ERA 2023 Benchmarking and Rating Scale Consultation Paper – IRU submission

Overview

The Innovative Research Universities (IRU) welcome the opportunity to comment on the proposed options to revise the Excellence in Research for Australia (ERA) 2023 rating scale. The retainment of the "world standard" terminology in both of the proposed ERA rating scale Options is welcomed. This helps provide administrative certainty to the university sector by maintaining consistency with TEQSA's higher education standards. The progress made towards implementing the field of research 45 Indigenous studies is also encouraging. The IRU also strongly supports the intention to revise key benchmarks to ensure ERA evaluations are consistently applied in citation and peer review disciplines, and the published data remains informative to stakeholders. However, the IRU questions the rationale for changing the ERA rating scale and the efficacy of revised guidance for peer review disciplines.

The current ERA rating scale was acknowledged in the ERA EI Review Final Report 2020-2021 to be well suited for its purposes and aligned with current international standards for research evaluation. ERA has provided a credible evaluation of the quality of our university research and has met its objectives as an evaluation framework and national stocktake of research. Over its decade of four assessments, the ERA has encouraged universities to ensure research investment is well directed to produce quality outcomes. The success of universities in achieving this is not a sufficient rationale for changing a rating scale. Although growth in lowly cited research from outside Australia may have shifted the world benchmark for citation disciplines, this does not appear to be the case for peer review disciplines

Both Options introduce a new "world leading" standard that places great faith in citation-driven impact as a valid proxy for research excellence in citation disciplines and the capacity for reviewers in peer review disciplines to differentiate between "world leading" research units and those merely "well above world standard". Changing the ERA rating scale carries credible risks that the ERA 2023 results will exacerbate inconsistencies and inequities between citation and peer review disciplines and be misunderstood by stakeholders as indicating a regression in research excellence due to a shift in the scale or scale mid-point. At a minimum it will require considerable stakeholder engagement to avoid these impressions.

The IRU preference, as outlined in the IRU submission to the ERA and EI review in October 2020, is to retain the current ERA rating scale, but publish volume metric data to allow differentiation between institutions of similar ratings. The ARC may also consider revising the above world standard benchmarks exclusively in citation disciplines where there are a large number of units achieving the highest rating. This would maintain consistency over time in the ERA rating scale and "world standard", while increasing the ability to differentiate between top performing units.

1. Retain the current ERA rating scale, but publish volume metric data

The ERA rating scale is intuitive, appropriate and should be retained in its current five-tier rating scale, with minor adjustment. The rating scale has an appropriate mid-point (3) of "at world standard", and a symmetrical range from (1) "well below" to (5) "well above" world standard. The ERA EI Review Final Report 2020-2021 acknowledged that the current scale is "broadly consistent with the approach taken in research evaluation processes in other countries and is well suited to represent a broad spectrum of performance ranging from low to high."

The ARC should be cautious about modifying a rating scale that is well suited for its purposes and aligned with current international standards for research evaluation. Demonstrating quality and informing stakeholder decisions are two of the core objectives of ERA. Stability in ERA rating scale and methodology over time has helped ensure information on university research excellence is understood by stakeholders within the university community, as well as in industry, government, the community and internationally.

The main rationale for revising the ERA rating scale outlined in the ERA EI Review Final Report 2020-2021 was that Australian university research excellence has improved so rapidly against the world standard that it has become difficult to differentiate between top performers. 90 per cent of university research now achieves a rating of 3 or higher, and if this was to continue, "there is a risk that the ERA rating scale will become less capable of differentiating between high performers as university ratings cluster at the top of the scale."

