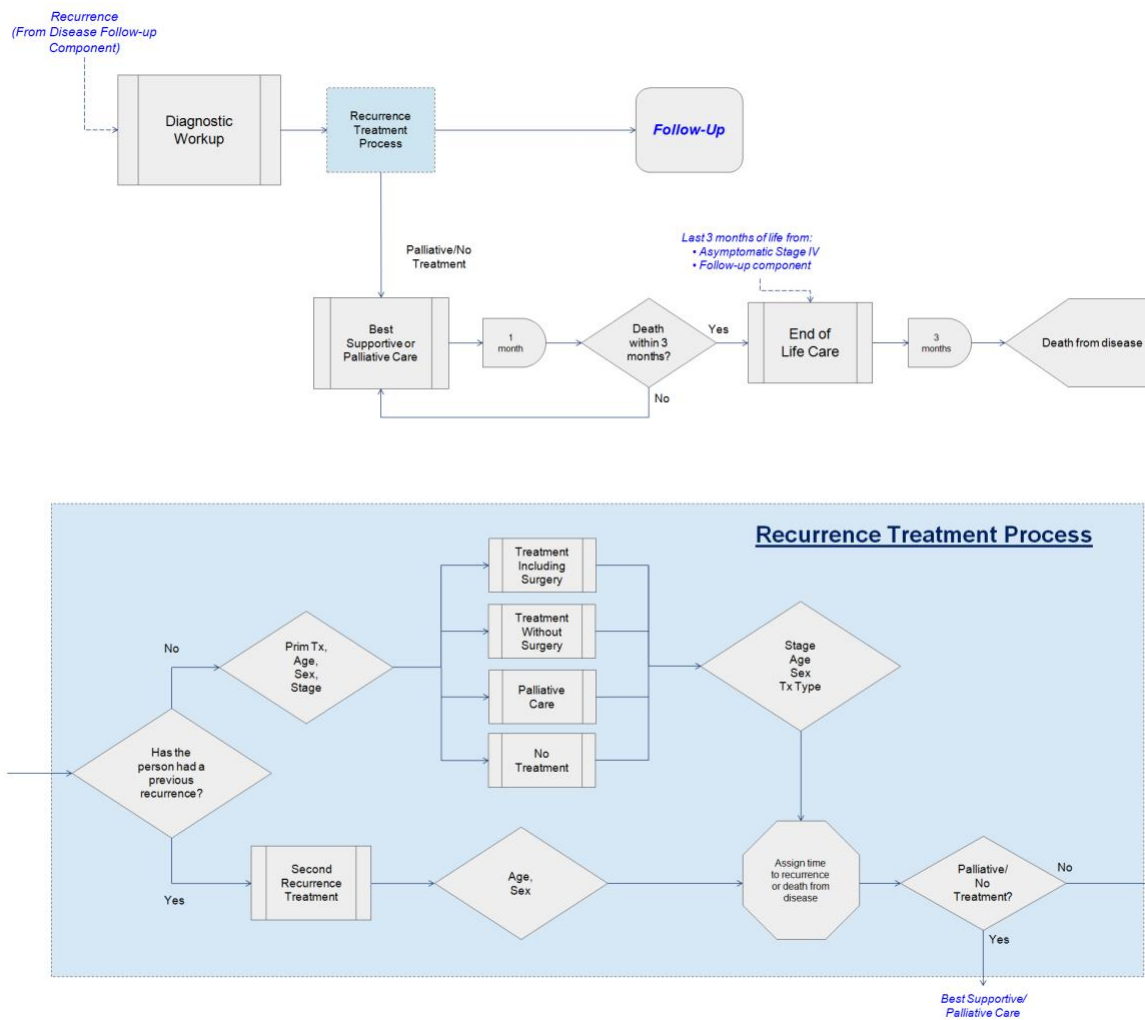


Death from disease or recurrence occurring during the follow-up period is managed within the Terminal Disease component.

1.4.7 Terminal Disease Component

Entities may enter this component either as a result of a detected recurrence, an undetected Stage IV cancer that is within three months of death, or a cancer undergoing follow-up that is within three months of death. The component is described graphically in Figure 4.6.

Figure 1.6 – Terminal disease management component



Recurring cancers are diagnosed and then undergo a Recurrence Treatment Process. Treatment falls into four categories: curative including surgery, curative not including surgery, palliative, or none. These categories are based on the parameter estimation process described in section 4.5. A time to death from

disease or second recurrence is calculated based on the entity's demographic and clinical characteristics (stage at first diagnosis, age at recurrence, sex, recurrence treatment type). If the entity has received curative treatment (surgery, non-surgery) they return to the Follow-up component.

Entities with a second recurrence undergo a similar Recurrence Treatment Process, but treatment is assumed to be identical for all patients, and time to additional recurrence or death is based only on age and sex, based on the parameter estimation process described in section 4.5.

Entities receiving palliative or no treatment are managed on a monthly basis until they are within three months of death, at which point they receive End-of-Life (EOL) care and die from disease. It is possible for someone to undergo full remission with best supportive or palliative care within the model.

Entities entering the Terminal Disease component as a result of terminal undetected stage IV cancer, or whose cancer is within three months of death from disease, receive EOL care and will die from disease. The model assumes that it is not possible for these cancers to go into full remission.

1.5 Stage 3: Implementation Modeling

An individual sampling model using a time-to-event approach was chosen[26]. Individual sampling models are a type of discrete event simulation model in which entities do not interact. Interaction between entities is of particular importance when a model needs to account for queuing for scarce resources, or in the case of infectious disease modeling where entities can influence each other's disease status. The WDMOC does not consider queuing, and assumes that all resources (including specialist care) are available instantaneously when the entity needs them. This is a common assumption in health economic decision modeling.

The WDMOC was programmed in the Python language (Python Software Foundation, Delaware, USA). Python was chosen in favour of the commercial software used to create Tappenden's original Whole

Disease Model. Python is an open-source and web-ready language that is free to license. While further reasoning behind this choice is detailed in Chapter 7, The primary rationale for this choice was to make a model that could easily be adapted, updated, and re-used by different researchers in different contexts. No model, no matter how well-designed, can adequately reflect the full breadth and depth of any decision environment. Whole disease models are still designed to reflect a single jurisdiction (i.e., a province, a health authority, a country), and simplifying assumptions that may be valid within one jurisdiction may not apply in another. Accordingly, in order to be useful beyond its original context, these models should be easy to edit and share across health care policy making jurisdictions, which implies the need for open-source and free software. Designing models in this way allow them to be adapted to quickly and seamlessly reflect not only differences between jurisdictions, but also technological and policy innovations that may be developed in the future.

1.5.1 *Approach to simulation modeling*

Individual sampling models simulate the movement of ‘entities’ (simulated people, in this case) through an environment (the parts of the health care system relevant to oral cancer, in this case). Entities experience changes to their characteristics and use resources over the course of their simulated time; these changes and resources occur during particular ‘events’ that occur at each step of the process. A population can be simulated by generating several entities and running them through the same environment.

