



Human  
Fertilisation &  
Embryology  
Authority

**Information for Quality**  
**Beta website and Choose**  
**a Fertility Clinic**  
**Feedback findings**



# 1. Executive Summary

---

## 1. Introduction

- 1.1.** The aim of our Information for Quality (IfQ) programme is to review the data clinics submit to us, how that data is submitted, the use to which we put that information, and how we then publish it through the website or Choose a Fertility Clinic.
- 1.2.** As part of this IfQ work, we have now developed the new version of the website up to the beta stage. This was launched for the public to view on 12 August 2016, so that they could provide us with feedback on the new design and content.

---

## 2. Background

- 2.1.** We sought feedback on the beta (draft) website and CaFC to discover how well the content was received and where it could be improved before launching the live version.
- 2.2.** The feedback consisted of three elements:
  - an online survey
  - a workshop for clinic staff
  - one on one user testing with patients and others.
- 2.3.** All respondents were self-selecting, so the responses are not representative of all stakeholder groups and interests, but they give a sense of the thoughts of a variety of individuals, clinics and some professional stakeholders. The different methods of seeking feedback were designed to accommodate different preferences.
- 2.4.** The online survey ran for eight weeks, between 12 August and 7 October 2016.
- 2.5.** The workshop with clinic staff was held on 29 September. We particularly asked for feedback on detailed questions about the presentation of clinic success data. The focal point of this session was how age and treatment types should be used as factors in presenting clinic data.
- 2.6.** The user testing was carried out by Reading Room, our design and development agency, between 10 and 14 October 2016.
- 2.7.** This report sets out the findings from each of these feedback strands and makes recommendations on possible changes we can make in response to these before the website goes live in 2017.

---

## 3. Summary of findings

3.1. Below is a summary of the questions we asked and the feedback findings.

### Usability ratings

3.2. We asked how people rated different aspects of the design and content of the beta site.

- The respondents' ratings of aspects of the beta site are overwhelmingly positive but do reflect that the website is a work in progress with room for development.
- There is particular space for improvement in the range of information and how easy it is to find.

### Finding information

3.3. We wanted to get a sense for how we might need to support the website if people couldn't find what they needed. This is called our 'assisted digital support'

- Individuals will take a wide range of approaches when they cannot find what they want and they need a variety of support.
- This 'assisted digital' measure will be a useful baseline for future website development.

### CaFC search

3.4. We wanted to hear how easy this was to use.

- It is clear that the CaFC search is not yet easy enough to use.
- One clear reason seems to be that the search is more complicated to get to in the first place than the current website. We may want to consider how to address this.
- Patients seem to be looking for other facilities in the search tool than are currently available, such as clinic comparison by success rates.

### Headline information

3.5. We present some headline summary information at the top of clinic pages and in search results. We wanted to hear what users thought of this.

- Responses suggest that many people including clinicians, agree that the concept of having headline information is helpful, particularly to compare across clinics.
- The results suggest that many people do not think the way we calculate the headline IVF birth statistic for the beta site is right as it is not meaningful with so many factors aggregated. In the clinic workshop this was the consensus.
- To make the birth rate statistic more meaningful we could consider adjusting this for other factors.
- We should continue to make the limitations of statistics and particular measures clear to users.

### Importance of headline figures

3.6. We asked people to rate the three measures we have chosen as the headline figures to see which they thought was the most important.

- Most respondents believe that a birth statistic is the most important headline figure.

- Though it is clear that birth statistics information is very important, respondents agree that other figures should also be highlighted.

### Age breakdowns in the headline figure

**3.7.** The headline figure in the beta CaFC site shows all ages grouped together. We wanted to hear whether people thought this was right.

- Most believe that all ages should not be aggregated for the headline figure.
- The clinic workshop highlighted three main alternative approaches we could consider:
  - Presenting the success rate for one, more comparable age bracket
  - Presenting the success rate for a 'gold standard' patient (eg, patients meeting the same categories for age, fertility etc. based on no. eggs collected)
  - Presenting a number of headline figures for different age brackets

### Understandability of clinic birth statistics

**3.8.** We wanted to know if people could understand the statistics that we provide for each clinic.

- The majority of users found the birth statistics understandable, though a significant minority found understanding them hard.
- Explanations available with the statistics seem to be widely appreciated and we could consider providing more of these to improve user understanding.
- We could consider making it clearer to users exactly what the measures can indicate to them.
- Because the graphs are non-standard, we may want to consider whether users need more guidance to familiarise themselves with the new presentation.

### Reliability range

**3.9.** We show a reliability range so that users can see how confident we are that a clinic will reproduce its past performance. We wanted to know whether people understood this.

- The current format of the reliability range is not generally understood by those it is intended to inform.
- We may need to further consider the purpose and presentation of this range.
- The meaning of the range may need to be explained more clearly.
- Going further to highlight the number of cycles that data is based on may provide a clearer indication of the reliability of the data.
- We may also wish to consider providing information on CaFC statistical methodology for a more expert audience.

### Division of data at 38

**3.10.** We provide data split into two age categories, under 38 and 38 and above. We wanted to hear whether people agreed with dividing the data like this.

- The two age breakdown at 38 was not a popular way of presenting high level birth statistics, some felt that this could mislead patients.

- An option to consider could be providing high-level statistics covering more age categories and doing away with the two group split further down the clinic page. Another could be providing more of a breakdown at the second level where we currently show just the two ages.

### Grouping treatment types in the success rate headline figure

**3.11.** In the beta CaFC we group all IVF treatment types in the headline figure. We wanted to hear if people thought that this was right.

- Most people disagreed with grouping all treatment types together.
- There was general agreement for grouping standard fresh IVF and ICSI and excluding other treatment types.
- However, there was more debate around how to deal with natural cycle and frozen transfers. We may want to consider these issues further.
- Different treatment types should be clearly displayed in the detailed statistics.
- We should provide national data for each treatment type as this would make it clearer to patients what the average success rate was before they get into the detail for individual clinics.

### Balance of detail between main page and detailed statistics

**3.12.** We wanted to know if people thought the balance of detail was correct.

- The survey responses provide little clear reasoning for the respondents' dislike of the balance between the detailed statistics and the main page.
- Clinic workshop attendees felt that many users may not access the more detailed statistics page. This raises the question of how much information should be provided on the main clinic page - perhaps the main clinic page is not providing enough information for users in its current format.

### Overall opinions - Likelihood to recommend

**3.13.** We asked people to tell us how likely it was that they would recommend the site to others and also to provide us any other comments and suggestions for the service.

- We now have a baseline figure for how likely people are to recommend the site. They are somewhat likely to recommend it. Ongoing measurement of this will be helpful for improving and developing the website once it goes live.
- Positive feedback is helpful, but the particular constructive critical comments provided can be reviewed as we consider what changes we may wish to make before the new site goes live and assist in improving the service.

## 2. Who responded?

### 1. Feedback participants

#### Survey respondents

**1.1.** During the eight weeks it was open we received 210 responses to the beta survey.

**1.2.** The table below breaks the survey respondents down by respondent type – not all who started the survey answered each question.

Type of respondent	Percentage	Number
currently seeking fertility treatment	11.5%	23
currently seeking treatment for reasons other than infertility (eg, to freeze eggs or sperm or have embryo testing)	7.0%	14
currently having fertility treatment	7.0%	14
previously had fertility treatment for reasons other than infertility (e.g. froze eggs or sperm)	5.5%	11
previously had successful fertility treatment	7.5%	15
previously had unsuccessful fertility treatment	9.5%	19
donor conceived person	14.5%	29
a parent of a donor conceived person	7.5%	15
has donated or plans to donate sperm, eggs or embryos	0.5%	1
a fertility doctor	6.0%	12
a fertility counsellor	2.0%	4
another member of clinic staff	15.5%	31
other	6.0%	12
Did not disclose	-	10
<b>Total</b>		<b>210</b>

**1.3.** Separated into three broad groups we can see that patient groups made up 48 per cent of responses, 22.5 per cent were donor groups and 23.5 per cent were clinic groups.

- 1.4.** A further 17 responses, not listed here, were responses to notify us of errors, since the survey had this dual function over the period of the detailed beta survey.
- 1.5.** Of the twelve people who described themselves as 'other', we can see that four were specific types of clinic staff or medical professional and the others were patients, donors, members of the public or other professionals related to the IVF field (ie, an academic and an ex-HFEA staff member). These included a digital lead in health and a statistician currently researching IVF outcome measures. For the purpose of analysis, these people have been considered under the 'other' group that they selected.
- 1.6.** Additionally, we had written responses on behalf of the British Fertility Association and two clinicians. These are included in the annexes to this paper.

### Workshop attendees

- 1.7.** The workshop for clinic staff was held on 29 September 2016. 38 clinic staff attended the workshop. A summary of discussions and points raised has been included in this report at relevant points and the write up of the workshop is included in full at annex 5.
- 1.8.** The attendees of the workshop represented around twenty different fertility clinics and many different types of clinic staff, including persons responsible, embryologists, nurses, clinic and medical directors and admin and office managers. Because this workshop was London based, inevitably the majority of attendees were from London and often larger clinics. The majority also represented private clinics, less than twenty per cent of those attending came from NHS centres.

### User testers

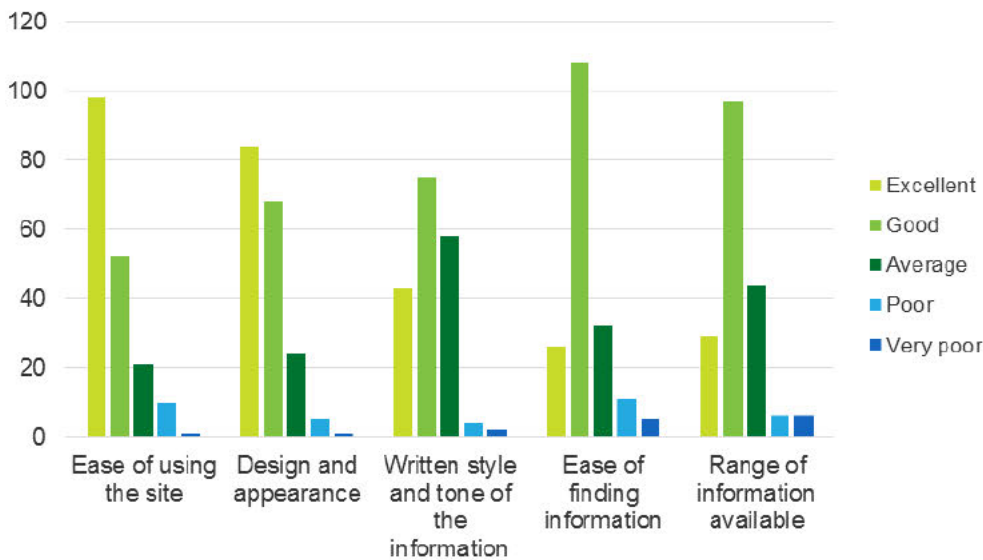
- 1.9.** The user testers came from a variety of different groups. 12 people participated and gave their feedback on the beta service. These people covered a variety of different situations including:
- women undergoing treatment
  - partner undergoing treatment
  - partner in a same sex couple
  - partner in a heterosexual couple
  - women who have donated eggs (egg sharing)
  - single women who have undergone treatment
  - a GP (being interviewed as a fertility patient rather than as a doctor)
- 1.10.** Further user testing of the beta service will be ongoing. A report on these initial findings has been completed by Reading Room, this will be provided to the Authority in full, but a sense of the feedback is noted at the relevant points of this report.

# 3. Using the website

## 1. User ratings

1.1. We asked users to rate different aspects of the beta website against a five-point range, from excellent to very poor. The graph below shows the number of people selecting each rating.

Overall, how would you rate our beta website in the following areas?



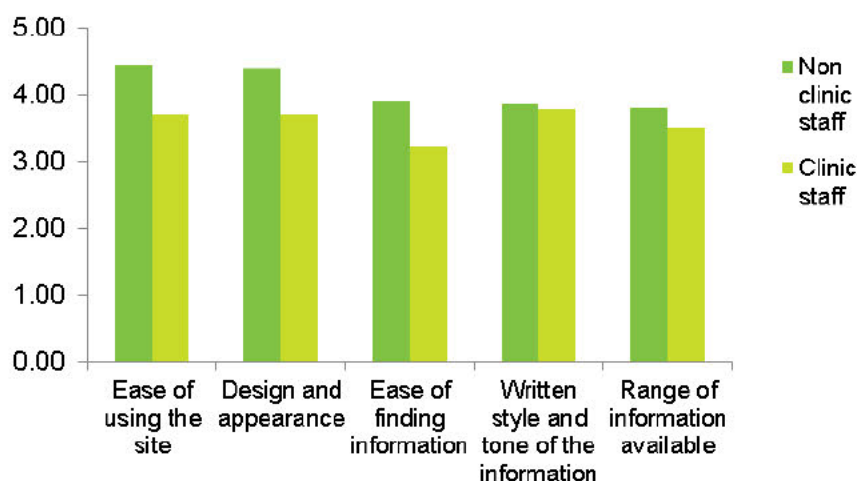
182 respondents. 28 skipped this question

1.2. We can see that a clear majority found the site excellent or good in each area.

1.3. If the answers are given a rating on a five-point scale, where excellent equals five and very poor equals one, we can compare the average rating of clinic staff responses to the average answer by non-clinic staff (excluding 'others' and 'undisclosed'), as in the graph below.



### Graph plotting average clinic versus non-clinic ratings ('others' and 'undisclosed' not included)



137 non-clinic responses, 37 clinic responses. 15 skipped this question

- 1.4.** The graph suggests that clinic staff may have given slightly lower ratings on average than other groups.
- 1.5.** This finding is somewhat limited as there was a much smaller group of clinic staff, so the outliers will have had more of an impact on the average rating.

#### Positive responses

- 1.6.** The general positive feedback is clear when looking at the clinic and non-clinic average ratings, which ranged between 3.22 and 4.45, which represents the average to excellent range. Most respondents found the site easy to use and liked the new design and appearance.
- 1.7.** There was no option to provide comments alongside this rating question, so we have no reasons for these ratings.

#### Areas for improvement

- 1.8.** A significant minority of respondents felt that the written style and tone of the website was 'average' or worse. This may be due to respondents not liking the more personable tone of voice or it could be that the tone is not personable enough.
- 1.9.** In user testing we found that content was seen as welcoming and well written, with the conversational tone coming across very well with patients. It may be that 'average' ratings from the survey responses mainly reflect no strong preference either way, rather than demonstrating a pressing need to change anything about the tone and style.
- 1.10.** 'Ease of finding information' and 'the range of information available' scored the lowest. Since this was in response to the beta site and there was not yet a full range of information uploaded, this may explain why people had difficulties finding information and felt the range could have been better.

