BMJ Open

Exploring physiotherapists' and occupational therapists' perceptions of the upper limb prediction algorithm PREP2 after stroke in a rehabilitation setting: a qualitative study

Camilla Biering Lundquist ![ORCID],[1] Hanne Pallesen,[1,2] Tine Tjørnhøj-Thomsen,[3] Iris Charlotte Brunner[1,2]

[1]Research Department, Regional Hospital Hammel Neurocenter, Hammel, Denmark
[2]Aarhus University, Aarhus, Denmark
[3]University of Southern Denmark, Odense, Denmark

Correspondence to
Camilla Biering Lundquist;
camilla.lundquist@midt.rm.dk

ABSTRACT

Objective To explore how physiotherapists (PTs) and occupational therapists (OTs) perceive upper limb (UL) prediction algorithms in a stroke rehabilitation setting and identify potential barriers to and facilitators of their implementation.

Design This was a qualitative study.

Setting The study took place at a neurorehabilitation centre.

Participants Three to six PTs and OTs.

Methods We conducted four focus group interviews in order to explore therapists' perceptions of UL prediction algorithms, in particular the Predict Recovery Potential algorithm (PREP2). The Consolidated Framework for advancing Implementation Research was used to develop the interview guide. Data were analysed using a thematic content analysis. Meaning units were identified and subthemes formed. Information gained from all interviews was synthesised, and four main themes emerged.

Results The four main themes were current practice, perceived benefits, barriers and preconditions for implementation. The participants knew of UL prediction algorithms. However, only a few had a profound knowledge and few were using the Shoulder Abduction Finger Extension test, a core component of the PREP2 algorithm, in their current practice. PREP2 was considered a potentially helpful tool when planning treatment and setting goals. A main barrier was concern about the accuracy of the algorithm. Furthermore, participants dreaded potential dilemmas arising from having to confront the patients with their prognosis. Preconditions for implementation included tailoring the implementation to a specific unit, sufficient time for acquiring new skills and an organisation supporting implementation.

Conclusion In the present study, experienced neurological therapists were sceptical towards prediction algorithms due to the lack of precision of the algorithms and concerns about ethical dilemmas. However, the PREP2 algorithm was regarded as potentially useful.

BACKGROUND

Stroke is a leading cause of long-term disability in the western world.[1 2] Upper limb

Strengths and limitations of this study

► To ensure successful implementation, healthcare providers have to regard an upper limb prediction model as useful. A strength of this study is the focus on the therapists' perceptions.
► A strength of this study is the use of focus group interviews as these are an appropriate method to stimulate discussion between participants and illuminate their diverse perceptions.
► Generalisability of results may be compromised, as perceived barriers and facilitators for implementation will differ between sites.

(UL) impairments are common, resulting in functional limitations affecting daily life activities.[3 4] Accurate prediction of recovery of UL function after stroke is desirable since it can lead to targeted rehabilitation in times of limited resources in healthcare.[5–7] Some researchers claim that accurate prediction can provide patients and therapists with realistic expectations for UL function and help to set goals for rehabilitation.[5]

From a clinical point of view, a prediction algorithm may be needed most in patients with severe UL impairment. These patients represent a particular challenge for therapists, as it is difficult, based on clinical measures alone, to distinguish patients who regain UL function from those who remain paralysed.[6 8 9] In patients with severe UL impairment, the use of a biomarker may improve prediction accuracy for motor recovery.[7 10 11] A biomarker widely used to assess corticospinal excitability is motor-evoked potentials (MEPs), assessed with transcranial magnetic stimulation (TMS).[7 10 11]

An UL prediction algorithm that combines clinical assessment with the use of a
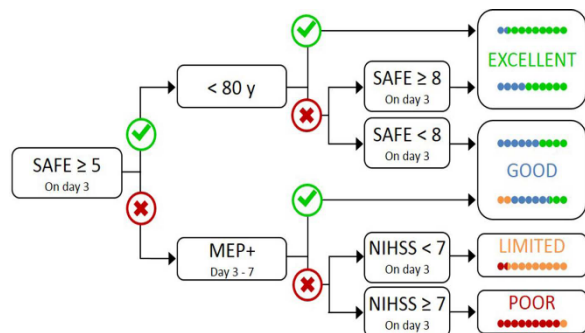
Figure 1 Predict Recovery Potential (PREP2) algorithm. Excellent: potential to make a complete, or near complete, recovery of hand and arm function within 3 months. Good: potential to be using their affected hand and arm for most activities of daily living within 3 months. Limited: potential to regain some movement in their hand and arm within 3 months. Poor: unlikely to regain useful movement in their hand and arm within 3 months. MEP+: motor-evoked potentials present; NIHSS, National Institute of Health Stroke Scale; SAFE, Shoulder Abduction and Finger Extension. Figure copied from the PRESTO homepage.[13]

biomarker is the Predict Recovery Potential algorithm (PREP2), displayed in figure 1. PREP2 involves several steps, depending on the severity of paresis. The first step encompasses a clinical assessment of UL function, using the Shoulder Abduction and Finger Extension (SAFE) test. In addition, for patients with low levels of UL function, the motor pathways between the stroke-affected side of the brain and the affected UL are examined using TMS. Information on age and severity of stroke further contributes to predicting UL recovery[7] The PREP2 algorithm has an overall accuracy of 75% and its prediction accuracy for patients with severe UL impairment exceeds the accuracy of other prediction algorithms.[5–7] Hence, research indicates that PREP2 is a promising tool for clinical application; and in the setting where PREP2 was developed, it was found to increase therapist confidence and rehabilitation efficacy.[12] To facilitate the implementation of PREP2 in other settings, the researchers behind the algorithm are hosting homepages that explain the rationale behind the algorithm and provide relevant instructions to therapists and patients.[13 14] In addition, a recent paper by Connell et al discusses barriers to implementation and how the implementation of PREP2 can be facilitated.[15] However, before we commenced the present study, we were not able to identify reports on clinical implementation of PREP2 outside the setting where it was developed.