A crucial step in any economic evaluation is assessing the impact that uncertainty (i.e., unknown information that is relevant to the decision being evaluated) has on its outputs. Uncertainty in economic evaluation is described within four categories[181]:

1. **Stochastic uncertainty** (or 'first-order' uncertainty) concerns the the random variability in outcomes that occurs between people with identical characteristics (e.g., people of the same age, sex, disease type, etc.). It is analogous to random error in a regression analysis.
2. **Parameter uncertainty** (or 'second-order' uncertainty) concerns the variability that surrounds each value the model uses to estimate its outputs. the level of uncertainty or imprecision in the estimation of a particular model parameter (e.g., cost of a resource, time to developing symptoms, probability of a false positive, etc.). It is analogous to the standard error of a coefficient estimate in a regression analysis.
3. **Heterogeneity** concerns characteristics of the population being modeled that may impact the magnitude of costs and outcomes. Rates of recurrence may be influenced by age and sex, and will differ between subgroups within each population.
4. **Structural uncertainty** concerns the assumptions that are inherent to the design of the model. For example, even a good model may necessarily exclude potential outcomes, simplify the relationships between events, and use data estimated independently from different populations. These structural decisions may affect the output of the model in unknown ways.

By allowing model input values (i.e., the variables that inform the model) to vary between each entity, individual sampling models can reflect first- and second-order uncertainty. First-order uncertainty was reflected in the random draws from the uniform distribution used to evaluate the assignment of characteristics and the sequencing of events – these random draws mean that two entities with identical characteristics will not necessarily follow identical paths through the model.

Second-order uncertainty was reflected in the random draws for the values of each model parameter, from their underlying distribution – this process means that the value of a given parameter (e.g., probability of receiving a certain type of treatment) is different for each entity, based on the uncertainty around that parameter.

This process of probabilistic random sampling is known as Monte Carlo simulation, and is a commonly-used technique in decision modeling[182]. Heterogeneity and structural uncertainty were evaluated through the use of sensitivity analysis, which will be described in section 4.8.

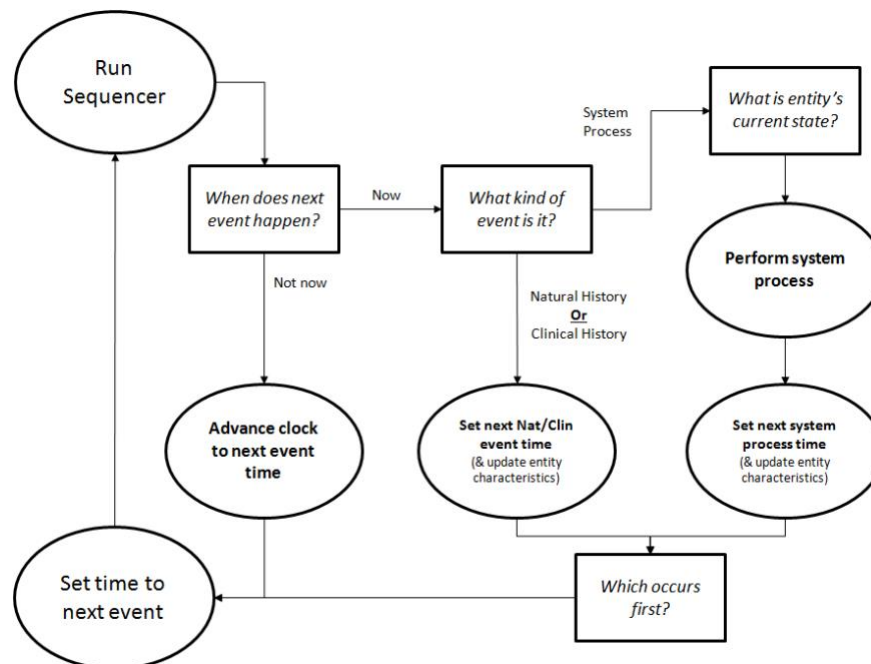
1.5.2 Time-to-event processing in the WDMOC

The general structure of the time-to-event approach used in the WDMOC relies on five principal types of programs. These are:

1. Sequencer: directs progress of entities from creation to a terminal condition (death)
2. Clock ('CheckTime.py'): identifies current simulated time and schedules next occurring event
3. Natural History processes: describe an entity's trajectory through the natural history of oral precancer and undetected invasive cancer
4. System Processes: describe an entity's trajectory through the health care system (screening, cancer treatment, follow-up)
5. Global processes: contain functions that are used by other programs

The relationship between these functions is described in Figure 4.7.

Figure 1.7 – Summary of simulation model implementation



The Sequencer creates a new entity, and that entity is assigned characteristics that will be used to determine its natural history and whether or not it will receive screening. Entities are then assigned a natural history – times at which different disease events will occur (development of precancer, progression to invasive disease, symptomatic detection, etc.). The next event to occur (either a natural history event or a system process event) is read by the CheckTime program and the system clock is advanced to that time. The Sequencer then runs the appropriate programs (system and/or global processes) that update the entity's characteristics and determine the next event to occur. Resource units and utilities are also appended to each entity as events occur and health status changes.

This process continues on a loop until the entity reaches a terminal state, at which point a new entity is created and the process restarts. The model runs until a user-defined number of entities has been simulated. The cohort of simulated entities can then be analyzed to determine survival, quality-adjusted survival, and cost (through resource unit costing) for each entity.

All model parameters are read in from a master spreadsheet containing estimates of the mean and standard deviation of each variable. By adjusting values in the spreadsheet, the model can simulate multiple cohorts moving through the same (or similar) policy environments. Incremental cost-effectiveness analysis can be conducted by comparing mean costs and outcomes in these cohorts. Because the master spreadsheet can be adjusted in ways that affect multiple decisions within the model's breadth, it is possible to analyze the cost-effectiveness of several policy decisions simultaneously.

Additionally, the approach used in the implementation of the WDMOC allows for entire subsections of the treatment pathway (e.g., a change in surgical management of early-stage oral cancers) to be programmed *de novo* and inserted into the appropriate place within the overall model. By making a

small adjustment to the Sequencer, entities can be routed through the newly-programmed subsection, allowing for the model to be updated and/or customized to a variety of policy-making settings.

An example of such a *de novo* substitution is presented in Chapter 5. An example of multiple simultaneous evaluations is presented in Chapter 6. Chapter 7 will discuss how these two methods can be easily incorporated.

1.5.3 *A description of model functions*

There are five basic types of functions used in this type of simulation modeling. A discussion of these nodes follows, and examples of each type are included in Appendix 4.3.

1. **Entity** – an entity (a simulated person) takes the form of a Python library that can be expanded to add any useful information such as age, treatment flags, current time and time to next event, among others.
2. **Parametric sampling** – all values in the model are generated probabilistically, based on a mean and standard deviation, as well as a value denoting the assumed parametric distribution of the parameter uncertainty (e.g., Beta distributed, Gamma distributed, etc.). These parametric assumptions follow ISPOR guidelines[30, 182]. Some model parameters are input as coefficients from generalized linear equations.
3. **Probability nodes** – values sampled through the above process are compared to a randomly generated number from a uniform distribution. If the randomly-generated number meets a given condition (i.e., is greater or less than the sampled value) then a model-defined outcome will arise (e.g., an entity's characteristics will change, an event will be scheduled, a natural history event will be encoded, etc.)
4. **Time** – time to next event is handled through the 'CheckTime.py' program. Briefly, the next event to occur (Natural History, System Process, Clinical History) is sampled from a parametric distribution based on the entity's characteristics at various points throughout the entity's simulated life. The various values of time are compared during each loop of the Sequencer, and the next to occur is scheduled. When that time is reached, the entity's characteristics are updated to reflect the occurrence of the event.
5. **Resources, utilities, and events** – lists that record the resource used or events occurring, and the system time at which they occurred. These can be compared at the end of the model run.

Once the specified number of entities has been simulated, the lists of resources, utilities, and events can be compared in incremental analysis. Resources are converted to costs through a unit costing approach, in which a monetary value is estimated for each resource unit.