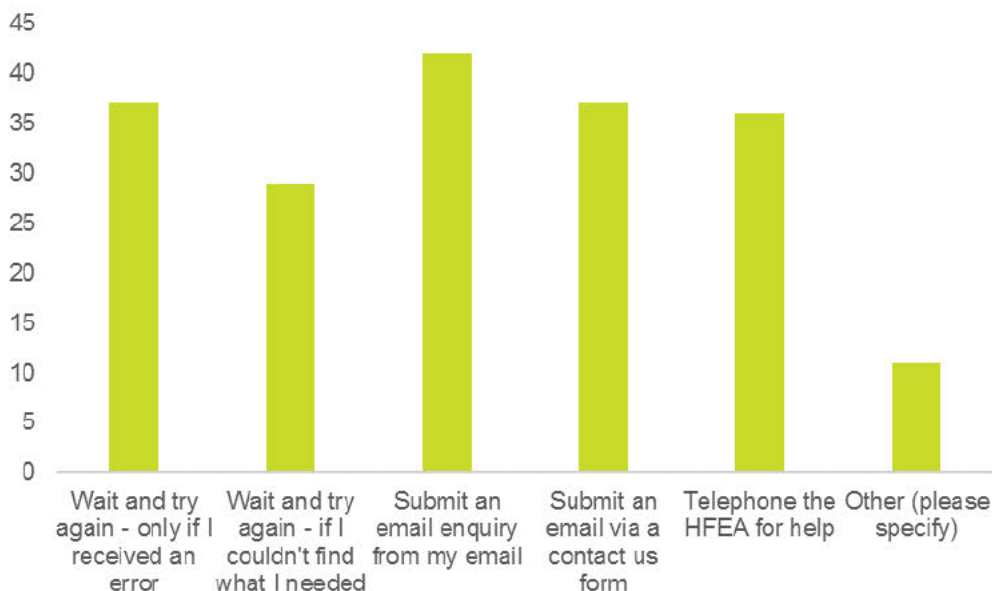
## Summary

- The respondents' ratings of aspects of the beta site are overwhelmingly positive but do reflect that the website is a work in progress with room for development.
- There is particular space for improvement in the range of information and how easy it is to find.

## 2. Finding information

- 2.1.** We asked participants what they would do if they were unable to find the information they wanted, so that we could see what the next line of enquiry would be.
- 2.2.** When building a digital service, you must provide help for people who need to use the service but don't have the skills or access to do so on their own. This is called 'assisted digital support'. We wanted to know which methods of assisted digital support people might use, in case this meant we needed to better support particular methods (ie, provide more telephone support).
- 2.3.** We did not have a baseline measure for assisted digital before this beta feedback.

**If this were the live service (not the beta one) and you were unable to find the information you needed, what might your next step be? (please select all that apply)**



**173 respondents. 37 skipped this question**

- 2.4.** The answers to this question demonstrate that people would take a range of different actions if they were unable to find information. There is no clear preferred next step, though email contact was a common choice.
- 2.5.** The 11 people who selected 'other' mainly indicated that they would look for the information elsewhere, through a search engine or often directly with a clinic.

**2.6.** Although the answers demonstrate a broad range of approaches, regularly monitoring the support methods that users need and use in support of the website will be an ongoing concern for our communications team. This will enable us to make sure that the website, and the support around it, provides for the needs of all users.

### Summary

- Individuals will take a wide range of approaches when they cannot find what they want and they need a variety of support.
- This 'assisted digital' measure will be a useful baseline for future website development.

# 4. Using Choose a Fertility Clinic and headline information

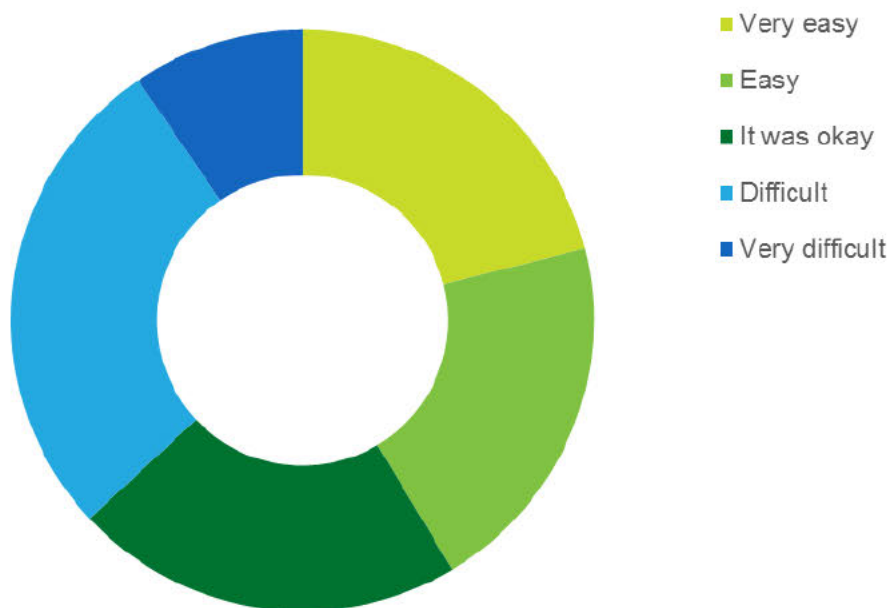
## 1. Introduction

- 1.1. Of those who answered the survey, 170 people, about 80 per cent, also indicated that they had used Choose a Fertility Clinic (CaFC). We asked people a number of detailed questions about the way we had chosen to present statistics in CaFC. Between 157 and 166 people answered these detailed questions.
- 1.2. The first section of questions was about general usability of CaFC and the high level headline information at the top of clinic pages and in search results.

## 2. Ease of search

- 2.1. The search function on CaFC allows users to search by postcode or clinic name. We asked people to tell us how easy it was to use this function.

When you used the search to find a clinic, how was it to use?



157 respondents. 53 skipped this question or did not use the search

**2.2.** Over 60 per cent of those who had used the search found it to be okay or better. Nonetheless, this means a significant number (58 respondents) reported that they found the search difficult or very difficult to use. If the clinic groups are excluded and only donor and patient groups are measured, the result is around five percentage points lower, with over 40 per cent finding it difficult or very difficult.

**2.3.** Since only 18 respondents gave comments and most of these came from clinic staff, the reasons that non-clinic users found the search less user friendly are not clear from the survey.

### Positive responses

**2.4.** Of the comments, a number stated they found the search function clear and did not have any difficulties finding clinics using the postcode search:

‘The layout is good and the tool is very simple to use. The information delivered about each clinic is clear and displayed well.’ **another member of clinic staff.**

‘gave the postcode and several options were immediately available, very clear.’ **a fertility counsellor.**

### Areas for improvement

**2.5.** The main criticism in the comments was not about the search itself, but that it was difficult to get to the first page of the CaFC search because you must first scroll through other information and then click through to CaFC at the bottom of this information:

‘very difficult to find in the first place’ **another member of clinic staff.**

‘When you select ‘Choose a Fertility Clinic’ from the main menu you have to scroll right to the bottom of the page to find the button to start your search so its [sic] not immediately obvious; although it could just be that I am used to it being right at the top of the page on your current site though.’ **another member of clinic staff.**

‘...The user has to scroll down to the bottom of the page to click on the link, which is allows you to access the list of the clinics, which is too long to navigate’ **another member of clinic staff.**

**2.6.** However, others saw the benefit of the information users need to scroll through:

‘I think people tend to expect the link to be at the top of the page but the information you have to go past to get to the link is good’ **a nurse.**

**2.7.** It is worth noting that the design of the beta CaFC intentionally provided background information before the CaFC search page, so that those accessing the tool were properly informed before using it. There is also a link directly to the search page as a sub-menu item at the top of the page, though this may not be obvious to users, as noted by one user when asked for recommendations for improvements:

‘Navigation to "choose a Fertility Clinic" should move to the top of the page rather than the bottom’ **another member of clinic staff.**

### Clinic comparisons

**2.8.** One potential patient wanted to be able to compare clinics and search based on factors other than proximity or name, a point which was also made at the workshop:

'Want to reorder clinics by success rates' someone currently seeking fertility treatment.

**2.9.** This position was supported by suggestions for improvements given by patients and others at other points in the survey:

'To have a summary comparison page, comparing all uk [sic] clinics' **someone currently having fertility treatment.**

'To look for success rates information for clinics across the uk[sic]. This information [sic] was not available as the information is only available if you search for one clinic at a time.' **someone currently having fertility treatment.**

'Should create a page where patients can store the various pregnancy rates of different centres on the same page so that they can easily compare the different centres.' **a fertility doctor.**

'It doesn't seem possible to evaluate ALL the licensed centres. ... What if a prospective [sic] patient wanted to "Choose a Fertility Clinic" from the all the licensed centres [sic], based solely on inspector and patient ratings, and results?' **another member of clinic staff.**

### User testing

**2.10.** The user testing showed that users had a number of issues with the search function, many of which were the same as those highlighted in the survey. These included:

- Finding the CaFC search at the outset was not easy – some found the link too hard to get to
- Some users missed links to detailed pages from the search results
- The treatments list on the search listing is not exhaustive, and some users pointed out omissions
- Some users didn't understand certain parts of search pages eg, how the ratings are calculated

**2.11.** The user testing report shows that there are a number of improvements that can be made to CaFC search in order to make it more useable. A key consideration will be how different information is shown to patients to make them aware of what things mean and how to get to other more detailed information.

### Summary

- It is clear that the CaFC search is not yet easy enough to use.
- One clear reason seems to be that the search is more complicated to get to in the first place than the current website. We may want to consider how to address this.
- Patients seem to be looking for other facilities in the search tool than are currently available, such as clinic comparison by success rates.

---

## 3. **Headline information**

**3.1.** Individual clinic pages show three headline figures, which we think gives an overview of the clinic's performance in different key areas. These figures are:

- IVF birth rate (per embryo transferred), including all IVF treatments (which is currently includes ICSI, IVF, PGD/PGS and egg donation treatments; for IUI only clinics we show their IUI pregnancy rate)
- inspection rating (based on length of licence)
- patient rating (based on a survey).

**3.2.** We asked respondents whether they agreed that this headline information should be provided at the top of the clinic page, in the most prominent position.

---

**Do you think it's right to have this headline information at the top of the page?**



**163 respondents. 47 skipped this question**

**3.3.** The majority, more than three quarters of respondents, did not agree.

**3.4.** In hindsight, this question is ambiguous, since it could be interpreted as meaning one of two different things. On the one hand it could be read as 'is it the right *place* for the information', ie, should it be at the top of the page? Or it could be 'is it the *right information* to have in that place' ie, should this or something else be at the top of the page? We can call the first the concept of having headline information and the second the execution of that principle.

**3.5.** The text responses to the question give a better indication of the issues people had with the headline information. There were a number of responses agreeing with the *concept* of having some form of headline information:

'I think all this essential information is available clearly and quickly and is easy to understand.' **a nurse with a specialist interest in fertility and maternity services and women's health.**

'Really helpful to see top information up front' **a service manager.**

'These are the things patients want to know about a fertility clinic' **another member of clinic staff.**

- 3.6.** These responses may suggest that the principle of having headline data was not the problem for people; there appears to be agreement that these are areas that are very important for patients. But in the execution of the headlines there seems to be a problem; this could explain the largely negative response. A potential patient who agreed with the headline figures expressed the nub of the issue directly:

'Useful to compare across clinics – although is dangerous as top line statistics don't always tell the full story – e.g. age / health conditions of patients.' **someone currently seeking fertility treatment.**

- 3.7.** A statistician expressed similar concerns:

'In terms of clarity, transparency and ease of interpretation for patients, this strikes me as a considerable step backwards compared to the existing Choose a Fertility Clinic site. [...] by obfuscating the characteristics of patients/treatments that go into this overall figure, the headline statistic actually becomes more difficult to interpret. Clearly, as this very website points out on the preceding Choose a Clinic information page, overall figures may be very misleading due to the differences in characteristics of patients treated at each centre. It therefore seems bizarre that HFEA would make a change that actually exacerbates this problem.' **a statistician currently researching IVF outcome measures.**

- 3.8.** Some respondents proposed changes which also suggest that the problem was mainly with the way in which the success figure was calculated, rather than the principle of a headline measure:

'However, I do think that the birth or success rates need to be split in to age groups as on[sic] over all figure can give patients a false understanding on their chances.' **an admin manager of fertility clinic.**

'The birth rate is confusing. It lumps too many factors together. I appreciate simplicity is required but there is the potential for being very misleading depending upon the characteristics of each clinic. Combining data from different treatments / ages can lead to misleading data.' **a fertility doctor.**

- 3.9.** This is also the thrust of the argument from a clinic director who wrote to us criticising the concept of a single success headline figure that was not broken into different age groups (annex 1):

'HFEA should stop publishing a single headline figure per clinic. The continued transparency of clinic data reporting is essential. Clinic outcomes are not simple and dependent on a number of clinical, demographic, funding and financial factors. The HFEA has to accept and indeed promote that data publication and success rates is a complex matter. The duty of the HFEA is to educate both patients and the wider public of this and outline why a single figure is inappropriate.' **a clinic director.**

- 3.10.** Similarly, adding a breakdown for age to the headline and so avoiding having a single headline figure entirely was a suggestion made in a number of comments to improve this headline rate, this is discussed in more detail at section 5 of this chapter.

- 3.11.** There were no comments explicitly disagreeing with the other two headline figures.



### Workshop responses

- 3.12.** In the clinic workshop, clinic staff broadly agreed that having some form of headline figures at the top of the page was helpful and that more than anything, patients wanted high-level indicators of the quality of a clinic. Many saw the benefits of providing information other than just success rates, since that isn't the only measure of a clinic.
- 3.13.** There was a minority of attendees who felt that the concept of having any headline success rate at the top of the page was inappropriate. They suggested that stating whether a clinic was in line with the national average or not might be fairer to clinics and clearer for patients. A smaller group still said no to a single success headline.
- 3.14.** The majority view was that they would support having a single headline figure provided it enabled meaningful and fair comparison between clinics. There was clear agreement that grouping by all ages and treatment types to show a single success figure was inappropriate. These two issues are discussed directly later in this report.