One of the barriers to implementation described by Connell et al was the use of TMS, which requires special equipment and trained staff.[15] Another barrier to the use of PREP2 in a rehabilitation setting may be the first step of PREP2, the early administration of the SAFE test. To perform a SAFE test within the first 72 hours may not be possible at a rehabilitation setting where patients are admitted at a later point in time. The present project took place at Hammel Neurorehabilitation Centre and University Research Clinic (RHN), Denmark. From June 2018 to

October 2019 patients were admitted to RHN a median of 10 days (IQR 6–27) after stroke. Thus, the majority of patients were admitted too late to obtain a SAFE test and a PREP2 prediction while in rehabilitation. However, prediction may still be relevant to decide on rehabilitation focus and goals. Connell et al suggest that future research may determine if the time windows for SAFE and TMS can be expanded.[15] It would ease implementation if the PREP2 could be applied 2 weeks post-stroke with satisfactory accuracy. Other factors of importance for implementation may exist. To ensure successful implementation in a clinical setting, a crucial first step is identifying and describing potential barriers and facilitating factors.[16–18] Healthcare providers have to regard a UL prediction model as meaningful and useful for themselves and their patients.[18 19] Physiotherapists (PTs) and occupational therapists (OTs) responsible for UL treatment are the clinicians most likely to obtain and use the PREP2 predictions.

The aim of this study was, therefore, to explore how therapists perceive UL prediction with the help of the PREP2 algorithm in a stroke rehabilitation setting and to identify potential barriers to and facilitators of implementation.

METHODS
Study design
This was a qualitative study. We used an implementation framework to develop the interview guide, performed focus group interviews and applied a thematic content analysis. Focus groups are appropriate to illuminate both shared experiences and different perspectives of the group.[20 21] Group interaction was expected to stimulate discussion of thoughts, beliefs and attitudes towards UL prediction.[20 21] The interviews were explorative and focused on feasibility, acceptability and perceived usefulness of UL prediction algorithms.[17]

The Consolidated Framework for advancing Implementation Research (CFIR) was applied as a guiding framework to develop a semistructured interview guide and structure data collection.[15–17] The CFIR is composed of five major domains: intervention characteristics, outer setting, inner setting, characteristics of the individuals involved and the process by which implementation is accomplished.[16–18] The domains from the CFIR most thoroughly explored in this study are intervention characteristics, inner setting and characteristics of the individuals involved. The participants' views and attitudes within these three domains are expected to be important to a future implementation. On the contrary, the structure and organisation of the fourth domain, outer setting, will not be influenced by the views and attitudes of the participants; and the final domain, implementation process, is still in a preliminary phase.

The focus group interviews were centred around UL prediction algorithms, in particular the PREP2 algorithm. PREP2 is a three-step process (see figure 1). The first step is a calculation of the SAFE score by scoring

shoulder abduction and finger extension strength separately between a minimum of 0 and a maximum of 5. SAFE is based on the medical research council grades for limb power and two subscores are added to form a SAFE score of a maximum of 10. The second step of PREP2 depends on the SAFE score. For patients with an initial high degree of UL function reflected in a SAFE score of 5 or above, information on age is used. For patients with a SAFE score below 5, the function of motor pathways between the stroke-affected side of the brain and the affected arm is examined using TMS to elicit MEPs. For patients in whom MEPs cannot be elicited a measure of stroke severity, the patient's National Institute of Health Stroke Scale (NIHSS) is used. If this scale is not available, the equivalent Scandinavian Stroke Scale (SSS) score may be used.[21] PREP2 predicts UL function at 3 months in one of four categories, from 'poor' to 'excellent'. Patients who are predicted to fall in the category 'poor' are unlikely to regain useful movement in their hand and arm within 3 months, while patients in the category 'excellent' have the potential to make a complete, or near complete, recovery of hand and arm.

Study setting

The interviews were performed at RHN, Denmark. The RHN is distributed across three physically distinct rehabilitation units. Patients are admitted to one of the three units. Unit 1 has approximately 70 beds, and units 2 and 3 have 30 and 15 beds, respectively. While adult patients with stroke attend all three units, some of the beds at unit 1 are allocated to patients with severe (traumatic) acquired brain injury. A research department is placed in connection to unit 1. Clinical staff at all three locations work in teams, and in total 67 PTs and 67 OTs, involved in the treatment of patients, are employed. Some of the therapists have key positions, for example, specialist PTs or specialist OTs, and are responsible for professional development.

In recent years, therapists at the RHN have developed an interest in using UL prediction algorithms in clinical practice at a patient level. The evidence, local relevance and potential implementation have been examined and discussed by a group consisting of two OTs and four PTs assigned to positions within research or professional development. Based on these discussions, the most relevant algorithm for clinical use on an individual level appeared to be the PREP2 algorithm. The main reason was that the predictive value of PREP2 for patients with severe paresis exceeded the accuracy of other prediction algorithms.[5 6 12]

Several organisational obstacles prevented an implementation of PREP2 at a local level in the rehabilitation unit. The first part of PREP2, the SAFE test, should be performed within the first 72 hours, while patients are frequently admitted to RHN at a later point in time. Another barrier was the use of TMS, which requires special equipment and specially trained staff. Despite a desire to attain systematic prediction as a clinical routine, implementation of PREP2 in its current form was not possible.

However, two steps to ease a future implementation were taken: first, a prospective longitudinal cohort study was commenced to examine the accuracy of PREP2 when the SAFE score and TMS examination were obtained at a later point in time than originally proposed (clinicaltrials.gov number: NCT03632499). Results of this study have been published recently by Lundquist et al.[22] Second, the current study was conducted to explore facilitators and barriers for future implementation.

Participants and procedure

Before the actual data collection, the interview guide was tested for comprehensibility in an interview with an OT and a PT who were both involved in research and implementation. After this, a pilot focus group interview was performed with three PTs invited by the first author, CBL, a PT and a PhD student. The test interview and pilot focus group interview resulted in minor corrections: the number of questions was reduced, some questions were merged and information about UL prediction algorithms was simplified. For the complete interview guide, see table 1. Information posters displaying relevant illustrations about the topic, for example, the PREP2 algorithm, were produced in order to explain and facilitate discussion in the subsequent interviews.

The ward managers at each of the three hospital units were asked to invite participants based on the following criteria: a mix of PTs and OTs, involved in the treatment of patients, at least 1 year of experience in neurorehabilitation and from different wards. Experience with UL prediction algorithms was not a requirement. The intention was to achieve maximal variation in profession, clinical experience and degree of specialisation.[23]

After being appointed for the interview, an information letter was sent to the participants, in which the purpose of the interviews and the background for UL prediction algorithms for patients with stroke were presented. The participants were specifically informed about the PREP2 algorithm and were instructed to perform step 1 of the algorithm, the SAFE test, on at least three patients before participation. For this purpose, they were given a written scoring instruction. Performance of the SAFE test should ensure practical experience with the test and qualify the discussions during the interviews.

The interviews took place at the participants' work site. The interviews started with a few broad questions about what the participants considered important factors for UL prognosis. The purpose of these broad questions was to make the participants relax and feel comfortable and get their spontaneous opinions. Afterwards, the questions were more specifically about prediction algorithms, the use of tests and attitudes towards evidence-based practice. Finally, the PREP2 algorithm was introduced and discussed.

