HOUSE OF COMMONS

Work and Pensions Committee

Oral evidence: Health assessments for benefits, HC128

Wednesday 22 June 2022

Ordered by the House of Commons to be published on 22 June 2022.

Watch the meeting

Members present: Sir Stephen Timms (Chair); Debbie Abrahams; Shaun Bailey; Siobhan Baillie; Neil Coyle; Steve McCabe; Nigel Mills; Selaine Saxby; Dr Ben Spencer; Chris Stephens; Sir Desmond Swayne.

Questions 387 - 424

Witnesses

I: Professor Ben Barr, Professor in Applied Public Health Research, University of Liverpool; and Dr Ben Baumberg Geiger, Reader in Sociology and Social Policy, University of Kent.

Written evidence from witnesses:

Professor Ben Barr, Professor in Applied Public Health Research, University of Liverpool - HAB0055

Dr Ben Baumberg Geiger, Reader in Sociology and Social Policy, University of Kent – HAB0099

Examination of witnesses

Witnesses: Professor Ben Barr and Dr Ben Baumberg Geiger.

[*This evidence was taken by video conference.*]

Q387 **Chair:** Welcome, everybody, to this meeting of the Work and Pensions Select Committee, an evidence session in our inquiry on health assessments for benefits. A warm welcome to Professor Ben Barr, who is joining us virtually this morning. Could you very briefly introduce yourself?

Professor Barr: Good morning, everyone. My name is Ben Barr. I am a professor in applied public health research at the University of Liverpool and head of the WHO Collaborating Centre for Policy Research on Determinants of Health Equity. I lead a research group that specialises in evaluating the health effects of social and welfare policies, including assessing some of the health impacts of the introduction of the work capability assessment.

I am not specifically an expert on disability assessments, per se; the other Ben, who will hopefully be joining us, has more expertise in that area. The assessment of health impacts of policies is my area of expertise.

Q388 **Chair:** Thank you. When Dr Geiger joins us, also virtually, we will ask him to introduce himself, too.

I will start by putting a question to you: do you think that the work capability assessment and the PIP assessment tell the Department for Work and Pensions what they are supposed to tell it—that is, the impact of a health impairment on somebody's ability to work or on their daily life?

Professor Barr: It is very difficult to determine how well the assessment processes are working outside the context of the wider disability benefits system. To judge whether the disability assessment system is working, we need to decide on the outcomes that we want that process to achieve. There are three key outcomes that we should be looking at. One is improving the employment chances of people with disabilities. The second is reducing or eliminating the risk of poverty among people who are unable to work due to disabilities. The third is improving their health and wellbeing.

A well-functioning system with the right thresholds for assessment is one that improves employment, reduces the risk of poverty and improves wellbeing for people with disabilities. Unfortunately, the Department does not routinely assess the impact of its disability benefit policies on those three outcomes. What can we say about those outcomes? We assessed the employment effects of the introduction of the work capability

assessment between 2010 and 2013, in particular the reassessment of people on incapacity benefits, and found that that had not had a positive effect and had not increased the chances of employment. It had moved people off benefits, but it tended to move them into unemployment benefits such as jobseeker's allowance, and many of them subsequently ended up back on employment and support allowance. The OBR recognised at the time that that was one of the reasons why it had not realised some of the cost savings that had been predicted from that policy.

We have also conducted a systematic review of international evidence of the employment effects of changing disability benefit assessment criteria. We found that there is very little consistent evidence internationally that changing those criteria has much of an effect on the employment of disabled people. There is some evidence from the US historically, where the expansion and relaxation of some of those criteria led to a reduction in employment, but there is no evidence from the UK that more stringent assessment increased employment chances.

If we look at it in terms of poverty, we see the disability poverty gap increasing since 2013, particularly with rising poverty among people with disabilities who are out of work. It is likely that that is due to reduced access to and adequacy of disability benefits.

On the third outcome, health and wellbeing, we conducted a study in 2016 looking at the health impact of the work capability assessment and identified some major negative impacts on mental health. If we look more recently, we see that the prevalence of mental ill health among people with disabilities has increased since 2014. Wellbeing measures such as the ONS wellbeing measures show anxiety increasing among people with disabilities quite considerably since 2013.

Across those three outcomes, we see that the assessment process seems to be failing, in that it does not seem to be improving the employment chances of people with disabilities particularly. We see increasing risks of poverty and potentially adverse effects on mental health and wellbeing.

Q389 **Chair:** On your point about increasing employment prospects, the Government often make the point that there has been a significant increase in the number of disabled people in employment. Does that not suggest that something positive is happening?

Professor Barr: There has been an improvement in the disability employment gap, but whether the changes to the benefits system have contributed to that is unclear. This Committee and work from the Office for National Statistics have pointed out that to a large extent that is due not only to the overall increase in employment across the economy, but to increases in the prevalence of disability, in particular the prevalence of mental health with disability.

While we have seen an improvement in the disability employment gap, it does not seem that the changes to the benefits system have contributed to it. One of the other things that is likely to contribute is a reduction in the disability educational gap—an improvement in the gap in educational attainment between people with disabilities and people without disabilities.

Dr Baumberg Geiger: Could I come in on that point? Can you hear me?

Q390 **Chair:** Dr Geiger, good morning. Yes, we can hear you and see that you are in an attractive-looking park somewhere. Thank you very much for joining us. I will ask you to tell us for the record who you are.

Dr Baumberg Geiger: I am Benjamin Baumberg Geiger. I am at the University of Kent and am shortly to be at King's College London. I have done a lot of research on health assessments for benefits internationally, including a stint on secondment in the DWP a few years ago.

I add my apologies: I came from Canterbury to London overnight so that I could be with you in person, but I have failed to make the last step into the centre of London because of the traffic this morning. I am glad to be able to join you in some capacity.

Q391 **Chair:** Thank you for joining us. Let me put to you the question that I put to Professor Barr: do you think that the work capability assessment and the PIP assessment, the two assessments that we are considering in this inquiry, tell the DWP what they are supposed to tell it—that is, the impact of a health impairment on somebody's ability to work or on their daily life?

Dr Baumberg Geiger: I will answer that, but I just want to flag that it would be good to come back to the disability employment gap issue. I will return to that point in a second.

