Good practice guidance

Framework to review models

NAO
National Audit Office

January 2022

We are the UK's independent
public spending watchdog

Contents

The National Audit Office (NAO) scrutinises public spending for Parliament and is independent of government and the civil service. We help Parliament hold government to account and we use our insights to help people who manage and govern public bodies improve public services. The Comptroller and Auditor General (C&AG), Gareth Davies, is an Officer of the House of Commons and leads the NAO. We audit the financial accounts of departments and other public bodies. We also examine and report on the value for money of how public money has been spent. In 2020, the NAO's work led to a positive financial impact through reduced costs, improved service delivery, or other benefits to citizens, of £926 million.

If you are interested in the NAO's work and support for Parliament more widely, please contact:

Parliament@nao.org.uk
020 7798 7665

Introduction

Government departments and agencies rely on models for their day-to-day activities including estimating costs; distributing funding within organisations; and testing policy options. They routinely develop and use models to generate insight into a question or to better understand a problem related to their business. These models can vary in complexity from relatively simple spreadsheets to detailed forecasts using specialist software. Outputs from models underpin decisions made by departments and arm's-length bodies that often have very real impacts on people's lives and can involve large amounts of money and resources.

In our audit work across government, we continue to find weaknesses in models such as:

● limited or poor-quality data;

● unrealistic assumptions and optimism bias; and

● inadequate sensitivity and scenario analysis.

The framework provides a structured approach to review models, which organisations can use to determine whether the modelling outputs they produce are reasonable, robust and have a minimal likelihood of errors being made.

Evidence base

The framework to review models builds on the findings, conclusions and recommendations from our 2022 report on Financial modelling in government. It is also based on the evidence and guidance available from:

● HM Treasury's review of quality assurance of government analytical models (2013)

● HM Treasury's Aqua Book (2015)

● The Department for Business, Energy & Industrial Strategy (BEIS) Modelling Quality Assurance tools and guidance

● International Standard on Auditing, ISA, (UK) 540 (Revised) Auditing Accounting Estimates and Related Disclosures (2018)

● Office for Statistics Regulation Quality Assurance of Administrative Data

● The government's Uncertainty Toolkit for Analysts in Government
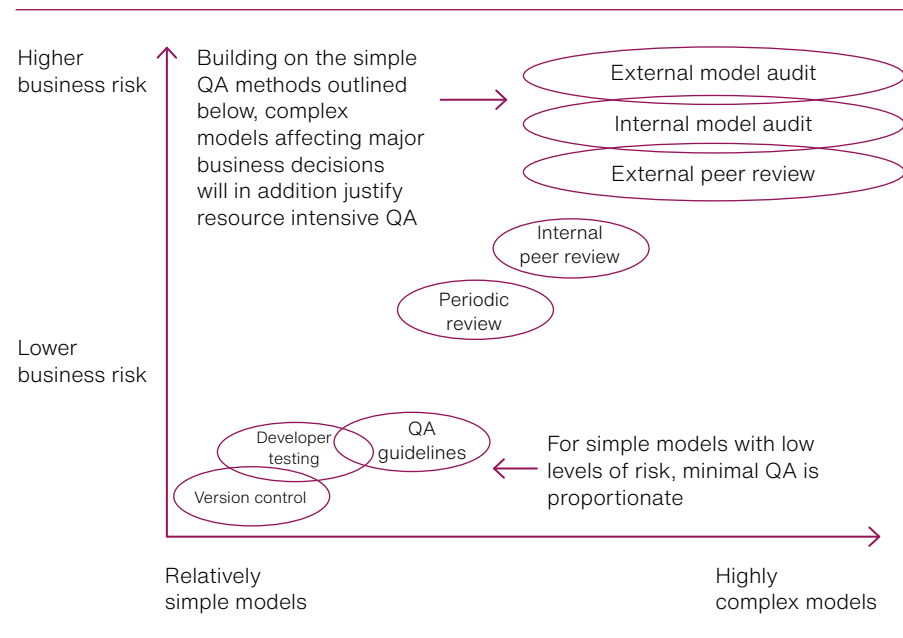
How to use the framework

This framework is aimed at people who commission analysis, provide analytical assurance and deliver the analysis itself.