The IRU view is that the ERA rating scale is intended to present the quality of Australian university research against international benchmarks. It is not intended to provide a fine grained differentiation between Australian universities at the upper (or lower) end of the scale. Broadly defined categories are also appropriate for an evaluation exercise that contains elements of subjective judgement. Although few units of evaluation (UoEs) in ERA 2018 were rated at the lower end of the scale, changes to ERA methodology, such as a by-line method for output eligibility or new guidance to Research Evaluation Committees (RECs), may affect distributions in future rounds. The ERA ratings are also intended to guide assessments for emerging fields of research and new universities. Differentiation between institutions of similar ratings at the upper end could be more effectively achieved by making the volume of outputs and contextual data publicly available. Transparency on quantum of output would improve interpretation of ERA scores, particularly for low volume fields at the top of the scale.

The IRU has also previously recommended increasing the low-volume threshold to 100, with universities retaining the option to opt-in for fields between 50 to 99 outputs. Increasing the threshold would increase reliability and minimise privacy concerns in relation to publicising data. It may also resolve some of the concerns over differentiating between a growing cluster at the top of the scale. Providing an opt-in would maintain inclusion of smaller universities' fields of strategic importance and maintain ERA's objective of identifying excellence and emerging excellence across the full spectrum of research. The opt-in would also counteract the introduction of the ANZSRC 2020, which increases the number of 4-digit codes from 157 to 213 and will dilute outputs in some fields.

2. Consider revising above world standard benchmarks exclusively in citation disciplines

The Report acknowledged that, while Australian research excellence has genuinely improved, the "world standard" may no longer offer a useful benchmark due to an increase in the number of uncited or lowly cited publications. Presumably this is due to a growth in outputs in countries Australia does not wish to benchmark performance against, skewing citation metrics that guide REC evaluations.

The Report identified no concerns about granularity or ability to differentiate between high performers in peer review disciplines, or the ongoing suitability of the current world standard for RECs in peer review disciplines. The main concern was that ERA may not apply the world standard consistently between citation analysis and peer review disciplines, leading to inequity between humanities and social sciences (HASS) disciplines and science, technology, engineering and mathematics (STEM) disciplines.

As outlined in the Australian Academy of the Humanities' [submission](#) to the ERA and EI Review, the average ERA 2018 rating for peer review disciplines was 3.16 versus 4.17 citation disciplines. The difficulty of differentiating between top-rated fields is almost exclusively a problem for citation disciplines, particularly in STEM where 59 per cent of 4-digit fields were rated a '5' compared to 15 percent for others.

An alternative to changing the rating scale and world benchmark would be to adjust the citation benchmarks for "Above world standard" and "Well above world standard" for citation disciplines. This may include defining "Well above world standard" or "world leading" performance around an RCI benchmark indicating an institution was within the top 10% in the world, with a lower but statistically significant benchmark for "Above world standard". This would be a simpler way to:

- Allow for clear differentiation at the top between citation disciplines – which is where the issue lies – and leave peer review and the ERA rating scale alone;
- Remove the issue of being assessed against "high performing institutions" which will have their own spectrum of outputs and may not be reflective of the composition of Australian universities.
- Redress the inequity of outcomes between peer review and citation disciplines.

3. Supplement peer review discipline evaluation with citation data

Peer review remains the most appropriate method of evaluation, particularly for disciplines where citation analysis is not available (e.g. journal articles are not the main or most important output) or inappropriate (e.g. where excellent research takes a long time to be accepted or become highly cited). The additional guidance is welcomed, but peer review expert panels are likely to continue to struggle to reach consensus to differentiate between research that is "at world standard", "well above world standard", "world leading" or other variants contained in the Options. This potentially leads to more conservative (lower) evaluations, exacerbating the gap between peer review and citation based disciplines.

