Introducing requirements on the type and quality of evidence needed for scoring fisheries

Impact Assessment Report

April 2021

Tim Davies

Marine Stewardship Council

Table of Contents

1. Purpose 5

2. Impact Assessment Framework 5

3. Problem Statement 6

3.1. Need for intervention 6

3.2. Business-as-usual scenario 6

4. Objectives and intention 7

4.1. Objectives 7

4.2. Intended effects 7

5. Options 7

5.1. Overall approach 7

5.1.1. Applicable SIs 7

5.1.2. Associated revisions to some scoring guideposts 8

5.2. Alternative framework designs 8

5.2.1. Version A: Scoring bound by information quantity 8

5.2.1.1. Conceptual design 8

5.2.1.2. Worked example: Primary species information 9

5.2.2. Version B Balancing information quality and quantity 14

5.2.2.1. Conceptual design 14

5.2.2.2. Response to STAC working group feedback 15

5.2.2.3. Worked example: Primary species information 16

5.3. Approach to setting information thresholds 18

6. Summary of impacts 18

6.1. Impacts of the business-as-usual scenario 18

6.2. Impacts of Version A 18

6.3. Impacts of Version B 19

7. Impacts 21

8. Discussion and conclusion 27

The views and opinions expressed in this report do not necessarily reflect the official policy or position of the Marine Stewardship Council. This is a working paper, it represents work in progress and is part of ongoing policy development. The language used in draft scoring requirements is intended to be illustrative only, and may undergo considerable refinement in later stages.

How to refence this report: Davies, T. 2021. Introducing requirements on the type and quality of evidence needed for scoring fisheries Impact Assessment Report. Published by the Marine Stewardship Council [www.msc.org], (https://www.msc.org/docs/default-source/default-document-library/stakeholders/consultations/impact-assessments/msc-impact-assessment-report_evidence_requirements_april21.pdf), 27 pages.

1. Purpose

This report presents a summary of the impact assessment undertaken for alternative policy options developed for the project 'Introducing requirements on the type and quality of evidence needed for scoring fisheries', which is part of the MSC's Fisheries Standard Review (FSR) project Ensuring effective fisheries management systems are in place.

This report provides a description of the options under consideration at the time of the impact assessment January-February 2021) and a summary of the likely impacts for each of the different options.

The results of the impact assessment were used to inform the choice of recommended options, which were presented to the MSC's governance bodies in March 2021. This report was also presented as supporting background material.

2. Impact Assessment Framework

The aim of impact assessment is to provide clear information on the impacts of the options developed to sort out the policy issues identified in the project inception. It serves as a basis for comparing options against one another and against the business-as-usual scenario, and identify a preferred option if possible. It does not replace decision-making but is used as a tool to support the decision-making process and underpin evidenced based decision-making; increasing transparency, making trade-offs visible and reducing bias.

Impact assessment should help to:

- Specify how proposed options will tackle the identified issues and meet objectives
- Identify direct and indirect impacts, and how they occur
- Assess impacts in both qualitative and quantitative terms.
- Help find perverse or unintended consequences before they occur.
- Where possible, make risks and uncertainties known.

This is achieved by following MSC's Impact Assessment Framework that outlines when and how to undertake Impact Assessment. This ensures an efficient, systematic and consistent approach to policy development to underpin a responsive, robust and credible program. In particular, the Impact Assessment Framework defines the different types of impact (see below) and a suite of methodologies best suited to assessing each type.

The impact types used in the Impact Assessment are defined as follows:

- **Effectiveness:** The extent to which the change is deemed likely to be successful in producing the desired results and resolving the issue(s) originally identified.
- **Acceptability:** The extent that the change is considered tolerable or allowable, such that the MSC program is perceived as credible and legitimate by stakeholders.
- **Feasibility:** The practicality of a proposed change and the extent to which a change is likely to be successfully implemented by fisheries within a given setting and time period.
- **Accessibility & Retention:** The extent to which the change affects the ability of fisheries (both currently certified and those potentially entering assessment in the future) to achieve and maintain certification (i.e. changes in scores, conditions and pass rates).

- **Simplification:** The extent to which the change simplifies and does not further complicate the Standard such that it can be easily and consistently understood and applied.
- **Auditability:** The extent to which the change can objectively be assessed by Conformity Assessment Bodies (CABs) and Accreditation Services International (ASI) to determine whether the specified requirements are fulfilled, and CABs can provide scores.

The Impact Assessment report presents the results of this process, whereby each of the options for proposed changes to the Fisheries Standard are tested to understand their potential effects across the six defined impact types.

3. Problem Statement

Good information is crucial to providing confidence in the assessment of a fishery's impact or the performance of a management measure. As such, it is important that the adequacy of information available for a fishery assessment is taken into consideration. Similarly, by setting requirements on how much information is needed, and how good it must be, to meet certain scoring guideposts (SGs), the MSC Program can push forward improvements in fisheries monitoring.

3.1. Need for intervention

A number of scoring issuess (SIs) require assessment teams to make a determination on the adequacy of information that is available in the assessment, for instance, that information from or regarding the Unit of Assessment (UoA) is adequate to assess its impact, determine compliance or support management measures. This requires consideration of the quality and quantity of available information and a judgement on its adequacy in the context of the assessment. There is only limited guidance to support assessors in this task. This has resulted in differences in teams' judgement of what is adequate, and in the transparency of how they have reached their judgement.

Inconsistency in how the adequacy of information is judged creates inequality in the program, as fisheries may be held to a higher or lower bar at the discretion of the assessment team. This creates uncertainty in what quality and quantity of information is needed to in order to perform well in the MSC Program.

3.2. Business-as-usual scenario

The MSC Fisheries Standard contains limited and fragmented instruction on how assessment teams should consider information adequacy. This is either provided as requirements at the SI/SG level, or as supporting guidance. This guidance is verbose in parts, without providing clear direction to assessors. Instruction is most developed for Principle 1 (P1) SIs, and to a lesser extent for certain SIs of Principle 2 (P2) related to UoA impact. Instruction related to the determination of compliance is very limited and largely superficial.

4. Objectives and intention

4.1. Objectives

The objectives of this project are:

- To ensure that fishery assessments are based upon a high and consistent standard of information, in terms of its quantity and quality; and
- To ensure that the determination of information adequacy is consistent and transparent.

4.2. Intended effects

In terms of our Theory of Change, the intended effect will be to incentivise the collection and provision of best available information by fisheries, and so encourage improvement in fisheries monitoring and surveillance. This will be done by reducing the disincentive for a fishery to generate less or lower quality information on its activities, as to do so would risk information being determined as inadequate.

In terms of implementation of the Standard, the intended effect is to improve the consistency of its application and increase the transparency of assessors' decision making with respect to the information adequacy SIs.

5. Options

5.1. Overall approach

This report compares two versions of an evidence requirements framework. Both versions build on the existing instruction already provided to assessment teams, but go further to provide for systematic evaluation of information and sets clear expectations of what information should be considered as 'good enough'.