Costs are discounted to account for society's preference for goods now rather than in the future – a concept known as future time preference[5, 19]. Discounting is a distinct concept from currency inflation. Discounting was applied according to the following formula:

$$Present\ Value = Future\ Value \cdot (1 + i)^{-t}$$

Where i is an annual discount rate, and t is the amount of time in the future that the cost occurs, expressed in years. An annual discounting rate of 1.5% was used, based on guidelines established by CADTH[5]. Unit costs are applied to resources utilized by each entity, and are then discounted and summed to estimate the total costs experienced by that entity over the course of their trajectory through the model.

Simulating multiple entities and estimating costs for each produces an estimate of mean costs for the simulated population. Mean costs from different simulated populations can be compared in incremental analysis to produce the incremental cost term (ΔC) of an Incremental Cost Effectiveness Ratio (ICER – see Chapter 1).

1.6 Model Parameterization

The probabilities, times to event, resource unit costs, and utilities that govern the model's behavior are represented as model parameters – numerical estimates of variables used within the model.

Parameterization of the model was undertaken using both primary data and secondary data.

Primary data is data that was collected directly from participants in research studies and retrospectively-collected datasets. This stochastic data allowed for several model parameters to be estimated in ways

that make the operation of the WDMOC reflect multiple aspects of clinical realities at a depth that is not typically possible with published (secondary) data.

Secondary data is collected from sources in the scientific literature. It is common to need to convert the best available data into a form that is useable by the model through statistical methods. The ways in which both these types of data were converted into useable model data are described below.

Primary Data

Primary data for this project was derived from two sources:

1. A Retrospective Oral Cancer Cohort (ROCC), comprised of the electronic medical records of 864 patients previously treated for oral cancer in the province of British Columbia between January 1, 2001 and December 31, 2015 (diagnosed between January 1, 2001 and December 31, 2009). These records were identified by a member of the Oral Cancer Control Program's research team through the Provincial Cancer Registry. A chart review was conducted by this same researcher to identify relevant clinical dates (e.g., diagnosis, treatment, recurrences, death) for each person within the cohort. This dataset was linked through Popdata BC and the BC Cancer Agency's Information System (CAIS) to identify resources used from diagnosis to death, censoring, or loss to follow-up.
2. Anonymized data from the clinical trial conducted by Zhang and colleagues (2012) exploring the role that LOH plays in oral cancer development. Times between lesion detection and progression to cancer were observed within this cohort, as well as some basic demographic information (age, sex, tobacco use, alcohol use).

These sources were analyzed to determine statistical associations between individual entity characteristics and times to events of interest. Parameter inputs were derived from these data sources using linear regression methods, described in section 4.6.

Secondary Data

A variety of secondary data sources were used to build the WDMOC, each with its own appropriate method of being incorporated into the simulation process. These parameterization methods are described in section 4.6.

The parameters used in each component of the WDMOC are described in the following sections. The methods used to value each parameter are based on assumptions about the statistical distribution that each parameter takes. The process by which estimates of each parameter were derived for each entity are described in section 4.6.

1.6.1 Entity Creation

Newly-created entities are assigned demographic and disease characteristics, derived from four sources: population statistics published by Statistics Canada, figures published by the Canadian Dental Association, values published in the Speight *et al* oral screening model, and values published in the Zhang *et al* trial. Parameter inputs are summarized in Table 4.1.

Table 1.1 – Parameter estimates: entity creation

Parameter	Mean	SD	Distribution	Source
Smoking prevalence (British Columbia)			Beta	[183]
Men	0.174	0.0003		
Women	0.113	0.0002		
Has access to a dentist (British Columbia)	0.688	0.0002	Beta	[184]
Prevalence of oral premalignancy			Lognormal	[124]
Men				
<50	-3.132	0.258		
50-59	-2.817	0.234		
60-69	-2.788	0.261		
70-79	-3.040	0.370		
80+	-2.670	0.520		
Women				
<50	-4.132	0.311		
50-59	-3.817	0.283		
60-69	-3.788	0.310		
70-79	-4.040	0.403		
80+	-3.671	0.545		
LOH Risk Score	Count [†]		Dirichlet	[61]

Low	130
Medium	120
High	28

† – counts, rather than means, were used to calculate the probabilities and uncertainty using the Dirichlet distribution.

Smoking prevalence

Statistics Canada publishes smoking rates for each province by sex. These values were expressed as probabilities assuming a Beta distribution.

Access to a Dentist

The percentage of British Columbians with regular access to a dentist was estimated from rates published by the Canadian Dental Association. This value was expressed as probabilities assuming a Beta distribution.

Prevalence of Oral Premalignancy

Newly-created entities are assigned a probability of starting the model run with an undetected premalignancy, based on the underlying prevalence of OPLs in the population. This prevalence, by age and gender, was estimated based on values published in the Speight *et al* model. These values were originally estimated from an opportunistic screening study conducted in the general population of the UK. Individual-level data was analyzed using logistic regression. The published coefficients of the log odds from the regression were used to calculate probabilities in the WDMOC assuming a lognormal distribution.

Progression risk score

Entities with an OPL were assigned an LOH risk score based on the prevalence of high, medium, and low-risk LOH profiles published in the Zhang *et al* trial. Counts of each risk group were converted to probabilities assuming a Dirichlet distribution.

1.6.2 *Natural History*

Parameters for the WDMOC's natural history processes were derived from two sources: secondary analysis of the Zhang *et al* trial cohort, and the values published in the Speight *et al* oral screening model. Parameter inputs are summarized in Table 4.2.

Table 1.2 – Parameter estimates: natural history

Parameter	Mean	SD	Distribution	Source
Time to OPL progression to Stage I cancer			Weibull GLM	Zhang et. al cohort
Intercept (β_0)	9.296	0.744		
Sigma (σ)	0.801	0.106		
Age	0.0013	0.010		
Sex				
Men	Ref.			
Women	-0.0804	0.257		
LOH Risk Score				
Low	Ref.			
Medium	-0.868	0.449		
High	-1.249	0.391		
Time to symptomatic detection of cancer			Weibull	[124]
Stage I	0.27	0.19		
Stage II	0.56	0.27		
Stage III	0.68	0.28		
Stage IV	0.71	0.3		
Time to undetected cancer progression			Weibull	[124]
Stage I to Stage II	0.53	0.27		
Stage II to Stage III	0.59	0.25		
Stage III to Stage IV	0.67	0.25		
Time to death from Stage IV			Weibull	[124]
Men				
<50	0.378	0.251		
50-59	0.439	0.251		
60-69	0.487	0.251		
70-79	0.670	0.251		
80+	1.00	0.01		
Women				
<50	0.320	0.224		
50-59	0.372	0.224		
60-69	0.412	0.224		
70-79	0.566	0.224		
80+	0.873	0.224		

Time to OPL progression to Stage I cancer

These values were estimated from a secondary analysis of data from the Zhang *et al*/ trial. Time to progression was fit to a parametric Weibull survival curve. Details on this process are available in Appendix 4.4. Coefficients from a GLM regression with a Weibull link function of observed survival with

multiple covariates (age, sex, LOH risk score) were used to estimate survival times and probabilities for each entity.

Time to Symptomatic Detection of Cancer

The time between developing an invasive cancer and that cancer being detected through symptomatic presentation was estimated using transition probabilities published in the Speight *et al* model. The authors calculated these values by eliciting expert opinion from health care practitioners concerning the proportion of people whose cancer would be detected in the absence of a routine screening program. These assumptions were reviewed by the committee and were determined to be appropriate within a BC context. These probabilities were converted to time-to-event values using the Weibull method of moments.