The focus group interview was moderated by CBL, who was aware of ensuring a confident atmosphere that welcomed a diversity of opinions. It was emphasised that there were no right or wrong answers. A senior researcher, HP, functioned as an observer, providing feedback to the

Table 1	Interview guide
Main categories	**Questions**
General questions	In patients with paresis of arm and hand: Which factors do you consider relevant for future arm and hand function? (important elements)
	What is relevant for your own approach to the treatment of the arm and hand? (write down three to four issues/ things)
Thoughts on prediction	What are your thoughts about the prediction of arm and hand function at an early point in time? What are the likely consequences?
	Which patients/groups of patients would benefit from knowledge of prognosis (eg, paralytic UL)?
	UL prediction models: to whom will it not make sense?
	Does age matter for prognosis (in general and for UL in particular)?
	Severity of UL impairment at stroke onset is relevant for UL prognosis. Where do you seek this information (eg, ward round, patient record, looking for particular scores as NIHSS or SSS)?
	Do your expectations of future UL function influence your approach and choice of UL treatment?
SAFE score	Before participation, you were asked to perform a SAFE score on at least three patients. How was it?
	What are your thoughts on using specific UL tests for (all) patients with reduced strength in arm and hand (eg, SAFE score, Fugl-Meyer score)?
	Are you aware of other hospitals focusing on UL prediction (eg, if they use SAFE)?
Knowledge of evidence	How do you get knowledge updates on UL treatment?
	Do you have the time and opportunity to get updated on new knowledge?
	Exercise: I explain the PREP2 model and show pictures of the elements: What are the pros and cons of a UL prediction model similar to the PREP2?
	What should it take for you to use a UL prediction model?
	Do you see patients for whom a prediction model would make no sense?
	Would the use of a UL prediction model change your approach to a patient?
	PREP2 can predict future UL function with approximately 75% accuracy. What is your opinion on that?
	Transcranial magnetic stimulation—can it be used in this setting?

Continued

Table 1	Continued
Main categories	**Questions**
Summarising	What we have talked about. Do you have anything you would like to add?

NIHSS, National Institute of Health Stroke Scale; PREP2, Predict Recovery Potential algorithm; SAFE, Shoulder Abduction and Finger Extension; SSS, Scandinavian Stroke Scale; UL, upper limb.

moderator and observing interactions in the focus group. HP asked clarifying and supplementary questions during the interviews. Directly after the interviews, the overall impression of the interview and any spontaneous reflections and considerations were noted. The interviews were audio-recorded and transcribed verbatim by CBL.

Analysis
The interview transcripts were imported to the qualitative research software program NVivo V.12 to facilitate coding and make data analysis more manageable. The pilot focus group interview was considered to add interesting dimensions to the topic, and data from this interview were included and analysed along with data from the succeeding three focus group interviews. A thematic content analysis of the interviews was performed by CBL.[23 24] The qualitative data analysis was both a deductive and an inductive process.[23 24] Deductive as we used the CFIR framework and sought to answer the specified research question regarding barriers and facilitators for implementation (theory-based coding). Inductive as we let the material talk (data-based coding) because attitudes towards UL prediction algorithms have not previously been explored, and knowledge of how to implement algorithms into the clinic setting is scarce. First, the four interviews were individually open-coded in NVivo and meaning units were identified. Second, the interviews were compared for similarities and differences and based on the meaning units, subthemes were formed. Finally, information gained from all four interviews was synthesised, and four main themes, considered of great importance to the participants and relevant for implementing prediction algorithms, emerged (figure 2).

The coding and interpretation of results were continuously discussed with coauthors. According to Malterud, this triangulation between authors with different positions and perspectives will increase the understanding of complex phenomena.[25] The four interviews revealed a broad array of relevant considerations, many of them appearing repeatedly, indicating data saturation.

In the Results section below, quotations are used to display from what kind of original data categories are formulated, thereby documenting and substantiating findings and increasing the trustworthiness of the study.[24–26] Where cited, the context is quoted in parentheses with anonymised participant initials and focus group origin, in accordance with table 2, for example, participant E from
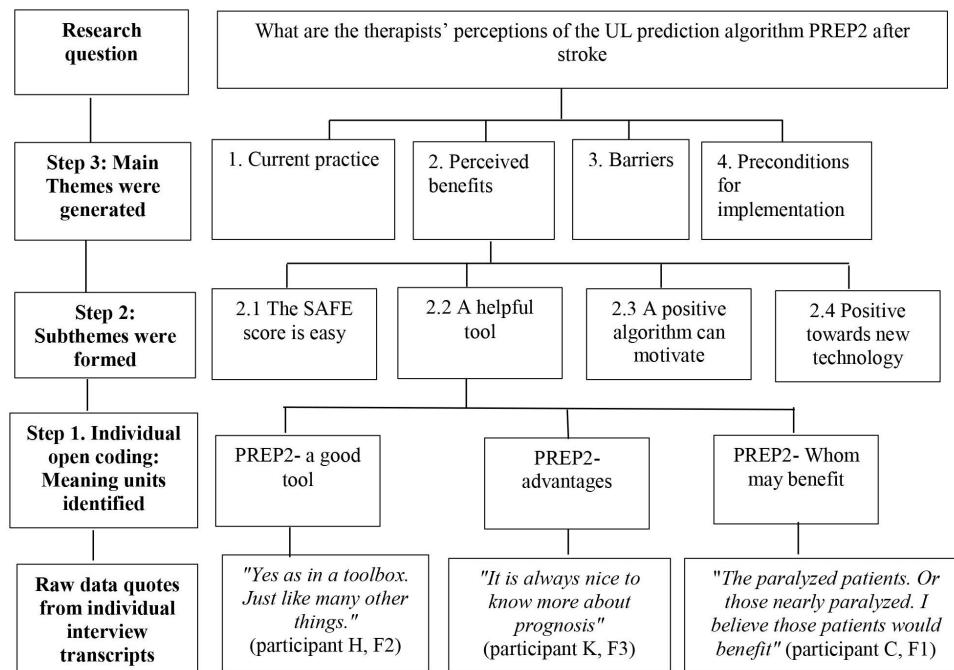
Figure 2 Diagram showing example of theme formation. PREP2, Predict Recovery Potential algorithm; SAFE, Shoulder Abduction and Finger Extension; UL, upper limb.

focus group 2 would be quoted as (participant E, F2). To ensure credibility, a participant from each focus group interview has reviewed the interview transcripts and the interpretation of the findings.[23] The participants agreed to the transcripts, recognised themselves in the descriptions and provided further nuance to the findings.