It is not intended to be a checklist; instead it is a flexible framework which can be tailored, taking into account:

● the amount of time and resource available;

● the complexity and risk associated with the model; and

● the level of assurance needed to reach an overall judgement.

This proportional approach is in line with HM Treasury's review of quality assurance of government analytical models (see diagram overleaf). Additionally, please note the framework does not represent a comprehensive methodology for audits of models carried out to test compliance with the International Standard on Auditing (UK), where further requirements will apply.

Schematic showing indicative types of quality assurance (QA) that might be expected given different levels of risk[1]

Higher business risk

Building on the simple QA methods outlined below, complex models affecting major business decisions will in addition justify resource intensive QA

External model audit

Internal model audit

External peer review

Internal peer review

Periodic review

Lower business risk

Developer testing

QA guidelines

Version control

For simple models with low levels of risk, minimal QA is proportionate

Relatively simple models

Highly complex models

Deciding on whether a model and its outputs are robust, reasonable and used appropriately requires a proportionate, evidence-based judgement. It will often be the case that a review will identify issues and weaknesses in some aspect of how the model was designed, built and used. Crucially, the objective of a model review is to identify, in your opinion, whether those issues had an impact on the quality of the model, and whether there is a risk of material impact on the outputs and how they are interpreted and used for decision making, planning and/or disclosures.
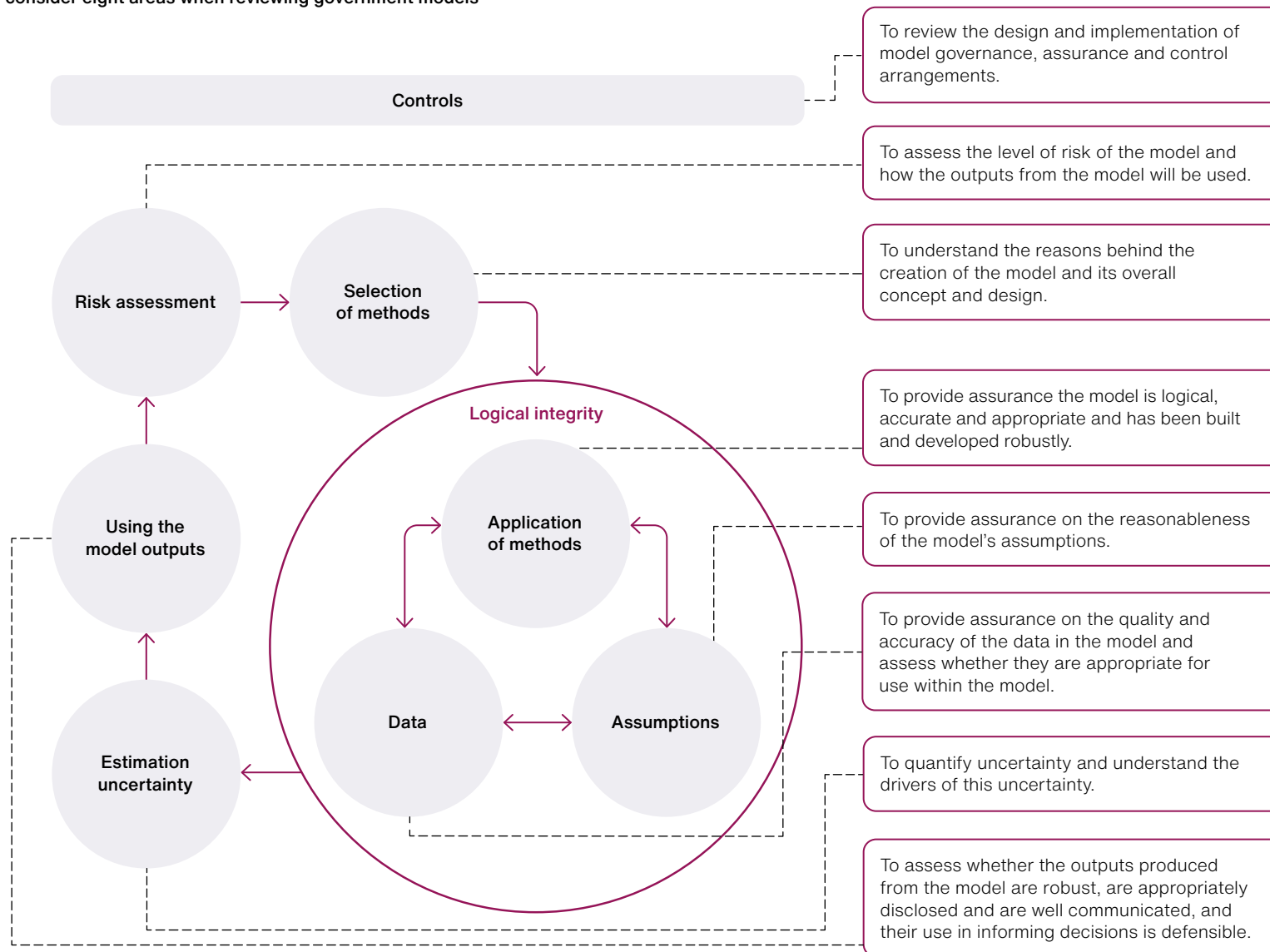
How the NAO can help