5.1.1. Applicable SIs

This framework approach would be applicable to seven SIs that require assessors to make a determination of information adequacy with respect to impact on main P2 species, endangered, threatened and protected (ETP) species and habitats, and shark finning (see Table 1). Note that some SIs will share the same evidence requirements, e.g. shark finning SIs (SIs), and would only need to be applied where necessary. It is possible that the approach may also be applied to new SIs created as part of the Fisheries Standard Review (FSR), or removed from others that become redundant, i.e. main/minor designations in P2.

In addition, the part of the framework that focuses on evaluation of information quality would be applied to a further eight SIs that relate to impact on minor species, supporting the management strategy and compliance. In this case no information thresholds would be applied. Instead, scoring would be similar to business as usual, but the assessors' decision making and reporting would be structured around the quality evaluation criteria.

This distinction is based on the nature of the available information. Thresholds for both options are designed to be applied to quantitative information, such as catch data. They are much less

appropriate for qualitative information, such as records of non-compliance, or where information is used to support management (as importance, relevance or interpretation of information may vary).

Table 1: SIs to which the evidence requirements framework would apply, either in full or in part.

Full evidence requirements framework	Evaluation of information quality only
PI 1.2.1 Harvest strategy; SI (e) shark finning PI 2.1.2 Primary species management; SI (d) shark finning PI 2.1.3 Primary species information; SI (a) impact on main primary species PI 2.2.2 Secondary species management; SI (d) shark finning PI 2.2.3 Secondary species information; SI (a) impact on main secondary species PI 2.3.3 ETP species information; SI (a) assessment of impacts PI 2.4.3 Habitats information; SI (a) assessment of impacts	PI 2.1.3 Primary species information; SI (b) impact on minor primary species PI 2.1.3 Primary species information; SI (c) information for management strategy PI 2.2.3 Secondary species information; SI (b) impact on minor primary species PI 2.2.3 Secondary species information; SI (c) information for management strategy PI 2.2.3 ETP species information; SI (c) information for management strategy PI 2.5.3 Ecosystem information; SI (d) information to infer consequences PI 2.5.3 Ecosystem information; SI (d) information to detect risk increase PI 3.2.3 Compliance & enforcement; SI (x) information for assessment of compliance [new SI]

5.1.2. Associated revisions to some scoring guideposts

In addition to the introduction of an evidence requirement framework, it will be necessary for several SIs to revise the language used in the SGs. Some SGs specify or imply the amount or type of information that is required, e.g. that 'qualitative' or 'some quantitative' information is needed. This overlaps with the function of the evidence requirements framework and gives rise to the possibly of inconsistency and confusion.

5.2. Alternative framework designs

5.2.1. Version A: Scoring bound by information quantity

5.2.1.1. Conceptual design

In this version, assessors' judgement of information adequacy is structured by an evidence requirements framework where the scoring outcome is primarily driven by the amount of information available. The framework initially uses thresholds, based on the amount and type of information available, to determine the maximum theoretical score that can be achieved for the SI. These thresholds are set at a different level depending on the UoA's risk of causing negative impact, whereby more information would be needed for UoAs considered to have high impact risk. The quality of the information is then evaluated to finalise the score, which could be awarded at or below the theoretical maximum.

Version A follows the four-step process in Figure 1. Each of the steps is further described in the worked example in section 5.2.1.2: Worked example: Primary species information.

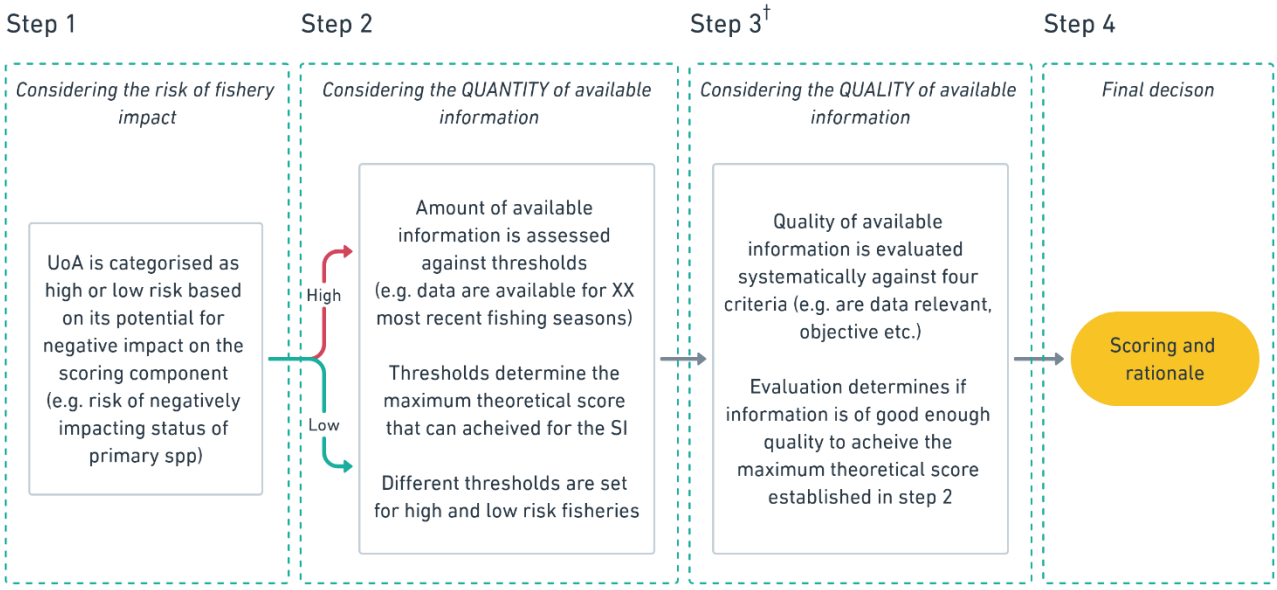

Figure 1: Steps involved in the evidence requirements framework (Version A) to determine the adequacy of information. This process is followed for each of the seven SIs to which the process will apply in full. †A reduced process involving only Step 3 will apply for a further eight SIs.

5.2.1.2. Worked example: Primary species information

The Primary species information performance indicator (PI 2.1.3) asks whether information on the nature and amount of primary species taken is adequate to determine the risk posed by the UoA. SI (a) is focused on information adequacy for assessment of impact on main Primary species. The evidence requirements framework is applied to this SI in the steps outlined below to illustrate the concept.

Step 1: Assessment of UoA risk

The assessment team first consider the risk of UoA having negative impact on the scoring component. In this example, it is the risk of the UoA negatively impacting the status of Primary species. The risk classification, either high or low, will determine which information thresholds are applied in the next step of the process. See an overview of the process in Figure 2.