RESULTS

A pilot focus group interview followed by three ordinary focus group interviews were performed from January to April 2019 and had a duration of 68–90 min.

In the pilot focus group interview, three PTs participated. All had clinical experience in either neurorehabilitation or acute neurology and were engaged in either a Master's degree or a PhD. In the succeeding three interviews, all participants were employed at neurorehabilitation wards. The number of participants in the focus group interviews corresponded to the size of the rehabilitation unit: six participants from unit 1, four from unit 2 and three from unit 3. Both PTs and OTs participated; three were specialists in the field, one was a student advisor, all were female,

Table 2 Characteristics of focus group participants

Group	Pilot focus group (F1)	Focus group 1 (F2)	Focus group 2 (F3)	Focus group 3 (F4)
Number of participants	3	6	4	3
Profession	3 PTs	3 PTs; 3 OTs	2 PTs; 2 OTs	1 PT; 2 OTs
Assigned position	1 specialist	2 specialists, 1 student advisor		
Educational level	2 Master; 1PhD	5 Bachelor; 1 Master	4 Bachelor	3 Bachelor
Gender	2 F; 1 M	6 F	4 F	3 F
Average years since graduation (range)	15 (12–18)	12 (5–17)	20 (13–23)	17 (9–23)
Average years of experience in neurorehabilitation (range)	11 (10–18)	10 (3–17)	17 (13–20)	12 (2–18)
Current unit of employment	Unit 1 and acute neurology	Unit 1	Unit 2	Unit 3
Anonymised initial of participant when quoted	A; B; C	D; E; F; G; H; I	J; K; L; M	N; O; P

F, female; M, male; OTs, occupational therapists; PTs, physiotherapists.
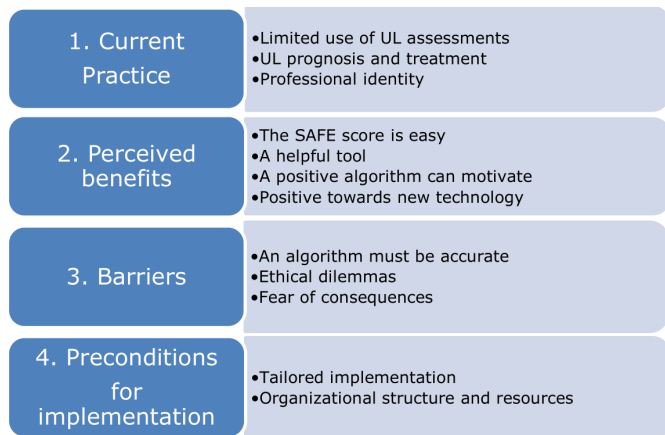
Figure 3 The four main themes and their subthemes. SAFE, Shoulder Abduction and Finger Extension; UL, upper limb.

had graduated 5–23 years previously and had 3–20 years of neurological experience (see table 2 for characteristics of participants).

Findings and quotations in relation to the four main themes: current practice, perceived benefits, barriers and preconditions for successful implementation are presented below and in figure 3.

Current practice

Knowledge of current practice is a precondition for understanding the participants' perceptions of barriers and perceived benefits. This first main theme concerned the participants' considerations on current practice and encompassed three subthemes: *limited use of UL assessments, considerations on UL prognosis and treatment,* and *professional identity.*

Limited use of UL assessments

As prediction algorithms comprise the performance of standardised assessments, information about the use of UL assessments was relevant. Participants in all four interviews agreed that UL tests were used, but on a limited scale. There was consensus that the UL test had to be clinically relevant for the specific patient and not a routine test for everyone. According to several participants, UL tests were primarily meaningful and used for patients with moderate to good UL function:

> Yes… MAS I believe I use a lot. With arms that can… do a bit more. (participant F, F2)

In addition, the test had to be quick to perform and easy to administer:

> One has to prioritize the time to do it. So it has to make sense to do it. (participant B, F1)

UL prognosis and treatment

Many factors were considered important for UL recovery. Some but not all aligned with factors highlighted in the literature. Pain was highlighted in all four interviews and had to be prevented and treated for UL function to occur:

> Well, I believe pain has a big say. Because…if they have pain, they don't move their arm. They just try to protect it… (participant F, F2)

Initial UL function and time since stroke were also mentioned in all interviews, but not stressed by the participants as important predictors:

> I think that having some function is important. We have a lot…I believe where the SAFE score is zero… because they are paralyzed… you cannot palpate any muscle activity. That has a huge importance for… whether they regain any function at all… (participant G, F2)

Other factors mentioned in the interviews as important for recovery were sensory motor deficits, time since stroke, location of stroke, type of stroke and initial medical treatment. According to all of the participants, cognition was vital, especially neglect and awareness of own disabilities:

> Yes. And cognition, all things considered…yes that matters. The ability to understand instructions. And maybe even to be able to perform self-training…that they understand the importance of focusing on arm and hand training. (participant K, F3)

Many participants mentioned the importance of past experiences, self-efficacy, motivation and inner drive:

> And the patients that have a good inner drive…they have a good prognosis [the group agrees]. (participant F3)

The PREP2 algorithm includes information on age and initial score on stroke severity. However, age was not considered particularly important for UL prognosis, and only a few participants were aware of initial scores performed in the acute units.

When planning UL treatment and choosing interventions, the participants took many of the same elements into account as when considering UL prognosis. Importantly, they found that the patients' individual goal should guide whether or not UL treatment was a main priority:

> The patient's priority counts. If the most important thing is to get that arm and hand going. Right now, I have a patient where eating was the most important issue and what I prioritized. (participant G, F2)

Professional identity

Professional identity concerns how the participants perceive themselves in relation to their profession and membership of their profession.

In all of the interviews, the participants agreed that use of UL assessment and algorithms such as the PREP2

aligned more with the PT profession than with the OT profession. Even though both professions use similar UL interventions and approaches, the PTs traditionally treat patients on an impairment level, while the OTs focus on activities and activity limitations:

> I believe our examinations differ. PTs have this…what can I say…very body-level examination, while we involve them during activities and in the bathroom or kitchen [the group agrees]. (participant K, F2)

> Well, if I have a patient I look for … because I am an OT… for activity limitations in relation to the use of arms and hands…because I am an OT. (participant H, F2)

Most of the participants considered themselves experienced neurotherapists; and according to many participants, prediction algorithms may make most sense for recently qualified therapists. Recently qualified therapists will need a simple tool, while the more experienced can draw on years of experience:

> I believe this PREP2 is for more recently qualified therapists…a lot easier to access…because then you can draw on the cold facts: this is what we have to guide us. And they are more schooled in that that the rest of us [the group agrees]. (participant M, F3)

Perceived benefits

This theme centres around how the participants thought an algorithm could aid and ensure UL treatment and rehabilitation. Subthemes were *the SAFE score is easy; a helpful tool; a positive algorithm can motivate* and *positive towards new technology*.