If you have any queries about this framework or suggestions for how it can be improved, please use the contact form on our website.

1   HM Treasury, *Review of quality assurance of government analytical models: Final report*, 2013, page 22, chart 2.C.

The framework is split into eight stages starting with and driven by an initial assessment of risk[2] (see diagram).

The National Audit Office's framework to review models

We consider eight areas when reviewing government models

To review the design and implementation of model governance, assurance and control arrangements.

To assess the level of risk of the model and how the outputs from the model will be used.

To understand the reasons behind the creation of the model and its overall concept and design.

To provide assurance the model is logical, accurate and appropriate and has been built and developed robustly.

To provide assurance on the reasonableness of the model's assumptions.

To provide assurance on the quality and accuracy of the data in the model and assess whether they are appropriate for use within the model.

To quantify uncertainty and understand the drivers of this uncertainty.

To assess whether the outputs produced from the model are robust, are appropriately disclosed and are well communicated, and their use in informing decisions is defensible.

Source: National Audit Office

2   The questions in the framework are not exhaustive, meaning there will be other checks that can be applied.

The framework
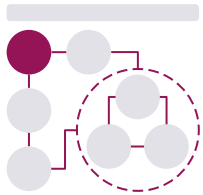
Risk assessment

Government departments face different levels of complexity when it comes to risk management. Some government departments have complex delivery models, and challenges in monitoring and overseeing arm's-length bodies, while others may be significantly less complex. To assess the level of risk of the model and how the outputs from the model will be used, the risk assessment below can be used to prepare and prioritise your review work. The following areas should be considered with a grading of risk defined against each element. The outcome of this exercise should form the basis for the quality assurance log.

Questions to consider	Examples of checks to make or evidence to look for
Is the model business critical[3] and/or does the model underpin activities, decisions or disclosures which are critical to the organisation?	Identify the model output and if appropriate, the 'materiality' or 'tolerable level of error' it is required to operate within. Identify the activities, decisions and/or disclosures the model supports.
Is there estimation uncertainty risk in the model's outputs?	Review how the model outputs performed historically when compared with outturns/actuals. Assess the range of plausible outcomes in the model outputs.

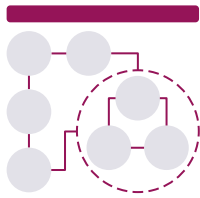| Questions to consider	Examples of checks to make or evidence to look for
Is there control risk in the management and use of the model?	Identify who is responsible for managing, producing and assuring the model (including any third parties). Identify what controls are in place to manage and assure the model. Identify whether the model and its outputs are made available to Parliament and the public.
Is there risk in the selection and application of the model's methods?	Identify the maturity of the model. Identify any standards or guidance the model must align to. Identify the methods applied in the model. Identify the software environment used to host and run the model (including its data). Identify whether any adjustments or overlays are made to the model outputs outside of the core modelling environment.
Is there risk in the model's data?	Identify the data used in the model. Identify how the data used in the model are verified and assured prior to use. Identify how the data used in the model are cleaned and/or joined together prior to use.
Is there risk in the model's assumptions?	Identify the assumptions applied in the model. Identify how the assumptions applied in the model are validated and assured prior to use. Identify the assumptions which require significant judgement and/or are highly uncertain.
Do you have the required skills to review the model?	Assess whether specialised skills or knowledge are required to review the model.