To deal with your immediate question, the work capability assessment does not assess people's capability for work. Somebody giving evidence before described it as pseudoscientific. It gives the appearance of objectivity and of assessing something meaningful, but it suffers from two particular problems that mean that it does not achieve what it really should.

The first is that it is set up—not in its own words, but very obviously—to assess the genuineness of someone's impairment, but it does so without any transparency or any evidence base about the situations in which it overrules people's accounts of their own lives. It does that very unreliably; I can see that in evidence that others have given in earlier Committee sessions.

The other thing that it does not do is tie impairments to someone's capacity to work. As we may come to, the Netherlands has a very comprehensive and systematic process that links the impairments that people may experience to the demands of the labour market, but there is

no evidence base that was created when the work capability assessment was created, and there remains no evidence base linking the things in the work capability assessment, and the thresholds for it, to people's capability for work. On both those grounds, it is not doing what it set out to do.

Chair: The PIP assessment?

Dr Baumberg Geiger: The PIP assessment I know slightly less about, but I have done some research on it. Again, it is not capturing the actual costs that people face in their daily lives. It is not set up to. It assesses some proxies; you could argue that those proxies are slightly better than those in the work capability assessment but, again, there is no robust evidence linking those proxies to the extra costs that people face. In fact, when people have done research on the level of extra costs that people will face, it can best be thought that PIP is a sort of crude gesture towards some of the very high costs that people face, but does not come close to covering them, particularly for those with more severe disabilities.

This may be something we will return to throughout the session: there is a lack of transparency, evidence and rigour in what goes on in these assessments and their link to the things that they are meant to be assessing. There is no evidence base that the DWP have produced to link the PIP criteria to the costs that people face. That needs to be rectified.

Chair: You wanted to make a point about the disability employment gap.

Dr Baumberg Geiger: This is something that I have done lots of work on. Some of it was with Vicki Wass, who has given written and oral evidence to the Committee and who continues to do work on this.

The problem with the crude measure of disability that drives the headline rates—I warned the DWP about this problem when I was on secondment there a few years ago, and have done since—is that it can show failure when there has been success and vice versa. The problem is that in a world in which barriers to inclusion for disabled people are going down, some people will no longer report that they have a disability, because their impairments do not affect their day-to-day lives in a way that they would report in a survey. I think we would all agree that that is success. However, that means that those people do not count as disabled in the definition, leaving only the more severely disabled people who will report they have a disability under any definition. Those people have more severe disabilities, and the disability employment gap for those people will be higher. You have this paradoxical situation where, if you have more inclusion, the disability employment gap might go up because there are fewer people with less severe disabilities reporting that they are disabled. I have warned about that for some time; it is the opposite of what you want a measure to do.

At the moment, we are seeing the reverse: disability rates are going up, including people who on average are less severely disabled than the people who have always reported having a disability. The disability employment gap looks as if it is going down, but that just reflects the fact that we have a bad measure, rather than anything that we could seriously call success.

Q392 **Selaine Saxby:** Good morning. We have heard that people sometimes travel hours at a cost to visit assessment centres. Do you have any evidence that more rural parts of the UK have greater difficulties in accessing assessments? Does that have any bearing on the long-term outcomes for claimants?

Dr Baumberg Geiger: I can come in very briefly on that. We have some data from a survey of claimants from June last year that includes people's preferences for telephone versus face-to-face assessments. I would be happy to link that up to data on whether that is a rural constituency, or even on the rurality of the small neighbourhood that people live in—that might be one way of telling us. What we find in practice is that people are very split in whether they want a face-to-face assessment or a telephone assessment. It varies by a lot of different criteria. There is no overwhelming majority one way or the other.

Given the known issues about access to assessments in rural areas—*[Inaudible]*—review it and send that to the Committee afterwards, if that would be useful.

Q393 **Selaine Saxby:** Thank you. Is there anything that you would like to add, Professor Barr?

Professor Barr: It is not an area that I have done particular research in. I think Ben probably has more evidence around that. All we can say is that the ease with which people can go through the assessment process and their access to it clearly influences their experience of it and the potential adverse consequences, so anything that can be done to improve that access is beneficial.

Chair: Dr Geiger, we cannot see you any longer, but I understand that you are still on the line. Is that correct?

Dr Baumberg Geiger: Yes, I am still here. I can put my camera on if that would help, but I am happy to do it just by voice as well.

Q394 **Sir Desmond Swayne:** Does it matter whether applicants trust the system? Is it not sufficient for the Department to believe that its methods of assessment are robust? I know that Ben has challenged that, but nevertheless, why do we need them to trust the assessment system?

Dr Baumberg Geiger: I think trust is crucial. One of the reasons—this was in one of the reviews of the work capability assessment—is that it is not enough for an assessment to be fair; it needs to be perceived to be fair, both for its legitimacy for the people going through it and for the

wider public legitimacy and accountability that we would expect of something that affects large numbers of people every year.

Even beyond those relatively obvious points, trust is essential for getting good outcomes from the benefits system as a whole. For most people with a health condition or disability, it is not immediately obvious to them or anybody else how they will get on if they try working, if they are not in work at the time. It depends a lot on the particular job and on a lot of unknowable things.

To get good employment outcomes out of the system, people need to be willing to take chances. They need to have a space for safe experimentation to see whether things work out and to know that if things do not work out, the benefits system will be there for them. That is the opposite of the system that we have at the moment. We have a system that—in the words of Paul Gregg, who did a review of conditionality in the late 2000s—lets people hunker down. The lack of trust means that people are more likely to cling on to the benefits that they have, because they do not understand why they were awarded them and they have no trust that they would get the benefits that they think they need if they were to try working.

That is part of the reason that it is driving negative outcomes of the system. For more positive experiences for claimants, to reduce the mental health consequences that the excellent research of Ben and colleagues has shown, and for more positive employment outcomes, we need trust. We can talk about how we might go about getting that.

Professor Barr: As Ben says, trust is crucial for the effectiveness of the assessment process. It is also an important part of having effective programmes to support people with disabilities into work. The breakdown of trust with the DWP potentially has an adverse effect on welfare-to-work programmes as well.