Availability of citation data should not determine appropriateness of citation analysis and nor should it replace peer review, but the expanding coverage of the publications in the Scopus database since

ERA 2018 has increased the feasibility for comprehensive, robust and transparent citation analyses in some previously peer reviewed disciplines. Where appropriate, citation data could supplement peer review evaluations. Evaluations are best made by disciplinary expert panels, but the following data may assist expert panels reach evaluation consensus in peer review disciplines:

- Number and percentage of outputs indexed by the citation provider. This will provide context and allow for the discipline variation between institutions where parts of the discipline are indexed and other parts are not.
- RCI of the indexed outputs in relation to the world and Australian standards only (not HPI's)
- RCI distribution (dynamic or static)

4. Transparency and public availability of ERA data

The value of ERA would be enhanced at little administrative cost by increasing transparency of evaluations and public availability of data feeding into 4-digit level evaluations, such as bibliometric results (RCI data). This would help identify the contribution of research at more granular (e.g. 6-digit) level where research strategies and partnerships are more practical. It would improve usefulness of data for secondary analysis. Currently there is no public data to indicate the size and composition of the research workforce or funding by field at an institutional level, or to identify research volumes by field and geographical and regional areas.

5. Further refinement of Indigenous research excellence criteria

It is essential that Indigenous research is recognised and evaluated in ERA 2023 for its excellence, informing and supporting stakeholder expectations. This is in addition to the need for a robust national stocktake of research that is Indigenous led and utilises Indigenous theory. However, there is a risk that the current ERA Peer Review Criteria mixes these two objectives and may not capture either sufficiently well.

The relatively small number of journals currently included in the ERA Journal list underestimates the extent to which Indigenous research is currently undertaken, damaging the esteem for Indigenous research. Considerable Indigenous research is published in journals that are "general to the discipline", some of which undertake citation analysis. Given that the assessment of Indigenous research will be by peer review with no supporting bibliometrics, supporting Indigenous research in ERA 2023 may mean diluting research in other areas. Without clear guidance, universities may make decisions about apportioning Indigenous research based on citations received. There is a risk that low cited journal articles, which would traditionally sit in a citation based field, could be diverted to an Indigenous code for peer review, whilst well-cited Indigenous research is retained in the "traditional code". The ambiguity also raises the risk that universities could apportion low cited research into fields where universities are under the low volume threshold, avoiding being assessed at either a 2- or 4-digit level.

IRU response to the consultation paper questions

Section 2: Options for a more granular rating scale

Q2.1 Which rating option (A or B) is preferred?

Option A.

In the absence of an option to retain the current rating scale or make adjustment to the current benchmarks for citation disciplines, Option A more closely reflects the IRU preference and is the only option validated by the expert working group for providing reliable and valid granular outcomes.

IRU member justification supporting Option A over Option B includes:

- Option A provides suitable granularity at the higher end of the scale, without compromising the assessment at the lower end of the scale. There is no benefit in the very granular upper scale of ratings in Option B.
- Option A assesses all levels using both the High Performing Indicator (HPI) benchmarks and the world standards unlike Option B which only uses *either* the HPI benchmarks or the world standard.
- The language in Option A, including the "rating" and the description of the rating, is much clearer and better defined than Option B.
- Option A will be more meaningful to the public, as it is simpler and easier to understand.

Q2.2 Are there particular features of either option that should be adopted or modified?

IRU member feedback includes:

- **Adopt "world standard" as a rating descriptor in both Options:** Although both Options provide consistency with TEQSA's higher education standards for research, the explicit reference to "world standard" as a rating in Option A is far clearer in alignment with the TEQSA requirements for "research that is 'world standard' measured using best practice indicators".
- **Retain positivity by adopting Option A descriptors:** The definition wording in Option A is much more positive and contains no negative connotations for institutions that do not receive the highest results. This is compared to Option B which uses the phrases "not at world class level" and "not among the highest performers" (Section 2.2, Option B, page 11). Additionally, rating descriptors should reflect the result rather than being labelled by a "grade". If Option B is selected, then it is our strong preference that the wording of both the rating and the definition be adjusted to be similar to Option A.
- **Ensure rating descriptors reflect world standard benchmarks:** The proposed changes relating to the introduction of HPIs and shifting of the world standard benchmark from 3 to 2 have the potential to cause reputational harm to the Australian university sector. If Australian universities appear sub-standard due to having below mid-point ratings in fields at or above world standard, this will have associated self-fulfilling international reputational impacts, including reluctance on the part of prospective staff and students to move to an institution. This is particularly the case

for Option B where a 'B' grade and 'C' grade do not intuitively reflect research that is 'above average' or 'around world average'.