3 HM Treasury's review of quality assurance of government models defines the criteria for judging if a model is business critical is based on the extent to which:
- the model drives essential financial and funding decisions;
- the model is essential to achievement of business plan actions and priorities;
- errors could engender serious financial, legal or reputational damage or penalties.

Controls

To review the design and implementation of model governance, assurance and control arrangements.

Questions to consider	Examples of checks to make or evidence to look for
Depending on the level of risk/criticality of the model, is the model compliant with requirements which would be expected of a model of this nature?	Assess whether the model is recorded on the organisation's register of business-critical models. Does the organisation have an appropriate QA framework defined for business-critical and other models? If so, assess whether the model has been produced and assured in line with the organisation's internal guidance for models. Review the activities, decisions and/or disclosures the model supports.
Is the model documented?	All models should be documented. Well-documented models will have at a minimum the following features documented: ● specified roles and responsibilities such as the appointment of appropriate and named senior responsible owner (SRO) and analytical assurer; ● concept and design of the model; ● technical guide for users; ● version control log; ● data and assumptions book; ● analytical assurance plan; ● analytical assurance log; ● model output reports; ● proof of SRO sign-off; and ● succession plan. In practice, these features may be consolidated into a smaller number of documents. Some of these features (for example version control) might also be automated into software which controls the management and use of the model.

| Questions to consider	Examples of checks to make or evidence to look for
Is the model independently assured?	Assess whether the model is subject to appropriate independent analytical assurance. Common analytical assurance techniques which might be applied to a model include: ● verification of the model (this might include specification, user and/or logical integrity testing); and ● validation of the model (this might include input data and assumptions testing, sensitivity analysis and/or backtesting). Assess whether the model's analytical assurer is active in leading and/or facilitating activities to ensure that the model is independently assured.
Is the model subject to appropriate scrutiny and challenge?	Assess whether the model's SRO is active in overseeing the model. For models with multiple stakeholders, assess whether an inclusive group is in place and actively used by all relevant stakeholders to challenge the development and use of model outputs. Assess whether there is evidence of bias. For example, is there evidence that users of the model have unduly influenced its operation or selected favourable assumptions to ensure its outputs meet set expectations? Where appropriate, assess whether a model's methodology, data, assumptions or outputs are published to facilitate scrutiny and challenge from Parliament and the public.
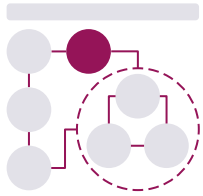
Selection of methods

To understand the reasons behind the creation of the model and its overall concept and design.