The SAFE test is easy

In the pilot focus group, participants had a general knowledge of prediction algorithms, but across the other interviews, knowledge of algorithms was less profound. All participants had heard of UL prediction algorithms and in one unit, some of the participants used the SAFE test. SAFE is step one in PREP2. All participants had been asked to perform a SAFE test on at least three patients before the interviews. Especially the PTs found the SAFE test easy to administer:

> The SAFE test is easy and quick and you can allow yourself to do it no matter what. (participant B, F1)

Some participants found the SAFE score insensitive, as the difference between score 2 (=limited range of motion without gravity) and score 3 (=full range of motion against gravity, but not resistance) was rather large. Despite this, the same participants considered the SAFE score to be appealing, because it was quick and could be performed everywhere and without equipment.

> But that big gap…we actually discussed it…. Actually, for some patients we would like to score 2½ [the group agrees]. (participant M, F3)

But apart from that, it is an easy score as…it doesn't need you to bring anything with you. And you can do it everywhere. (participant M, F3)

A helpful tool

PREP2 was considered a potentially helpful tool considering the prognostic potential of UL and for planning treatment. The algorithm might not be able to stand alone, but in combination with information from other sources, it could be used as a tool or an indicator to decide what way to go, for example, whether to intensify UL training or instead start the use of compensatory strategies.

> I believe an indicator is a good word. An indicator. Because it is not an answer to functions they will not achieve…or that it will be amazingly good. But it gives an indication. For this reason, we choose to go this way. But it does not mean that when the patient is discharged from RHN, we will write: The patient will never achieve any function. It is just a good tool. (participant F, F2)

> Yes as in a toolbox. Just like many other things. (participant H, F2)

> It is always nice to know more about prognosis. (participant K, F3)

Across interviews, there were different views on whether UL prediction algorithms would be a prognostic aid for all patients, mainly those with no or little function or those with moderate function. The predominant opinion was that it would be particularly relevant to patients with little or no UL function, reflected in a SAFE score below 5, as it was difficult for the therapists to predict UL function for these patients.

> The paralyzed patients. Or those nearly paralyzed. I believe those patients would benefit. (participant C, F1)

> Well… if so… it is only those with a SAFE score below 5. (participant N, F4)

A positive algorithm can motivate

All participants envisaged that a prediction algorithm could be used to motivate patients and therapists, given that the prediction was optimistic.

> Some indication… would be nice. It could be used to motivate when progression is slow and you think nothing is happening in an arm. If I could say: I KNOW if we do this exercise for the next four weeks every day, then it will come; that would motivate the patient. And me as a therapist. (participant P, F4)

Even though the participants preferred prediction algorithms to be as accurate as possible, many believed that an algorithm could be an aid without being 100% accurate. Several participants said that an accuracy of 75% would give an indication of whether your treatment plan was on the right track and what you could expect. It could still be

used by the team or individual therapist along with other indications and tools of prognosis.

> For me, it will be a tool to use in a team. I often believe that…with FIM and other functional measures…it is so interesting when…it does not fit. Then we get some beneficial discussions. (participant E, F2)

Positive towards new technology

In all four interviews, the attitude towards TMS and MEP was positive. The participants found using TMS to establish MEP status appealing since this could add information to UL prediction that could not be obtained by a clinical test. They imagined this information would motivate both patient and therapist:

> But what I find really interesting is that you can have this… MEP…? If there is a connection in the corticospinal tract. So you can have a SAFE below five and still expect a good function. (participant L, F3)

> There might be some people where you think they should have got some more (UL training)…because if we had that examination, TMS… (participant G, F2)

Barriers

This theme encompassed the participants' perceptions of the limitations of prediction algorithms and potential barriers to their implementation in clinical practice. Three subthemes emerged within this main theme: *an algorithm must be accurate, ethical dilemmas* and *fear of consequences.*

An algorithm must be accurate

All participants agreed that an algorithm should be as accurate as possible:

> Definitely, definitely [the group agrees]. (participant L, F3)

> It must, of course, be very precise for us to use it. (participant O, F4)

However, disagreement existed on whether the 75% accuracy of the PREP2 algorithm was precise enough. For some, a precision of 75% would be a barrier; and one participant stated that even if the algorithm was 100% accurate, she still might not follow it.

Ethical dilemmas

Whether or not to present and discuss the UL prediction with patients emerged as a dilemma for many participants. If a patient was predicted to have little or no function, this might depress the patient and would conflict with the participants' desire to motivate the patient:

> Yes. And what day do we tell the patient? Is it when they arrive and have been here in…? Well. I really don't know. On top of everything else? (participant N, F4)

Even if some participants were sceptical, they were still open for discussion and dialogue when other participants responded that informing the patient could make it easier to focus on other aspects of the rehabilitation where improvement seemed more realistic:

> I find it difficult to shatter someone's dream. You need to dream and believe this one will gain function. For some time. Of course, not for several years. (participant G, F2)

> Well…Well it is a balance isn't it. We have patients who come and tell us they are sorry that they weren't told…so the most important thing is to dare tell them, to be honest…well why should we treat an arm that we are nearly 100% will never function again? (participant H, F2)

> No…No… but… (participant G, F2)

> I don't believe we necessarily shatter someone's dream…necessarily… by letting the patient know how much this arm can improve. Instead, we consolidate and focus rehabilitation. (participant H, F2)

Fear of consequences

The participants believed that UL treatment affected future UL function and should be offered regardless of their initial function. The general view across interviews was that all patients deserved that therapists did their best to restore UL function.