The risk assessment considers the likelihood (level of probability) and the consequence (level of negative impact) of an event. In this example, it is the risk of fishing mortality negatively impacting the stock status of a Primary species. These attributes are estimated qualitatively, allowing for the fact that information and methods may not available to undertake quantitative estimation. It is proposed that likelihood is described as either likely, possible, or unlikely; and consequence as either minor, moderate or major. These terms are modified from those outlined in the Standards

Australia Risk Management Guidelines[1] and Fletcher (2005)[2], and would be clearly described in guidance.

The risk classification, assessed for each scoring element individually, will be based on the combination of likelihood and consequence. For example, a 'high risk' classification would be triggered by a possible severe impact, or a likely moderate impact (see Table 2). These trigger combinations would be defined in guidance. The limited selection of qualitative terms and the prescribed trigger combinations build precaution into the process. Furthermore, where there is uncertainty the assessment team should be precautionary when selecting the likelihood or consequence classification. When completed for all scoring elements, the overall risk classification would be based on the most precautionary outcome, i.e. a UoA is 'high risk' if one or more scoring elements has this classification.

Note that this is an assessment of potential risk that should consider the full range of possibilities, including hypothetical risks. This mean that the risk assessment is not limited to the available information, which may be incomplete, and allows space for expert judgement and reasonable argument. Also note that the assessment of risk is completed separately for each SI (i.e. risk of negative impact on ETP species, risk of shark finning occurring, risk of non-compliance and so on).

Table 2: Matrix to determine high or low risk from likelihood and consequence. Red is high risk; green is low risk.

		Consequence		
		Minor	Moderate	Major
Likelihood	Unlikely			
	Possible			
	Likely			

[1] Standards Australia, Environmental Risk Management: Principles and Process, 2000Homebush, NSWStandards Australia Based on AS/NZS 4360:1999, Risk Management. HB 203-2000 81 pp.

[2] W.J. Fletcher, The application of qualitative risk assessment methodology to prioritize issues for fisheries management, ICES Journal of Marine Science, Volume 62, Issue 8, 2005, Pages 1576–1587, https://doi.org/10.1016/j.icesjms.2005.06.005.

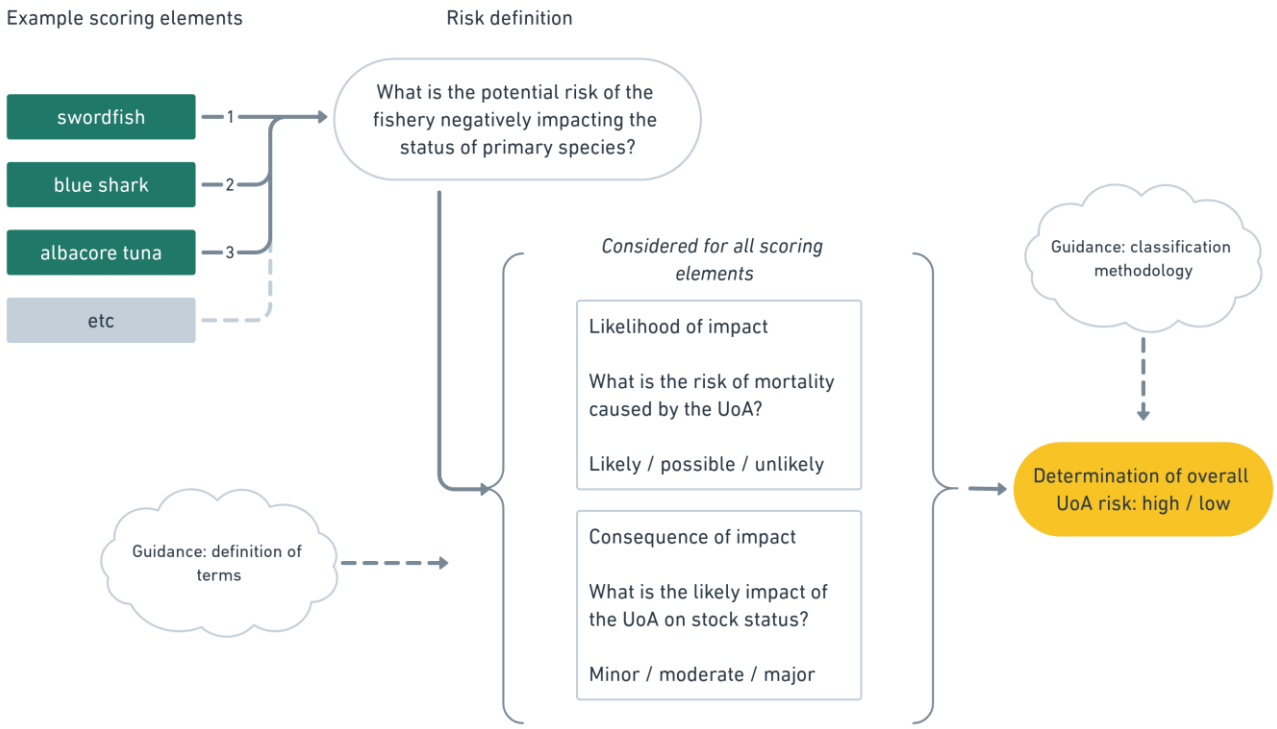

Figure 2: Process for UoA risk assessment using the Primary species example. This process applied to both options.

Step 2: Information thresholds

The type and amount of information available from the UoA is assessed against thresholds to determine which scoring level it can achieve. This will determine the maximum theoretical score it can achieve for the SI; the final score is also based on the quality of the information, which is evaluated in the next step of the framework (see Figure 3 for an overview of the thresholds process).

This step is intended to introduce clear instruction to assessment teams for what level of information should be considered as adequate to meet each of the three SGs. These thresholds describe the type of information needed (e.g. independently verified catch data), at what quantities (e.g. from X% of Unit of Certification (UoC) trips) and from what time period (e.g. available for X most recent seasons) for each of the SG levels. This format provides clear expectations regarding the objectivity, completeness and relevance of the information.

Thresholds require incrementally better information at the SG60, SG80 and SG100 levels, and are modified to apply either to high or low risk UoAs. This in effect creates two pathways: one for low risk UoAs, which are assessed against one set of thresholds (e.g. A1, A2 and A3); and high risk UoAs, assessed against a different set of thresholds (e.g. B1, B2 and B3). This is achieved by modifying the thresholds denoted with 'X' in Figure 3.

The setting of information thresholds is likely to involve substantial discussion and receive particular attention from stakeholders. In anticipation of this, a consultative methodology for developing and agreeing information thresholds is presented in section 5.3: Approach to setting information thresholds.