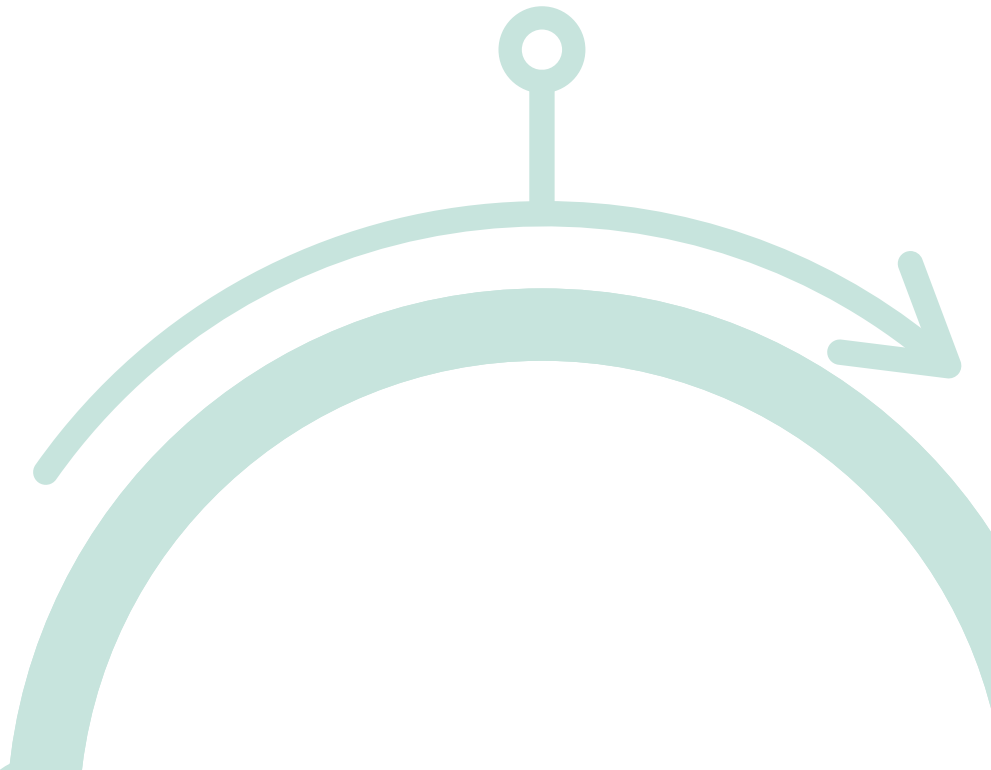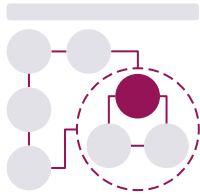| Questions to consider	Examples of checks to make or evidence to look for
What is the model's aim, or what decision is it supporting, and are the selected methods documented?	Model concept and design documentation which details: ● the aims and objectives of the model; ● any standards the model must comply with (for example, HM Treasury Green Book or applicable financial reporting standards); ● the logic of the model including its inputs, calculations and outputs; ● identification of model limitations; and ● the required precision of the model's output (offset against complexity).
Are the selected methods appropriate for the activity, decision and/or disclosure the model is designed to support?	Assess whether appropriate and reasonable methods for the model have been selected. Where appropriate, assess whether the model complies with the requirements of applicable modelling standards. Where appropriate, assess whether the application of novel or complex methods in the model is justified. Where appropriate, assess whether alternative methods for the model have been considered and whether the reasons for their rejection are reasonably justified.

Questions to consider	Examples of checks to make or evidence to look for
Has there been a change in the model's selected methods?	Assess whether there is a robust rationale for changing the selected methods in the model. Where appropriate, assess whether the change in the model's selected methods is aligned with new circumstances, information or knowledge.

Application of methods

To provide assurance the model is logical, accurate and appropriate and has been built and developed robustly.

Questions to consider	Examples of checks to make or evidence to look for
Is the application of the methods in the model documented?	Technical guide which details the 'nuts and bolts' of the model. This should be sufficiently clear to allow a model auditor or developer to understand how the model has been developed and to repeat the calculations if necessary.
Do you understand the model?	Are you able to draw a simple picture representing the model or can you describe it in lay terms?
	Assess whether the inputs, calculations and outputs in the model are separated.

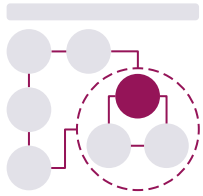| Questions to consider	Examples of checks to make or evidence to look for
Is the detail of the model accurate and robust?	Assess whether the calculations in the model are applied in accordance with the selected methods.
	Assess whether the calculations in the model are mathematically and logically accurate.
	Assess whether the integrity of the data and assumptions has been maintained in applying the model.
	Techniques to test the accuracy and robustness of the model could include:
	• independently recalculating the model; for example, in a different software environment, using the selected methods and comparing the results;
	• independently producing a simplified version of the selected methods and comparing with the original model's results; and
	• directly testing the calculations in the model to check they are accurately implemented. This could include reviewing the integrity and logic of the formula, and checking for alignment of the formulas to the selected methods.
	For Excel-based models, additional checks might be used to identify areas that might expose weaknesses in the model, such as:
	• circular reference warnings;
	• hard coding of values;
	• linking of data from other files; and
	• complexity of formulae.
	For code-based models, additional checks might include a review of whether the code runs free from error, whether data are correctly read into the software environment and whether the code produces reproducible results, and a review of the output logs.