In focus groups 3 and 4, concern was expressed that the use of a prediction algorithm would dictate which patients should receive treatment and which should not. If so, patients with a negative prediction would receive little or no UL treatment, and the algorithm would serve as a self-fulfilling prophecy. As such, the participants feared that introduction of a prediction algorithm would alter their approach to the patients:

> And then I might prioritize other issues instead. I am afraid so. And I hope I wouldn't. Because I believe that they need all the treatment they can get…. Because, truly, there is a chance in reality. (participant O, F4)

> If I had a diagram that could tell…your arm will never be good…. then I believe the patient should get the opportunity to prove this wrong. (participant J, F3)

In three of the interviews, it was mentioned that an algorithm could be used to stratify and prioritise which patients should receive treatment. In one interview, the participants regarded this as a positive consequence because it could be used to optimise treatment in times of limited resources:

> Because it is such a difficult matter already - and we do not have that many rehabilitation beds. So that would be an enormous help, I believe. (participant A, F1)

On the other hand, participants from unit 3 and especially unit 2 looked at algorithms in light of pressure from budget cutbacks; they feared that an algorithm would be used to accelerate and shorten rehabilitation periods, regardless of the patients' rehabilitation potential. In addition, some participants feared that an algorithm would introduce a too simplified view of humans.

> It depends how – if you can say so - our managers wish to use this tool…because we are under pressure. And will this be a tool to evaluate…which patients should be here? (participant K, F3)

> And I think it is like a tendency in society. That we need something that can be measured and recorded, and hard facts [the group agrees]. (participant M, F3)

Preconditions for implementation

Preconditions for future successful implementation were grouped in two subthemes: *tailored implementation* and *organisational structure and resources*.

Tailored implementation

The focus group interviews were performed at three different units. The overall impression was that despite being part of the same rehabilitation hospital (RHN), different cultures existed at the three units. Especially at unit 1, the participants (see table 2) were open to new ideas and implementing new knowledge seemed an integrated part of their culture. The participants at the other two rehabilitation units seemed open to new ideas, too, but were at the same time more sceptical. In all interviews, the participants discussed the importance of tailoring implementation to the specific unit, ward and patient. If something new had to be implemented, a persistent focus on the topic was needed:

> I believe that you must realize that implementation is just a lot more time-consuming and difficult than you imagine. A single day - when you present, discuss and maybe do something practical - is just not enough. (participant D, F2)

> I can say…in my ward…if something must be implemented, you have to take the specific patient and the patient's team to make it work. We cannot say something general about you having to….in all upper limbs…to do so and so. It has to be specific so they can relate to that. (participant I, F2)

> We aren't different from the patients. We, too, need a lot of repetition to implement something new and learn it [the group agrees]. (participant M, F3)

Organisational structure and resources

Time to get acquainted with new evidence and practise new skills was considered insufficient. All interviews showed that there was a sense among participants of being well-informed, while time to incorporate and practise new skills and routines was lacking:

> No. we don't even have the time to plan our daily treatments. So no, not at all. That is a real challenge [the group agrees]. (participant E, F2)

> The level of information is actually okay…it is more the time afterwards…to incorporate it. (participant J, F3)

> Yes exactly. (participant K, F3)

> True …to make it a routine. (participant J, F3)

Several participants mentioned that prioritisation and support from the ward manager were important for success. In all interviews, participants mentioned that weekly or monthly meetings were planned ahead and could be dedicated to specific issues. These meetings were valued and considered important by the participants when new knowledge had to be practised, modified and implemented. Flexibility from colleagues was acknowledged, and several participants considered this a prerequisite for attending a course. Generally, the participants felt that there was a culture of sharing the acquired knowledge with colleagues.

> It depends on how your ward works, I believe. How generous your colleague or ward manager is in relation to….well there is some economy in it too… but how much energy will we put into this? And are the rest prepared to run faster while someone is attending a course….? That is the culture and what you want. (participant D, F2)

Members of staff were assigned specific positions as a specialist OT or specialist PT. Specialists were considered a resource, capable of, and responsible for presenting and implementing evidence.

> And when some of the specialist therapists have been out in the wide world and return home and tell us about it, or some colleagues have been at a course… (participant M, F3)

> For me, it is about responsibility. Someone has to take responsibility. Because if all are responsible, nothing happens. I, as a specialist, can be the one responsible and say: Your patient, has he got an UL problem? (participant F, F2)

DISCUSSION
Summary of main findings

In the current qualitative study, four main themes were identified when exploring therapists' perceptions of prediction algorithms: current practice, perceived benefits, barriers and preconditions for implementation. Most participants knew of UL prediction algorithms. However, in practice, only some elements were applied and by a few therapists. Most participants considered themselves experienced neurotherapists and regarded UL prediction algorithms as particularly useful for more recently qualified therapists. The PREP2 algorithm was considered a potentially helpful tool when planning treatment

and setting goals. The perceived benefits centred on the SAFE test, a core component of the PREP2 algorithm. In addition, participants appreciated the use of TMS if it could add information to UL prediction. The main barriers were concern about the accuracy of the algorithm and dilemmas arising from having to confront patients with their prognosis. Preconditions for implementation encompassed tailoring the implementation to a specific unit, having sufficient time and being part of an organisation supporting implementation.

Comparison with previous findings

Current practice was characterised by limited knowledge and use of UL measurements and UL prediction algorithms. This result is corroborated by a recent Danish survey study by Kiær *et al* that revealed that prediction models for UL function after stroke are not yet a part of daily practice in Danish stroke rehabilitation.[27] This is not surprising as the Danish clinical stroke guidelines do not recommend the use of any particular UL measurement or UL algorithm.[28 29] Moreover, international recommendations of standard use of UL measurements have not yet been implemented in clinical practice.[30 31]

According to previous research, the initial UL function is the main predictor for UL recovery.[6–9 32–34] The participants in our study acknowledged the predictive value of initial motor function. At the same time, they considered several other aspects important, such as the patients' goals, their motivation and their self-efficacy. Studies have shown that other factors such as individual goals, motivation, self-efficacy, aphasia and depression influence rehabilitation outcomes and should be considered.[35–38] The experienced therapists in our study drew on their clinical knowledge and expertise. This is in line with their views on the PREP2 algorithm as a useful but supplementary tool that cannot stand alone. Therefore, they suggested that a UL algorithm would be particularly helpful for recently trained therapists. However, previous research indicates that prognoses based on clinical expertise are not superior to those of algorithms, even among experienced therapists.[39]

In the present study, participants were positive towards the SAFE test and found it easy to use. The PREP2 algorithm has three steps and requires information from few sources.[7] For approximately 2/3 of the patients, only the SAFE test is needed and according to Connell *et al*, simple prediction algorithms are more likely to be implemented.[15]

Prediction of UL function in severely impaired patients was considered a particular challenge. The participants therefore welcomed new technologies such as TMS, which could help to distinguish patients who regained function from those who remained paralysed. However, their views on prediction algorithms differed depending on the prediction outcome. A favourable prediction was considered motivating for both therapists and patients. By contrast, most participants found a negative prediction demotivating. This is in line with the findings of