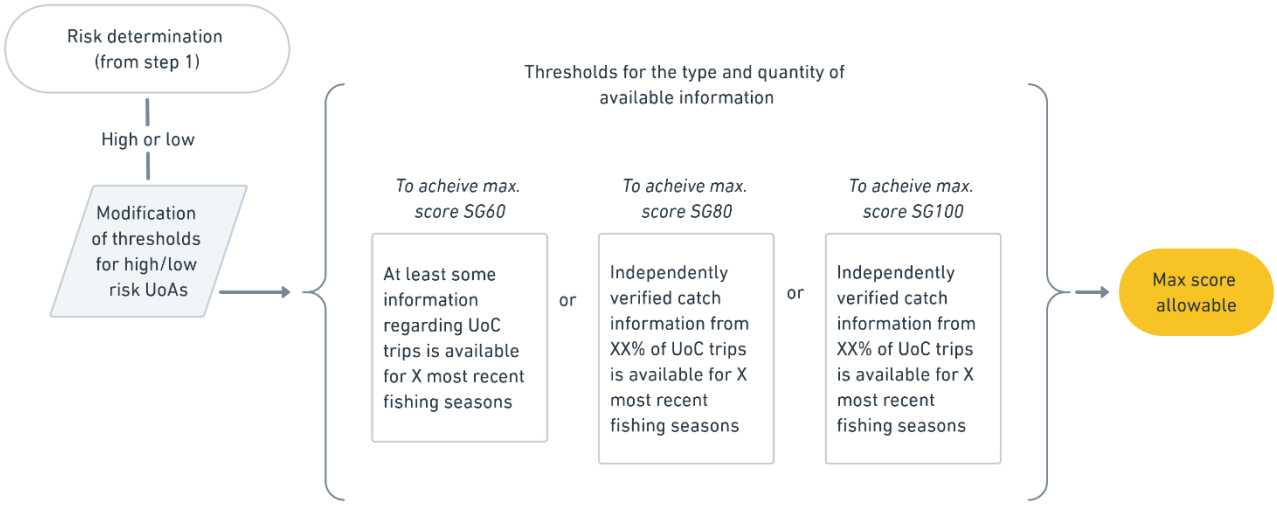

Figure 3: Process for applying information thresholds (Version A) using the Primary species example.

Step 3: Evaluation of information quality

The quality of the available information is evaluated against four pre-determined criteria. This evaluation provides opportunity for assessment teams to use their expert judgement on 'how good' the information is, taking into account the fishery context, best practices, nuances in the information system, and any other relevant considerations (see Figure 4 for an overview of the process).

The intention of this step is to ensure a rigorous and consistent evaluation of information by assessment teams. The four evaluation criteria have been selected with input by a panel with expertise in the appraisal of information, policy development and fishery assessment.[3] The criteria and their definitions, which would be clearly described in guidance, are as follows:

- Objectivity - the extent to which information is unbiased and free from conflict of interest;
- Relevance - the extent to which information is pertinent or connected to the matter in hand;
- Consistency - the extent to which information is accordant with itself and other comparable sources;
- Completeness - the extent to which information captures all elements and dimensions.

The output of this process is a qualitative description of information quality. Assessment teams would use this this to inform their determination of the score, and a summary of the evaluation would form the basis of the scoring rationale.

[3] Report of this consultation process to be made available at a later date.

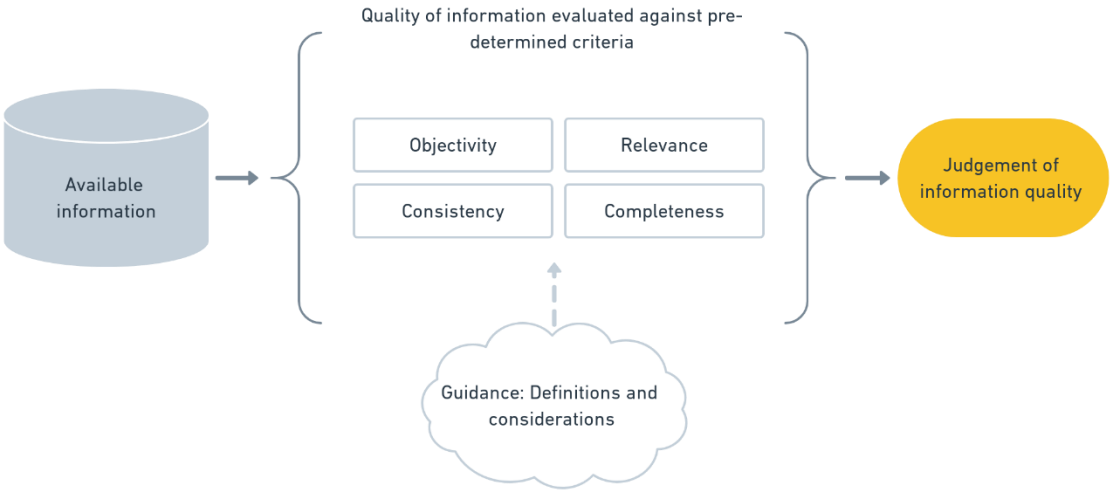

Figure 4: Process for evaluating information quality. The process is the same for all options and SIs.

Step 4: Final determination of score

The assessment team will need to use their expert judgement to decide which SG is met, informed by the preceding steps of the evidence requirements framework. The final score is set by considering both the maximum theoretical score, as determined using the information thresholds, and the quality of the information. The thresholds set the maximum score that the UoA *could* achieve for the SI, but the evaluation of information quality informs whether it *should* achieve that score (see Figure 5 for an outline of the process).

Using the Primary species example, it may be that the SG100 threshold is met, i.e. independently verified catch information from XX% of UoC trips is available for X most recent fishing seasons. This is the maximum score than can be achieved. SG100 requires that information "*is adequate to assess with a high degree of certainty the impact of the UoA on main primary species with respect to status*". The assessment team therefore need to consider if the available information is of good enough quality to achieve this statement.

It may be, for example, there are some concerns with the consistency and completeness of the available data that prevent SG100 being met. However, they may be confident that the data is of good enough quality that SG80 is met, i.e. that information "*is adequate to assess the impact of the UoA on the main primary species with respect to status*". In this example, the assessment team would determine that SG80 is met, but not SG100.

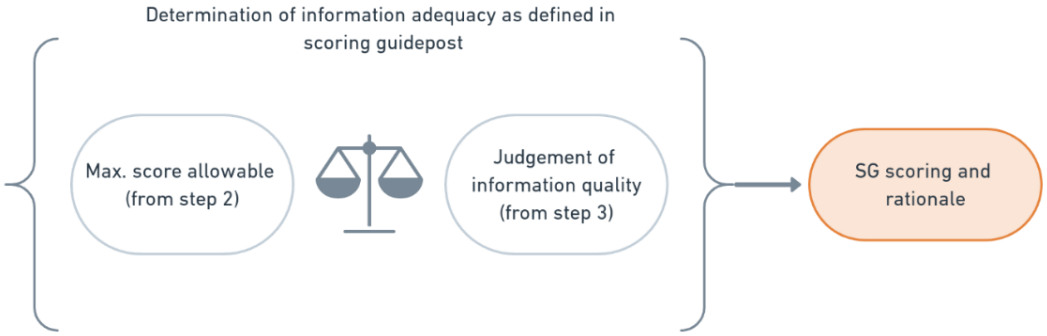

Figure 5: Process for determining the final score (Version A). This process is the same for all SIs.