Time to Undetected Cancer Progression

The time that it takes for an undetected cancer to advance in stage was estimated using transition probabilities published in the Speight *et al* model. These values were originally estimated by eliciting expert opinion from health care practitioners concerning what proportion of undiagnosed patients, at each stage of disease, would progress to the next stage within a year. These probabilities were converted to time-to-event values using the Weibull method of moments.

Time to Death from Undetected Stage IV Cancer

The time between developing an undetected Stage IV cancer and death from disease was estimated using transition probabilities published in the Speight *et al* model. These values were originally estimated from a retrospective analysis of survival in oral cancer patients undergoing treatment, and making an assumption of a more dire prognosis for undetected cancers, using an exponential regression process. These values were reported by age and sex. The probabilities were converted to time-to-event values using the Weibull method of moments.

1.6.3 Screening/Asymptomatic Component

Values concerning the passage of asymptomatic people through regular dental appointments were derived from three principal sources: the Speight *et al* oral cancer screening model, an opportunistic screening study from the US, and a similar study conducted in British Columbia. Parameter inputs are summarized in Table 4.3.

Table 1.3 – Parameter estimates: asymptomatic/screening component

Parameter	Mean	SD	Distribution	Source
Regular appointment interval (days)	180	30	Normal	Assumed
Return appointment interval (days)	21	3	Normal	Assumed
Sensitivity of visual screening	0.965	0.012	Beta	[124]
Specificity of visual screening	0.848	0.043	Beta	[124]
Probability of non-OPL lesion	0.0452	0.0071	Beta	[73]
Probability that non-OPL lesion resolves within interval	0.290	0.0408	Beta	[72]
Probability of non-OPL biopsy	0.111	0.0524	Beta	[72]
Probability of attending return appointment	0.350	0.0242	Beta	[72]

Appointment intervals

The WDMOC assumes that all entities are managed according to guidelines, returning for regular dental visits every six months (180 days). If a suspicious lesion (either OPL or another type) is found during the course of an appointment, an entity returns after 21 days for re-evaluation. Data on the number of days between appointments was not available, so parameter uncertainty was assumed and these assumptions were verified by the stakeholder committee. These values were assumed to be normally distributed.

Sensitivity and Specificity of Visual Screening

Values describing the ability of routine dental screening to accurately detect lesions in the mouth were derived from corresponding values published in the Speight *et al* oral screening model. These values were originally synthesized through a meta-analysis of multiple studies evaluating the effectiveness of routine screening. These values were expressed as probabilities assuming a Beta distribution.

Non-premalignant lesions

The WDMOC accounts for the presence of oral lesions that are not premalignant but will nevertheless trigger a response from a dentist if they are detected (i.e., they will be asked to return for re-evaluation). The rate at which these lesions occur was estimated based on values published in a population screening study in Boston.

Non-premalignant lesions may resolve spontaneously upon return screening. Non-resolving lesions may be biopsied if they are suspected to be premalignant. Some entities may choose not to return for re-evaluation. Values for each of these outcomes were estimated based on results published in an evaluation of an oral cancer screening education program in British Columbia. All parameters were expressed as probabilities, assuming a Beta distribution.

1.6.4 Oral Premalignancy Component

Parameter values concerning management of people with detected OPLs were derived from the Speight oral cancer screening model and values published in the Cromwell *et al* OPL management model (i.e., the inputs from Chapter 3). Parameter inputs are summarized in Table 4.4.

Table 1.4 – Parameter estimates: oral premalignancy component

Parameter	Mean	SD	Distribution	Source
Appointment interval (days)	180	30	Normal	Assumed
Sensitivity of visual screening	0.965	0.012	Beta	[124]
Specificity of visual screening	0.848	0.043	Beta	[124]
Probability of SCC at detection	0.678	0.038	Beta	[148]

Appointment interval

The WDMOC assumes that all entities return for regular follow-up visits every six months (180 days).

Values were assumed to be normally distributed.

Sensitivity and Specificity of Visual Screening

The accuracy of visual screening in OPL was derived from the same source as described in the Asymptomatic/Screening component.

Probability of SCC at detection

Premalignant lesions undergoing observation may be surgically resected before progression to invasive cancer (SCC). These high-grade lesions (HGL) are managed surgically but have a different prognosis to SCC. In order to reflect this, the WDMOC assumes that a proportion of OPLs under surveillance are detected as HGLs. This proportion was calculated in the Cromwell *et al* cost-effectiveness model, and was derived from a retrospective cohort of OPL patients undergoing follow-up at the Vancouver Cancer Centre. This value was expressed as a probability, assuming a Beta distribution.

1.6.5 *Incident Cancer Component*

Parameters concerning treatment type and outcomes were derived from the ROCC. Parameter inputs are summarized in Table 4.5.

Table 1.5 – Parameter estimates: incident cancer component

Parameter	Mean	SD	Distribution	Source
Treatment type	Count [†]		Dirichlet	ROCC
Stage I				
Surgery	235			
Surgery + RT	26			
Other	58			
Stage II				
Surgery	124			
Surgery + RT	56			
Other	120			
Advanced stage				
Surgery	46			
Surgery + RT	43			
Other	145			
Treatment time (days)	90	0	Normal	Assumed
Time to First Event			Weibull GLM	ROCC
Intercept (β_0)	11.034	0.334		
Sigma (σ)	1.320	0.045		
Age	-0.0424	0.005		
Sex				
Men	Ref.			
Women	0.364	0.117		
Cancer Stage				
I	Ref.			
II	-0.376	0.145		
Advanced	-0.874	0.154		
Treatment type				
Surgery alone	Ref.			
Surgery + RT	-0.796	0.191		
Other treatment	-0.869	0.130		
Time to First Event – Death			Weibull GLM	ROCC
Intercept (β_0)	11.216	0.328		
Sigma (σ)	1.272	0.044		
Age	-0.043	0.005		
Sex				
Men	Ref.			
Women	0.307	0.114		
Cancer Stage				
I	Ref.			
II	-0.376	0.142		
Advanced	-0.925	0.150		
Treatment type				
Surgery alone	Ref.			
Parameter	Mean	SD	Distribution	Source
Time to First Event – Death				
Treatment type cont.				
Surgery + RT	-0.739	0.184		
Other treatment	-0.817	0.126		

Treatment Type

Treatment type was simplified to encompass three basic approaches – surgery alone, surgery + RT, and ‘other’. The ‘other’ category is primarily comprised of people receiving chemotherapy and radiotherapy, but other combinations were seen in the data as well. Treatment type was counted by stage at presentation, and expressed as probabilities assuming a Dirichlet distribution.

Time to First Disease Event

Time to first recurrence or death was estimated for each entity using the competing events approach. Briefly, time to a first event (either recurrence or death) was sampled from a parametric Weibull curve fit to the observed survival data, and the probability of the event occurring at that time was calculated using the hazard function. The corresponding probability was calculated from a second curve of time to death. These probabilities were compared to a random draw, and the nature of the event (recurrence or death) was determined. Details on this process are available in Appendix 4.5. Coefficients from a linear Weibull GLM regression of observed survival with multiple covariates (age, sex, stage at detection, treatment type) were used to estimate survival times and probabilities for each entity.

1.6.6 Follow-up Component

The time between each follow-up appointment was derived from NCCN guidelines. The WDMOC assumes some variability around guideline adherence using a Normal distribution. Parameter inputs are summarized in Table 4.6.

Table 1.6 – Parameter estimates: follow-up component

Parameter	Mean	SD	Distribution	Source
Interval between follow-up appointments (days)				[47]
0 to 3 years post-treatment	90	10	Normal	
3 to 5 years post-treatment	180	20	Normal	
5 to 10 years post-treatment	365	50	Normal	

1.6.7 *Terminal Disease Component*

Parameter inputs are summarized in Table 4.7.