- **Restrict comparisons to universities:** The consultation paper uses a range of institutional comparator descriptors that are unnecessarily confusing and inconsistent. The inclusion of research institutes and private research labs is not useful, particularly as the remit of such organisations differs significantly to that of universities (e.g. narrow focus on research in a specialized field, rather than delivering research-informed teaching and PhD supervision across a range of broad courses).

  For the high performer benchmark, the Executive Summary explains that the aim is "to compare Australian universities with the top 10% of *universities* in the world in a given field" (emphasis added). However, the content of both Options use a broader definition, interchangeably comparing Australian universities against the "top 10% of world institutions", "high-performing (top 10%) research institutions worldwide", and "top 10% of organisations".

  For the world benchmark, comparisons are made only to universities, but sometimes these are restricted to "competent universities", "research-active universities" and in a limited number of comparator countries, including "global competitors", "comparable countries worldwide" and "international peers". These restrictions, contained primarily in the plain English descriptions, add confusion and inconsistency to the Key Benchmarks.

  To ensure like-against-like institutional comparisons, restricting the analyses and comparisons to "universities" (anywhere in the world) is simpler, intuitive and a more consistent "world standard".

Q2.3 How will the change in ratings shift university research efforts?

IRU members have identified the following possible implications, positive and negative:

- Increase collaborations with international high performing institutions;
- Increase participation in large global studies (e.g. Global Burden of Disease) where publications are known to be very highly cited;
- Decrease emphasis on research less likely to be highly cited within the timeframe of ERA, including:
  - "Basic", "fundamental" or "blue-sky" research
  - Research that focuses on benefits to small populations or local communities.
  - Industry collaboration on commercial research where results are not made publicly available.
- Research specialisation with (limited) resources, recruitment and funding directed towards already well-performing areas of research which are rated "well above world standard" in an attempt to get them to "world leading". Competitive pressure may generate invidious behaviour, such as highflyer poaching, that will see a weakening of performance among institutions with less purchasing power.

Q2.4 To what extent would the proposed options be more challenging for universities than the existing ERA rating scale?

The proposed ERA rating scales are more complex, less intuitive and will require considerable stakeholder guidance (particularly Option B). Challenges identified by IRU members include:

- **Less predictability and transparency:** Assessment of high-performance being made against outputs from a group of institutions (HPI's) rather than to all outputs worldwide, and lack of predictability of dynamic RCI (and associated lack of transparency).
- **Additional effort/cost:** Universities will be pressured to calculate/purchase benchmark data to enable predictions of ratings. Larger institutions with subscriptions to multiple citation providers may have an advantage in undertaking new modelling and applying that to benefit their submission.
- **Additional stakeholder guidance:** Universities will need to ensure understanding that a reduction in the number of "high-performing" fields is due to the increased granularity at the higher end of the rating scale, not due to poor performance of the institution.
- **Disadvantaging HASS:** Although some changes are suggested in the peer review assessment, the risk is that the suggested changes may exacerbate the divide between the performance of HASS and STEM, incentivising further disinvestment in HASS.

Q2.5 What changes, if any, are required to the characteristics that accompany each rating level?