### Other issues with the headline birth statistic

- 3.15.** In January 2015, following earlier consultation, the Authority agreed the policy of 'births per embryo transferred' as the primary IVF birth statistic. For this reason, the question of which measure we use was not asked again as part of this survey. Nonetheless, this was an issue that some wished to give their views on. These are noted below.
- 3.16.** A very small minority in the clinic workshop disagreed with using 'live births per embryo transferred' although the majority agreed with this measure. The main concern referenced was that it acted as a disincentive to replace the number of embryos that were clinically indicated.
- 3.17.** In response to the survey, a couple of respondents stated a preference towards births per egg collection as a headline figure, rather than the second level figure, as it is on the beta site.
- 3.18.** A detailed response to the survey by a statistician also raised concerns about live births per embryo transferred as a headline measure. The preference of this respondent was to present data by cycle started:
- 'presenting results that exclude patients who did not reach the transfer stage does [not] provide any indication of the chance of success to patients considering whether or not they should start IVF treatment. Success rates reported 'per transfer', 'per egg collection' or per 'embryo transferred' are likely to be higher than rates reported 'per cycle started', which includes patients who drop out in the earlier stages of treatment [...] results must be reported 'per cycle started' so as not to hide this fact from patients' **a statistician currently researching IVF outcome measures.**
- 3.19.** 'Births per embryo transferred' supports our policy to minimise the incidence of multiple births following IVF. The British Fertility Society repeated their support for this measure in their submission to us (annex 2). A detailed discussion of the benefit of this figure is also included in the submission from a clinic director (annex 1).
- 3.20.** In 2015, the Authority also agreed to make clear to users what the information is able to tell them. This includes explaining the reasons why we present certain measures. Some explanations have been provided for the measures, but it may be worth further considering these to address the concerns raised and clearly show the limitations of each measure.

### Summary

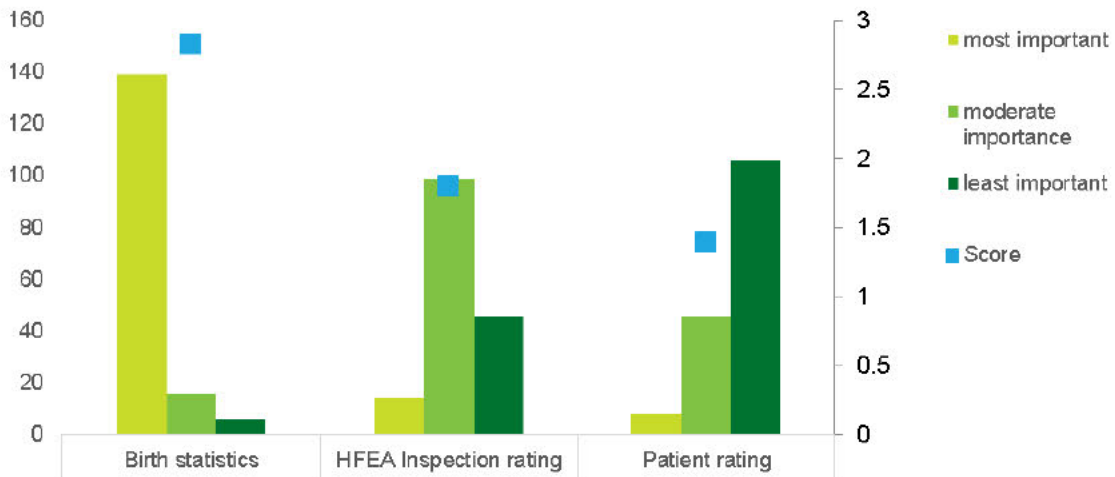
- Responses suggest that many people including clinicians, agree that the concept of having headline information is helpful, particularly to compare across clinics.
- The results suggest that many people do not think the way we calculate the headline IVF birth statistic for the beta site is right as it is not meaningful with so many factors aggregated. In the clinic workshop this was the consensus.
- To make the birth rate statistic more meaningful we could consider adjusting this for other factors.
- We should continue to make the limitations of statistics and particular measures clear to users.

## 4. Importance of headline measures

**4.1.** Although, as we have seen, respondents did not agree with the way we currently present the headline information, when asked to rank the headline figures there was broad consensus around what was more important. This can be seen in the graph below.

**Please rank these headline measures in order of importance to you, with one being the most important and three being the least important.<sup>1</sup>**

**Graph shows the number of people and the blue markers show the overall score of each measure**



**161 respondents. 49 skipped this question**

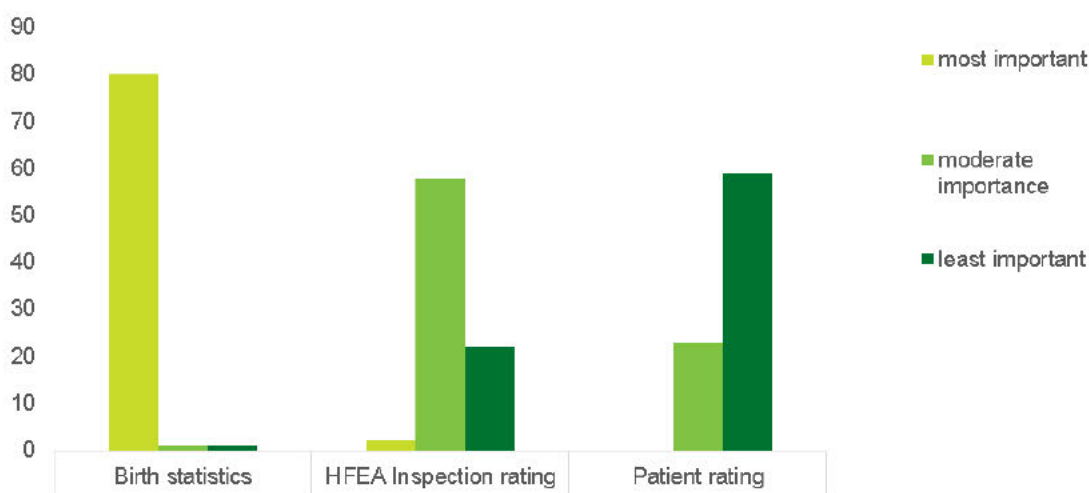
**4.2.** The score of each measure, calculated by weighting the ratings, shows that the most important headline was birth statistics. Over 85 per cent of those responding thought that figure was the most important of the three. Inspection rating was rated the second most important by over 60

<sup>1</sup> The wording of this question was clarified slightly on 25/08/16 at 13:40 pm from 'Please rank these headline measures in order of importance to you.' to 'Please rank these headline measures in order of importance to you, with one being the most important and three being the least important.' In order to provide more clarity for respondents.

per cent of responders, followed by patient rating, which about 65 per cent of people thought was the least important of the three ratings.

**4.3.** The responses from past, present and potential patients were more polarised than those of all responses grouped together, as can be seen in the graph below.

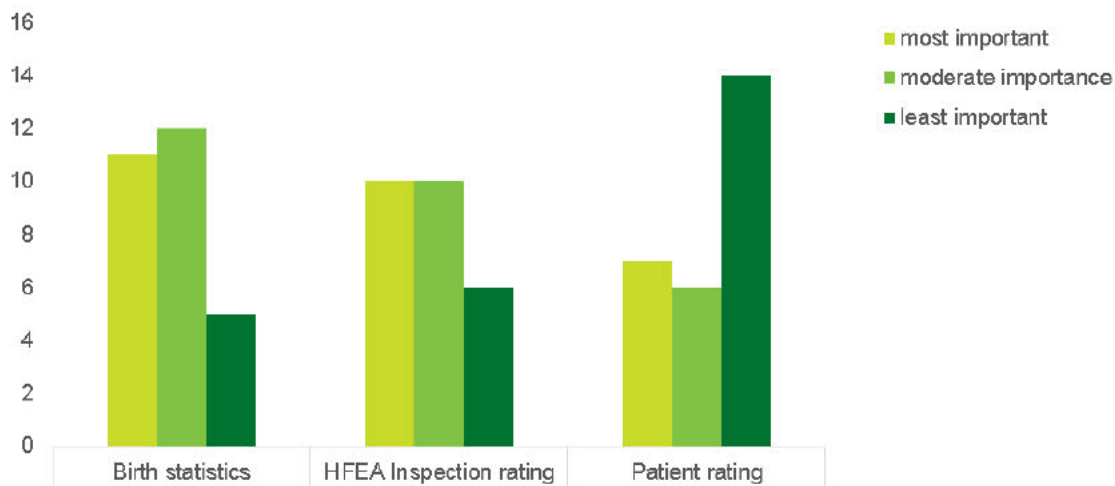
**Responses for past present and potential patients (excluding all clinic professionals, donor groups, 'others' and 'undisclosed'.)**



**82 respondents. 14 skipped this question**

**4.4.** Clinic staff responses were more balanced between the different measures, although the same trend emerges, as shown in the graph below.

**Responses for all clinic staff groups (excluding all patients, potential patients, donors, donor families, 'undisclosed' and 'others'.)**



28 respondents. 19 skipped this question

- 4.5.** This general consensus on the importance of the headline figures may support the finding that issues perceived with the birth statistic headline may have been a key motivator for those who disagreed with the headline figures.
- 4.6.** Although it is clear that for most the perceived importance of the measures was not equal, there is also evidence in other comments provided about CaFC that some strongly agree with presenting all three different headline figures, as they show distinct aspects of clinic effectiveness.

'Excellent to see patient feedback and really liked the advice given on choosing a clinic'  
**another member of clinic staff.**

'I think the consistent message to take into account all aspects of the clinic services and not just the data is excellent.'  
**a nurse with a specialist interest in fertility and maternity services and women's health.**

### Summary

- Most respondents believe that a birth statistic is the most important headline figure.
- Though it is clear that birth statistics information is very important, respondents agree that other figures should also be highlighted.

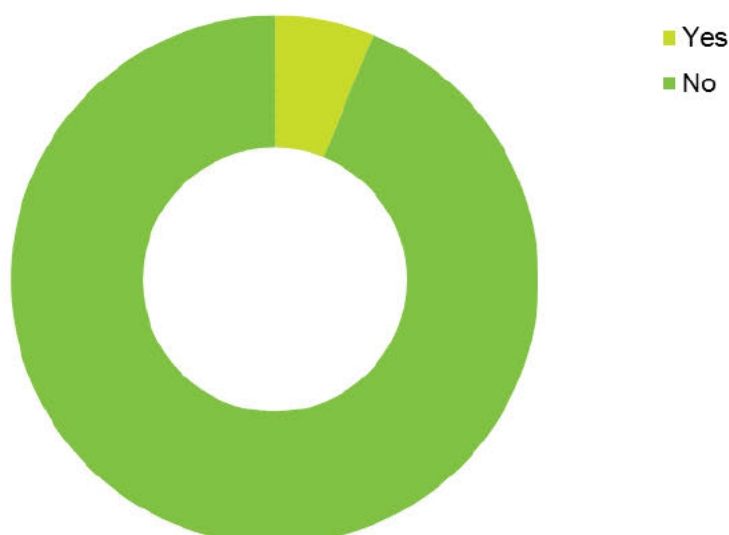
---

## 5. Grouping all ages

- 5.1.** In the beta version of Choose a Fertility Clinic, the headline success figure currently consists of all ages grouped together. One motivation for doing this was to provide a bigger sample size which could offer more meaningful information to patients.
- 5.2.** We asked people whether they thought doing this was right and requested reasons for their responses.

---

Is it right to group all ages together?



**164 respondents. 46 skipped this question**

**5.3.** Responses were overwhelmingly against grouping all ages together for the headline figure.

**5.4.** There were 26 comments explaining reasons for different replies. These responses fell into four broad themes:

- it is OK to do this or no strong view
- age is a key success factor so should be reflected in the headline
- it wouldn't be as meaningful to patients to group all ages
- it would be unfair to clinics to group all ages.

**5.5.** Most of the text responses came from clinic staff with only six coming from other respondent groups. The arguments are summarised below.

#### Grouping all ages is OK/ no strong view

**5.6.** Although it is clear from the overall responses, that this was a minority view, a few people did tell us that they thought it was reasonable for the headline figure to be for all ages, though only as long as detailed figures were available elsewhere by age:

'This is good for headlines, but would also be good to be able to filter search results by "people like me" **a service manager.**

'It is fine for an overall figure, as long as there is a breakdown of ages at a different section' **another member of clinic staff.**

'this will give an idea of the overall success of the clinics. As long as these figures are easily available by ages somewhere easily accessible.' **another member of clinic staff.**

'no strong view. makes it easy to compare but need to be clear whether IVF or IUI is being looked at.' **another member of clinic staff.**

**5.7.** In the beta CaFC users could see that these different age options are available elsewhere where clinic statistics are presented and age is not aggregated.

### Age as a key success factor

- 5.8.** Those who highlighted that age was really important and shouldn't be excluded as a factor in presenting the headline birth rate, focussed on the significant impact of age on success rates:

'Different ages have completely different statistics' **another member of clinic staff.**

'because success rates differ dramatically between different age groups eg 35 vs 39 yrs old' **another member of clinic staff.**

'Success rates depend so heavily on the patient's age that grouping them all together gives a misleading or inaccurate idea of a patient's chance of success. I think it would be more useful for patients if success rates were still broken down into age groups.' **another member of clinic staff.**

### Meaningfulness to patients

- 5.9.** A related view expressed in a number of responses was that when looking for clinics, patients want to be able to see figures that might reflect their own chances. They asserted that grouping all ages would not do this and could mislead patients.

'From our experience, patients very often request age and treatment specific statistics and therefore a headline figure may be of limited use for patients. [...]' **another member of clinic staff.**

'I think this should be split up into smaller age brackets to give women a true understanding of the average chances of a live birth based on their age. I think this displays the clinics in a better light because it adds clarity and the patient may be able to better understand the figures.' **a trainee embryologist.**

'No, the headline needs to be split by age as it is more meaningful for patients who want to know results specific to their age range and once again centres treating a higher age range are not misrepresented.' **another member of clinic staff.**

### Fairness to clinics

- 5.10.** The final argument of those who disagreed with grouping all ages in the headline figure was that it is unfair to clinics, as different clinics have different mixes of patients. Clinics who treat a higher proportion of patients over 38 could be disadvantaged as their success rate would naturally be lower due to their patient mix.

'Some clinics only deal with over 40's this will impact on their business' **another member of clinic staff.**

'In relative small unit with relative low number of cycles, the grouping of older patients with younger ones may distort the true success rates of the unit and its quality.' **a fertility doctor.**

- 5.11.** It is worth noting here that the issue of ages affecting clinic statistics may be more likely to affect clinics treating private patients, since within the NHS women over a certain age may not be considered for treatment.

## Workshop feedback

- 5.12.** At the clinic workshop, all of the arguments above were touched on in discussion. Clinic staff agreed that the current headline figure, grouping all ages is oversimplified. They suggested a number of different options for adjusting the headline success rate for age by either:
- Presenting the success rate for one, more comparable age bracket, such as under 35 year olds
  - Presenting the success rate for a 'gold standard' patient (eg, patients meeting the same categories for age, fertility etc. based on no. eggs collected)
  - Presenting a number of headline figures for different age brackets to be more meaningful to patients - perhaps either showing all 6 age brackets or alternatively 3 categories.
- 5.13.** Attendees noted that each of these approaches will have pros and cons but they felt that they would lead to a fairer comparison which was more meaningful for patients.
- 5.14.** One additional issue discussed was the impact of using a smaller sample for the success figure and the fact that this could make rates for smaller clinics less meaningful. Some attendees noted that it would not be appropriate to let the impact on small clinics sway the whole presentation of CaFC especially because most clinics would undertake enough treatments to allow for age breakdowns. An option suggested to address this was an explanatory note to be presented for small centres with less than 50 treatments in a given age group.