Table 1.7 – Parameter estimates: terminal disease component

Parameter	Mean	SD	Distribution	Source
Treatment type	Count [†]		Dirichlet	ROCC
Recurrence				
Surgery	62			
No Surgery	48			
Palliative	20			
No Treatment	17			
Time to Second Event			Weibull GLM	ROCC
Intercept (β_0)	7.687	0.632		
Sigma (σ)	1.185	0.083		
Age	-0.008	0.009		
Sex				
Men	Ref.			
Women	0.028	0.229		
Recurrence treatment type				
Treatment includes surgery	Ref.			
Treatment does not include surgery	-0.691	0.261		
Palliative Care	-1.656	0.325		
No Treatment	-1.414	0.378		
Time to Second Event – Death			Weibull GLM	ROCC
Intercept (β_0)	8.080	0.647		
Sigma (σ)	1.182	0.086		
Age	-0.008	0.009		
Sex				
Men	Ref.			
Women	-0.244	0.233		
Recurrence treatment type				
Treatment includes surgery	Ref.			
Treatment does not include surgery	-0.814	0.271		
Palliative Care	-1.957	0.332		
No Treatment	-1.165	0.389		
Time from Second Recurrence to Death			Weibull GLM	ROCC
Intercept (β_0)	5.946	2.434		
Sigma (σ)	1.249	0.304		
Age	-0.036	0.041		
Sex				
Men	Ref.			
Women	-0.244	0.233		

[†] – counts, rather than means, were used to calculate the probabilities and uncertainty using the Dirichlet distribution.

Recurrence Treatment Type

Treatment type was simplified into three categories – treatment including surgery, treatment not including surgery, and treatment flagged as ‘palliative’. A fourth category was included for people whose charts indicated that they were not prescribed any treatment, either curative or palliative. In the retrospective cohort analysis described in Section 4.5, this fourth category had a statistically significantly different survival curve (see Appendix 4.5), and so was treated as a distinct population within the model. Treatment type was expressed as probabilities using a Dirichlet distribution.

Time to Second Disease Event

Time to second event (second recurrence or death following recurrence) was calculated through the same process as time to first event (see above). Coefficients from a linear Weibull GLM regression of observed survival with multiple covariates (age, sex, recurrence treatment type) were used to estimate survival times and probabilities for each entity.

1.6.8 Unit Costs

The costs of treatment were estimated via a linkage exercise between the retrospective cohort (ROCC) and data held by the BCCA and the Ministry of Health. The linkage exercise is described in Appendix 4.7. Briefly, unit costs were applied to retrospective records of tests, appointments, hospitalizations, chemotherapy drugs, radiotherapy, and provincially-insured drug prescriptions for each member of the cohort over a three-month period from each clinical event (initial treatment, recurrence, second recurrence) and preceding death. Estimates for unit costs were derived from Medical Services Plan (MSP) reimbursement rates[185], sources in the published literature, and expert opinion where necessary (see Appendix 4.7 for full description). All costs were expressed in 2017 Canadian dollars, adjusted for inflation using the Consumer Price Index for Health Care[186].

Unit costs for resources aside from cancer treatments were estimated primarily from MSP fee-for-service (FFS) reimbursement rates[185].

Unit costs are presented in Table 4.8. Parameter values were estimated for each entity assuming a Gamma distribution, except for FFS values which were assumed to be equal for all entities.

Table 1.8 – Parameter estimates: unit costs

Parameter	Mean	SD	Distribution	Source
Dental Appointment	\$43.10	N/A	FFS	[187]
Specialist Appointment	\$254.91	N/A	FFS	MSP – 03770
Dental Screening	\$0.00	N/A	FFS	Assumption
Biopsy	\$250.40	N/A	FFS	MSP – 03773
OPL surveillance appointment	\$59.51	N/A	FFS	MSP – 03785
Diagnostic Workup	\$591.29	N/A	FFS	MSP†
Treatment			Gamma	ROCC
Stage I				
Surgery	\$9,268.55	\$10,758.64		
Surgery + RT	\$21,219.00	\$17,525.96		
Other	\$7,630.33	\$9,051.55		
Stage II				
Surgery	\$19,299.16	\$25,503.44		
Surgery + RT	\$26,059.51	\$18,567.71		
Other	\$10,279.56	\$8,660.23		
Advanced Stage				
Surgery	\$38,185.45	\$40,438.87		
Surgery + RT	\$34,925.65	\$22,256.47		
Other	\$14,848.61	\$16,238.45		
Parameter	Mean	SD	Distribution	Source
Treatment				
Recurrence				
Includes Surgery	\$29,262.14	\$42,730.52		
Does Not Include Surgery	\$17,066.59	\$20,314.99		
Palliative	\$20,778.66	\$31,106.02		
No Treatment	\$11,119.52	\$8,518.00		
Second Recurrence	\$16,616.98	\$27,427.74		
End of Life	\$17,930.25	\$21,977.09		
Follow-up – 1 to 3	\$154.00	\$75.00	Gamma	MSP‡
Follow-up – 3 to 5	\$80.67	N/A	FFS	MSP – 33512
Follow-up – 5 to 10	\$80.67	N/A	FFS	MSP – 33512
Follow-up appointment – final	\$154.00	\$75.00	Gamma	MSP‡

Death from Natural Causes	\$0	N/A	Assumption
---------------------------	-----	-----	------------

FFS – fee for service; **MSP** – medical service plan; **ROCC** – retrospective cancer cohort; **N/A** – these are fixed costs that do not have any parameter uncertainty

†- Diagnostic workup includes ‘Diagnostic Examination and Consultation’ (MSP 03770), CT Scan (MSP 08693) and a PET Scan[188]

‡- This figure is a frequency-weighted estimate of appointments based on treatment type, reflecting the fact that patients may see one or multiple members of their medical team (surgeon, radiation oncologist, medical oncologist) at a given follow-up appointment.

1.6.9 *Health State Utilities*

Health state utilities were retrieved from sources cited in a recent systematic literature review[103].

Exercises performed in oral cancer or general head & neck cancer patient cohorts were considered eligible. The EQ-5D-5L was the most commonly-used utility measure in the literature review, and was therefore used in this exercise. Three published studies were included from this review, based on similarity of the health states in those exercises to those found in the WDMOC. Parameter inputs are summarized in Table 4.9.

Table 1.9 – Parameter estimates: health state utilities

Parameter	Mean	SD	Distribution	Source
Well (no disease)	1.0	N/A		Assumption
Undetected OPL	0.92	0.18	Beta	[124]
Detected OPL	0.92	0.18	Beta	[124]
Undetected cancer				
Stage I	0.84	0.02	Beta	[106]
Stage II	0.84	0.02	Beta	[106]
Stage III	0.82	0.14	Beta	[104]
Stage IV	0.82	0.14	Beta	[104]
Detected Cancer				
Undergoing Treatment	0.65	0.19	Beta	[104]
During Follow-up	0.82	0.18	Beta	[109]
Recurring cancer				
Undergoing treatment	0.65	0.19	Beta	[104]
During Follow-up	0.82	0.18	Beta	[109]
Incurable/Terminal disease	0.68	0.33	Beta	[124]
End of Life	0.68	0.33	Beta	[124]
Cancer in Full Remission	1	N/A		Assumption

Utility values for detected and undetected OPLs were derived from the same source as the Speight model[124, 178]. These values were taken from general population evaluations of health states related to oral conditions including cancer. The WDMOC assumes no difference between health state utility for detected and undetected OPLs.