Connell *et al*, who point to that individualised prediction as a new field for therapists, and that negative predictions may be particularly challenging.[15] According to Connell *et al*, therapists may need assistance in delivering negative predictions[15] and the PRESTO homepage as well as the PREP2 training homepage by Two homepages provide suggestions on how to phrase and deliver negative predictions.[13 14]

Some participants considered the 75% accuracy of the PREP2 algorithm to be insufficient to improve clinical decision-making. To enable more precise predictions, participants in the present study proposed combining PREP2 with other sources of prognostic information. However, many factors the participants claimed to be relevant for UL prognosis have, in fact, already been examined.[6 7] During the development of the PREP2 algorithm, Stinear *et al* showed that the most important predictors to incorporate in PREP2 were MEP status, SAFE score and NIHSS.[7] The predictors irrelevant to include were sex, hemisphere affected, hand affected, stroke classification, thrombolysis and previous stroke. UL outcome was not predicted by these factors nor was it modified by UL therapy dose.[7] In the present study, participants also mentioned the importance of personality traits that are not easily quantifiable, for example, inner drive and the approach to life. Such personality traits have so far received little attention in research and may need further investigation.

Although the three units were part of the same rehabilitation hospital, different cultures existed, for example, expressed in participants from units 2 and 3 being more sceptic to implementation of a UL algorithm. A reason for this might be greater focus on UL prediction algorithms prior to the interviews at unit 1 resulting in more knowledge and experience with UL prediction algorithms than in the other two units. Furthermore, some of the observed differences in culture may be attributed to the characteristics of the participants and the site. Participants who had graduated more recently who were employed at the largest site and in proximity to the research unit were more prone to a positive attitude towards prediction algorithms. Differences in culture stress the importance of tailoring a future implementation to the particular setting in which it is intended to be used. This is in accordance with similar studies that focus on the importance of translating and implementing research into practice.[16–18]

The present study was the first step in an implementation process. By identifying potential barriers and facilitators, these can be addressed. CFIR was used as an overarching framework, ensuring that aspects relevant to future implementation were systematically captured. Emphasis was on the participants' perspectives and the CFIR constructs to which they could easily relate: intervention characteristics, inner setting and characteristics of the individuals involved. This approach was chosen as the beliefs of healthcare staff about an intervention are often more influential in implementation than other

factors such as the strength of evidence supporting the intervention.[16 17 19]

Limitations and strengths

The scientific trustworthiness of the present study was evaluated using the concepts credibility, confirmability and transferability.[23–25] Also, a checklist for focus group interviews was used to assure that important aspects considering the research team, study methods, context of the study, analysis and interpretations were addressed.[26]

Credibility was ensured by involving several researchers with different positions and perspectives who could supplement and challenge each other in the analysis. According to Malterud, multiple researchers can in this way strengthen a study.[23 25] To further ensure credibility, a participant from each focus group interview reviewed transcripts and findings. To assure confirmability, we aimed not to let our preunderstandings influence the interpretation of the findings.[23] Being aware of our own preconceptions is essential as this enabled us to analyse the interview transcripts with open minds.[24 25]

Transferability or generalisability concerns the application of the study findings beyond the context in which the study was undertaken.[25] Perceived barriers and facilitators for implementation will differ between sites, depending on the characteristics of the clinical setting and the people involved.[15 17] As a consequence, findings from the current study will not necessarily be generalisable to other settings. Nevertheless, the systematic use of CFIR as a framework even before the start of implementation can be transferred to other contexts.

A limitation of the present study might be that the ward managers selected participants to volunteer for the present study. Therapists with an interest in UL algorithms or implementation may be more eager to participate. If so, the expressed perceptions towards UL prediction algorithms may well be more positive than what is the case among therapists in general.

Another limitation of the present study is that the participants did not try to perform the complete PREP2 before attending the interviews. PREP2 is comprised of the SAFE test, information on age, NIHSS score and MEP status. While the NIHSS score or the equivalent SSS score is always performed at the acute units and can be found in the medical records, knowledge of MEP status is on the contrary not easily obtainable. The performance of TMS to obtain MEP status requires a longer period of training. Using this amount of resources was not feasible for the present study, thus, only the first part of PREP2 was practised. For this reason, the participants' thoughts on TMS are merely theoretical. However, TMS is expensive to purchase and requires ongoing training of staff to operate and therefore constitutes an obvious barrier to a future implementation of PREP2. Still, the participants practised the most essential part of the PREP2, the SAFE test, before attending the interviews.

Future directions

The present study reveals that the perceptions of the participants only partly align with current scientific evidence, reflecting a lack of translation from evidence to applied knowledge. Connell *et al* state that the beliefs of healthcare staff about interventions are often more influential than other factors such as the strength of evidence for the intervention.[19] As a consequence, the evidence behind UL prediction algorithms should be presented and discussed in more detail with the therapists prior to implementation. In this context, it is important to consider that prediction algorithms should not be implemented in clinical practice until both development and validation studies have been conducted.[6] Before implementation of a UL prediction model in our local setting, the results from a recent study on the accuracy of PREP2 when obtained 2 weeks post-stroke[22] should be combined with the knowledge obtained from the current study.

Conclusion

In the present study, we found that experienced neurological therapists knew about UL prediction algorithms. However, only a few had a profound knowledge and few were using the SAFE test. The participants regarded algorithms as potentially useful tools and particularly relevant for recently qualified therapists and for patients with little or no UL function. They were positive about using the two main components in the PREP2 algorithm, the SAFE score and TMS.

Performance of the SAFE score aligned more with the physiotherapy profession than the occupational therapy profession. If PREP2 is to be implemented, PTs may be the ones performing the algorithm. A future implementation strategy should address how to support therapists in handling and delivering predictions, especially if they are negative.

accordance with Danish legislation on research ethics, and due to the nature of the study, approval by the Research Ethics Committee was not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplemental information. The Danish Data Protection Agency's terms and conditions were complied with, and the data were deleted at the end of the study.