## Summary

- Most believe that all ages should not be aggregated for the headline figure.
- The clinic workshop highlighted three main alternative approaches we could consider:
  - Presenting the success rate for one, more comparable age bracket
  - Presenting the success rate for a 'gold standard' patient (eg, patients meeting the same categories for age, fertility etc. based on no. eggs collected)
  - Presenting a number of headline figures for different age brackets

# 5. Using Choose a Fertility Clinic: birth rates

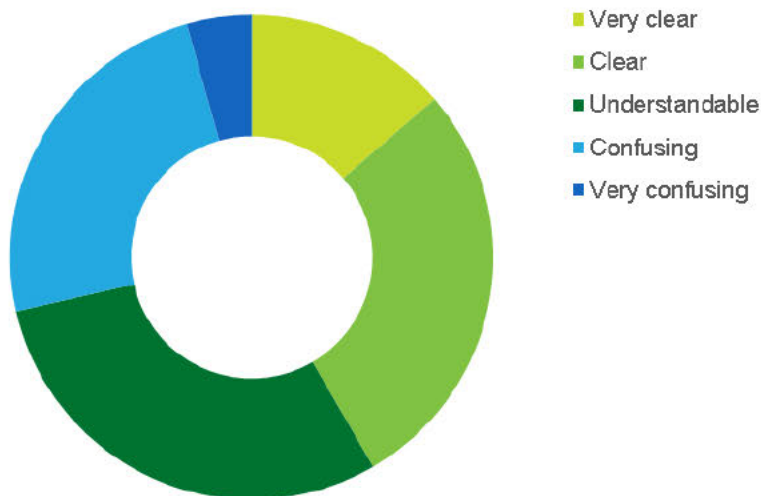
## 1. Introduction

- 1.1. After asking questions about the high level, headline data, we asked more detailed questions on what people thought about the clinic pages themselves and particularly about the presentation of clinic birth statistic data.
- 1.2. The data used in beta CaFC site is largely the same as the current CaFC dataset but the site presents clinic data in significantly different ways to the current site.
- 1.3. In the beta site we present the headline data, which is then supplemented further down the clinic page with data divided into two age brackets. There is then a separate detailed statistics page where the user can specify the age, treatment type, source of eggs and time period they would like to see.

## 2. Understandability of birth statistic presentation

- 2.1. We wanted to hear how easy users found the beta presentation of birth statistics to understand, so we asked a general question about this.

Generally speaking, how understandable was the presentation of the clinic’s birth statistics?



161 respondents. 49 skipped this question



- 2.2.** It is clear that the majority of people gave positive answers and found the statistics ‘understandable’, ‘clear’ or ‘very clear’. However, 30 per cent of respondents found the clinic birth statistics to be ‘confusing’ or ‘very confusing’.
- 2.3.** Clinic staff responses were slightly more likely to be critical than the overall responses, though a greater proportion also found the statistics very clear.
- 2.4.** For past, present and potential patients, the level of understanding was generally a little lower again, with more than 40 per cent either finding the statistics ‘confusing’ or ‘very confusing’.
- 2.5.** Although the patient group results were slightly lower, the difference between them and the overall responses or clinic ones is not significant. This may demonstrate that there is not a barrier to understanding that is particular to patients. The same barriers may affect everyone.
- 2.6.** These responses suggest that improvements are needed to these pages to make them more understandable. The text responses to this question give some indications of suggestions that might help in doing this, however, it is notable that very few past, present or potential patients, donors or donor conceived families chose to provide additional explanations for their answers.
- 2.7.** The two comments from these groups give a small indication of what might help them understand:
- ‘Explanatory text is useful’ someone currently seeking fertility treatment.*
- ‘Reliability range?’ someone currently having fertility treatment.*
- 2.8.** These comments suggest that being very transparent about what is being displayed and providing clear explanations may increase user understanding. The reliability range is examined in detail in the next section.
- 2.9.** When answering this question, clinic staff gave a range of comments, both positive and negative. These are summarised below.
- Positive responses**
- 2.10.** Many responses praised how clear the presentation was. People generally liked the demonstration of the national average, as well as the explanations of various components. The new graph style also received a positive mention:
- ‘I think it is extremely helpful to have the national average so clearly presented. I like the consistent reminder that small variations in the statistics are not significant.’ a nurse with a specialist interest in fertility and maternity services and women’s health.*
- ‘Showing the clinic rates and national average as a graph is much clearer and easier to understand at a glance than all of the tables which the data is currently presented in.’ another member of clinic staff.*
- ‘Each element is explained and defined well.’ a trainee embryologist who is also considering oocyte freezing.*
- 2.11.** These comments suggest that putting the clinic statistics in the context of the national average and their significance has helped increase understanding.

## Areas for improvement

- 2.12.** However, others did not feel the site went far enough to make the information understandable. This particularly centred around the type of graphs used which some did not think were simple enough:

*'This way of presenting data is good for people who are statistics savvy, but may not be for everyone Tom, Dick & Jerry'* **a fertility doctor.**

*'It is not easy to understand immediately, especially for patients not used to looking at graphs etc as this is not a standard graph like a bar graph'* **another member of clinic staff.**

*'Didn't understand Concept of lines first time I looked at it'* **another member of clinic staff.**

- 2.13.** Two other respondents, reporting concerns with CaFC under a different question noted similar thoughts about explaining terms and clarity of information, and gave some ideas for how this could be made clearer:

*'The statistics pages about each clinic are great - really helpful visualisations, but there are a lot of terms in there presented without explanation or glossarisation [sic].*

*Would be helpful if I was trying to choose a clinic to have hyperlinks or tooltips which explain each technical term for me like ""Proportion of all embryos transfers that were blastocysts"" - no idea if this is good or bad!*

*As this is for the public too, personally you might be better saying ""Treatments resulting in the birth of one baby"" or ""Treatments resulting in the birth of two or more babies (twins or triplets)"" rather than ""singleton births"" and ""multiple births"". It would make it more accessible to the general user' **a service manager.***

*'Need to explain more about the meaning of each type of success rates. For instance, pregnancy rate per embryo transferred would indicate a) the quality of the laboratory service and therefore exclude units that transfer high number of embryos in order to artificially elevating their pregnancy rate, b) the adherence of the unit to eSET policy and reduction in risks of IVF.*

*Also need to clear to patients that the cumulative pregnancy rates per egg collection relate to treatments that were carried out 4 years ago and therefore do not reflect current success rates, but give idea of how good the unit is with respect to both fresh cycles and frozen cycles.'* **a fertility doctor.**

## User testing

- 2.14.** The user testing with patients showed that most users did not instantly understand the graphs used to present statistics. Most users missed the explanatory text associated with the graphs. There were also features of the graphs and statistics that some users were particularly unclear about such as the 'national rate' and 'reliability range'.
- 2.15.** It is clear that elements of the presentation need to be clearer to help users to understand features of the statistics and to make them more meaningful.

## Summary

- The majority of users found the birth statistics understandable, though a significant minority found understanding them hard.
- Explanations available with the statistics seem to be widely appreciated and we could consider providing more of these to improve user understanding.
- We could consider making it clearer to users exactly what the measures can indicate to them.
- Because the graphs are non-standard, we may want to consider whether users need more guidance to familiarise themselves with the new presentation.

## 3. The reliability range

- 3.1.** The reliability range has been designed to show how confident we are that a clinic will repeat its success rate in the future – the narrower the range, the more confident we can be.
- 3.2.** Large clinics normally have a narrower reliability range because their rate is less likely to be affected by small changes in the number of births in one year. Small clinics aren't worse but their success rate is more likely to be affected by these kinds of changes.
- 3.3.** On the current CaFC site we describe this as the 'predicted chance'. We believe that providing this range is important, since it highlights the impact of small sample sizes, however we know from user testing the current site that patients find the concept of 'predicted chance' confusing.
- 3.4.** We asked respondents about whether the range made sense to them.

**We show something called a reliability range for each rate. Did that make sense to you?**



161 respondents. 49 skipped this question

- 3.5.** A significant majority, more than 85 per cent, could not understand the reliability range. More telling still is the result when the answers of clinic staff, fertility doctors, fertility counsellors and 'others' are removed, as in the graph below.

**Responses of patient and donor groups (excluding fertility doctors, fertility counsellors, other clinic staff 'others' and 'undisclosed'.)**



126 respondents. 15 skipped this question

- 3.6.** Only one person from any of the patient or donor groups understood the reliability range. A lack of understanding of the reliability range among patients and the public is something that the clinicians themselves predicted when they told us:

'It makes sense to us professionals but it might be a bit more difficult to understand for the public' **another member of clinic staff.**

'I do [know] what this data is but the explanation needs to [be] clearer or[sic] patients and general public' **another member of clinic staff.**

- 3.7.** Some suggestions for how to make the information clearer were provided by respondents. One suggestion was adding an indication of the number of cycles to the headline figure. This is actually available however, which suggests that it may not be prominent enough:

'...I think it may be useful to give an idea of the average number of cycles per clinic in the relevant areas so that patient's [sic] can consider the reliability range with actual figures.' **Another member of clinic staff**

'No data [on] how many cycles have been completed I [sic] total, therefore its [sic] not very clear.' **Another member of clinic staff**

- 3.8.** Of the other respondents, a statistician expressed other technical concerns. One of these was around the intention of the reliability range and how it should be presented:

'It is unclear what [the reliability range] actually is; it may be a confidence interval or a prediction interval (the text indicates that the range can be used as a guide to how the clinic will perform in the future – this is reasonable if the range is calculated as the latter but not if it is calculated as the former).' **a statistician currently researching IVF outcome measures.**

**3.9.** It is clear that this respondent felt the intention of the range could be better explained, although the explanation is designed for a less expert audience, so may not touch on the statistical detail that this expert wanted. The same statistician was also concerned that the correct interval should be used to effectively present the HFEA confidence in the assessment of a clinic's performance.

#### Summary

- The current format of the reliability range is not generally understood by those it is intended to inform.
- We may need to further consider the purpose and presentation of this range.
- The meaning of the range may need to be explained more clearly.
- Going further to highlight the number of cycles that data is based on may provide a clearer indication of the reliability of the data.
- We may also wish to consider providing information on CaFC statistical methodology for a more expert audience.

## 4. Division of data by under 38 and 38 and above

**4.1.** Underneath the headline success rate we split the data into two broad age categories. This was done to give patients more relevant information whilst keeping the presentation simple.

**4.2.** We chose age 38 as the cut off because the success rate is significantly lower after this age. Data split by six age categories can be found on the detailed statistics page.

**Do you think we have got the right balance of age detail between this page and the detailed statistics page?**



159 respondents. 51 skipped this question

**4.3.** Over 85 per cent of respondents disagreed with the balance of detail between the detailed statistics page and the first level of detail where data is broken down into two age groups, with 38 as the dividing point for the two.

**4.4.** However, comments from the survey do not present a coherent reason for this. Some comments actually supported this age cut off and the balance between the high level and detail:

‘Yes, agreed 38 as the cut off because the rate is significantly lower after this age’ **another member of clinic staff.**

‘Yes, the main page is basic and gives a brief outline. The detailed statistics page is still clear and the layout is appropriate. A significant amount of information is listed here and I think it is user friendly.’ **another member of clinic staff.**

‘I feel there is adequate information given’ **a fertility counsellor.**

‘Probably but need to make it clearer that patients can drill down further for more detail.’ **another member of clinic staff.**

**4.5.** Some comments suggested various alternative ages to use as a cut off:

‘should be below 36 and above’ **another member of clinic staff.**

‘The cut off should be below 35, 35 - 39 and 40 or above.’ **a fertility doctor.**

‘may be better for patients to link with the NICE age cut off guidance so that NHS funding might make more sense’ **another member of clinic staff.**

**4.6.** Meanwhile, some gave other suggestions indicating that more detail was desired, although it is not clear if this was because they wanted more at an earlier stage (before the detailed statistics) instead of the two groups:

‘Good to show in detail’ another member of clinic staff.

‘Other factors such as low amh not considered which are very key to someone looking into an IVF cycle.’ **someone currently having fertility treatment.**

**4.7.** This final comment about AMH is not applicable to this discussion however, since we do not collect information on patient AMH levels.

**4.8.** The detailed response from the clinic director (annex 1) outlines stronger arguments against using this two age breakdown, stating that this could be overly reassuring for those who are at the upper limit of the lower category and could be misleadingly positive for those in the older category where individual chances of pregnancy vary greatly. In brief:

‘Published data for outcomes by the current 6-tier age bands (adopted by the HFEA for more than a decade) make it abundantly clear how outcome drastically changes from band to band even within the <38 or >38 group and as such the simplification to 2 simple bands is therefore misleading’ **a clinic director.**

### Workshop feedback

**4.9.** The discussion on this topic at the clinic workshop was more illuminating about the reasons some didn’t like the two groups with the dividing age of 38.

- 4.10.** Though some could see why 38 was chosen as the cut off, some clinic staff felt that this was a fairly arbitrary figure and some asked whether it had been informed by user testing. Some people noted that patient prognoses could be highly variable in the above 38 category, so this grouping wouldn't be very helpful.
- 4.11.** Others raised concerns that the over 38 figure could give false reassurance to older patients who could think their chance of success was higher than it really was.
- 4.12.** There was discussion of other age division options, overall, some felt that the more detailed groupings were better and that two sub-categories should not be used at all, since they would always be too broad, especially when it came to the chances of those in the higher age category.

### Summary

- The two age breakdown at 38 was not a popular way of presenting high level birth statistics, some felt that this could mislead patients.
- An option to consider could be providing high-level statistics covering more age categories and doing away with the two group split further down the clinic page. Another could be providing more of a breakdown at the second level where we currently show just the two ages.

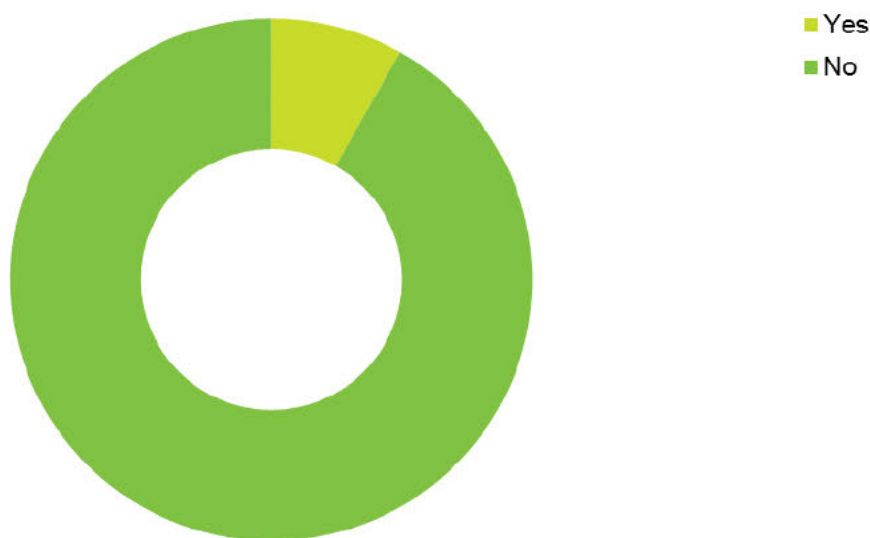
---

## 5. Division of data by treatment type

- 5.1.** For the headline figure on the beta site, all IVF treatment types (IVF/ICSI/egg donation/PGS/PGD/fresh/frozen and natural cycle) have been grouped because the figure is showing the success rate once you have an embryo (births per embryo transferred).
- 5.2.** We asked people if they agreed that was right.