5.2.2. Version B Balancing information quality and quantity

5.2.2.1. Conceptual design

This option is a modification of Version A with more equal consideration of information quality and quantity in determining the scoring outcome. This option has been developed in response to feedback on Version A received from the MSC Stakeholder Advisory Council (STAC) Principle 3 (P3) working group in early February 2021 (see section 5.2.2.2: Response to STAC working group feedback below).

In this alternative version of the framework, the amount of information available is not used to set a maximum score. Instead, assessors must consider both the quality and quantity of available information together when determining a score. Rather than split the steps of the process into consideration of quantity and then quality of information, the framework instead makes a distinction between assessment of information system quality[4] (e.g. is information relevant? is it complete?) and statistical quality (e.g. is it biased? is it precise?).

To achieve this, Version B introduces the option of describing information thresholds based on statistical properties of the observed variable, rather than specifying the number of observations required. This responds to the concern that information thresholds in Version A could be unjustly arbitrary and subjective.

The overall process has been simplified in Version B. Risk assessment has been made more manageable for multiple scoring elements, and the information threshold process is simplified by avoiding the need to develop separate pathways for high and low risk UoAs. The move towards considering quantity and quality together also shortens the process overall.

Version B follows the three-step process in Figure 6. Each of the steps is further described in the worked example in section 5.2.2.3: Worked example: Primary species information below.

[4] System quality refers to the quality of the array of available information. The concept of "system quality" was introduced in this version of the framework and needs some work to be fully developed. For example, system quality is higher if there are few or no gaps in the spatial and temporal coverage of fishing activity, if information pertains directly to the UoC, if information is up to date, and so on.

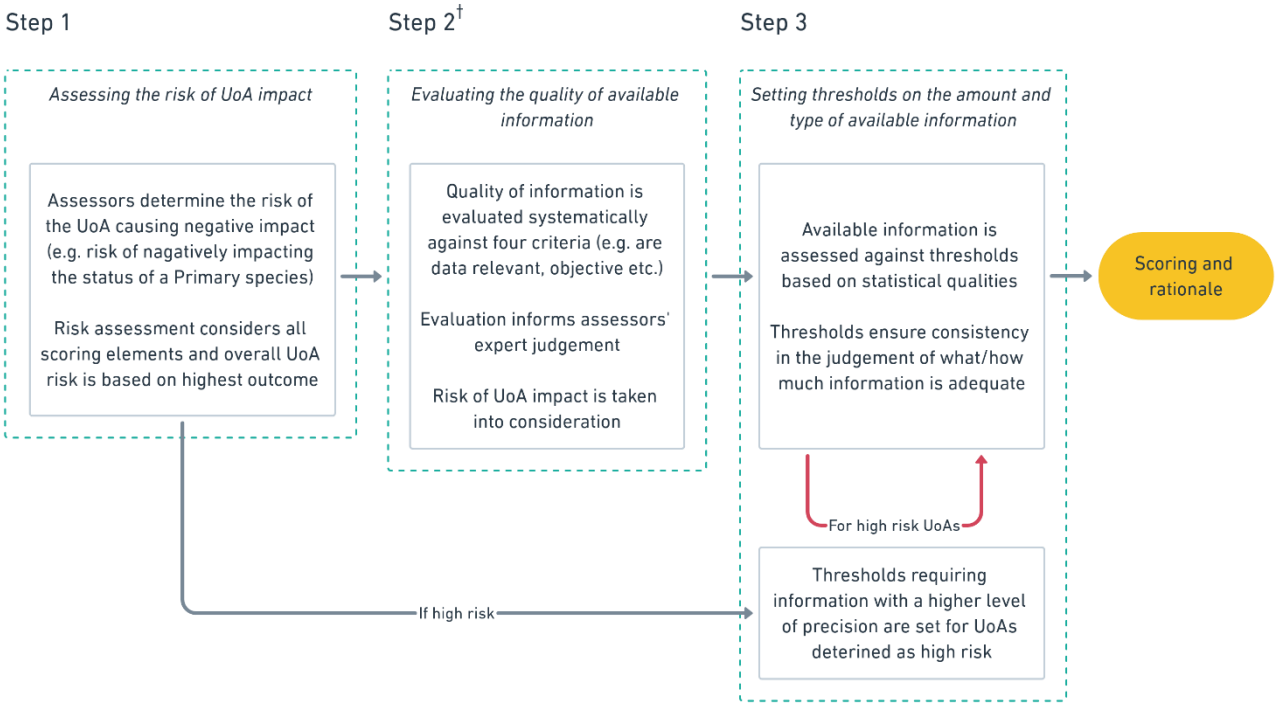

Figure 6: Steps involved in the evidence requirements framework (Version B) to determine the adequacy of information. This process is followed for each of the seven SIs to which the process will apply. †A reduced process involving only Step 2 will apply for a further eight SIs.

5.2.2.2. Response to STAC working group feedback

Version A of the evidence requirements framework was presented to the STAC Principle 3 working group on 3 February 2021. Version B has been developed in response to feedback received from the working group. A summary of key feedback received, and how it has been responded to, is presented in Table 3.

Table 3: Summary of STAC P3 working group feedback on Version A and how it has been considered in Version B.

STAC P3 working group feedback	How Version B addresses feedback
Concern on ordering of information quality and information quantity in the framework, and on the setting of maximum theoretical scores	Version B does not use information quantity to set maximum scores; instead, quality/quantity are considered together, and more equally, in the process Version B no longer includes the concept of maximum theoretical scores
Noted risk of information thresholds being arbitrary or subjective, and questioned how precision and bias are reflected in the information thresholds	Version B proposed alternative thresholds based on statistical properties (e.g. providing a level of precision equivalent to a coefficient of variation of XX)
Supported focus on simplicity of the process, including how scoring elements affects scoring or risk classification	Version B streamlines the process by simplifying risk assessment for multiple scoring elements and how risk modifies information thresholds

STAC P3 working group feedback	How Version B addresses feedback
Suggestion that precaution should be built into the framework	It is clarified in both options where precaution is built in, i.e. designing thresholds to be precautionary, use of precautionary approach in classifying risk
Question about definition of risk and suggestion to ensure all terms are used consistently in papers	Terminology and definitions are made clear in both options, with methodology references provided

5.2.2.3. Worked example: Primary species information

The Primary species information PI 2.1.3 asks whether information on the nature and amount of primary species taken is adequate to determine the risk posed by the UoA. SI (a) is focused on information adequacy for assessment of impact on main Primary species. The evidence requirements framework is applied to this SI in the steps outlined below to illustrate the concept.