We conducted a systematic review back in 2010 of welfare-to-work programmes. One of the crucial aspects contributing to their effectiveness is the trust between clients and case managers within those programmes. Also, the evidence indicates that integrating support across the health and welfare system is important for helping people into work. One of the problems with the breakdown in trust is that that integration becomes much harder to work with. In 2005, I was working with jobcentres and GP practices, trying to join up support. That work is very difficult to do now because of the breakdown in trust between people with disabilities and the welfare system.

Q395 **Sir Desmond Swayne:** Given that you have both identified trust as critical, what can the Department do to try to build an acceptable level of trust?

Dr Baumberg Geiger: There are several things that can be done. One of them is around the transparency of the system and people's

understanding of what is going on. The DWP is not transparent and does not encourage scrutiny and evaluation of what is going on in the assessment. I do not think that that helps.

You could loosely regard what is going on as process changes and more fundamental changes. There are things that you can do on a small level. I know that you have had health assessment providers talking about all the things they do to try to make this a better process, where people have a better understanding of what is going on and the reasons for decisions are better explained. There are further improvements we could do there, particularly in explaining to people why a decision has been taken afterwards. That is something that is more in the control of the Department, rather than of the private sector contractors.

More fundamentally, it is hard to trust a system that says that it will assess your capability for work, but does not. There needs to be a more fundamental overhaul of the system so that it assesses the thing it is meant to assess. Gradually, as people experience that and word travels around, that will have consequences far beyond what you can do by improving the process of the current assessment.

Professor Barr: Transparency in the process and its procedural justice is clearly crucial for developing trust. Part of that should be having a systematic assessment of the benefits and harms of the process as they change. Alongside that, it is about ensuring that the organisation that is providing the assessment is trusted by disability organisations and people with disabilities. Developing that process with the engagement of people with disabilities and of organisations that represent them will be crucial.

The DWP tends to work slightly in isolation from other organisations. If some roles could be devolved to local communities and local government and more decentralised, that could help to build trust in the process with local communities.

Q396 **Sir Desmond Swayne:** On transparency and its importance in building trust, can there be any plausible reason why the Department will not publish its work capability assessment data on universal credit applicants?

Professor Barr: No, I cannot see any reason why that could not be published. In general, while some data is published fairly openly by the Department for Work and Pensions, there should be complete transparency on the data that is available on the assessment process and any potential harms from that process. There have also been issues with the transparent publication of information on reviews of cases where there is evidence of harm. The more transparency and open publication of data on the process there is, the better we can enable the process to be trusted.

Dr Baumberg Geiger: I will just expand on that a bit. I think the publication of WCA statistics in UC is a major issue, and it is great that

the Committee is drawing attention to it. It should also be linked to a wider culture of lack of transparency around the assessments in the wider benefits system. It would be very helpful for the Committee to challenge that, not least because I do not think it helps the DWP in doing its work effectively.

There are many examples. The WCA statistics in UC is one; the unwillingness to publish some reports, which I know the Committee has worked hard to try to get into the public domain, is another. Researchers trying to use the DWP's administrative data to do research is another issue, which I know has been in the news recently. I know that the Glasgow team involved was trying to work with the DWP for years to get that data.

More generally, there are a number of us who would like to go into jobcentres and observe what happens there—with the consent of the people taking part, of course—to see which bits seem to be working well and which bits seem to be working badly. The same goes for observing the assessments themselves. Is there a problem that people are not saying things about how their health condition or disability affects them? Is that a problem because of the structure of the conversation that is happening? Is there a problem with the interpretation of the assessor? And so on.

None of us has been able to get permission to do this research. In some cases, that even harms the Department because everybody suspects the worst sometimes. I have been very surprised in emerging research that I am doing about the extent of conditionality that people, particularly disabled people, face in universal credit at the moment—it was actually much less than I was anticipating, but because the DWP makes it so difficult for us to go in and observe what is happening, I think people fear the worst.

It would be very helpful if the Committee could continue its excellent work in pushing the Department on that. The Department's reputation would not be harmed if there were more transparency and openness; indeed, it would be helped, because otherwise people fear the worst. It would enable us to try to work towards a better system all round.

Q397 **Nigel Mills:** Do you accept that some sort of assessment has to take place—that there is no alternative to these benefits having some assessment done by somebody, at least?

Dr Baumberg Geiger: Yes. I have just had a conversation with the OECD team about whether you need two assessments or just one; I think we probably need two, but various people think there could just be one, and there are different arguments in favour of that. I do not think that there is any plausible system in which you have no assessments.

Q398 **Nigel Mills:** Do you think it is ever possible to find an assessment process and a contractor that people will have confidence in and trust, or is it

inevitable that there would be some concern, some cynicism and enough mistakes made that some level of distrust will always be there? Is that just the way these assessments will always be portrayed and viewed?

Dr Baumberg Geiger: I am glad you ask, because to my mind that is the single biggest barrier towards making the system better, including in the Treasury, as I understand it. The title of a report I wrote a few years ago was "A Better WCA is Possible!", which is the mindset we need going into this.

I have looked at many different countries' disability assessments in benefits systems. There are undoubtedly things that are difficult about it in all countries, and there are undoubtedly some people who will not be perfectly happy about it in all countries. The UK's assessment does stand out as being particularly bad, causing particular unhappiness and distress, and being particularly difficult to link to any sensible conception of what it should be assessing. I do not think that this is ever easy, but there are definitely concrete ways in which we could make it a better assessment.

Professor Barr: Clearly a perfect system is not possible, but a better system probably is. With any assessment process, you will have some misclassification: some people will be categorised as fit for work who are not, and some people will be categorised as not fit for work who are. That is the same for any kind of screening programme. The important thing is that we have methods to assess the impact of the assessment processes and ways to determine how well they are working. Unfortunately, we have not been applying those approaches and developing that evidence base to determine which changes would make it better.

Q399 **Nigel Mills:** I am trying to work out if what you are trying to say is that this could be done well, but not by the people who are doing it, who are just not up to the job; or that the best approach would be to try to get them to improve their standards further; or that the whole thing can only work if you take it back in-house and the private sector profit motive is not believed to be a factor at play. Which of those do you think is the strategy that the Department should follow?