Because we use births per embryo transferred as one of our three headline measures, we don't think it's relevant to separate the different treatment types. That's because, once you have an embryo ready for transfer, how it was created is less important. So, we have included IVF, ICSI, PGS and PGD. Do you think it's right to group treatment types together in this way?



159

respondents. 51 skipped this question

- 5.3.** Over 90 per cent of those who responded disagreed with having a headline figure including all IVF treatment types.
- 5.4.** Some in the minority who agreed with the proposal commented and stated that they agreed that it didn't matter how the embryo was created:
- 'The chance of an embryo implanting is not hugely affected by how it was created - I think?' a nurse.
- 'Agree with your reasoning' another member of clinic staff.
- 'I agree that it is not relevant to separate the different treatment types' a nurse.
- 5.5.** Among the responses of the majority there were several particular treatment types that many people thought should not be included in the headline figure. These suggestions for exclusion were PGD, PGS, natural cycle, and egg donation.
- 'Egg Donation, Natural IVF and PGS/PGD has to be kept separate' a clinic admin manager.
- 'We think that the most helpful way to present the data would be as follows:  
IVF/ICSI fresh and frozen transfers within a two-year period

Egg donation, PGS, PGD, surrogacy and natural cycles should be presented separately, as they may be artificially high or low which could skew the data and be misleading for patients' **another member of clinic staff**.

**5.6.** The possible reasons given for excluding each technique are summarised below where these were provided in responses:

#### PGS/PGD

**5.7.** Those who thought PGD and PGS should be excluded said:

- these patients usually don't have a fertility problem.
- centres performing a higher proportion of these treatments will have better statistics.
- conversely, some felt that embryos from PGS and PGD are associated with lower success rates.
- clinicians have genetic information about PGS and PGD embryos and not for IVF/ICSI embryos. It is possible that PGS/PGD embryos could be more likely to result in a live birth due to the known genetic stability.

#### Natural cycle IVF

**5.8.** The arguments for excluding natural cycle IVF were:

- The complexity of the natural cycle issue is highlighted in the submission by the BFS (annex 2) which stated it is 'a little vexed' by this issue. It highlights that the term 'natural cycle' is unclear since different amounts of medication are provided to patients. In essence it suggests that the natural group should not be included:  
  
'Whilst it is difficult to argue against implantation rate in a defined group including the natural group [in] this may hide a multitude of sins.'
- one person felt that centres performing a higher number of these cycles will be negatively represented.
- A detailed response was provided by a clinic addressing the issue of natural cycle IVF in detail, (annex 3). This submission argues that natural IVF and conventional IVF are:  
  
'two fundamentally different treatment methods and are offered to different patient populations.'
- Consequently, the authors believe natural cycles should not be included with standard IVF in the same headline statistic but should be shown separately.

**5.9.** The arguments against excluding natural cycle IVF were:

- Conversely, the clinic director's response (annex 1) expresses a view that both natural and standard IVF should be grouped together in the headline success rate:  
  
'every egg collection performed should be included in the denominator when results are expressed as LB\_EC. This is true whether the cycle was unstimulated or stimulated or with batching embryos.'
- The respondent agrees with the view that the HFEA should highlight the different success rates for natural and stimulated cycles in the detailed statistics.

### Other factors raised

- 5.10.** Freeze-all cycles were seen by one commentator as not being properly presented to patients with this headline metric:

‘What about freeze all embryo cycles done to achieve 1 pregnancy. Patients should be able to evaluate the data based on how much money they will need spend to achieve the pregnancy, therefore this data doesn't help with the money side.’ **another member of clinic staff.**

- 5.11.** Some other commenters were not keen to amalgamate IVF and ICSI success rates:

‘It is still useful to present figures on percentage [sic] of ICSI as some units do unnecessary ICSI’ **a fertility doctor.**

‘People want to know what is the difference in success rates between IVF/ICSI or if they are having a specific treatment they want to know what are their chances using that intervention’ **another member of clinic staff.**

- 5.12.** We should note that the data for IVF and ICSI success figures are available separately on the detailed statistics page so that they could be compared at that point.

### Workshop feedback

- 5.13.** The clinic workshop discussed a number of issues with grouping all treatment types and as noted earlier, clinic staff universally disagreed with combining different treatments in the headline figure. The attendees reached the following conclusions:
- 5.14.** The headline figure would be more meaningful if it included only fresh stimulated IVF/ICSI cycles with the patient's own eggs.
- 5.15.** Because these treatments are fundamentally different, it should exclude:
- Egg recipients
  - Frozen (though there was debate around this issue and some felt the first frozen transfer should be included for freeze-all cycles and some were more relaxed on this)
  - Surrogacy
  - PGD/PGS
- 5.16.** They also discussed natural cycle IVF and strong and differing views were expressed about how this information should be presented. The overall feeling was that we should consider excluding natural cycle IVF from the headline figure since this is a different treatment type. Some suggested that to include it would mean not comparing like with like. Although a minority of clinicians disagreed with this and thought it should continue to be included in the headline.
- 5.17.** The majority in the workshop felt natural cycle results should be presented separately and clearly in the detailed statistics to facilitate informed patient choice.
- 5.18.** They also recommended that we should make national success data available for different treatment types before patients see the clinic statistics.
- 5.19.** This last point was also previously recommended by the Authority in January 2015.

### Summary

- Most people disagreed with grouping all treatment types together.
- There was general agreement for grouping standard fresh IVF and ICSI and excluding other treatment types.
- However, there was more debate around how to deal with natural cycle and frozen transfers. We may want to consider these issues further.
- Different treatment types should be clearly displayed in the detailed statistics.
- We should provide national data for each treatment type as this would make it clearer to patients what the average success rate was before they get into the detail for individual clinics.

## 6. Detailed statistics – is the balance right?

- 6.1. The first clinic page includes high-level statistics, but the majority of the available statistics are included on the detailed statistics page for each clinic.
- 6.2. We asked whether the first and detailed clinic statistic pages had the correct balance of information.

**If you had a look at the detailed statistics page, do you think overall we've got the right balance of information about statistics between the first clinic page and the detailed statistics page?**



**157 respondents. 53 skipped this question**

- 6.3. Of those who responded, over 85 per cent disagreed with the balance between the statistics available on the first clinic page and the detailed statistics page.

- 6.4.** Unfortunately, the comments do not give a coherent overview of why people did not think the balance was right, since we only received nine. Many of these comments were actually positive, noting the information was clear and helpful to patients and that it went into sufficient detail. Some also just noted the person had no additional comment.
- 6.5.** One person gave a particular pointer that they wanted live birth per embryo transferred on the first page, although had not realised that this is actually already there. This suggests some users may not be able to find information in the beta presentation.
- 6.6.** Other comments on the detailed statistics pages and on CaFC as a whole reiterated that some people were most concerned that different treatment types should be separately presented, perhaps on the main clinic page rather than just on the detailed page:

‘Different treatments should be separated’ **another member of clinic staff.**

‘[patients should be able to see data] for cycles where no ET have been performed’  
**another member of clinic staff.**

### Workshop responses

- 6.7.** The workshop attendees thought that the detail available was valuable, particularly for displaying the success rates for different treatments, but some people felt that many users of CaFC may never access this detailed level of information.
- 6.8.** In such cases the headline and first page information is particularly important. They therefore recommended that the detail should be as accessible as possible to encourage its use and the higher level data to be as meaningful as possible for users.

### User testing

- 6.9.** The user testing showed that some users were happy with the level of detail provided and others felt that they didn’t need it.
- 6.10.** Although the detailed statistics were too much for some people, others thought it was very good, with one even commenting that this was the only data that really mattered as it meant she could access data for people in a much more similar situation to her own.

### Summary

- The survey responses provide little clear reasoning for the respondents’ dislike of the balance between the detailed statistics and the main page.
- Clinic workshop attendees felt that many users may not access the more detailed statistics page. This raises the question of how much information should be provided on the main clinic page - perhaps the main clinic page is not providing enough information for users in its current format.

## 7. Overall opinions - Likelihood to recommend

**7.1.** We asked those responding to the survey to give us an idea of their overall opinion of the site by providing a score for how likely it was for them to recommend it to others. The results are displayed in the table below:

Would you recommend this site to a friend, relative or colleague requiring similar information?

Not at all likely	1	2	3	4	Neither 5	6	7	8	9	Extremely likely 10
4	2	6	3	15	26	42	32	26	8	10

174 responded, 36 skipped this question

**7.2.** On the 11-point scale (including 'not at all likely') the weighted average rating was 6.14.

**7.3.** The likelihood of recommending did vary very slightly for different groups and average ratings of different groups is summarised in the table below:

Average recommendation ratings for different groups

Group:	All patient groups	All responses	All clinic groups	Donor groups
Average rating	5.97	6.14	6.16	6.18
Number of responses this is based upon	88	174	32	45

**7.4.** It is clear that opinions on the website are broadly similar for all users and the average rating has limited significance since it is based on only a few responses in each category.

**7.5.** These figures are not especially valuable for analysis in isolation, but they will form a baseline against which the website could be measured in the future.

**7.6.** Other general feedback comments are summarised below.

### Patient views

**7.7.** As one clinic staff member added, the patient view of improvements to the service are very important:

'It is really for the patient and their understanding and information need, therefore the most important feedback would be from the patients themselves.' **admin manger of fertility clinic.**

**7.8.** Unfortunately, again, there were not many patient comments regarding recommendations on the site. Those that there were did not necessarily give feedback on what could be done better, as in:

'I found this site hard to understand' someone who previously had successful fertility treatment.

**7.9.** Meanwhile, some gave very particular comments which we can consider about the content and presentation of the site:

'I note you have no advice for single men. While this may be a small demographic, this seems an obvious oversight.' **someone who previously had successful fertility treatment.**

'the original site has links to enable downloading the forms to sign, and this was the greatest help to us from the whole site.' **someone currently having fertility treatment.**

'The colours you combine and use when text is written inside box (like the orange box) in the section where you speak about same sex couples and options are a bit annoying for the eye.' **someone who previously had successful fertility treatment.**

### User testing

**7.10.** The user testing was another source of qualitative information and showed that:

- The website experience is very well aligned with user expectations and needs, and participants were unanimous that it is a huge improvement on the current version.
- CaFC was highly praised, although there are still usability improvements that can be made to the search, search results and clinic detail pages.

### Clinic staff views

**7.11.** Clinic staff acknowledged there was work to do and gave specific comments for improvements on additional features than those discussed already in this paper, particularly on content issues:

- two counsellors noted that counsellor details were not available under the staff heading on clinic pages
- one person thought the term 'research' a clinic should be more carefully used since patients will be spoken to about scientific research later
- one person felt we should provide more up to date figures
- someone expressed dislike of having to scroll lots.

**7.12.** Positive responses covered a number of aspects:

- clarity and quality of information
- ease of navigation
- helpful presentation – particularly graphics and flow charts
- user friendliness
- providing the right points for choosing a clinic
- an improvement on the current site.

### Summary

- We now have a baseline figure for how likely people are to recommend the site. They are somewhat likely to recommend it. Ongoing measurement of this will be helpful for improving and developing the website once it goes live

- Positive feedback is helpful, but the particular constructive critical comments provided can be reviewed as we consider what changes we may wish to make before the new site goes live and assist in improving the service.



---

## Annex 1

### Submission from a clinic director in response to the HFEA beta website survey.

Received via email on 28 September 2016

Dear Sirs,

The objective of fertility treatment is to achieve a healthy live birth event preferably a singleton full term birth. This is universally considered by all professional and advisory bodies as the optimal outcome of all forms of fertility treatment. On occasion, clinicians transfer more than one embryo because they are not sure which of them has the best chance to implant. When clinics are judged only on the basis of the fresh cycle, the pressure to transfer more than one has often been irresistible. The thinking has historically been “After all, buying two lottery tickets will improve your odds rather than only one”. The price for that is a risk of multiple birth around 30% with all its associated complications for mother and child.

The use of “Live Birth per embryo transferred” (LB\_Emb) is not new as the HFEA has been reporting this since 2010 when the latest format of data publication was launched. The only difference is that it was published in a table that comes second to the table reporting live birth per cycle. The new publication will give priority to this measurement.

The human embryo is the final product of all IVF laboratories. It is the competence and quality of these embryos that what will determine the pregnancy and live birth rate. LB\_Emb could therefore be seen as the measure that assesses the quality of what we produce to achieve what we want (a live birth event). So whether this event is singleton or twin it will be counted as one. The denominator, however, will be the number of embryos transferred.

The LB\_Emb provides a headline rate that corrects for under reported cycles started, differences in embryo transfer policy and has a commendable public health advantage of encouraging the transfer of the lowest number of embryos that would be effective in achieving a pregnancy, thus reducing multiple pregnancy rate. The move to this measure should, therefore be celebrated by all.

Using this measure (LB\_Emb), Abdalla et al 2010, found that the apparent significantly higher live birth rate per fresh cycle in the USA (37.6% with multiple rate of 36% of which 4.4% were triplets) compared to the UK (29.6% of which 30% multiple and 0.2% were triplets) was almost solely explained by the embryo transfer policy (See Table below). There was a significantly higher number of embryos transferred in the USA (an average of 2.4 compared to 1.9 in the UK). The LB\_Emb however, was very similar (18.1% and 17.6% respectively).