Utility values for undetected early-stage cancer (stage I, II) were derived from an exercise by Govers *et al* [106], from a population of patients with early-stage cancer whose disease was managed with watchful waiting. No data was available in the literature regarding health state utilities for early-stage cancers prior to diagnosis, but this estimate is similar to baseline values estimated in the COOLS Trial for early-stage SCC and HGLs.

Utility values for stage III and IV cancers were derived from the baseline (pre-treatment) values taken from a cohort of clinical trial participants with stage III/IV head and neck cancers[104].

Utility values for cancers undergoing treatment were derived from the same Truong *et al.* clinical trial, using EQ-5D-3L values reported by patients at the end of their treatment[104]. Use of this estimate assumes that treatment-related utility does not differ across disease stage.

Utility values for cancers in remission less than ten years after treatment were derived from a study of a consecutively-recruited cohort of previously-treated head and neck cancer patients between 3 months and 3 years following treatment[109].

Utility values for terminal and end-of-life stages of disease were derived from the same source as the Speight model[124, 178].

People with no disease or with disease in full remission (after ten years follow-up) were assumed to have a utility of 1.0. Utility values were applied to each entity assuming a Beta distribution.

1.7 Derivation of Parameter Estimates from Data Sources

The following section describes the statistical approach used to produce individual estimates of each model parameter, based on the type of data used for the parameter estimates.

Generalized Linear Modeling (GLM) techniques were used to derive time-to-event values from primary data sources. GLM is a mathematical expression of the relationship between two or more variables through the statistical fitting of a linear equation. The predicted value of some dependent variable (Y) can be estimated through a linear predictor of a number of independent variables ($X\beta$) and a link function that estimates the mean that is derived from assuming a given statistical distribution:

$$E(Y) = g^{-1}(X\beta)$$

GLM regression functions can be fit assuming a variety of statistical distributions, including the Weibull distribution, when using the corresponding link function. The Weibull function is highly flexible, and can

be used to approximate a number of time-to-event functions. The Weibull probability density function can be expressed in terms of two parameters, σ and λ (referred to as the scale and shape parameters respectively):

$$f(t; \sigma, \lambda) = \sigma(t)^{\lambda-1} \cdot e^{-\left(\frac{t}{\lambda}\right)^\sigma}$$

where t is the time whose probability is being described.

The scale parameter σ can be expressed as e^{y_i} , allowing for mathematical predictions of time t elapsing between two events to be calculated as a function of other independent variables (x_n) and their coefficients of association (β_0, β).

GLM regressions were performed on time-to-event data from the Zhang trial data and the ROCC to determine the associated coefficients and Weibull parameters (σ and λ). This approach was used to take advantage of the stochastic nature of these datasets, and to reflect multiple entity characteristics simultaneously.

A Python function was written to generate random draws from a Weibull distribution based on an entity's characteristics, the regression coefficients associated with those characteristics, and Weibull parameters from a best-fitting curve. Regression analysis was conducted using the LIFEREG function in SAS 9.4 (SAS Institute, USA), which returns estimates of β_0, β_n , and λ .

Given these parameters, it was possible to use linear regression methods to generate predicted time-to-event values for an entity with a given set of characteristics using the function `numpy.random.weibull(σ, λ)` in Python.

Beta Distributed Values

Time-independent probabilities are estimated using the Beta distribution. This distribution is bound between zero and one and its probability density function can be expressed as a function of two parameters α and β :

$$f(x; \alpha, \beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}$$

where $B(\alpha, \beta)$ describes the bounds of the distribution between zero and one.

Probabilities can be sampled from the Beta distribution given the value of a mean probability (p) and standard deviation (σ) using the following equation, derived from the method of moments for the Beta distribution:

$$\alpha = \left(\frac{(1-p)}{\sigma^2} - \frac{1}{p} \right) p^2$$

$$\beta = \alpha \left(\frac{1}{p} - 1 \right)$$

This pair of equations can be applied to a mean and standard deviation to sample a random probability, which can be called using the function `numpy.random.beta(α, β)` in Python.

Normally (Gaussian) Distributed Values

Normally distributed variables are sampled from the Gaussian distribution. This distribution describes continuous values bound between negative and positive infinity and its probability density function can be expressed as a function of its mean (μ) and standard deviation (σ):

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

This equation can be used to sample a random value based on the mean and standard deviation, which can be called using the function `numpy.random.normal(μ, σ)` in Python.

Gamma Distributed Values

Estimates of unit cost are samples from the Gamma distribution. This distribution is bound by 0 and infinity and its probability density function can be expressed as a function of a shape parameter α and a scale parameter β :

$$f(x; \alpha, \beta) = \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)}$$

where $\Gamma(\alpha)$ is a complete gamma function.

Values can be sampled from the Gamma distribution given the value of a mean (μ) and standard deviation (σ) using the following equations, derived from the method of moments for the Gamma distribution:

$$\alpha = \frac{x^2}{y^2}$$

$$\beta = \frac{y^2}{x}$$

This pair of equations can be applied to a mean and standard deviation to sample a random value, which can be called using the function `numpy.random.gamma(α, β)` in Python.

Dirichlet Distributed Values

The Dirichlet distribution is related to the Beta distribution, and can be used to randomly sample probabilities (values between 0 and 1). The Dirichlet distribution is multinomial, meaning it can be used to sample multiple random probability values that sum to 1 (i.e., mutually exclusive probabilities that

cumulatively represent all possible outcomes) The distribution is bound by 0 and 1 and its probability density function can be expressed as a function of a multivariate Beta function $B(\alpha)$ and a vector of k integers $(\alpha_1, \dots, \alpha_k)$:

$$f(x_1, \dots, x_k) = \frac{1}{B(\alpha)} \prod_{i=1}^k x_i^{\alpha_i-1}$$

Random probabilities can be sampled from a Dirichlet distribution using the function `numpy.random.dirichlet($\alpha_1, \dots, \alpha_k$)` in Python.

Log-normal Distributed Values

The log-normal distribution is related to the Normal distribution, and can be used to randomly sample continuous values bound between negative and positive infinity. Its probability density function can be expressed as a function of a logarithmic mean (μ) and associated standard deviation (σ):

$$f(x; \mu, \sigma) = \frac{1}{x} \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}$$

Random values can be sampled from a log-normal distribution using the function `numpy.random.normal(μ, σ)` in Python.

Transition Probabilities

Transition probabilities that are published as a mean and standard deviation can be used to generate time-to-event values randomly sampled from a one-parameter Weibull distribution using a two-step process. In the first step, a transition probability value t_p is sampled from a Beta distribution using a given mean probability (p) and standard deviation (σ). In the second step, the Weibull function is

assumed to have a shape parameter λ equal to 1[180]. In the presence of this assumption, the Weibull distribution is equal to the exponential distribution:

$$f(t; \beta) = \frac{1}{\beta} \cdot e^{-\left(\frac{t}{\beta}\right)}$$

where $\beta = \frac{1}{-\ln(1-t_p)}$.

Random values can be sampled from an exponential distribution using the function `numpy.random.exponential(β)` in Python.

The above methods were used to derive parameter estimates for each entity as it moved through each component of the model.

1.8 Calibration of the WDMOC

In order to draw useful conclusions from a model, its outputs must match values seen in the population it is intended to represent – in this case, the population of British Columbia. Even in cases where the model's inputs may be drawn from representative sources, it is reasonable to expect that parameter interactions within the structure of the model may result in output values that differ from those observed in the real world. Accordingly, it is typically necessary to 'calibrate' a model by making reasonable adjustments to its parameter values and/or structural assumptions[24].