ORCID iD
Camilla Biering Lundquist http://orcid.org/0000-0001-6836-2233

REFERENCES

1 Feigin VL, Krishnamurthi RV, Parmar P, *et al*. Update on the global burden of ischemic and hemorrhagic stroke in 1990-2013: the GBD 2013 study. *Neuroepidemiology* 2015;45:161–76.
2 Benjamin EJ, Blaha MJ, Chiuve SE, *et al*. Heart disease and stroke statistics—2017 update: a report from the American heart association. *Circulation* 2017;135:e146–603.
3 Faria-Fortini I, Michaelsen SM, Cassiano JG, *et al*. Upper extremity function in stroke subjects: relationships between the International classification of functioning, disability, and health domains. *J Hand Ther* 2011;24:257–65.
4 Patel AT, Duncan PW, Lai S-M, *et al*. The relation between impairments and functional outcomes poststroke. *Arch Phys Med Rehabil* 2000;81:1357–63.
5 Stinear CM, Byblow WD. Predicting and accelerating motor recovery after stroke. *Curr Opin Neurol* 2014;27:624–30.
6 Kwah L, Herbert R. Prediction of walking and arm recovery after stroke: a critical review. *Brain Sci* 2016;6:E53.
7 Stinear CM, Byblow WD, Ackerley SJ, *et al*. Prep2: A biomarker-based algorithm for predicting upper limb function after stroke. *Ann Clin Transl Neurol* 2017;4:811–20.
8 Winters C, van Wegen EEH, Daffertshofer A, *et al*. Generalizability of the proportional recovery model for the upper extremity after an ischemic stroke. *Neurorehabil Neural Repair* 2015;29:614–22.
9 Byblow WD, Stinear CM, Barber PA, *et al*. Proportional recovery after stroke depends on corticomotor integrity. *Ann Neurol* 2015;78:848–59.
10 Hayward KS, Schmidt J, Lohse KR, *et al*. Are we armed with the right data? Pooled individual data review of biomarkers in people with severe upper limb impairment after stroke. *Neuro Image Clin* 2016;13:310–9.
11 Kim B, Winstein C. Can neurological biomarkers of brain impairment be used to predict poststroke motor recovery? A systematic review. *Neurorehabil Neural Repair* 2017;31:3–24.
12 Stinear CM, Byblow WD, Ackerley SJ, *et al*. Predicting recovery potential for individual stroke patients increases rehabilitation efficiency. *Stroke* 2017;48:1011–9.
13 Presto PREP2. Available: https://presto.auckland.ac.nz/prep2-overview/ [Accessed 9 Sep 2020].
14 Prep2 training. Available: https://preptraining.auckland.ac.nz/ [Accessed 1 Jul 2021].
15 Connell LA, Smith M-C, Byblow WD, *et al*. Implementing biomarkers to predict motor recovery after stroke. *NeuroRehabilitation* 2018;43:41–50.
16 May C. Towards a general theory of implementation. *Implement Sci* 2013;8:18.
17 Damschroder LJ, Aron DC, Keith RE, *et al*. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement Sci* 2009;4:50.
18 Varsi C, Ekstedt M, Gammon D, *et al*. Using the consolidated framework for implementation research to identify barriers and facilitators for the implementation of an Internet-based patient-provider communication service in five settings: a qualitative study. *J Med Internet Res* 2015;17:e262.
19 Connell LA, McMahon NE, Harris JE, *et al*. A formative evaluation of the implementation of an upper limb stroke rehabilitation intervention in clinical practice: a qualitative interview study. *Implement Sci* 2014;9:3.
20 Krueger R, Casey M. *Planning the focus group study. focus groups: a practical guide for applied research*. 5th edn. SAGE Publications, 2015: 19–37.
21 Wilkinson S. Focus groups in health research: exploring the meanings of health and illness. *J Health Psychol* 1998;3:329–48.
22 Lundquist CB, Nielsen JF, Arguissain FG, *et al*. Accuracy of the upper limb prediction algorithm PREP2 applied 2 weeks poststroke: a prospective longitudinal study. *Neurorehabil Neural Repair* 2021;35:68–78.
23 Öhman A. Qualitative methodology for rehabilitation research. *J Rehabil Med* 2005;37:273–80.
24 Elo S, Kyngäs H. The qualitative content analysis process. *J Adv Nurs* 2008;62:107–15.
25 Malterud K. Qualitative research: standards, challenges, and guidelines. *Lancet* 2001;358:483–8.
26 Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *Int J Qual Health Care* 2007;19:349–57.
27 Kiær C, Lundquist CB, Brunner I. Knowledge and application of upper limb prediction models and attitude toward prognosis among physiotherapists and occupational therapists in the clinical stroke setting. *Top Stroke Rehabil* 2021;28:1–7.
28 Sundhedsstyrelsen. *Anbefalinger til nationale redskaber til vurdering af funktionsevne - hos voksne med erhvervet hjerneskade*, 2020.
29 Sundhedsstyrelsen. *Anbefalinger for tværsektorielle forløb for voksne Med erhvervet hjerneskade*, 2020.
30 Pohl J, Held JPO, Verheyden G, *et al*. Consensus-based core set of outcome measures for clinical motor rehabilitation after stroke—a Delphi study. *Front Neurol* 2020;11:875.
31 Kwakkel G, Lannin NA, Borschmann K, *et al*. Standardized measurement of sensorimotor recovery in stroke trials: consensus-based core recommendations from the stroke recovery and rehabilitation roundtable. *Int J Stroke* 2017;12:451–61.
32 Gebruers N, Truijen S, Engelborghs S, *et al*. Prediction of upper limb recovery, general disability, and rehabilitation status by activity measurements assessed by accelerometers or the Fugl-Meyer score in acute stroke. *Am J Phys Med Rehabil* 2014;93:245–52.
33 Prabhakaran S, Zarahn E, Riley C, *et al*. Inter-individual variability in the capacity for motor recovery after ischemic stroke. *Neurorehabil Neural Repair* 2008;22:64–71.
34 Buch ER, Rizk S, Nicolo P, *et al*. Predicting motor improvement after stroke with clinical assessment and diffusion tensor imaging. *Neurology* 2016;86:1924–5.
35 Winstein C, Varghese R. Been there, done that, so what's next for arm and hand rehabilitation in stroke? *NeuroRehabilitation* 2018;43:3–18.
36 Jones F, Pöstges H, Brimicombe L. Building Bridges between healthcare professionals, patients and families: A coproduced and integrated approach to self-management support in stroke. *NeuroRehabilitation* 2016;39:471–80.
37 Mayo NE, Fellows LK, Scott SC, *et al*. A longitudinal view of apathy and its impact after stroke. *Stroke* 2009;40:3299–307.
38 Subramanian SK, Chilingaryan G, Sveistrup H, *et al*. Depressive symptoms influence use of feedback for motor learning and recovery in chronic stroke. *Restor Neurol Neurosci* 2015;33:727–40.
39 Nijland RH, van Wegen EE, Harmeling-van der Wel BC. Early prediction of functional outcome after stroke I. accuracy of physical therapists' early predictions of upper-limb function in hospital stroke units: the EPOS study. *Phys Ther* 2013;93:460–9.