Dr Baumberg Geiger: There are some sensible arguments for taking things back in-house in the medium term, partly because it is not an area in which it is very easy to make good contracts. The only way you can get people to take this on is by spending more money than you would if you just did it in-house, because of the inherent risk and complexity involved. I am not necessarily sure that in the short term the Department could just take it back in-house, although it has made sensible steps to build some in-house capacity and learning, which I think is a valuable step towards making it a possibility.

More fundamentally, while it is sensible to look at that, I do not think that that is the major issue. If anything, the fact that we have the private sector contractors has enabled DWP, particularly in its WCA days, to

deflect some of the responsibility for what is fundamentally an assessment that needs to be changed, rather than it all being the fault of the contractors and the fact that it is not done by the Government.

Professor Barr: I agree that the primary issue is not who is providing it, but there is an issue with what the incentives are and what the approach is to commissioning that support. As we have said, what is needed is an evidence-based approach to developing the assessment. It may well be difficult to ensure that that is happening within a profit organisation. It could be more effectively carried out if in the initial period it were in-house, to develop a more evidence-based approach.

Q400 **Chris Stephens:** I have some questions about the Scottish system, but I will set them aside just now. In your introduction, Dr Geiger, you touched on some international work. Since your 2017 paper on disability assessments, have there been any international examples that the Department should learn from to reform the work capability assessment?

Dr Baumberg Geiger: I will answer that in two parts. First, it is helpful for me to restate some things in the work that I did in 2017 and 2018. Broadly, there are three ways in which you can assess work capacity for benefits. One of them is to have experts make an individual assessment, although it is not very clear exactly what they are doing.

The better two ways of doing it are what can be loosely called the Dutch system and the Danish system. In the Dutch system, you have a better version of a functional impairment-based assessment—similar to what we have now, but linked through a large database to the demands placed on people in the workplace. That means that when somebody is found fit for work, they are given three jobs that they would be able to do based on their impairment profiles and given the current state of the Dutch labour market, which adds much greater transparency and much greater legitimacy. That is a structured assessment, where you are trying to take what we have now but make it work.

The other way of doing it that I think is worth looking at is the Danish system, which, rather than trying to make it very structured, tries very hard to get people back to work. They have multidisciplinary team meetings where they are trying to think of what steps they could take to help somebody back to work. When they reach a point where they think that there is nothing they can do by working with somebody to help them move back to work, they move on to the more permanent disability benefit. That is a more revealed work capacity through trying to get people back to work. We can come back to both of those possibilities for the British system, but they are important.

There are a number of more recent developments that I have not been able to look at in detail, but that are promising. Estonia has just introduced a new working ability allowance covering people out of work and in work, although it is early days and I have not seen an evaluation of quite how well it works.

The Dutch are trying to do something in between what I described as the Dutch system before, but for young people, where one of the things that I think is worth looking at is structuring the process and the thinking on whether somebody is capable for work. The person doing the assessment must go through a very transparent process of thinking and justification about why they have made their decision. That marks a stark contrast with the way in which, for example, assessors can decide that someone is not being honest about their impairment, but in such a way that there is no real transparency over the evidence they have used to make that judgment and the reliability of it, which makes it very difficult to scrutinise and challenge as well as being very unreliable. That is a particular thing that I would draw attention to since the work that I did before.

Q401 **Chris Stephens:** Do you have any views on the Scottish Government's plans for the adult disability payment? Is that likely to offer an improvement on PIP?

Professor Barr: It is not something that I have looked at in detail. My understanding of the assessment process for the Scottish plan is that it basically takes out the functional assessment process and is based more on the first of the types of process that my colleague has just mentioned. Potentially, that reduces some of the harmful effects from the assessment process. I think that we can potentially learn from that through evaluating what the impact is and what potential adverse consequences might come from that for employment, for example. Ben, you may well have further things to add on that.

Dr Baumberg Geiger: I mostly echo what Ben said. I think one of the priorities should be to do comparative research to understand exactly what is happening in the two systems and, where there are differences between them—which there will be, although some of them will be revealed by the implementation—to understand what the consequences are for claimants. I think that that would enhance our knowledge.

It has been interesting to see from afar the process of the Scottish Government wanting to make a better assessment but coming up with the challenges of exactly how to do that and whether they have the capacity in place, and to see the learning exercise that seems to have taken place through it. It is great that they are trying to minimise the number of assessments and do more paper-based assessments, but in some core way it will recognisably be the same sort of assessment. It would be nice if over time the Scottish Government could think about whether they need to do something more fundamentally different, even if they need to go through the process of learning how to do something similar first before venturing further afield to an entirely new assessment.

Q402 **Chris Stephens:** Commenting on the Government's Green Paper, the Disability Benefits Consortium said that the UK Government offered case studies "of other countries, including…Australia, New Zealand, France and Switzerland, yet failed to consider reforms being made by the Scottish

Government." Do you have any views on why Scotland's example was omitted?

Dr Baumberg Geiger: I am not going to comment on some of the more political aspects of that, on which I defer in many respects to the Committee's greater political expertise. What I will draw attention to, as someone who has tried very hard to bring international examples into DWP thinking, is that that sometimes works well, but it often feels as if international examples are very superficially thrown in to make a point that people were going to make anyway, rather than thinking about all the different ways in which countries around the world do these things and what we can learn from them. It is never a case of simply copying them, but there is a lot that we could find out about.

Perhaps this is just an example of the Government not engaging sufficiently with different ways of doing things and what they can learn, but rather throwing in a small number of international examples in an unambitious Green Paper, particularly in respect of things to do with the health assessment.

Q403 **Chris Stephens:** Thanks for that, Dr Geiger. The written evidence from interest groups is pretty unanimous in its support for the new adult disability payment system in Scotland. Are there any changes that you think the DWP should adopt? Are there any changes that should not be adopted from what is happening in Scotland?

Dr Baumberg Geiger: I think that the lived experience advisory boards are fantastic. Again, a lot depends on their having genuine power in the system, so that when the payment is introduced, policy makers are listening actively to people's experiences and trying to improve things as quickly as possible, but that has certainly been very promising to start with.

The written commitment to dignity rather than stigma within the system is important. Seeing how that translates into lived experience is important, but having that commitment there is great.

On the commitment to trying to minimise the number of assessments—this goes back to a past question—it is very hard to make disability assessments a really positive experience for people, at least broadly within the confines of the current system. Accepting that, one of the best things the Government can do is to put as few people through the assessments as possible, given other objectives. It is great that that is seemingly central to the way the Scottish Government are approaching it.