Possible changes identified by IRU members include:

- **Clarify a "significant proportion of the unit's outputs":** The characteristics for both Options refer to "a significant proportion of the unit's outputs" as a key characteristic for assessing each rating. The reference to "significant" has connotations to statistical significance for differentiation from the broad world standard/benchmark category that is "*around* the average standard of universities worldwide". However, what constitutes a significant proportion is unclear. It will also differ by rating level. For example, a majority of outputs "below world standard" may be sufficiently "significant" to be categorised in the bottom ERA rating, whereas a small minority of outputs "at the forefront of the field" may be sufficient for units in the top ERA rating.
- **Reconsider the appropriateness of "forefront of research" terminology:** There is no guarantee ERA will be able to identify this type of research, particularly in metrics based FoRs where the content of output is not reviewed. Not all highly cited output is at the forefront of research and there are plenty of examples of cutting-edge research not being well received at the time it is published. The ARC may reconsider if the term "forefront of research" is appropriate, particularly in relation to metrics assessed FoRs.

Q2.6 Would it be feasible for expert reviewers to draw meaningful distinctions between each rating points using the characteristics provided?

IRU members believe meaningful distinctions will be:

- More reliable under Option A where there are fewer categories to distinguish between; and
- Easier in citation disciplines where the additional HPI and dynamic RCI benchmarks will assist reviewers to draw meaningful distinctions between rating points.

IRU members also suggest the ARC review how the new benchmarks have been used across the different REC panels before public release of results to ensure that different REC panels have used similar interpretations in their assessment.

Q2.7 What kind of additional training or guidance may be required in ERA 2023 to support the revised rating scale?

The key feedback from IRU members is the need for stakeholder guidance on the substantive changes to the ERA rating scale and inconsistency with prior rounds. This includes bespoke communication to university executives, government staff, media, industry partners, university staff and academics. The is in addition to new training and guidance for REC panel members.

Section 3: How can the citation metrics support the options for a revised rating scale?

Q3.1 How appropriate is the HPI as a method of supporting the rating scale options?

IRU members are positive about the HPI method, but have identified a number of concerns and recommendations.

- **Restrict HPI to universities:** The inclusion of research institutes and private research labs is not useful, particularly as the remit of such organisations differs significantly to that of universities (e.g. narrow focus on research in a specialised field, rather than delivering research-informed teaching and PhD supervision across a range of broad courses). To ensure like-against-like institutional comparisons, the metric providers can limit their analysis to academic institutions (not systems).
- **Increase the minimum publication threshold:** The threshold of 50 papers set for determining the HPI top 10% is too low.
- **Exclude or limit impact of very large multi-authored papers:** The HPI may be skewed by highly cited co-authored papers. If the RCI was weighted for first, last or corresponding author attached to the institution, this may mitigate institutions being included due to a small number of exceptionally highly cited papers. The ARC may consider the option of excluding or limiting those papers with a very large number of collaborators.
- **Clear communication to REC panels and stakeholders that a HPI is not equivalent to world leading performance – but is below it:** A reference point to assess world leading performance may be required. Additionally, post-assessment validation of the use of this indicator should be conducted by comparing how the HPI indicator was used across different UoE's and where the "world leading" benchmark was drawn for each UoE.
- **Simpler alternatives:** It may be unnecessary to create new metrics. Citation providers already have a comprehensive range of metrics available. Using citation centiles would likely be a suitable metric, such as the Average Percentile from Clarivate's existing methodology to identify the top 10% of organisations in each field. Such an approach may be simpler, more transparent and more accessible to smaller universities and other research organisations.

Q3.2 How appropriate are the dynamic RCI classes as a method of discipline-specific benchmarking?

IRU members are mostly positive about the HPI method, but have identified some concerns.

- **Inconsistency and likely downgrading of ERA rating:** The change is likely to negatively impact many UoEs and the results of ERA 2023 will not be comparable to previous exercises. These changes will need to be carefully communicated to stakeholders.
- **Scale of research:** A number of submissions to the ERA and EI Review (including the IRU submission) noted the need to provide clarity in terms of the scale of research being undertaken. There is an important difference between a small research group with 60 highly cited papers compared to a large group with 600 papers achieving similar results. This information is vital for would be staff and students, and also for potential collaborators.