Step 1: Assessment of UoA risk

This step is only modified slightly from that described for Version A and shown in Figure 2. Risk is characterised in the same way (likelihood x consequence) and uses the same terminology and designations.

The key difference is that process is not repeated independently for each scoring element, allowing flexibility to group species together when considering likelihood and consequence. All scoring elements would still need to be considered, and the risks clearly reported in the scoring rationale to ensure transparency of the risk assessment. This would simplify the process significantly for UoAs that have a large number of scoring elements. The overall risk classification would still be based on the most precautionary outcome, i.e. a UoA is 'high risk' if one or more scoring elements has this classification.

Step 2: Evaluation of information quality

This process is unchanged from that described for Version A and shown in Figure 4. The key difference in Version B is that the risk classification from step 1 becomes an explicit consideration in the evaluation of information quality. It is noted that assessment teams do already consider risk to a greater or lesser extent when making a judgement on information adequacy; Version B would build on this to ensure that risk is considered systematically and consistently as part of the evaluation of information.

Step 3: Information thresholds

There are two key changes to the information thresholds process in Version B: how thresholds are expressed, and how they are applied.

In Version B, thresholds based on statistical properties describe what is required from the information, rather than how much is required (see Figure 7). Instead of requiring that a prescribed coverage of observer data is needed, for example, the threshold specifies what the available information must be able to achieve in statistical terms, e.g. that information for UoC trips is available that guarantees an unbiased estimate of catches with moderate precision (equivalent to a coefficient of variation of XX% in a random sampling programme). The use of equivalents minimises

subjectivity in how the thresholds could be interpreted by assessment teams (cf. the use of "a level consistent with MSY" in PI 1.1.1).

The statistical property chosen for use in the threshold will be appropriate to the SI that it is applied to. For instance, using the Primary species example, the threshold might require that the available information that guarantees an unbiased estimate of catches with moderate precision at or equivalent to a certain coefficient of variation. This approach does not preclude the use of setting quantity-based thresholds on variables such as observer or vessel monitoring system (VMS) coverage, which may be useful proxies to aid accessibility of the approach, although this would be alongside corresponding requirements on the quality of these data.

This option also proposes a simplified way of applying thresholds. Both high and low risk UoAs have a set of basic thresholds they must meet (white boxes in Figure 7), but only high risk fisheries would have an additional layer of thresholds applied to them (red boxes in the same figure). These 'red box' thresholds specify the additional level of statistical quality that is expected given the high risk of negative impact. For example, it may be required that information for high risk UoAs is 100% more precise. In this case, if information from low risk UoAs must achieve a level of precision equivalent to a CV=0.5, whereas for high risk UoAs this level increases to CV=0.25 for the same SGs. These thresholds would be developed using the methodology presented in section 5.3: Approach to setting information thresholds.

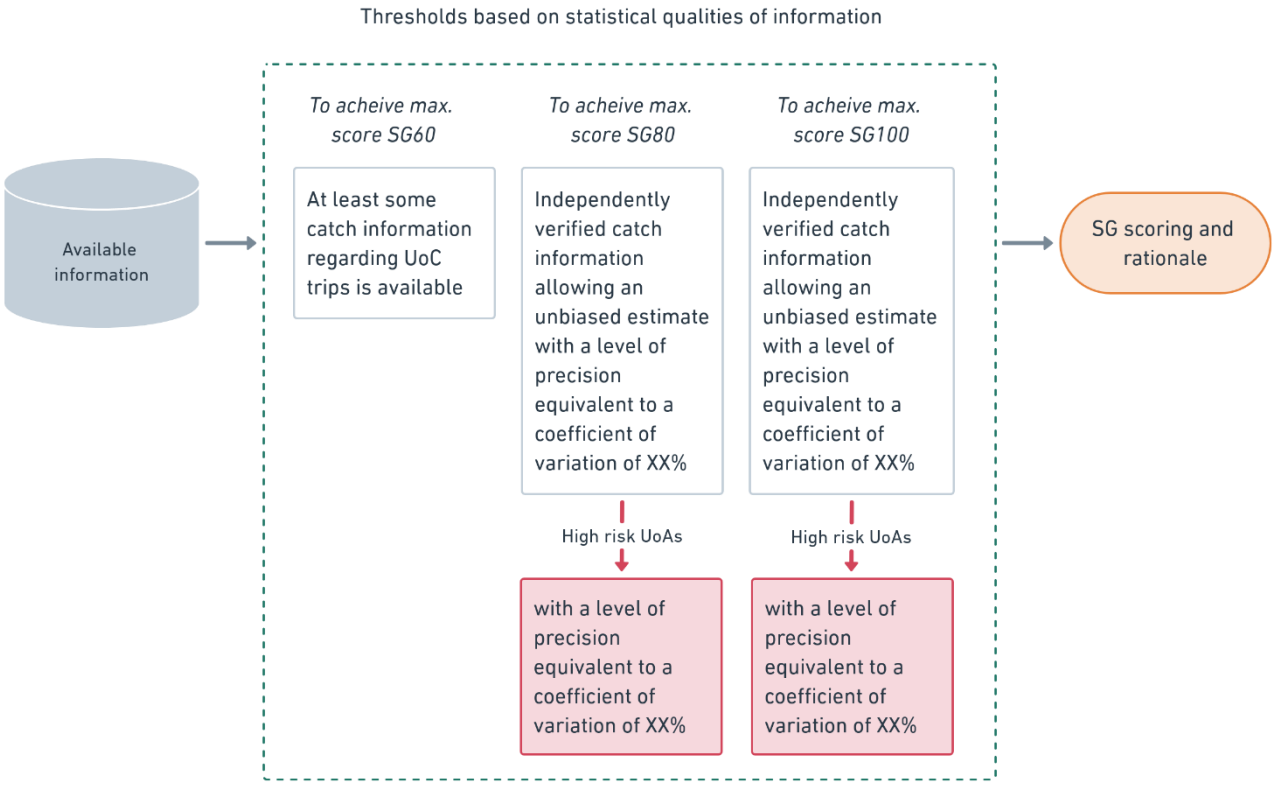

Figure 7: Process for applying information thresholds based on statistical qualities (Version B) using the Primary species example.

5.3. Approach to setting information thresholds

It is proposed that the setting of information thresholds, whether based on quantities or statistical properties, is done in the following stages:

1. Gathering the knowledge base
   a. Commission an independent study to review evidence on optimal levels of fishery monitoring, including the scientific basis for different coverage levels, and global differences in best practice.
   b. Compile and review appropriate metrics for describing bias and precision in fishery monitoring information, including how these can be estimated and how their use varies globally.
2. Workshop bringing together experts to recommend appropriate information thresholds for each of the information adequacy SIs; this should involve participants with expertise in fishery monitoring design, statistics and fishery assessment, as well as representation of the MSC's stakeholders.
3. Review of information threshold recommendations by the MSC Executive, the STAC and the Technical Advisory Board (TAB), passed to the Board of Trustees for approval.