Q404 **Chris Stephens:** Finally, one of the consequences of the Scottish changes is increased spending: the estimate is that in the next five years the spending will go from £2.5 billion to £3 billion, compared with staying with the PIP system as it is now. Do you consider that a cost-effective use of resources?

Dr Baumberg Geiger: I will answer that question, but the first thing that I should point out is that estimates of the cost of reforms to health assessments have tended to be pretty inaccurate. We can take the DLA-to-PIP reform as something that was intended to be a cut, but it did not turn out that way in practice, and there are many other reforms to health assessments and disability benefits that have not worked out as intended. It is important to take the projections with a pinch of salt and pay close attention in real time to what is happening.

Having said that, I think that additional spending is a worthwhile investment, partly given what we know—Ben will definitely have more to say on this—about sick and disabled people's quality of life. Everybody is struggling at the moment, but disabled people in particular have been struggling for a long time, and things have become worse over the past year. An investment in people's quality of life when people are really struggling is valuable, but also it is at heart a brilliant payment to enable people to participate in society and help them to make sure that work is accessible and wider contributions are accessible. An investment in the inclusion and participation of disabled people, rather than their exclusion, is always an investment worth making.

Professor Barr: I think that that hits on one of the fundamental issues in debates on welfare policy, which is that there are huge potential benefits particularly from investment in disability benefits. To answer your question about cost effectiveness, we need to know what the benefits are from that investment. The benefits system should not be seen purely as a cost; we need to effectively measure the benefits and demonstrate their cost-effectiveness. There is growing evidence, as Ben has said, that investing in benefits for people with disabilities has huge benefits for their inclusion, wellbeing and health. That potentially realises savings in other parts of the system as well, while we can also see the potential for harms resulting in onward costs to the health sector, for example. Assessing the benefits as well as the costs is crucial to understanding the effectiveness of those policy choices.

Dr Baumberg Geiger: If I can add a specific piece of evidence, one thing I think people often expect to see when they look at different countries is a trade-off between the employment rates for disabled people and the generosity of disability benefits or the number of people claiming them. This is not what we see when we look internationally: if anything, countries with more generous disability benefits and greater numbers of people claiming disability benefits also have higher employment rates for disabled people. There are a number of possible reasons behind that, which we are trying to explore and research, but it is valuable to see that it is not an either/or thing: it is not that the more you put in, the less people are going to work. Particularly with disability, that is just not how it works. You need to invest in inclusion.

Q405 **Chair:** Dr Geiger, can I go back to what you were saying about the systems in the Netherlands and in Denmark, which you have picked out

as particularly attractive? From your description of the Danish system, an attempt to adopt it here would require a big reorganisation to deliver. The Dutch system, from your description, sounds as if it might be something that could be adopted in the UK, but there would need to be a big up-front investment in establishing the database and describing the UK labour market. Do you have a view on whether the UK ought to adopt one of those systems?

Dr Baumberg Geiger: Again, this may be something that we will return to in a later question, but what I think we should have immediately is a three-track approach with different timescales and levels of ambition. The first is to try to understand and improve the way the system judges whether people's impairments are genuine; it is deeply problematic and lacking evidence and scrutiny at the moment. That is something that we could do by investing in looking in the UK, observing what is going on and trying to create a clear process for improving it.

The second set of things that we could do is like the Dutch system. It would take a little longer, but it is not unfeasible: we could have a look at what is required in the British labour market with people's capacities and try to find something that looks a bit like the current assessment, but is much more closely linked to the demands of the labour market. That would take three or four years from inception to conclusion, at a guess, so it is short to medium-term, but it is very doable and would not be particularly expensive, given the consequences for people's lives that it would have and the amount we spend on health assessments and disability benefits more broadly.

At the same time, I think it is worth looking at a more fundamental reform of the benefits system to look more like the Danish system. My characterisation of the benefits system is that it is primarily about transferring payments to people, with some employment support loosely attached around bits of it, which does not always work particularly well. Instead of that, what if we have a system that is fundamentally organised around trying to help people to get back into the labour market and that also supports them during those periods? Obviously in the Danish system there are places for people where they cannot think of what else they could do, but everything, particularly for sick and disabled people, is oriented around occupational health expertise, education expertise, medical expertise and people coming together.

Thinking about whether we could do that in the UK would require us to change not just the health assessments, but the whole system. It would probably mean training a large cohort of people with the occupational health expertise necessary to deliver that kind of thing. That is something that for a long time people have been calling for, but as you will have seen from evidence from GP representatives and others, we still do not have that in the UK.

That is a long answer, but there are three things that we could concretely do to move towards a better system. With the one about genuineness,

looking at the UK and the short term, it should not take too long to make big improvements. Moving towards a more Dutch system with a programme of research on the labour market could be done relatively straightforwardly in the next few years. It is important to think about a more substantive and ambitious reform of what the benefits system could be like, but that will take longer and various aspects will need to be thought about. I want us to consider that, but it is a complex undertaking and is not something that we could move to in the next two years.

Chair: Interesting. Thank you very much.

Q406 **Steve McCabe:** Good morning. I guess what we are looking for is an assessment system that is fairly accurate in the judgments it makes and does not cause too much distress or difficulty for the people being assessed. Are there any examples you have come across in the DWP of existing assessments that you use in your area of work that you think are examples, right across the range of benefits they deliver, of good practice that it could build on?

Dr Baumberg Geiger: That is an interesting question. I would not necessarily restrict it to the DWP's assessments, but rather think about the broader set of capacity assessments, including disability-related assessments, that the state does, of which there are very many. For example, there are assessments within the education system and health and social care assessments. I have been involved a little in thinking about what we can learn from social care assessments, which are done very differently within the context of a very different system, but which also often assess many similar things to those that the WCA and particularly PIP are trying to look at.

While I do not think we necessarily can just copy one from the other, it would be helpful to have more of a dialogue about these things. If you look at social care assessments, for example, there is a lot of worry about the state of social care in the country and the rationing of care, but in the context of that there is less concern about the quality of social workers' assessments of social care needs. In some sense there is something about that that even in a very difficult context can often work better, although it is not without its problems. That would be helpful for us to learn from in a way that I do not think easily happens across departmental silos at present.