Application of methods *continued*

To provide assurance the model is logical, accurate and appropriate and has been built and developed robustly.

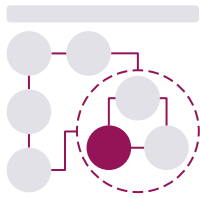| Questions to consider	Examples of checks to make or evidence to look for
Does the model respond logically to basic changes being made to the model inputs?	Review how changing basic model inputs impacts the model outputs, for example by: ● simplifying settings to the most basic scenario; ● examining the initial (starting) conditions for the model; ● sensitivity analysis with realistic input variations; and ● sensitivity analysis with extreme or implausible input variations.
Where appropriate, are adjustments (sometimes referred to as 'overlays') made to the output of the model reasonable and appropriate?	Assess whether any adjustments made to the outputs of the model outside of the core model environment are reasonable and appropriate. (It is worth noting that a strong rationale is needed for overlays as the main model and its outputs should capture estimation uncertainty.)
Do the outputs of the model agree to reported content?	Review documentation which reports the outputs of the model. Review published information which reports the outputs of the model.

Data

To provide assurance on the quality and accuracy of the data in the model and assess whether they are appropriate for use within the model.

Questions to consider	Examples of checks to make or evidence to look for
Are the data used in the model documented?	A data log should detail all data used in the model. For each source of data, the log should also include information on: ● key data characteristics (for example, description, units and source); ● appropriateness of data; and ● accuracy, quality, strengths and limitations of the data (including any adjustments made to the data).
Are the data used in the model appropriate?	Assess whether the data selected are appropriate for the model's selected methods. Assess whether alternative sources of data for the model have been considered and the reasons for their rejection are well reasoned.
Are the data used in the model reliable?	Review the quality of data and sources. Consider the following: ● Are the data up to date and do they agree to source? ● Are the data free from bias or error? ● Where appropriate, are the steps taken to cleanse or prepare the data reasonable? ● Where appropriate, are the data based on a robust sample?
Is there a robust relationship with data supply partners?	Assess whether there is a robust relationship between model producers and data suppliers. This might include: ● regular communication between model producers and data suppliers; and ● a written agreement between model producers and data suppliers.

| Questions to consider	Examples of checks to make or evidence to look for
Are the data the model using coming from other models?	Assess whether outputs from other models (and therefore their selected methods) need to be part of the scope of the model review.
Are the data the model using coming from a source of expertise?	Assess whether specialist expertise might be required to interpret and verify the data used in the model.
Does the model robustly handle input data through the software environment?	Review how input data are included in the model. This could include considerations such as how data are cleaned or transformed from the original source, and how easily this is repeated when the model is refreshed. Review how data are applied throughout the model software environment.
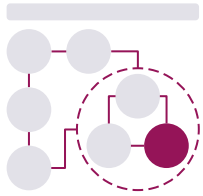
Assumptions

To provide assurance on the reasonableness of the model's assumptions.

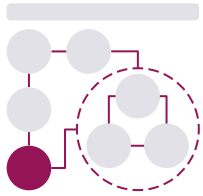| Questions to consider	Examples of checks to make or evidence to look for
Are the assumptions used in the model documented?	An assumptions data log should detail all assumptions used in the model (including both explicit and implicit assumptions). For each assumption the log should also include information on: ● suitability of selection based on the aims of the model; ● underlying evidence – source and quality, and the extent to which assumptions have been tested with experts, including external stakeholders; and ● level of complexity of each assumption's derivation.
Do you know which assumptions are most significant to the model outputs?[4]	Sensitivity analysis can be used to support the identification of significant assumptions. Options include: ● varying independent assumptions by a standard unit, that is, +/-1%; and ● varying independent assumptions by the minimum and maximum values of their plausible range.