6. Summary of impacts

This section summarises the findings of the impact assessment, which are also presented in section 7: Impacts.

6.1. Impacts of the business-as-usual scenario

The business-as-usual scenario is problematic when a conformity assessment body's (CAB) determination of adequacy is inconsistent with the MSC's intent and stakeholders expectations for what is reasonable. While the status quo is simple in concept, the existing guidance is wordy and the fragmentation of instructions throughout the Standard is confusing and inefficient. There are no major auditability concerns with the existing requirements, in part because most allow for broad interpretation of what CABs need to demonstrate. The majority of stakeholders perceive this to be an issue and accept the need for improvement.

6.2. Impacts of Version A

Version A would introduce a framework that would incentivise better data for fishery assessments and promote transparent and consistent scoring, therefore broadly achieving the objectives of this project. Providing assessors with a clear decision-making framework would ensure a high level of consistency in the judgement of information adequacy, both within and between assessments. Similarly, being prescriptive on the information required at the different scoring levels would eliminate much of the subjectivity in how information is assessed. This would level the playing field in terms of the information required from fisheries, and bring assessors' decisions on what is (and is not) adequate more in line with stakeholders' expectations. This approach is made more targeted by the incorporation of risk into the framework. The result is that higher levels of information are required when the likelihood and consequence of impact is greater. This allows the Standard to be adaptable in its information requirements, pushing harder in those situations where there is need for it.

The use of the framework would reverse a disincentive to provide more and better information, which may sometimes exist in the Program. For SIs focused on impact or compliance, the risk of an issue being detected increases with more information. Given there are penalties associated with lower scores, this creates a theoretical incentive for fisheries to provide only the minimum level of information needed to achieve an acceptable scoring outcome. This is problematic in cases where subjectivity in the requirements and latitude in expert judgement allow for higher scoring in spite of high levels of uncertainty. Minimum requirements set at the at the lower end of the information spectrum would remove this unintended refuge provided by uncertainty, and the logic of better information resulting in higher scores would be more firmly established.

An area of concern for Version A is the quantitative nature of the information thresholds, which risk being unjustly arbitrary and subjective. Setting thresholds on the quantity of information required creates a set of very narrow requirements, which are unlikely to be appropriate across all fisheries and regions. This may discount differences and innovation in best practice, and creates accessibility challenges. Moreover, choosing what those values should be is inevitably a subjective process that will leave some parties unconvinced with the outcome, undermining the legitimacy of the process. Finally, there is a risk that putting emphasis on quantity of information may drive the generation or more data but at the expense of quality.

Another concern with Version A is the complexity of the process. Efficiency is vital, as the evidence requirements framework would need to applied to multiple SIs, and duplicated for each UoA. While the process for judging information adequacy would, broadly speaking, be more clearly structured and easier to navigate than the status quo, parts of the framework are convoluted. For instance, the risk assessment becomes burdensome for SIs multiple scoring elements, and there may be more efficient procedures than modification of information thresholds for high and low risk UoAs respectively. There is a risk these inefficiencies would accumulate over the course of an assessment to hinder the FSR's overriding objective for simplification.

6.3. Impacts of Version B

Version B shares the positive impacts described in Version A with respect to ensuring consistency, minimising subjectivity and levelling the playing field for fisheries information. The high level concept is essentially the same, but with some important differences in how the component steps are designed. These adjustments address most of the core concerns with Version A.

Information thresholds based on statistical properties have a number of advantages over those based exclusively on quantity of information. They allow for a tighter, more sophisticated level of prescription on what information is adequate. Controlling for levels of bias and precision in the information, for example, will influence not just the amount of information provided but also its quality. There remains latitude for expert judgement on the circumstances by which information achieves a prescribed level, but not on the level itself. This retains the advantages of expert judgement where differences in fishery context and situation are important, while minimising subjectivity in how assessors determine what is adequate. This allows the framework to be flexible enough for a global programme while ensuring consistency in scoring from one fishery to the next.

The consideration of information quality and quantity together also offers a more precautionary approach to scoring than if quantity-only thresholds were used. Where SGs are pegged to an amount of information there is risk that assessment teams would defer to those thresholds in the absence of

suitable evidence to determine adequacy more broadly.[5] This could result in exaggerated and inconsistent scoring. This risk is mitigated where information quantity and quality do not influence the scoring outcome in isolation, and where score is not initially set at a maximum and revised downwards.

A key advantage of thresholds based on statistical properties is that they are flexible to the range of novel as well as more conventional monitoring and statistical approaches used in fisheries. For instance, some management authorities employ risk-based strategies for setting observer coverage levels, which may range from comparatively low to full coverage levels, to achieve a level of precision appropriate to risk. In other examples, modelling or mathematical approaches are used to improve the precision of estimates. Where these strategies achieve an adequate level of information quality, regardless of observer coverage levels, this could be recognised by thresholds based on statistical properties. However it would not necessarily be the case for thresholds based on quantities, which are more rigid and as such risk supressing innovation.

There are, however, some potential disadvantages of using statistical qualities that would need to be mitigated. Statistical terminology can be highly technical, with many stakeholders likely to be unfamiliar with metrics such coefficient of variation. If these terms and concepts are not clearly communicated there is risk of evidence requirements being perceived as a black box. There is also an accessibility risk if thresholds are computationally demanding or if statistical properties are not readily available; for instance, the CV of observer data estimates is rarely reported. This risk will be addressed by considering feasibility and accessibility as part of the consultative process for developing thresholds described in section 5.3: Approach to setting information thresholds.

Version B also reduces some of the complexity in the process, particularly for assessment with many P2 scoring elements. The main simplification is to the information thresholds process, which integrates high and low risk thresholds rather than creating distinct process pathways. In this way, all UoAs must achieve a basic set of information thresholds, with high risk UoAs needing to provide information of a higher level of quality. This approach creates an efficiency in terms of how the process is applied by assessors, as well as in how thresholds are developed and presented in the Standard itself.

[5] An example of this can be seen in existing critical guidance on information adequacy. GSA3.6.3.1 gives examples of data collection methods that have higher/lower validity, and recommends that at least one method associated with higher validity is required to meet SG80 and SG100, e.g. observer programmes, electronic monitoring, independent research. This allows assessment teams to point to observer data as evidence for information meeting the SG80 or SG100 level without any further consideration of uncertainty or bias in catch estimates.

7. Impacts

The impact assessment presented in Table 4 below is based on expert judgement of the project and outreach leads, senior colleagues, feedback provided by outreach co-readers and responses to a public consultation workshop.

Table 4: Impact assessment reporting table.