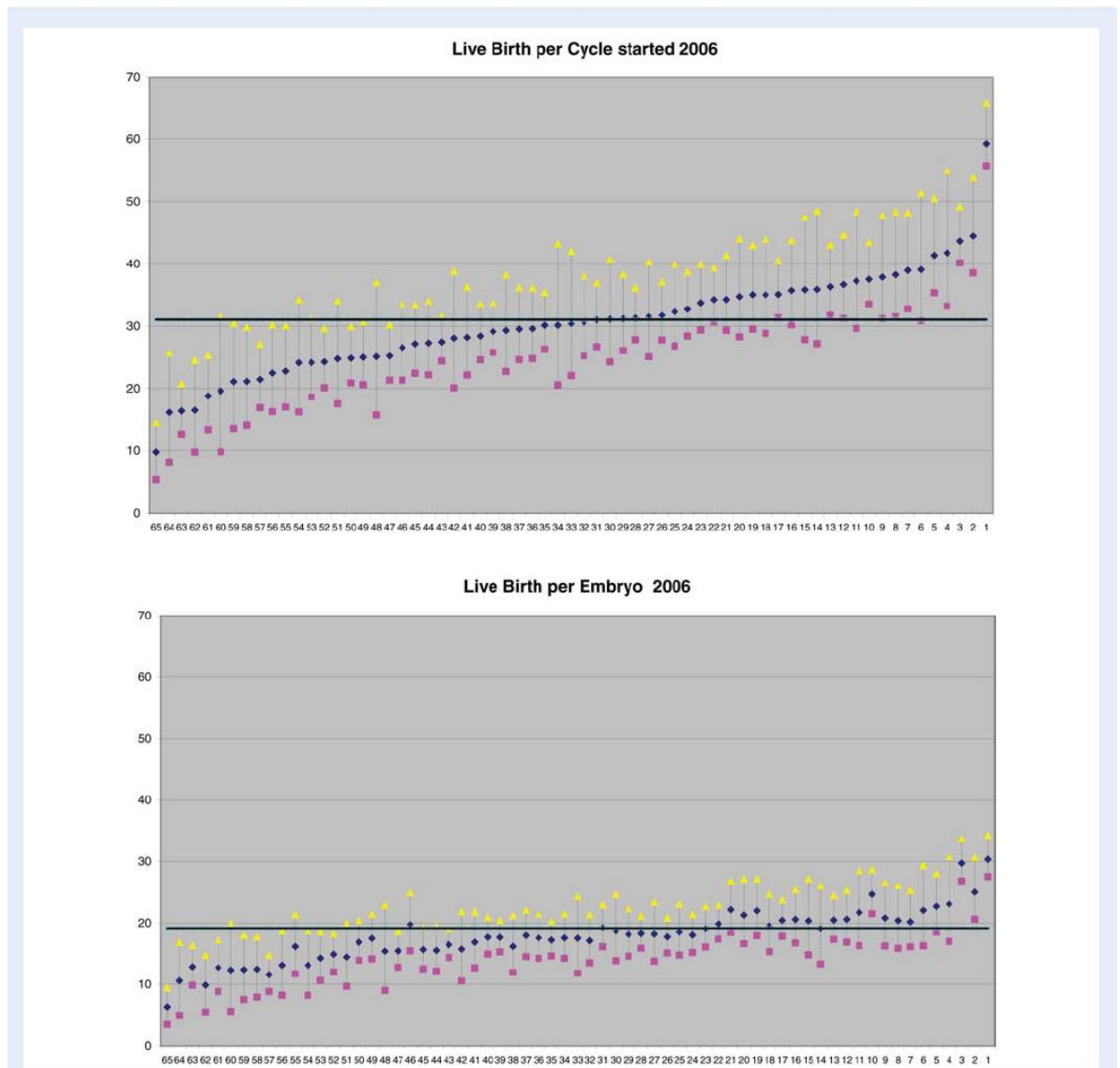
**Table II Comparison between cycles performed in USA\* and UK† (2005) for women under the age of 35 years (Abdalla, 2009)**

	USA	UK
Cycles Started	41 302	13 947
Live birth	15 406	4131
Live birth per cycle (%)	37.3	29.6
Multiple rate (%)	36	30
Triplets or more (%)	4.4	0.2
Cycles reaching embryo transfer	35 497	12 135
Average number of embryos transferred	2.4	1.93
Total number of embryos transferred	85 193	23 413
LB_Emb (%)	18.1	17.6
Emb_LB	5.5	5.7

\*CDC, 2005.

†<http://www.hfea.gov.uk>.

Indeed, in the same paper, the apparent differences between clinic within the UK disappeared using this measure. (see graph below)



**Figure 2** (A) Ranking of clinics based on live birth per cycle started for women under 35 years of age as taken from The Human Fertilisation and Embryology Authority's find a clinic (<http://www.hfea.gov.uk/>) with confidence interval calculated for 2006 data. (B) Results for live birth rate per embryo (with corresponding confidence interval) for the same patient group appearing in the same order of that seen in (A). Solid line represents national average.

As early as 2004, Thurin et al in a randomized multicenter trial in the *New England Journal of Medicine* demonstrated that the transfer of 2 fresh embryos had a significantly higher LBR (43%) compared to an elective fresh single embryo transfer (30%). However, when a single frozen embryo was transferred into those that did not get pregnant in the eSET group the LBR became 39% which is no longer significant. The multiple birth rate in the double transfer group was significantly higher (33% vs 1% in the single transfer group). Kalu et al in 2008 (*BJOG*. 2008 Aug;115(9):1143-50) showed that the LBR following elective single blastocyst transfer and additional frozen transfer in those that did not achieve a LBR in the fresh cycle was equivalent at 68% compared to those with double blastocyst transfer of 69%. The multiple birth rate however was 5% and 46% respectively.

Given the above data, if we are comparing two clinics; Clinic A transfers the best two embryos in the fresh cycle and Clinic B transfers one and freeze's the other. In looking

at the fresh results (as is the case with the current data reporting) Clinic A will look significantly better than clinic B, but with a multiple birth rate of 30 - 50% dependent on age and embryo quality. When Clinic B transfers the other embryo a month or two later it may achieve similar results but with negligible multiple birth rate. The introduction of Live birth per collection (LB\_EC) as an additional frontline measure negates this effect as clearly in this example there is no difference in the quality of either clinic. The difference in success rate is totally dependent on the embryo transfer policy of either clinic and the way the data is reported. (We have a problem with the time span for the data used – see later)

The use of LB-Emb and LB\_EC will play some role in minimising the effect of embryo transfer policies and the under reporting of started cycles and will certainly improve the safety and health of babies born following IVF through the reduction in multiple birth rate. However, for such changes to truly benefit both patient and clinic, the published information must provide clarity to all and avoid any potential to mislead.

It is important, however, to understand that the use of these measures or any other will never be sufficient to prevent clinics from playing the system. The other way is how to treat patients with reduced ovarian reserve.

#### **Treating mainly patients with good ovarian reserve**

For the same age group, a patient with good ovarian reserve is bound to do better than her counter part of similar age who has a reduced ovarian reserve. The former will produce a good number of eggs / embryos to choose from and the latter will produce a much smaller number of eggs or embryos so that the one transferred may probably be the only embryo available. It is well documented that the chances of pregnancy in the former will be higher.

Therefore, clinics who primarily treat patients with high ovarian reserve will have higher LB\_Emb even with a similar number of embryos transferred by virtue of more embryos to select from. Clinics who treat patients with reduced reserve should accept this. However, the data for patients should be transparent; for example, the average number of eggs collected (a reflection of ovarian reserve) per patient of different age groups ought to be published. The HFEA should be able to provide national results for success rate per eggs collected e.g. group the data 1 egg, 2-3 eggs, 4 – 6 eggs, 7 -10 egg, 10 - 15eggs and so on. The use of the national data is very informative and can and should guide treatment.

#### **Diverting patients with low reserve to other modalities not reported in the headline figure**

##### a) Natural cycle IVF

- Advising patients with reduced ovarian reserve to undergo natural cycle IVF when the results of those cycles are not published in the headline figure is another method of driving patients with low reserve outside published headline data. It is because of this the inclusion of unstimulated cycle in the outcome is paramount – This however must also be included in the results per egg collection procedure. Obviously, some patients elect to use unstimulated cycles but we believe that they must be made aware exactly of the significant difference in the outcome even if the ovarian reserve is low. In London alone in 2014 data there were 750 natural cycles reported mainly

from 3 clinics. The total live birth from them was 26 making the success rate per cycle of only 5%. This is significantly less successful than the overall outcome for stimulated cycles not merely 'it can be less successful' as you mention in the information you provide. (Please see Appendix II)

b) Batching eggs or embryos

- In this situation, patients are advised to undergo 2-3 cycles of stimulation where eggs are collected and frozen, to all be thawed and inseminated later and all and the best embryo is transferred. The problem here is that the patients pay for 3 egg collections which with current methods of publishing are not included in the stats of the clinic.

c) Advising egg donation

- This of course can be appropriate advice, particularly if low reserve is coupled with advanced female age.

### **PROBLEMS WITH THE WAY THE HFEA PROPOSES TO PUBLISH DATA**

As our licensing authority and independent regulator, your strategy statement sets out your duty to *increase and inform choice* by “ensuring patients have access to high quality meaningful information”. The unmistakable drive to reduce a clinic’s output to a **single headline figure** is anathema to this philosophy. This approach to data publication ignores the complexities of the treatments we perform and the plethora of factors that affect outcome and can skew data.

The desire to publish a single headline figure that encompasses all ages, all types treatments, on the front page of any clinic belies a desire to provide the definitive ANSWER when everybody knows that there isn’t one. Such a single figure is against the advice of all statisticians, who have in the past recommended to the HFEA the use of a range with no central point.

This approach will inevitably mislead patients, inaccurately assess clinics and go further towards creating a “league table” approach to data interpretation; a consequence that the HFEA have long denied was their goal. In its quest to simplify and achieve a simple headline figure, the HFEA rendered valueless the new approaches it has adopted. This is particularly pertinent when you take into account academic research that highlights ***that over 85-90% of people delve no further than the first page/level of an internet search.*** Therefore, a single “headline” piece of data on the first page, may be the only data at which a patient glance.

We will outline our concerns in more detail below.

### **Age bands**

“All age groups” is a meaningless data point as highlighted by the HFEA many years ago and has never been used by any body as different clinics are likely to treat differing patient populations and the average age of the population treated can be significantly different.

When the HFEA first published data regarding clinic outcomes in 1995, success rate was expressed as adjusted live birth per cycle. This was based on an unpublished

statistical model developed by HFEA statisticians, taking into account factors such as age, embryo freezing, number of previous attempts, cause and duration of infertility and number of embryos transferred. This approach was soon abandoned and data was published crudely for ages below and above 38. This in turn was criticized and thus HFEA changed to adopting the internationally recognized and widely used data range to represent assisted conception outcomes ( $\leq 34$ , 35-37, 38-39 40-42 etc.). This is because age is the most important factor in determining the success rate for IVF regardless of the denominator used.

Furthermore, to then also revert to the simplification of only using  $<38$  and  $\geq 38$  age bands beggar's belief. Published data for outcomes by the current 6-tier age bands (adopted by the HFEA for more than a decade) make it abundantly clear how outcome drastically changes from band to band even within the  $<38$  or  $>38$  group and as such the simplification to 2 simple bands is therefore misleading; (See appendix 1)

- A 37 year old may will be given an artificially higher expectation of livebirth success by virtue of her inclusion in a  $<38$  group that includes all ages below this mark.
- Those above 38 are an even more heterogeneous group and success rate will be substantially different dependent on age distribution for patients above that age.
  - o Those well above the 38 year cut off may be given misleadingly high expectations
  - o Those just above the 38 year cut off may be given misleadingly low expectations

### **Time Span for Live Birth per Collection (LB EC)**

We really do not understand why should the results for LB\_EC be five years earlier. It is understandable to have an extra year, after all, once a patient has a live birth she is not included in the analysis. So, if the last egg collection was at the end of June, say 2013, then the last potential transfer from frozen embryos should be 2014. The data therefore should refer to cycles performed between 2012 – 2013 and not as currently published (2011-2012) So, one extra year is more than enough for the absolute majority of patients who have frozen embryos and did not get pregnant from the fresh transfer to come back and use them (Your statisticians should be able to confirm this). Waiting for a full two years is not understandable and really makes the published data very old.

It should only include data from patients undergoing fresh egg collection for the purpose of IVF or ICSI using their own eggs including natural cycle and PGS and the subsequent frozen transfers for extra year from those treatment types.

Mixing all of this with egg freezing, donated eggs and so on makes the data difficult to verify and understand and subject to changes between different clinics so we may not be comparing like with like.

### **Misleading Headline Figure (ALL IVF),**

As we mentioned before, livebirth per egg collection should be one year earlier than that being published for LB\_Emb from fresh cycles.

It should only include data from patients undergoing fresh egg collection for the purpose of IVF or ICSI using their own eggs including natural cycle and PGS and the subsequent frozen transfers for extra year from those treatment types. Mixing all of this with egg freezing, donated eggs and so on makes the data difficult to verify and understand and subject to changes between different clinics so we may not be comparing like with like. The results from egg freezing, from egg donation, should be published separately for each clinic as well as nationally.

The collective use of all sorts of treatment such as fresh, frozen transfer, egg freezing and egg donation into one single headline figure (ALL IVF), demonstrates a total misunderstanding of the difference between these modalities and the effect they can have on the apparent success of a clinic.

1. It is one thing to look cumulatively at live birth from fresh and subsequent frozen cycles as is the case with LB\_EG. It is another thing to include all frozen and fresh cycles performed in the same period, whether a pregnancy already resulted or not is simply wrong. This may make a clinic that carry on transferring single frozen embryos repeatedly looks worse than one that mainly use fresh.
2. A clinic with a high proportion of egg donation cycles may appear to have a higher success rate than a clinic with no donation program. The results from egg donation, should be published separately for each clinic as well as nationally.
3. The inclusion of cycles where eggs are frozen within the LB\_EC is hard to understand. They should either be included in both birth per embryo transferred and per collection or not at all. We believe that egg freezing cycles should be treated as a separate entity alone and not included in either due to a lack of data on outcomes that may distort success in those clinics with a high number of such cycles. The results from egg freezing, like that from egg donation, should be published separately for each clinic as well as nationally.

### **Natural or Unstimulated Cycles**

We welcome the inclusion of unstimulated cycle in the figures for LB\_Emb as it is long overdue. However, there should be consistency here whatever the denominator is. It is very important that every egg collection performed should be included in the denominator when results are expressed as LB\_EC. This is true whether the cycle was unstimulated or stimulated or with batching embryos. These are all cycles of treatment that the patients undergo and pay for; either directly or through the NHS.

Therefore, the HFEA should publish data for stimulated and unstimulated cycles together whatever the denominator as well separately highlighting the success rate from natural cycles both nationally and at the level of each clinic. This is important since many such cycles are performed with the belief that the outcome of that treatment will be similar to the overall success rate of that clinic.

Furthermore, given the debate regarding unstimulated cycles we advise that the HFEA publish national results outlining success rate related to the number of eggs collected. This is paramount as it will help a lot of patients with reduced reserve to understand what can and can't be achieved. (See Appendix III – Lister Fertility clinic data related to number of eggs)

### **Choose or Find a Fertility Clinic**

The information for quality group twice voted that the title of the new data publication would be “Find a Fertility Clinic”. Yet this was ignored by the executive and I am not sure whether this discrepancy was put to the authority or whether the authority supported the executive’s view.