Professor Barr: I am sure that there is a lot that could be learned from the different assessment processes. Within each of those assessment processes, as Ben has outlined, there is a tension between an assessment being about identifying the needs of the individual, and the rationing and prioritising of access to resources to support those needs. One of the fundamental differences between things in the Department and in social care assessment is that in the social care process there is a focus on working in the best interests of the person with a disability. There is a reorientation of attitude in the assessment process to determine the best

outcome for individuals and how the assessment can support us to make decisions in the best interests of people with disabilities.

Q407 **Debbie Abrahams:** Good morning, both Bens. It is nice to see you, or hear you. I want to start by asking Ben Barr about the 2015 research that you undertook, which was published in the peer-reviewed *Journal of Epidemiology and Community Health*: "'First, do no harm': are disability assessments associated with adverse trends in mental health?" Could you explain what you did and what you found?

Professor Barr: In the 2015 study, we looked at assessing the mental health impact of the programme of reassessment of people on incapacity benefits with the work capability assessment. The first thing to recognise is that it was a period when a large number of people were rapidly put through the work capability assessment between 2011 and 2013—over a million people within England, which is the population that we looked at—but the assessment process did not proceed at the same extent or rate across the country. Some places moved forward faster. A pretty large proportion of the population were going through this assessment: in Knowsley, near where we are, 6% of the working-age population went through that assessment process during those few years.

Our study used the variation across the country in the roll-out of the reassessment programme to investigate whether there was an increase in mental health problems and particularly suicide in those areas that had progressed to a greater extent, where more people had been through that process. We found a very strong association between those places where more people had been through the process, and a rise in mental health problems and suicides. We estimated that during that period, across England, the process had led to an additional 600 suicides, 300,000 additional cases of mental health problems and a large rise in the prescribing of antidepressants. We looked at whether it could be explained by other factors or other economic trends, but there was quite a unique pattern in the increase in mental health problems, and the most likely explanation was that it was due to the reassessment process.

That evidence is also backed up by a number of qualitative studies that have been conducted with people who have been through the assessment process. Those highlighted some of the potential mechanisms—the uncertainty around the process; the threat, particularly in the application of conditionality; the way the process was seen not as taking into account people's lived experience, but as stuck in a never-ending cycle of bureaucracy—as well as the potential loss of income for those people.

There is also evidence from several individual case studies and coroners' inquiries that have highlighted a number of tragic suicides in the benefits process. There is clear evidence of the potential for the assessment process to cause some major adverse effects on mental health, particularly with so many people going through that process.

Q408 **Debbie Abrahams:** Thank you so much. It caused quite a bit of

controversy when it was first published. Some people made the fair point that it showed not a causal relationship, but an association. There are those who still say that it is not about the association with the work capability assessment—or the reassessment process, as you rightly clarify—and that it was down to other factors. What would you say to them?

Professor Barr: It is impossible to prove with 100% certainty the effect of these processes on people's health with any design of research, but what we can do is identify what is likely. The pattern of mental health problems and the lack of explanation by any other cause means that it is likely that the process did lead to those outcomes. We need to make policy based on the evidence of what is likely and what the potential adverse effects are.

Q409 **Debbie Abrahams:** I am sorry to push you on this, but this was metadata. This was not interviews; it was multivariate regression analysis. You have said that it is likely. What confidence would you give in that statement?

Professor Barr: It is very likely, given that it is also backed up with qualitative evidence, case studies and numerous reports on the adverse effects. It is not surprising, with a group of highly vulnerable people, a large majority of whom had mental health problems, being put through a stressful process. When you are putting such large numbers of people through a process, you only need to increase the average risk by a small amount to have a major impact on adverse health outcomes and outcomes such as suicide. It is extremely likely and extremely plausible that the assessment process led to those outcomes.

Q410 **Dr Spencer:** Thank you, Professor Barr. It is great to have an opportunity to look at your paper and ask you a few questions about it. I read it with a great deal of interest.

I have a couple of questions on how you have done the research. With the self-reported mental health problems data that you have used, could you have used coded diagnoses from the local authority, for example from GP data? You pulled the antidepressant data. I am just wondering why you went for the self-reported mental health data, as opposed to diagnosis of depression and so on.

Professor Barr: We did use data from primary care in terms of antidepressant prescribing. One of the recommendations that we had from that paper is that the Department uses secure, linked, anonymised data to look at the specific outcomes of individuals going through the process. Unfortunately, as my colleague has mentioned, there has been a lack of engagement from the Department in enabling access to that kind of data for researchers. The data that was used in that study was the best available data at the time, which included primary care antidepressant prescribing. Data on diagnoses from primary care at that level was not available. The other data we used was from the annual population survey, which is the data that the DWP uses for monitoring

things like the disability employment gap and in relation to mental health. It is the same outcome that is used for national statistics in looking at the prevalence and changes in mental health across the country.

Q411 **Dr Spencer:** I have another question on the self-reported mental health problem and the antidepressant prescribing correlation data you have had. What immediately jumps out at me is whether there is a possibility that the reassessment process itself triggers an appraisal by the person going through it about what symptoms they have and their experience, whether that could be conceptualised as a mental health problem and whether it could trigger somebody going to see their GP to get antidepressants if they are suffering from low mood symptoms. Is it possible that increased antidepressant prescribing, for example, which is your strongest relationship that I can see in terms of correlation, is indicative of people who have depression following a reassessment going to get treated, as opposed to people having depression generated through going through the assessment process?

Professor Barr: It is unlikely that would create such a big increase in people accessing health services for those issues and reporting in surveys that they have developed mental health problems. Clearly it would not explain the very strong relationship with suicides that we found: the policy implementation was clearly connected to an increase in suicides, which would not be related to people's access to healthcare.

Q412 **Dr Spencer:** I will come on to the suicide data in a second, but I am just wondering about the reassessment process and what is being measured. I suspect that some people may think, "If I am on an antidepressant or I have a diagnosis, that may lead to my being taken more seriously or to people taking my symptoms more seriously when I am going through the assessment process." This is an argument that it may be a harm in itself, depending on how you want to think about it, but that could be what is going on. Do you not think that that is the case?