	Description	Business as usual	Version A	Version B
Effectiveness	*Is the change effective at meeting the MSC's intent?*	a. The status quo is not effective when a CAB's determination of adequacy is inconsistent with MSC's intent of the information adequacy SIs and stakeholders expectations for what is reasonable.	a. Framework will promote transparent and consistent scoring, although this depends on clear reporting in scoring rationales. b. Prescriptive information thresholds narrow down assessors' discretion; this will promote consistent scoring and allow the MSC to have more control in assuring its intent. c. Risk that quantitative thresholds are arbitrary and create accessibility challenges / unintended impacts. d. Putting greater emphasis on quantity of information may drive the creation or more data but at the expense of quality.	a. Same as Version A on points a and b. b. Thresholds based on statistical properties allow more room for assessors' judgement and are more flexible to the context of the UoA; this will reduce possible accessibility challenges and reduces the politicisation of setting thresholds.

	Description	Business as usual	Version A	Version B
	The option seems effective at resolving the issue(s) consistently and reliably.	*2 = Disagree*	*4 = Agree*	*4 = Agree*
Acceptability	*Is the change acceptable to stakeholders?*	a. Majority of stakeholders perceive this to be an issue and accept the need for improvement.	a. General support for the conceptual framework at public consultation and with MSC governance bodies. b. Concern over ordering of quality and quantity in the framework, and on the setting of maximum theoretical scores, as this could incentivise collection of a lot of data but of low quality. c. Concern that quantitative information thresholds are arbitrary and subjective, and would lead to accessibility challenges and unintended outcomes.	a. Same as Version A point a. b. Alleviates concerns on weighting information quantity above quality. c. Thresholds based on statistical properties minimise the risk around setting arbitrary quantitative thresholds.
	The option seems acceptable to stakeholders	*2 = Disagree*	*3 = Neither agree nor disagree*	*4 = Agree*

	Description	Business as usual	Version A	Version B
Feasibility	*Is the change feasible to fishery partners?*	a. The status quo is feasible.	a. Evidence requirements per se should be feasible, although there is a risk that arbitrary thresholds would raise the bar for some fisheries. b. Similarly, a risk thresholds may not be technically feasible for a fishery, creating accessibility challenges; this would be problematic where requirements are not appropriate in the context of a particular fishery or not consistent with best practice. c. The risk assessment and evaluation framework steps are not changes that need to be implemented by a fishery, and are therefore considered to be feasible and affordable.	a. Similar to Version A points a and b, although risk is greatly reduced if thresholds based on statistical properties are used; these better allow assessors to take into account the context of the fishery and relevant best practice. b. Same as Version A point c.
	The option seems technically feasible for fishery partners	*5 = Completely agree*	*3 = Neither agree nor disagree*	*4 = Agree*
	The option seems affordable for fishery partners	*5 = Completely agree*	*4 = Agree*	*4 = Agree*

	Description	Business as usual	Version A	Version B
	The option seems possible given the management contexts of fishery partners	*4 = Agree*	*4 = Agree*	*4 = Agree*
	The option seems doable within 5 years for fishery partners	*4 = Agree*	*4 = Agree*	*4 = Agree*
Accessibility and retention	*Does the change affect the accessibility and retention of fisheries in the MSC Program?*	a. The status quo may arguably have a positive impact on accessibility and retention as CABs have a lot of latitude in how they can determine adequacy.	a. There is a risk ill-judged quantitative thresholds not be technically feasible for a fishery, creating accessibility challenges; this would problematic where requirements are not appropriate in the context of a particular fishery or not consistent with best practice.	a. Similar to Version A, although risk is greatly reduced if thresholds based on statistical properties are used; these better allow assessors to take into account the context of the fishery and relevant best practice.
	The option seems accessible to fisheries seeking certification in the future	*5 = Completely agree*	*3 = Neither agree nor disagree*	*4 = Agree*
	The option seems accessible to currently certified fisheries	*5 = Completely agree*	*3 = Neither agree nor disagree*	*4 = Agree*

	Description	Business as usual	Version A	Version B
Simplification	*Does the change simplify the Standard?*	a. The status quo is simple in concept, although existing guidance is wordy and the fragmentation of instructions throughout the Standard is confusing and inefficient.	a. The process for judging information adequacy is more clearly structured and easier to navigate than the status quo. b. The quality evaluation step is procedurally simple; this does not add any substantial process. c. Information thresholds require modification for high/low risk UoAs respectively, which adds complexity. d. Risk assessment will add a layer of complexity, especially in UoAs with multiple scoring elements.	a. Same as Version A points a and b. b. Information thresholds are less complex than Version A as high risk thresholds are a 'bolt-on' rather than a distinct process pathway. c. Similar to Version A point d but risk assessment is slightly less complex for multiple scoring elements.
	The option seems to simplify the Standard	*2 = Disagree*	*3 = Neither agree nor disagree*	*4 = Agree*

	Description	Business as usual	Version A	Version B
Auditability	*Is the change auditable by CABs?*	a. Requirements with respect to determining adequacy in P1 are generally clear. b. Some requirements in P2 and P3 are clear, but most allow for broad interpretation of what CABs need to do or demonstrate. c. Some parts of the guidance are clear on how assessors should consider information in different situations, but these instructions are not normative.	a. Framework will promote transparent and consistent scoring, and clear definitions and guidance should also improve auditability. b. Prescriptive requirements should provide greater clarity of the MSC's intent to CABs, potentially improving auditability.	a. Same as Version A point a and b.
	The option seems to be auditable by CABs	*3 = Neither agree nor disagree*	*5 = Completely agree*	*5 = Completely agree*

8. Discussion and conclusion

Both versions presented here would introduce a framework that would incentivise better data for fishery assessments and promote transparent and consistent scoring. This framework would ensure a high level of consistency in the judgement of information adequacy, and improve the transparency of assessors' decision making. Being prescriptive on the information required at the different scoring levels would eliminate much of the subjectivity in how information is assessed, which would level the playing field in terms of the information required from fisheries. Assessors' judgement on what constitutes as 'adequate' would be more in line with stakeholders' expectations. The incorporation of risk into the process ensures higher levels of information are required when the likelihood and consequence of impact is greater, meaning the Standard will be able to drive improvement hardest where it is most needed.

Information thresholds based on statistical properties have a number of advantages over those based exclusively on quantity of information. They allow for a more sophisticated level of prescription on what information is adequate setting requirement on not simply the amount of information provided but also its quality. They retain the advantages of expert judgement where differences in fishery context and situation are important, while minimising subjectivity in how assessors determine what is adequate. This allows the framework to be flexible enough for a global programme, and crucially for innovative monitoring approaches to be recognised and rewarded. However, there are some potential disadvantages of using statistical qualities that would need to be mitigated to avoid problems of accessibility or feasibility.

Version B also reduces some of the complexity in the process, particularly for assessment with many P2 scoring elements. The main simplification is to the information thresholds process, which integrates high and low risk thresholds rather than creating distinct process pathways. This approach creates an efficiency in terms of how the process is applied by assessors, as well as in how thresholds are developed and presented in the Standard itself. Efficiency is vital, as the evidence requirements framework would need to applied to multiple SIs, and duplicated for each UoA.