“Choose a Fertility Clinic” directs the reader to believe that the information provided by the HFEA (especially results) is of a definitive nature. So, we, the “Authority”, will help you (the patient) to choose between the clinics. As oppose to, we the “Authority” providing you with information about different clinics and allowing you to decide where to go.

Past publications by the authority used the title ‘Guide to Fertility clinics’. Although, this all appears to be semantics and all these publications are ultimately used as a league table, the emphasis in the name implies how the HFEA thinks of its publication.

### **Other technical comments about website and data presentation .**

1. Going to the detailed stats section, the system asks 4 question, which is fine. The defaults in the choices is really is not what it should be.
  - a. You ask first, about the time period, the default, however is the oldest period in this case, 2011/2012. There is nothing wrong in involving previous years but the default should be the latest data publication i.e. 2013-2014 in the current published data. If the observer wants to look for earlier year or all combined as you provide them well and good. This is not difficult to programme the default choice should be latest years with live birth, followed by the most recent data for pregnancies. Other choices including combined data can follow that. This, we believe, would reflect the most recent activity in any clinic and will not be confusing in make the choices (as is the case with current publication).
  - b. Moving on to the type of treatment, normally we believe that should be combined IVF / ICSI. Other choices can follow so the observer can look at IVF alone or ICSI alone or egg donation alone it becomes easier (as is the case with current publication). Unfortunately, the default choice is what is called “all IVF”. This can be the last choice if at all, as it encompasses all sorts of treatments including frozen transfers, fresh transfers, egg donation and in this case they are not even related to the cycles so whether a patient became pregnant or not. We believe it is an inappropriate choice but if you insist on keeping it , it ought not be the default.
  - c. Age the defaults to 35 – 37 we believe going chronologically is more appropriate with under 35 is the first choice. You should also have all ages as the last choice (as is the case with current publication)
  - d. Finally, the choices in whether what type of embryos, eggs, fresh or frozen, again, it should really default on the most common which is fresh embryo, patient’s eggs followed by Frozen embryos patients eggs then



followed similarly with donor eggs (as is the case with current publication).

2. There is inconsistency in the way the data is presented. In the front page LB\_Emb is displayed first. In detailed analysis page, the default choice is live birth per cycle. We believe that LB\_Emb should have the same priority in detailed stats section
3. You provide single live birth per cycle and multiple live births per cycle. We would have thought that the multiple live births should be referred to the total number of births rather than the percentage of those from the cycle as it starts.
4. In the tap that address the proportion of Blastocyst to embryos transfer please add (%)
5. Finally, it is impossible to go back from the detailed stats to the front page

### **Summary**

Publishing LB\_Emb is not new. It has been published by the HFEA for the last 6 years at least. The only difference that the HFEA gave it priority over LB per cycle started.

I propose that the HFEA publish the front page exactly as before but putting LB\_Emb in the top table followed by LB\_EC, both of them broken down into the standard age bands in the same page format as it is published now. The HFEA can also add the star system etc.

Every egg collection procedure – whether stimulated or unstimulated - MUST be counted when data is published, whether that related to LB\_EC or the current standard of per cycle started. Although there should be the ability to tease out data per clinic for stimulated or unstimulated cycles, overall results for any clinic, whatever the denominator, should include stimulated and unstimulated cycle and whether the eggs used were fresh or batched.

HFEA should stop publishing a single headline figure per clinic. The continued transparency of clinic data reporting is essential. Clinic outcomes are not simple and dependent on a number of clinical, demographic, funding and financial factors. The HFEA has to accept and indeed promote that data publication and success rates is a complex matter. The duty of the HFEA is to educate both patients and the wider public of this and outline why a single figure is inappropriate.

### **Annex I**

#### **Inaccurate assessment of clinic quality**

***Changing to the age bands suggested may inaccurately portray the quality of clinic, which we can demonstrate using a worked example.***

#### **Age Distribution Of Treated Patients ≥38 and effect on LB/embryo:**

The data below confirm that nationally 34.3% of cycles are performed in women  $\geq 38$  and shows the decline in outcome as age group increases.

#### HFEA website data

<u>Age group</u>	<u>Proportion of Transfer cycles : National Average</u>	<u>LB / Embryo transferred</u>
18-34	43.9%	26.9%
35-37	21.7%	21.9%
38-39	14.6%	14.9%
40-42	13.4%	8.3%
43-44	4.0%	3.0%
45+	2.3%	0.9%

If we therefore compare the outcomes of 3 clinics.

#### CLINIC A (“HFEA National Average Clinic”):

- National average for age distribution
- National Average for LB/embryo transferred in all current age bands
- Transfers 1.75 embryos in all women  $\geq 38$  (National HFEA data)
- **Headline combined  $\geq 38$  LB/Embryo: 10.0%**

<u>Age group</u>	<u>Distribution of cycles (National Average)</u>	<u>Number of Transfer Cycles</u>	<u>Embryos transferred</u>	<u>Livebirths</u>	<u>LB/Embryo (National Average)</u>
38-39	14.6%	146	256	38.1	14.9%
40-42	13.4%	134	235	19.5	8.3%
43-44	4.0%	40	70	2.1	3.0%

45+	2.3%	23	40	0.4	0.9%
Combined		343	601	60.1	<b><u>10.0%</u></b>

**CLINIC B :**

- Minimal cycles in older age groups
- Highest proportion of ≥38 group in the current 38-39 band than any clinic
- Headline combined **≥38 LB/Embryo: 14.0%**

<b><u>Age group</u></b>	<b><u>Distribution of cycles</u></b>	<b><u>Number of Transfer Cycles</u></b>	<b><u>Embryos transferred</u></b>	<b><u>Livebirths</u></b>	<b><u>LB/Embryo (Lower than Clinic C)</u></b>
38-39	12.1%	125	192	29	15.1%
40-42	2.2%	23	37	3	8.1%
43-44	0%	0	0	N/A	N/A
45+	0%	0	0	N/A	N/A
Combined		148	229	32	<b><u>14.0%</u></b>

**CLINIC C:**

- Less selective policy with even distribution in all age groups
- Headline combined **≥38 LB/Embryo: 11.1%**

<b><u>Age group</u></b>	<b><u>Distribution of cycles</u></b>	<b><u>Number of Transfer Cycles</u></b>	<b><u>Embryos transferred</u></b>	<b><u>Livebirths</u></b>	<b><u>LB/Embryo (Higher than Average)</u></b>

38-39	16.0%	328	477	99	20.8%
40-42	22.2%	455	726	74	10.2%
43-44	9.8%	200	307	6	2.0%
45+	3.4%	69	93	0	0%
Combined		1052	1603	179	<b><u>11.1%</u></b>

The data above clearly demonstrate that the proportion of patients treated in each age band above 38 can significantly influence the outcome of a single headline figure of  $\geq 38$ .

The clinic with the lowest LB\_Emb outcome in all age groups appears to have the highest success rate using a combined figure for all ages  $\geq 38$  by virtue of their patient group, either as a consequence of perhaps clinic selection criteria, demographics or funding differences. Similarly, the clinic with the highest LB\_Emb in all age bands appears to have the lowest success.

- National Average clinic: LB/embryo  $\geq 38$  10.0%
- Clinic B: **Poorer outcomes** in all bands LB/embryo  $\geq 38$  **14.0%**
- Clinic C: **Higher outcomes** in all bands LB/embryo  $\geq 38$  **11.1%**

Similarly distorted outcomes can be shown for LB\_EC when comparing such clinics.

Such differences in clinic demographics are not uncommon and could mislead a significant proportion of patients as well as unfairly impact on clinics.

## **Annex II**

### **Stimulated and Unstimulated Cycles**

The following example highlights the importance of inclusion of unstimulated cycles in both LB\_Emb and LB\_EC. Often, in such cycles, there may be no embryos to transfer due to no oocytes being collected or failed fertilization or cleavage.

Worked Example: This is HFEA data 2014 for 3 London-based clinics (one of them receives all its eggs from a sister clinic nearby). This clearly outlines the effect of including all cycles on outcome, but in particular LB\_EC (using cycle as a surrogate for collection).

#### **Clinic A**

	Stimulated	Natural	All
--	------------	---------	-----

	Cycles			Cycles			Cycles		
	Cycles	Embryos	Livebirths	Cycles	Embryos	Livebirths	Cycles	Embryos	Livebirths
	313	510	81	379	353	26	692	863	107
LB/Embryo	15.9%			7.4%			<b><u>12.4%</u></b>		
LB/Cycle	25.9%			6.9%			<b><u>15.5%</u></b>		
	<b><i>Current Headline Data Reported</i></b>			<b><i>Currently not included in Headline Data</i></b>			<b><u>Correct Data analysis</u></b>		

**Clinic B**

	Stimulated			Natural			All		
	Cycles			Cycles			Cycles		
	Cycles	Embryos	Livebirths	Cycles	Embryos	Livebirths	Cycles	Embryos	Livebirths
	558	856	238	164	88	7	722	944	245
LB/Embryo	27.8 %			7.9%			<b><u>25.9%</u></b>		
LB/Cycle	42.6 %			4 %			<b><u>33.9%</u></b>		
	<b><i>Current Headline Data Reported</i></b>			<b><i>Currently not included in Headline Data</i></b>			<b><u>Correct Data analysis</u></b>		

**Clinic C**

	Stimulated			Natural			All		
	Cycles			Cycles			Cycles		
	Cycles	Embryos	Livebirths	Cycles	Embryos	Livebirths	Cycles	Embryos	Livebirths
	546	1136	141	198	204	4	744	1999	145
LB/Embryo	12.4%			2 %			<b><u>7.2%</u></b>		

LB/Cycle	25.8 %	2 %	<b><u>19.5%</u></b>
	<b><i>Current Headline Data Reported</i></b>	<b><i>Currently not included in Headline Data</i></b>	<b><u>Correct Data analysis</u></b>

### **Annex III**

#### **Livle birth per cycle (Lister Fertility Clinic 2011 -2014) related to number of eggs collected**

	1 egg	2 eggs	3 eggs	4-6 eggs
< 38	(7/66)	(16/113)	(33/151)	152/500
	8%	14%	22%	30%
38 - 44	6/197	16/291	40/309	135/812
	3%	6%	13%	17%

---

## Annex 2

### Submission from the British Fertility Society in response to the HFEA beta website survey.

Received via email on 7 October 2016

We support the idea of a headline figure presented for each clinic, but are concerned that that figure should as closely as possible represent a figure that gives a fair indication of that clinic's performance.

In order to achieve this, we suggest including in the headline figure only patients and treatments that fall within set, agreed criteria.

We suggest that those criteria must include:

- (i) Age: female partner e.g. <38 years old
  
- (ii) Treatment type: include all ICSI/IVF with own eggs, ejaculated semen, no egg or sperm donation cycles, no PGD or PGS cycles

We support the intention to present figures on outcome per embryo

We are a little vexed about the issue of “natural cycles” as different clinics use the term in different ways and some also do give medication. Whilst on the one hand with a truly natural cycle there is never the option of transferring more than one embryo, so one cannot compare with stimulated cycles. Yet, on the other hand, including natural cycle treatment in the headline is the “intention to treat” figure and hence cancellation rates. Whilst it is difficult to argue against implantation rate in a defined group including the natural group this may hide a multitude of sins. Perhaps secondary highlights should be multiple pregnancy rate, cycle cancellation rate (including where no ET performed) and OHSS rate.

## Annex 3

### Submission from a clinic in response to the HFEA beta website survey.

Received via email on 7 October 2016

6 October 2016

For the attention of: Mr. Peter Thompson and Mr. Nick Jones

#### Natural and Stimulated cycles should not be combined in headline success rates

This short submission is to explain why we at Create Fertility believe that Natural cycle IVF should be regarded as a separate treatment methodology compared to stimulated IVF and consequently that success rates from these two methodologies should not be combined in a single headline figure. It should be read in conjunction with our previous submissions and correspondence with the HFEA in this regard.

#### Context

IVF is performed with oocytes collected in natural and stimulated cycles. Different approaches to ovarian stimulation are employed in IVF. Following the concerns regarding multiple births and Ovarian Hyperstimulation Syndrome, there has been a revival of physiological approaches to IVF in order to reduce health risks to mother and baby and to make treatments cheaper and women-centered. Natural cycle and Mild IVF have been increasingly used in selected patient populations to optimise health outcomes and reduce the burden of treatment.

In order to establish consistency in the terminology and protocols used, the International Society for Mild Approaches in Assisted Reproduction published a consensus paper in Human Reproduction (1). Subsequently, a glossary was published jointly by ICMART and WHO in Fertility and Sterility (2). The scientific publications have clearly recognised that stimulated IVF and natural IVF are two different entities with different aims, methodologies and outcomes that are applied to different patient groups. Hence, the scientific literature does not combine results of natural and stimulated IVF cycles in a single group.

The principles and definitions of Conventional stimulation, Mild stimulation and Natural IVF approaches could be found in publications mentioned above (1 & 2). Although different in concept from conventional IVF, Create are content to have our mild IVF data included with the conventional IVF data in the Headline Success Rate because they are both classified under stimulated IVF.

Our Scientific team at CREATE consists of Prof Campbell, Prof Frydman, Prof Nargund, Prof Chian and several experienced consultants and embryologists. Many members of our team have developed protocols to make natural and physiological approaches to IVF successful and our commitment to giving women the choice to have a baby with their own eggs and to reduce costs and health risks of IVF treatments are well known. Nearly 60% of our patients have natural IVF and we are extremely concerned about the proposal to combine success rates of natural and stimulated IVF under one headline figure. We cannot see any valid scientific reason to change the existing model wherein they are separated. We are grateful for the opportunity to explain our position as to why natural and stimulated IVF cycles should not be combined under one headline figure.



#### Rationale behind Natural Cycle IVF

Natural cycle IVF was originally applied to women who did not wish to have stimulating drugs or were advised to avoid such drugs because they had an oestrogen dependent cancer or were at high risk of such a cancer. However it became apparent in the early 2000's that it was also a more effective method of treating older women over 40 years of age and younger woman with low ovarian reserve who were only capable of maturing one or two oocytes and in whom the use of stimulation was ineffective or counterproductive. Traditionally such women would be offered egg donation or at best a stimulated cycle which, if it did not result in at least 3 mature follicles, would be cancelled before egg collection took place.

Natural cycle offered hope to such women because although it was accepted that their eggs were biologically less efficient than those of women with normal ovarian reserve, there was evidence that the natural approach and the natural selection of the oocyte increased their chance of conception (albeit significantly lower than with women with normal ovarian reserve). In our practice, women undergoing Natural Cycle IVF were usually advised to undergo up to 3 cycles to optimise their chances of having a baby.