Professor Barr: It is unlikely. Physicians and GPs in primary care would be applying standard diagnostic criteria in assessing those patients and determining their medication.

Q413 **Dr Spencer:** Yes, but the patient needs to go in front of them in the first place even to have that discussion. I wonder whether the reassessment is acting as a trigger for someone to go and get seen.

Professor Barr: Yes, there are possibilities there, but I think it unlikely that that would explain the large effect.

Q414 **Dr Spencer:** Obviously the most concerning relationship in the data that you have is the suicide data. In making sense of it, how do the correlation and association compare with other stressor events in people's lives? I do not know anything about epi studies in this area, but clearly reassessment will be a stressful event. It should not be, but it will, and we need to do everything we can to make it non-stressful. How does this data look in terms of effect size, compared with other life stresses that

people go through?

Professor Barr: Six hundred suicides is a large number, given the total number of suicides over that period. If you think of other stressful events, the previous study of ours looked at the effect of the financial crisis. There were 10,000 additional suicides in the financial crisis across Europe; I think there were 1,000 additional suicides in England during that period. It is potentially quite a large effect compared with other stressful shocks.

Q415 **Dr Spencer:** That is my lingering concern. You are closer to the data of your study, but it seems almost too big an effect for it to be believable that going through the work assessment process could, on its own, have created such a large number. Are you certain that there is not a hidden confounder going on somewhere that explains that relationship?

Professor Barr: This was over a three-year period and a million people were going through it. It is entirely plausible that putting a million people through such a process potentially risks leading to 600 additional suicides. Given the size and scale of the numbers of people exposed, it is entirely plausible that there could be that size of effect.

Q416 **Chair:** You made an interesting point earlier that you were able to look at the number of suicides with a geographical variation because this process proceeded at different speeds in different areas. In Knowsley, what was the percentage of the working population that you said went through the reassessment?

Professor Barr: It was 6% of the working-age population.

Q417 **Debbie Abrahams:** Professor Barr, has the Department been in touch, or have you been in contact with the Department, about what Dr Spencer has just said? Six hundred suicides over a three-year period, down to a Government policy for a work capability reassessment, is a hugely alarming number when you are saying that the equivalent for an international financial crash was 1,000. Surely the Department must have been in contact with you, as soon as it saw your data, about the alarm that it must obviously have caused.

Professor Barr: At the time, it was not directly in contact with us, but we recommended that the Department should actively monitor adverse outcomes among people going through the benefits service. We had some conversations with officers in the Department about that. As yet, nothing has systematically been put in place to monitor the impact and provide active surveillance of adverse outcomes for people going through the assessment process.

Q418 **Debbie Abrahams:** It is disappointing that that was not taken up.

What if you were able to do a comparable study again? I appreciate that there were quite unique circumstances around transferring from IB and IS and so on to ESA. There have been a number of changes over the last few years in the criteria for work capability reassessment and so on, so

what sort of changes would you envisage seeing in a repeat study?

Professor Barr: I think your colleague's point about whether we could have better evidence is a good one. We could have better evidence of the adverse impacts. Have things continued to cause harm? Have the changes made to the assessment process reduced those harms? The truth is that we cannot answer that question, because the Department has not been collecting the data to answer it and has not enabled researchers and others to access the data to enable it to be answered more robustly.

Something that the Committee could push for, as well as the reporting and investigation of individual cases, is an active system for the monitoring and surveillance of adverse outcomes from the assessment process. It could be structured along similar lines to the national confidential inquiry into suicidal safety in mental health, where an independent body monitors suicides and self-harm within mental health institutions, for example, and advises on reducing the adverse risks in those processes within the NHS. Having something along similar lines within the Department would be hugely beneficial.

Dr Baumberg Geiger: Can I come in quickly on the evidence of whether things have changed? Definite improvements have been made. I know what the representatives from Maximus, as well as the private sector PIP contractors, have said about how high the satisfaction rates are and how much things have improved.

That may all be true, but in a survey that we ran last June of several thousand benefit claimants, the proportion of people who said "The assessor listened to me talking about my disability properly" was still about 58%, so there is a substantial minority of people who do not feel listened to. In a survey that we have just run, the very provisional results among 7,000 benefit claimants are that over half of claimants who have been through a WCA say that it made their mental health worse.

It is not sufficient to say that this is a historical problem and that everything is fine now. If there were more transparency, it would be easier to know a bit more about it, but the evidence suggests that there are still major problems with the WCA that could lead to an increased risk of poor mental health.

Q419 **Debbie Abrahams:** My final question is about vulnerable claimants. The point was made in previous evidence sessions that claimants are vulnerable by virtue of relying on financial support from the Government. You may be aware of the hundreds of deaths of claimants we have been discussing in the Committee and elsewhere. Do you think that the Government or the DWP are able to identify the claimants who may be vulnerable through the assessment process? At what stage in the assessment process should they be identified, and using what criteria?

Professor Barr: As you say, clearly we need to recognise that everyone going through the process is potentially vulnerable. A large proportion of people with mental health problems are generally at a point of crisis when they are accessing these services and seeking help. They are likely to be vulnerable. That means that the whole process needs to be sensitive to the identification and support of people, particularly people experiencing mental health problems.

As I understand it, the Department's current approach is to try to identify a group of people. It frames that as being about offering them additional support to access services and identifying that by focusing on life events and personal circumstances. I do not think that it is possible to have a list of criteria that enables you to identify people at risk; I am thinking particularly of the risk of suicide. That is one of the reasons why NICE does not recommend that health professionals use screening approaches to identify people at risk of suicide, because it is very unpredictable.

What is needed is a shift in approach to focus on providing a duty of care to people. It is not just about providing adjustments, but about working in the best interests of people with disabilities. That means developing a trusting and supportive approach and an engaging relationship with claimants, and being aware of stigma and discrimination. The training of people involved in the process will be crucial for that, enabling them to recognise triggers and signs of mental health issues, and giving them the confidence to step in, whether that is providing reassurance or signposting claimants to support. Mental health first aid training is probably a crucial part of the process, rather than necessarily having a checklist of criteria of vulnerability.

Dr Baumberg Geiger: I echo everything that Ben Barr said. The other thing to come back to is the lack of transparency. We know that there are safeguarding procedures in place, both for the DWP and for private sector contractors, and we know that there are policies around vulnerable claimants, but we also know of horrific cases where there has been an issue and things have not been followed up. What we do not know, because researchers do not have access to being part of the system and following vulnerable claimants through it, is how common these things are or what can be done to improve them.