Q3.3 How would the proposed citation methodologies impact research planning?

As outlined above regarding the changes to the rating scale (Q2.3), the specific changes to citation methodologies will likely drive increased collaboration with HPI's or reduced funding towards certain areas of research based on future citation perceptions. However, aligning recruitment and funding towards citation impact will be more complex under dynamic RCI.

Q3.4 Do the new citation metrics support the drive for increased performance (especially in already high-performing disciplines)?

Possibly. The inclusion of the HPI's provide a clearer description and differentiation of excellent performance at top of the rating scale.

Q3.5 Is any additional criteria or information required in the citation disciplines to support the ratings at the highest end of each rating option?

The two main criteria IRU members have suggested for differentiating between UoEs approaching the highest rating are outlined below.

- **Treatment of very highly cited multi-authored outputs:** The presence of one or more very highly cited outputs as a result of a large multi-national, multi-author study such as Global Burden of Disease, can potentially skew results.
  - If the HPI indicator is to truly represent leading institutions in a field, an additional caveat could be put in place, such as at least 30% of the output within a FoR for an institution either being first, last or corresponding author affiliated (depending on disciplinary authorship customs).
  - There should be consideration of the option of excluding or limiting those papers with a very large number of collaborators. This exclusion is used by Clarivate Analytics when determining Highly Cited Researchers, where highly cited papers with more than 30 authors or explicit group authorship are excluded from the calculations.
- **Data on scale of research:** The importance of publishing scale metrics extends beyond the need for transparent and valid differentiation at the highest end, but it will help mitigate against institutions being included due to a small number of exceptionally highly cited papers (see above: Q3.2).

Q3.6 Please provide any additional comments on the proposed citation methodology.

IRU members offer the following examples of information that could be made available to the RECs and universities:

- World, Australian and HPI average output.
- Where the UoE ranks in terms of its volume of output (at least in 5% brackets).
- World, Australian and HPI average citations per paper.
- Where the UoE ranks in terms of its citations per paper.
- Australian and HPI average centile distribution.
- The UoE's centile distribution of eligible output.
- The HPI's for each UoE, either prior to or after the assessment.

Section 4: How can the peer review indicator support the options for a revised rating scale?

Q4.1 To what extent are the proposed changes to peer review guidance likely to result in reports that are useful, informative and relevant for assessment panels?

Whilst the additional guidance is welcomed, IRU members generally do not believe that the revisions will be sufficient to help peer reviewers make an accurate distinction between "above world standard" and "world leading". Possible revisions identified by IRU members are listed below.

- **Remove disciplinary characteristic descriptions:** The main issue with the peer review assessment is ensuring that reviewers and REC panel members are suitably qualified to assess the UoE, whilst not having a conflict of interest. Although methodologies may differ within a single discipline of research, ERA needs to rely on RECs having sufficient breadth of expertise in their field to be alert to "disciplinary practices or perspectives on research" and "disciplinary characteristics that are specific to this field". Rather than include disciplinary characteristic descriptions, we recommend that either panels are enlarged to ensure proper disciplinary representation, or panels are split to ensure that their focus isn't too wide.

- **Assess research against standards at the time it was conducted:** Items being assessed by peer-review can be up to seven years old. The language of the review questions is in present tense and questions relating to timeliness do not appear to take into account the fact that a work could have been both timely and used up to date research methods and/or theoretical foundations at the time it was produced, but the discipline could have moved in a different direction since the time of publication and the work would no longer be as relevant as it first was. In assessing a 2016 research output the reviewers will need to consider whether the research was "world leading" or "displayed a world leading approach" at the time it was conducted and published, not in context of the present day. The following clarification text may be considered:

  "In your assessment, please consider the context of the discipline at the time of publication. Research Outputs should be assessed against the standards and practices at the time of publication, not those of today."