#### Low Ovarian Reserve

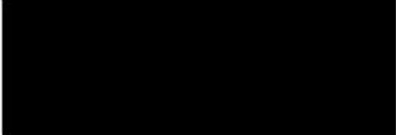
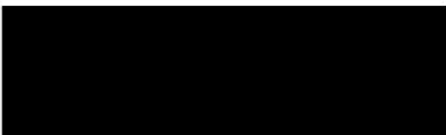
There are currently two accepted tests of a woman's ovarian reserve i.e. Antimullerian Hormone (AMH) and Antral Follicle Count (AFC). AMH is secreted from small antral follicles in the ovary and is measured in the woman's blood. Levels below 3 pmol/ml. are accepted as indicating low ovarian reserve i.e. low number of oocytes and low percentage of biologically efficient oocytes (3). An AFC below 4 (i.e. combining the count from both ovaries) also indicates a low ovarian reserve of oocytes. As explained above, conventional stimulation is not appropriate for such ovaries that are not able to respond in the conventional way (i.e. falling into the "poor responder") category and high doses of gonadotrophins might render biologically inefficient oocytes even less efficient. That is why a natural and physiological approach is required to give such women a chance to conceive a child with their own eggs. It has been proven to be more successful than high dose stimulation in such women giving them an opportunity to have a baby with their own eggs (4).


#### Why the results of Natural and Stimulated IVF Cycles should not be combined in Headline Success Rate tables

In summary, Natural and Stimulated cycles are two fundamentally different treatment methods and are offered to different patient populations. Women who receive Natural Cycle IVF have low success rates compared with stimulated IVF because they have biologically less efficient oocytes and embryos. Natural cycle offers them some hope but to combine success rates between Stimulated and Natural cycle IVF is unfair, illogical, unscientific and misleading. Such publication would go against the principles of fair, clear and transparent public information.



We also make the following observations:

- Women with very low AMH are offered Natural cycle IVF
  - Women with low AMH have a choice, either to have Natural IVF to have a baby with their own oocytes or to use donor oocytes. The success rates are evidently different but the success rate with a woman's own eggs is not negligible and many women would at least like to feel they have tried to conceive with their own genetic oocytes.
  - Women who are offered natural IVF because of low AMH and AFC do not benefit from stimulated IVF (as indicated by previous repeated failures and cancellations with stimulated cycles and as evidenced by scientific publications). Our aim is to promote patient choice and cost-effective options for these women with their own oocytes
  - Oocytes from women with low AMH behave like oocytes from older women who have lower success rates per embryo transferred because of lower quality oocytes. AMH is a reliable biomarker of oocyte quality (5). The mean implantation rate is lower in younger women under 35 years with AMH levels less than 1 ng/ml suggesting lower biological efficiency per embryo related to lower quality oocytes.
  - The studies have also shown that there is an association between oocyte pool and aneuploidy (6). It confirms that women with low AMH and AFC who undergo natural cycle IVF have less biologically efficient oocytes and embryos.
  - Furthermore, studies have also indicated that AMH levels in the follicular fluid of pre-ovulatory follicle are a predictor for oocyte fertilization and embryo quality (7). The oocyte regulates granulosa cell AMH expression. This is another physiological explanation for lower quality embryos in women with low AMH who undergo natural IVF.
  - Natural IVF cycles require a 7-day a week service and special expertise (such as advanced ultrasound, Doppler and 3D) in assessing the maturity and quality of follicles and to monitor closely in order to avoid spontaneous ovulation and to time the egg collection. Clinics with no experience or expertise in Natural Cycles IVF can attract patients because of their overall headline rate. This will mislead couples seeking Natural Cycle IVF who will be attracted by a headline success rate inappropriate to their needs.
- 
- 



#### References

1. Nargund, G.; Fauser, B.C.J.M.; MacKlon, N.S.; Ombelet, W.; Nygren, K.; Frydman, R.; Rotterdam ISMAAR Consensus Group on Terminology for Ovarian Stimulation for IVF (2007). "The ISMAAR proposal on terminology for ovarian stimulation for IVF". *Human Reproduction*. 22 (11): 2801–4.
2. F Zegers-Hochschild Adamsson GD, De Mouzon J et al: International Committee for Monitoring Assisted Reproductive Technology (ICMART) and the World Health Organization (WHO) revised glossary of ART terminology (2009): *Fertility & Sterility* 92, 5, 1520-1524
3. Irez T1, Ocal P, Guralp O, Cetin M, Aydogan B, Ashmawy S. Different serum anti-Müllerian hormone concentrations are associated with oocyte quality, embryo development parameters and IVF-ICSI outcomes. *Arch Gynecology Obstet*. 2011 Nov; 284(5): 1295-301: 10.1007/s00404-011-1979-6. Epub 2011 Jul 12.
4. Lianas TG, Sfontouris IA, Venetic CA et al: Live birth rates after modified natural cycle compared with high dose FSH stimulation using Gn RH antagonists in poor responders: *Human Reprod* (2015) 30, 10, 2321-2330
5. Lehmann P, Velez MP, Saumet J et al: Anti Mullerian hormone (AMH): a reliable biomarker of oocyte quality in IVF. *J Assist Reprod Genet* (2014) 31, 493-498
6. Fu X, Cheng J and Zhu S: The association between the oocyte pool and aneuploidy: a comparative study of the reproductive potential of young and aged mice; *J Assist Reprod Genet* (2014) 31: 323-331
7. Kim JH, Lee JR, Chang HJ et al: Anti- Mullerian Hormone levels in the follicular fluid of the Preovulatory follicle: A Predictor for oocyte fertilization and quality of embryo: *Obstetrics and Gynaecology JKMS* (2014) 29, 1266-1270

---

## Annex 4

### Further submission from a clinic director in response to the HFEA beta website survey. (additional to annex 2)

Received via email on 14 October 2016

15-10-2016

Dear Juliet

Further to my previous letter regarding data publications and following the meeting on the 29<sup>th</sup> September, I am writing to summarise what I believe the HFEA should do to improve the publication of IVF data in such a way that helps the patients and the rest of the stakeholders.

If you must have a single headline figure, then this must be age dependent and therefore, the results should be published according to all the agreed upon 6 age groups (less than 35, 35-37, 38-39, 40-42, 43-44, 45+) data. Nevertheless, as this data is optimised to mobile applications and that it will be very difficult for any patient to keep browsing into data of different age groups that are not of interest to them, I propose that once the patient press the button for choose a fertility clinic, after reading the information that you are providing, she is asked one question;

#### **What is your age? Or enter your age**

Once the patient enters her age, the system automatically assigns her to the appropriate age group out of the standard six. Thereafter, the headline figures for that age group is what the patient sees (exactly the same way you are providing currently for all ages), whether this is live birth per embryo or per egg collection procedure. This way you have a headline figure but specific for the observer.

I do propose, however, that once the patient is assigned to age group, the system should provide the patient with the National average detailed results for (IVF/ICSI combined including stimulated and unstimulated cycles) for that age group published as

- Life birth per embryo (LB\_Emb).
- Life birth per EC (LB\_EC)

Detailed info should also be provided after the above for

- IVF alone
- ICSI alone
- Natural cycle alone
- Frozen Embryo Transfer cycles alone

Going to the national data can be made obligatory before going to separate clinics (preferable). Alternatively, it can be available by choice; two buttons beside each other “Go to National Data” and “Go to a Fertility Clinic”.

**Principles I believe agreed upon in the meeting:**

**All IVF** is totally meaningless and should be abandoned (replaced by default of IVF/ICSI) at all levels of choices i.e. from headline data and when you go to “have a closer look”

In LB\_Emb the denominator includes all IVF/ICSI cycles using own eggs fresh (and in case of all frozen, then the first transfer of the frozen thawed embryos is included). Cycles included regardless of method of stimulation; including unstimulated cycles). Excluded from this are egg donation, frozen transfers and PGD/S cycles.

In LB\_EC the denominator includes same as the above types of patients from the moment they go to egg collection (so whether egg collection successful or not). Frozen transfers are included till the woman has a live birth or the embryos are exhausted in one more extra year. Excluded here egg donation or PGD cycles (the results are published separately). You go back only one year extra than in the case of LB\_Emb (not 2 years as it is currently proposed - data looks very old).

As long as natural cycle IVF data are not included in the headline figures, there is always the potential for using it for patients with reduced reserve with significantly reduced chances of success so as to hide their poor outcome from the headline figures. Alternatively, it is sold to patients as the treatment that depends on quality rather than quantity, therefore convincing patients who prefer not to take medications that this is an alternative and equivalent route for success. I can assure you that none of the clinics that use a high number of natural cycles, highlight the poor outcome of the natural cycle on their websites.

I would, therefore, like to refer you to a most recent publication (Sunkara et al 2016)\*, this paper studies the outcome from natural unstimulated cycles and compares it, to stimulated cycles for all treatments performed in the UK between 1991 and 2011 (HFEA data). The overall live birth rates were 4.7% per cycle following unstimulated fresh IVF versus 22.5% following stimulated fresh IVF. They estimated that to achieve one live birth across all ages you need to perform 21 natural cycles as opposed to 4.5 stimulated cycles. In other words, stimulated IVF cycles are 5 times more successful than a natural cycle. For all these reasons it is paramount that the HFEA publishes the

National figures for natural cycles and includes them in the overall stats for each clinic.

In all cases providing detailed analysis utilising national data is perhaps substantially more important than individual clinic success. That is why; patients are advised to refer to national results. National data is more robust and given the numbers used is more accurate. Finally given the increasing number of patients with reduced ovarian reserve, I suggest publishing national data related to number of eggs collected. This will help these patients understand their potential success (for example if they produce 1, 2, 3, 4 or more eggs) and what method of treatment to adopt.

I hope that the above helps in optimising data presentation and helps patients en rout to an important step in their lives. Please do not hesitate to contact me if you need to discuss any further details.

██████████

*\*Sesh Kamal Sunkara, Antonio LaMarca, Nikolaos P. Polyzos, Paul T. Seed, and Yakoub Khalaf. (2016) Live birth and perinatal outcomes following stimulated and unstimulated IVF: analysis of over two decades of a nationwide data Human Reproduction, Vol.31, No.10 pp. 2261–2267*

Advanced Access publication on September 2, 2016 doi:10.1093/humrep/dew184



---

## Summary of plenary sessions and table discussions

### Age stratification

There was broad consensus in the room that a headline success rate figure was important as patients really want this, though aggregating all ages was flawed and could be misleading for patients. Age has a huge impact on outcome and needs to be reflected.

Though some felt a success rate for comparison should not be included as a headline figure (and instead an indication of whether the clinic was above, below or in line with the national average), the majority agreed with having a headline figure, since patients want this and they accepted that the Authority wishes this to be births per embryo transferred. A small minority felt there are flaws with this particular measure and that other figures, for instance, 'births per cycle started' were better.

Attendees understood that grouping all ages was intended to create bigger sample sizes but noted that most clinics would undertake enough treatments to allow for age breakdowns. An explanatory note could be included to account for small centres with less than 50 treatments in a given age group.

Many groups suggested alternative approaches in presenting the headline. A number noted that showing the rate for a 'gold standard' patient would be a better approach but there were questions on how to identify the best comparator or what patient mix the gold standard would be. In lieu of this, a younger patient group would be a better comparison.

Some did feel that choosing to present a single headline figure was in itself misleading for patients. They felt that the concept of a single headline figure was too simplistic, suggesting either showing all 6 age brackets or alternatively 3 broad categories.

Discussing the age breakdown (into under 38 and 38 and over) further down the clinic profile page, some attendees felt this was fairly arbitrary and could be counterproductive in terms of clarity for patients about their chances of success. Some felt this could be unnecessary if more age breakdowns could be provided in the headline.

Conclusions for the HFEA to consider:

1. Present some indication of birth rate alongside the other headlines – but make sure this is meaningful.
2. The current headline figure, grouping all ages is too oversimplified. Adjust success rate for age in the headline by either:
  - a. Presenting the success rate for one, more comparable age bracket such as under 35 year olds
  - b. Presenting the success rate for a 'gold standard' patient (eg, patients meeting the same categories for age, fertility etc. based on no. eggs collected)
  - c. Presenting a number of top-line figures for different age brackets to be more meaningful to patients



Any of these approaches will have pros and cons, but they would lead to a fairer comparison which was more meaningful for patients.

3. Detailed statistics should continue to be available, with an understanding that some patients may not access this and will only see the headline rate. Two broad age bands were not thought to be a helpful sub figure by some and potentially a little arbitrary; the more detailed groupings were better.
4. Consider showing the average age of patients the clinic treats to give patients an indication of the usual mix for that clinic.

### Treatment types

There was agreement around the room that grouping the range of treatment types together was a problematic approach and some changes were needed to present a more meaningful headline figure.

There was detailed discussion of the benefit of including frozen embryo cycles in the overall headline success rate. Freezing currently occurs for a wide range of reasons, both clinical and patient preference, so to include these treatments could be misleading. This could be misleading and so the majority felt it could be better not to include this. Equally there are different factors such as length of freeze and type of freeze methodology ie, vitrification, which are important but are not reflected in a single 'frozen' category. Some thought an option could be to include the results from the first embryo transfer for a freeze all cycle, but this was not a universal view. Others were more relaxed about including frozen cycles and were happy for these to be included with fresh.

They also discussed natural cycle IVF and strong and differing views were expressed about how this information should be presented. The overall feeling was that we should consider excluding natural cycle IVF from the headline figure since this is a different treatment type. In natural cycle IVF only one egg may be produced which is very different from most standard IVF. Some suggested that to include it would mean not comparing like with like. Some raised that the term 'natural cycle' could be used for a spectrum of different treatments – the term was not universally consistently applied. Although a minority of clinicians disagreed with this and thought it should continue to be included in the headline.

The clear feeling from the workshop was that natural cycle results should be presented separately and clearly in the detailed statistics to facilitate informed patient choice.

Participants generally agreed that donor egg recipients are a different patient group and since the eggs come from healthy, fertile donors the inclusion of these treatments is misleading in the overall success figures. It could mean clinics encourage more patients to use donor eggs to improve success rates. The same is true for surrogacy which is likely to be undertaken for different reasons than standard IVF/ICSI.

Clinicians were worried that including PGD/PGS in the figures this might encourage the practice of referring all patients to have PGS, even when it isn't needed. Also, those having these treatments might be doing so for entirely different reasons, so one would not be comparing like with like.

Conclusions for HFEA consideration:

1. The HFEA headline figure would be more meaningful if it included:

- a. Only fresh stimulated IVF/ICSI cycles with the patient's own eggs
  - b. Because these treatments are fundamentally different, it should *exclude*:
    - i. Egg recipients
    - ii. Frozen (though some felt the first frozen transfer should be included for freeze-all cycles and some were more relaxed on this)
    - iii. Surrogacy
    - iv. PGD/PGS
2. HFEA should consider excluding Natural Cycle IVF from the headline figure and presenting this separately and clearly in the detailed statistics to facilitate informed patient choice.
  3. HFEA should publish national success data for different treatment types before patients get to the clinic statistics.