While there is some merit in continuing to improve those procedures, we need more transparency. Some have called for an independent group of people with lived experience—people from charities such as Mind and researchers who work in the area—who are able to apply some scrutiny to the delivery of these things on the ground. I think that that is essential for improving the system and so that everybody can have confidence that as much as possible is being done to make sure that the system is not placing unreasonable burdens on people.

Q420 **Debbie Abrahams:** Do you agree with what Professor Barr said about the duty of care for vulnerable claimants?

Dr Baumberg Geiger: Yes. One of the things that we really have to be aware of in the move to universal credit is that previously there was a distinction between ESA and JSA and other bits of the system. People on ESA are more likely, I suppose, to be capped as vulnerable—I prefer the term "complex needs", so I will start talking about complex needs from now on. There were more people on ESA who had complex needs, and then there is a period of waiting before the work capability assessment. For universal credit, everybody is part of the system. The demands are potentially imposed at the start before a work capability assessment, so you are relying on the safeguarding procedures to be good.

It is critical to conceive of this as a duty of care: with the large number of people going through the system, the Department has a responsibility to make sure that it is there to help and is not causing harm. It is not just about saying, "We have large numbers of people going through the system. We are going to make an attempt at doing it, but it is not going to be perfect." I do not think that that is good enough, given the burdens that the system places on people.

Q421 **Dr Spencer:** I have another question for Professor Barr. I am still trying to understand the comparator for the effect size of the reassessment on the suicide data that you have, because I cannot get my head around it and it looks so large. Something equivalent has happened to people recently: about 3.5 million to 4 million people went on UC during the pandemic. They lost their jobs and went on UC, which is a huge stressor. Is there any data on how many suicides resulted from that going through? How does the data that you have from your research compare to that?

Professor Barr: This brings us back to the Department having systems in place to monitor the impact. The work that we have done on universal credit has shown that that has led to an increase in mental health problems, but we do not have the data to look at the impact on suicides from universal credit, because the Department does not make that data available. It is within its power to do that, and it would be possible to assess those impacts.

Q422 **Chair:** Are you able to say whether the 600 additional suicides that you identified went through the reassessment? You do not have that data, if I understand correctly.

Professor Barr: No. What we show is that the geographical pattern in increases in suicides at that time closely mirrored the roll-out of the reassessment process, and we used that as a basis to estimate the increase in suicides. Again, it is a million people going through, so it is entirely plausible: the individual risk would not have to be that greatly increased to lead to an additional 600 suicides among a group of people who already have pre-existing mental health problems.

Q423 **Shaun Bailey:** I want to touch on the previous independent reviews, of PIP but also of work capability assessments. If we were to conduct

another review, particularly of work capability assessments, how do you think it should be framed? I know that in the past it has been done more broadly, trying to correlate it to policy objectives and whether they had been achieved, but do you feel that there needs to be a more nuanced approach in any potential future review, particularly in the light of our discussions around work capability assessments?

Dr Baumberg Geiger: The reviews were valuable, and it was great that they were set up as part of the reform processes, but—thinking particularly about the WCA reviews by Harrington and Litchfield—they ended up saying, by the fifth review, "We need to think about more fundamental changes. This is not about a review of the details of how this assessment is working. We have exhausted that as much as we can, and we now need to think more fundamentally about whether we need a different sort of assessment."

Trying to set up a further review would be great, but it should be tasked very specifically and given resources to think about what different assessments could look like in the UK and provide some evidence for them. I do not think that there is much mileage left in trying to make incremental changes to the existing system. Where we are at now, and where we were at by the end of these reviews, is the need to fundamentally change what we are assessing and why.

Professor Barr: I agree that we need to develop a fundamentally new approach, and that could be part of the review, but that means testing things out and evaluating the impacts. The Department is trying out some new things in the self-transition areas, but nothing has been published on how it is evaluating that and what outcomes it is looking at as a measure of success. The crucial outcomes are not just about numbers of people on benefits, but also looking at the impacts on poverty and evaluating the impacts on health and mental health. That has not been addressed in the previous reviews; they have not specifically looked at or assessed what the impacts are on health and mental health. That would be useful to include in a further review.

Q424 **Shaun Bailey:** You have pre-empted my next question. If we were constructing terms of reference for this review, what would the ideal terms of reference look like? What do you both think should be the goals? I know that you have just touched on them, but perhaps you could set out in a little more detail what you think a review should look to achieve.

Professor Barr: Reiterating that point, I think we really need to be clear about the benefits that we are aiming for from the disability welfare system, and that is what the review should focus on. How can it be developed or changed to better minimise risks of poverty when we see increasing poverty among people with disabilities? What can be done to minimise the adverse risks for mental health as well as promoting inclusion and employment among people with disabilities?

Dr Baumberg Geiger: The expanded terms that Ben talked about are important. What I think is most important is not to conceive of it as a review in exactly the same sense as before. In some sense, the review was an even longer and deeper process than you are doing as a Committee, which has produced an enormous amount of fantastic evidence already, even prior to your report, which will be incredibly valuable. But there is a need to do something beyond that—something a bit more like systematic research and testing, as Ben said.

I think a new review needs to have sufficient resources and access so that it can do a systematic study of the way assessments are currently being conducted, to get a random sample that is investigated in detail, both from the claimant's perspective and from the system's perspective. It also needs some additional resource to go out there and, in the case of the WCA, look at the world of work and see how it matches up, or, in the case of PIP, have a look at the extra costs that people face and how those match up to the assessment. Then it needs a testing and piloting strand to see what other forms of assessments could be done.

It would be a relatively small budget in the context of the money spent on health assessments, but I think it would still need a larger budget than the previous reviews had available. It needs to be a large process led by somebody independent, but with quite a large team of people involved. It needs to make a difference, rather than being what might first come to mind when you think of a review.

Chair: Thank you very much indeed. That concludes our questions to you both this morning. Thank you for being with us—and thank you, Dr Geiger, for staying with us, despite your difficulties with transport this morning. You have both given us some very useful information to reflect on, and a number of points that I think we will want to follow up.