- **Include citation data for peer review fields:** IRU members see value in provision of citation metrics for peer review disciplines. Whilst assessors use their own expert knowledge of the discipline to review individual outputs, the provision of metrics, similar to the provision of the "outlet list" (the names of Journals and Publishers which contain the largest number of outputs in the submission), can provide the assessor with context to assist in review. We suggest that a dashboard should include the following;
  - Number and percentage of outputs indexed by the citation provider. This will provide context and allow for the discipline variation between institutions where parts of the discipline are indexed and other parts are not.
  - RCI of the indexed outputs in relation to the world and Australian standards only (not HPI's)
  - RCI distribution (dynamic or static)

Q4.2 How feasible would it be for peer reviewers to address the proposed peer review guidance?

See above response to Q4.1.

Q4.3 How appropriate is the proposed guidance for Indigenous studies?

The proposed guidance is helpful for a robust national stocktake of research that is Indigenous led and utilises Indigenous theory, but lacking in criteria for the evaluation of research excellence against international benchmarks. Understandably this complicated because Indigenous studies research must first be identified consistently with the ABS definition (which requires research be Indigenous led and utilising Indigenous theory), then assessed for excellence. This is different to other fields and general peer reviewer guidance which focuses exclusively on the output, irrespective of the individual researchers or theory utilised ("world leading research in an ERA context refers to the quality of the output alone").

IRU members offer the following refinements that may be useful to gain insight into research excellence for Indigenous research.

- **An additional ERA Peer Review Criteria:**

ERA Peer Review Criteria	Possible Indigenous Assessor Considerations
Would you consider this research to be world-leading in Quality?	Please consider in your assessment this research in relation to other related Indigenous research in Australia as well as research with reference to other First Nations peoples worldwide.

- **Citation indicators:** The following citation indicators may be included for Indigenous codes which are based on the traditional citation disciplines**:**
  - 4503 – Aboriginal and Torres Strait Islander environmental knowledges and management;
  - 4504 – Aboriginal and Torres Strait Islander health and wellbeing;

- o 4506 – Aboriginal and Torres Strait Islander sciences; and
- o Related codes applicable to Maori and Pacific Peoples.
- **Additional guidance on inclusion of Indigenous research:** The relatively small number of journals currently included in the ERA Journal list underestimates the extent to which Indigenous research is currently undertaken, damaging the esteem for Indigenous research. Considerable Indigenous research is published in journals that are "general to the discipline", some of which undertake citation analysis. Given that the assessment of Indigenous research will be by peer review with no supporting bibliometrics, supporting Indigenous research in ERA 2023 may mean diluting research in other areas. Without clear guidance, universities may make decisions about apportioning Indigenous research based on citations received. There is a risk that low cited journal articles, which would traditionally sit in a citation based field, could be diverted to an Indigenous code for peer review, whilst well-cited Indigenous research is retained in the "traditional code". The ambiguity also raises the risk that universities could apportion low cited research into fields where universities are under the low volume threshold, avoiding being assessed at either a 2- or 4-digit level.

Q4.4 How would the proposed changes to peer review guidance impact universities and/or researchers?

The changes to the peer review guidance are unlikely to directly affect universities or researchers, other than those who serve on REC panels and the time commitment that this takes. The changes are unlikely to address the difficulty and lack of consensus in peer review disciplines to accurately distinguish between "at world standard", "above world standard" and "world leading".

Q4.5 Is any additional criteria or information required in the peer review disciplines to support the ratings at the highest end of each rating option?

As outlined above, IRU members support the provision of citation data for peer review fields (see above response to Q4.1).

Q4.6 Are there any other changes to peer review that the ARC should consider?

The ARC could consider transferring some FoR codes to citation analysis where are large percentage of outputs submitted are journal articles and the majority of these are indexed by a citation provider. This would reduce the pressure on peer reviewers and REC panels and would not compromise the assessment.

22 April 2022