Questions to consider	Examples of checks to make or evidence to look for
Are the assumptions used in the model reasonable?	An assessment of a model's assumptions may reasonably be prioritised to those assumptions which are most significant to the model's outputs. Options include: ● review model documentation and supporting evidence; ● compare the change in model assumptions over time (this might include assessing any changes to assumptions from a prior period); ● compare the model assumptions to third-party sources or industry norms; ● review whether there is a risk of interdependency in the model assumptions; and ● review whether assumptions should have been updated in light of any changes to circumstances and/or where relevant, in light of a comparison between previous outturn data and previous assumptions.

4   Sensitivity analysis is strongly linked to estimation uncertainty analysis; this is explored in greater detail in the next section.

Estimation uncertainty

To quantify uncertainty and understand the drivers of this uncertainty.

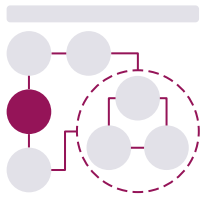| Questions to consider	Examples of checks to make or evidence to look for
Has uncertainty been assessed?	Assess whether there is a reasonable, systematic approach to identify and classify sources of uncertainty. This assessment might reasonably include: ● a list of modelling uncertainties, including both known and unknown sources of uncertainty such data quality, errors in assumptions, methodology or analysis; ● information used to inform the analysis and modelling of uncertainty; and ● the materiality or level of tolerable error which the model and its outputs need to work within. Assess whether the uncertainty of the model and its outputs are quantified. This assessment might reasonably include: ● a range consisting of high and low scenarios around a model's point estimate output; for example, by applying Monte-Carlo simulation to evaluate multiple independent assumptions at the same time; ● an assessment of confidence in the model's output; for example, by developing and applying scenario analysis to evaluate multiple independent assumptions at the same time; ● an assessment of whether appropriate techniques have been used to understand the model's range of uncertainty; and ● an assessment of whether the model's uncertainty analysis is up to date.

Questions to consider	Examples of checks to make or evidence to look for
Where estimation uncertainty is high, has a reasonable attempt been made to reduce uncertainty?	Review the impact of methods, data or assumptions applied in the model which drive high estimation uncertainty. See respective sections in this framework for associated examples of checks to make or evidence to look for.

Using the model outputs

To assess whether the outputs produced from the model are robust, are appropriately disclosed and are well communicated, and their use in informing decisions is defensible.

Questions to consider	Examples of checks to make or evidence to look for
What is the model output?	Identify the model output and, if appropriate, the 'materiality' or 'tolerable level of error' it is required to operate within.
Is the model and its outputs meeting its aims and the needs of its users?	Identify the activities, decisions and/or disclosures the model supports and assess whether the model is suitably serving its users.
Are you able to validate the model outputs?	Validate the outputs of the model by, for example: ● assessing whether the model produces 'logical' outputs; ● assessing whether the model outputs have been sense- or reality-checked and agreed with relevant stakeholders; ● back-testing the model to see if it matches historical results when using historical input data; ● comparing the outputs of the model with previous runs; or ● comparing the outputs of the model with independent sources.

Questions to consider	Examples of checks to make or evidence to look for
Are the limitations of the model output adequately communicated to users?	Review how the model outputs are presented to users; for example, how findings are presented in a business case, ministerial submission or disclosed as part of an organisation's Annual Report and Accounts or other published materials.
Is the uncertainty of the model output adequately communicated to users?	Review whether analysis of uncertainty is integrated into the presentation of the model outputs. This might reasonably include: ● a range consisting of high and low scenarios around a model's output; and ● an assessment of confidence in the model's output. Review whether unquantified uncertainties and their implications are communicated to users.
Are users of the model and its outputs considering and planning against the full range of plausible outcomes?	Review whether decisions/plans that are based on the model outputs include options against a range of outcomes. Review whether the model is used to monitor existing and/or emerging risks against decisions/plans. Review whether contingency arrangements for high impact outcomes or risks are built into decisions/plans.