The first step in model calibration involves choosing a set of 'target' values that are important to the model's predictive validity. It is most appropriate to choose a set of targets that reflect multiple points across the disease process being modeled, as using a single 'target' may obscure intermediate values that are not accurate, yet nonetheless produce reasonable values of the single target.

The following calibration targets were chosen:

1. Prevalence of oral premalignancy at time of diagnosis by age and sex.
2. Prevalence of cancer at time of diagnosis by age, sex, and stage
3. Age at death among entities with oral cancer

These targets represent initial, intermediate, and final values that are related to an entity's clinical trajectory within the model (i.e., age and sex are statistically related to premalignant progression; age, sex, and stage are related to treatment response and survival).

The second step in model calibration involves assessing how well the model's output fits reasonable values for the targets. While there are several methods that can be used to assess goodness of fit, the underlying complexity of the WDMOC, the heterogeneity in parameter data sources, and the relative lack of data for many parameter values suggested that a simpler approach was the most practical. Accordingly, an acceptable window approach was chosen. The acceptable window method compares the mean and variance of the identified target outputs to the mean of real-world values for those same targets. Adjusted parameter values were accepted when model output for the mean and standard deviation of the calibration targets were similar.

The values and data sources used in the acceptable window calibration process are described in Table 4.10. The model was run for one hundred iterations of 25,000 entities per iteration (for a total of 2,500,000 entities), to ensure at least 10 cancers diagnosed at each stage for each iteration. Values of the target outputs were calculated for each iteration, and the mean and standard deviation were calculated across all iterations (i.e., for the full run). The results of this process are provided in Table 4.11.

Table 1.10 – Calibration output target values

Target	Mean	Source
Prevalence of OPL	0.9%	[72, 73]
Men	55%	[61]
Women	45%	[61]
Age	59	[61]
At cancer diagnosis		

Men	58.5%	ROCC
Women	41.5%	ROCC
Age	65.6	ROCC
Stage I	37.3%	ROCC
Stage II	35.0%	ROCC
Advanced Stage	27.4%	ROCC
Age at death from cancer	70.6	ROCC

Table 1.11 – WDMOC baseline calibration outputs

Target	Reference	Model Output		Distance (z score)
		Mean	SD	
Prevalence of OPL	0.9%	0.908%	0.027%	0.22
Men	55%	63%	4.8%	1.67
Women	45%	37%	4.8%	-1.67
Age	59	59.0	2.31	-0.18
At cancer diagnosis				
Men	58.5%	58.6%	12.9%	0.008
Women	41.5%	41.4%	12.9%	-0.008
Age	65.6	62.2	1.9	-1.79
Stage I	37.3%	36.4%	8.0%	-0.11
Stage II	35.0%	31.1%	8.1%	-0.48
Advanced Stage	27.4%	26.7%	7.4%	-0.09
Age at death from cancer	70.6	64.96	9.71	-0.41

After each run, select parameters were manually adjusted in order to produce outputs that more closely matched the target values. This process was repeated until the model's outputs match the target values.

The following calibrating adjustments were made to the WDMOC:

- The mean and standard deviation for the 'starting age' parameter was adjusted to reflect the age distribution of detected premalignant lesions within the Zhang et al study of malignant transformation[61];
- The age and sex parameters for premalignancy prevalence were adjusted (from their baseline values within the Speight model) using the published Relative Rate values from a population prevalence study within British Columbia[72];
- A constant relative prevalence value was introduced to adjust the overall prevalence of premalignant lesions to values within the same study;
- A parameter representing the likelihood that an entity will receive regular screening from their dentist was created, in order to adjust the percentage of premalignancies and cancers that are detected symptomatically vs. through screening;

The output suggests that the calibrated WDMOC produces estimates of prevalence and age/stage distribution that approximate the target values. There is a noticeable gender difference at the time premalignancies are detected, and the age at which cancers are detected is slightly lower in the model output than in the target data. This early age at cancer detection carries forward to produce an earlier

death from cancer in the model output than in the target values. These discrepancies will be discussed later in this chapter.

1.9 Univariate Sensitivity Analysis

The final component of model calibration is conducting univariate sensitivity analysis, in order to investigate the impact that changes in single parameter values have on the model's outputs. In this exercise, univariate sensitivity was conducted in the following way:

- A group of model outputs was selected, chosen to represent key informative outputs of the model;
- A parameter input was adjusted to $\pm 10\%$ of its baseline value. This value was chosen arbitrarily, but allows for a constant and comparable level of change across all model parameters;
- An analysis set of 100 iterations of 25,000 entities (2,500,000 entities in total) was run with the adjusted value. The mean value of each model output was calculated for the analysis set.
- The resulting means were compared to mean output from a baseline set with all parameters at baseline values (100 iterations, 25,000 entities, resulting in a total of 2.5m entities)
- Parameter sensitivity was calculated as the sum of z scores between the means of the high (i.e., +10%) and low (i.e., -10%) analysis sets:

$$Sensitivity = \frac{X_{high} - X_{baseline}}{SD_{baseline}} + \frac{X_{low} - X_{baseline}}{SD_{baseline}}$$

The following output results were selected: OPL prevalence, cancer prevalence, cancer stage at detection, age at death from cancer, and mean cost and QALY per entity that began the model with an OPL. These values are summarized in Table 4.12.

Table 1.12 – Baseline outputs for sensitivity analysis

Model Output	Baseline value	
	Mean	SD
Prevalence of OPL	0.908%	0.027%
Prevalence of cancer	0.150%	0.02%
Stage at cancer detection		
HGL	5.75%	3.3%
Stage I	36.4%	8.0%
Stage II	31.1%	8.1%

Stage III/IV	26.7%	7.4%
Age at cancer death	64.96	9.71
Mean cost†	\$5,294.38	\$172.61
Mean QALY†	15.35	0.71

† – these values were calculated for entities who begin the model with an OPL

Not all parameter values underwent sensitivity analysis. Parameters that were unrelated to the research questions contained within this dissertation were excluded from this process. The implications of this exclusion are discussed further in Chapter 7.

Because the majority of entities in the model do not experience disease, the model was most sensitive to changes in the natural history parameters, particularly starting age. Table 4.13 lists the five parameters that exerted the largest influence on each output variable.

Table 1.13 – Univariate sensitivity results by model component

Model Output	Parameter	Model Component	Change (ΔZ score)
<i>Prevalence of OPL</i>			
	OPL prevalence conversion factor	Entity Creation	6.98
	Cohort starting age	Entity Creation	3.46
	Prevalence of oral premalignancy	Entity Creation	
	Men, 50-59		3.47
	Men, 60-69		3.33
	Women, 50-59		2.89
<i>Prevalence of cancer</i>			
	OPL prevalence conversion factor	Entity Creation	1.91
	Prevalence of oral premalignancy	Entity Creation	0.88
	Men, 60-69	Entity Creation	0.85
	Men, under 50	Entity Creation	0.85
	Women, 50-59	Entity Creation	0.81
	Time to OPL progression to Stage I cancer	Natural History	
	Smoking Status – Ever		0.82
<i>Stage at cancer detection</i>			
HGL	Time to undetected cancer progression	Natural History	
	Stage I to Stage II		0.72
	Prevalence of oral premalignancy	Entity Creation	
	Women, 70-79		0.59
	Return appointment interval	Screening	0.56
	Has access to dentist	Entity Creation	0.52
	Time to symptomatic detection of cancer	Natural History	
	Stage I		0.50
Stage I	Time to undetected cancer progression	Natural History	
	Stage I to Stage II		0.69
	Time to symptomatic detection of cancer	Natural History	
	Stage I		0.51
	Smoking prevalence – Women	Entity Creation	0.47
	Prevalence of oral premalignancy	Entity Creation	
	Women, 50-59		0.42
	Screening adherence	Screening	0.36
Stage II	Time to undetected cancer progression	Natural History	
	Stage I to Stage II		1.11
	Time to symptomatic detection of cancer	Natural History	
	Stage II		0.78
	Prevalence of oral premalignancy	Entity Creation	
	Women, 60-69		0.54
Model Output	Parameter	Model Component	Change (ΔZ score)
	Women, 50-59		0.39
	Time to undetected cancer progression	Natural History	
	Stage II to Stage III		0.48
Stage III/IV	Time to symptomatic detection of cancer	Natural History	0.69

	Stage II		
	Prevalence of oral premalignancy	Entity Creation	0.66
	Women, 60-69		
	Screening adherence	Screening	0.66
	Time to undetected cancer progression	Natural History	
<i>Age at cancer death</i>	Stage II to Stage III		0.50
	Return appointment interval	Screening	0.41
	Time to Second Event – Death	Terminal Disease	
	Sigma (σ)		1.20
	Age		0.99
	Cohort starting age	Entity Creation	0.99
	Time to First Event – Death	Invasive Cancer	
	Treatment type – Surgery + RT		0.98
	Cancer stage – Advanced		0.96
<i>Mean cost_T</i>	Time to OPL progression to Stage I cancer	Natural History	
	Sigma (σ)		9.52
	Intercept (β_0)		6.17
	Time to First Event	Invasive Cancer	
	Age		7.49
	Intercept (β_0)		3.85
	Time to First Event – Death	Invasive Cancer	
<i>Mean QALY_T</i>	Age		6.02
	Cohort starting age	Entity Creation	6.52
	Utility – Well (no disease)	Multiple	1.62
	Utility – Undetected OPL	Screening	0.88
	Prevalence of oral premalignancy	Entity Creation	
	Men, <50		0.76
	Men, 60-69		0.65

The model is not sensitive to changes in most individual model parameters, but can experience relatively large changes in important output values from even small changes in key parameters, particularly those concerning the age of the cohort and the prevalence and natural history of oral disease. Costs are highly sensitive to changes in the risk of premalignant transformation and cancer survival. The full implications of this parameter sensitivity will be discussed in greater detail in Chapter 7, but a brief discussion follows.

Pre-malignant disease is not as well-characterized in oral cancer as it is among other diseases where screening is widely practiced – cancers of the cervix, breast, and prostate. The sensitivity of the model to changes in natural history parameters seems to make a strong case for the value of future research in

the population prevalence, incidence, and development of novel premalignancies. It also suggests that the requirements of the Whole Disease Framework can best be met in disease areas where pre-symptomatic disease trajectory is well understood. The cost-effectiveness analyses that follow this chapter are focused on novel technology adoption that occur after premalignancies have been diagnosed, which limits the impact that natural history parameter sensitivity has on cost-effectiveness estimates. Nevertheless, it remains imperative to recognize this impact within the context of this dissertation as a whole.

1.10 Discussion

1.10.1 Strengths of the approach

The WDMOC was designed and implemented in accordance with guidelines set out for the WDM framework. It considered the full breadth of the disease pathway, from preclinical disease to death. The depth of the model was sufficient to reflect potential policy changes at multiple levels, guided by factors that were associated with preclinical disease progression and survival after diagnosis of invasive disease. The model's boundary was adults in British Columbia, calibrated to resemble a population with newly-diagnosed OPLs. These model characteristics make the WDMOC suitable for evaluating the impact of policy changes on a population of adults with newly-diagnosed OPL, but the decision node of the model can be moved to consider populations with invasive cancer or other disease sequelae.

1.10.2 Limitations of the approach

The nature of modeling requires simplifying assumptions to be made, and the WDMOC as described above is no exception. Parameters and pathways governing screening and treatment were limited by data availability. The individual sampling approach that was chosen for the model's design does not

reflect health care system wait lists. Assumptions were made in the model's calibration as well. The implication of these assumptions will be described in the following section.

1.10.2.1 Clinical Processes

Chapter 2 highlighted the relative lack of high-quality data on screening frequency and practice, particularly as it relates to how demographic factors like age and smoking status affect screening frequency. The screening process described in the WDMOC assumes that all entities with access to a screening dentist are treated identically at regular intervals, based on available guidelines and the opinion of the expert stakeholder group.

Cancer treatments were similarly simplified, although better data were available to guide these assumptions. Treatment options were collapsed into categories (surgery, surgery + RT, other) based on a regression analysis exercise (see Appendix 4.5), which consequently limits the ability of the model to reflect the full complexity of possible treatment approaches. The 90-day window for treatment costs was also chosen based on the available data and expert opinion, and does not reflect every potential treatment regimen that a given patient could experience. Local, locoregional, and regional recurrences were treated as synonymous events due to the limited number of events in the data.

These simplifications limit the ability of the model to reflect the full complexity of screening practice and invasive disease management, and the results produced by the WDMOC should be interpreted with these limitations in mind. It is worth noting, however, that the level of evidence used within the WDMOC is either at par with (i.e., directly drawn from) or a step above what is available in the extant literature.

1.10.2.2 Individual Sampling and Queuing

The model assumes that referrals are instantaneous and resources are always available. This is a simplifying assumption that is common in health economic decision models, especially since data on queueing is not typically available. In future cases where there is a compelling reason to believe that waiting is meaningfully associated with the cost and/or effectiveness of a new technology, the effect of queueing could be reflected in the model by adding a 'wait time' parameter.

The consequence of not including queueing means that referrals between specialists (e.g., referral to OPL management, referral to an oncologist, time between referral and treatment, etc.) happens on a time scale that may be days/weeks earlier than what would be observed in real-world practice.

1.10.2.3 Model Calibration

The calibration method chosen for the WDMOC produced a population that developed cancer slightly earlier than the calibration targets. This discrepancy was due to the data sets used to derive parameter estimates. Values related to OPL progression were derived from literature-published estimates and a clinical trial, while values related to oral cancer survival were derived from a population-based cohort. Because these values were drawn from separate populations, it is reasonable to expect that the values they produce would not necessarily match up exactly. The clinical trial population underwent more rigorous regular observation than is typical for people in the general population, and had

The committee agreed that it was important for prevalent cases of oral cancer to resemble the age distribution of the population from which parameter estimates were derived, given that age is significantly associated with time to progression to cancer (see Table 4.2). Because the statistical association between sex and cancer survival was stronger than the corresponding association to progression to invasive cancer, the model was calibrated to ensure that the sex distribution matched the cancer survival cohort.

When interpreting the TTE estimates for progression from OPL to invasive cancer, it is also important to note that the underlying study from which the estimates are drawn considered time from OPL *detection* to progression, rather than the time from *development*. The earlier age of cancer detection will also be a product of this discrepancy, and the lack of available data about OPL incidence.

Despite the limitations of the data sources, the model's outputs were broadly similar to values observed in the calibration targets.

1.11 Conclusion

This chapter described the process through which a Whole Disease Model of Oral Cancer was conceived and designed, using the guidelines set out by Tappenden. This description included a summary of the model's structure and implementation, as well as data sources and statistical procedures. The model's outputs were compared to expected values, and then the impact of changes in each parameter on key outputs was explored through sensitivity analysis.

The next chapter considers the same cost-effectiveness question as explored in Chapter 3 – cost-effectiveness of a risk-guided approach to premalignant management – but this time uses the WDMOC. Chapter 6 describes an exercise exploiting the ability of a Whole Disease Model to evaluate the 'upstream' impact of policy changes on the 'downstream' cost-effectiveness of new technologies.