Leiden University
Medical Center

iPluto – international Plexus Outcome Study Group

First Results of the Delphi surveys

Willem Pondaag MD PhD

dept Neurosurgery / Leiden Nerve Center

LEIDEN UNIVERSITY MEDICAL CENTER (NL)

ipluto@lumc.nl

Rationale

- OBPL is a rare condition
- for deciding on optimal treatment strategy pooling of data is necessary
- previous attempts:

Brachial Plexus Birth Palsy: Rationale for a Multicenter Prospective Study

Peter M. Waters, M.D.[1] and Donald S. Bae, M.D.[1] 2004

Outcome assessment

- many different and incompatible scoring systems in use
- all clinics have their own evaluation forms & follow-up structure

Hopefully in the near future we can:

- reach consensus on data collection (at least for a part of the data: shared minimal dataset)
- compare published results
- pool patient data to improve statistical analysis
- improve patient care with better outcome data

Systematic reviews

J Neurosurg Pediatrics 12:395–405, 2013

A systematic review of evaluation methods for neonatal brachial plexus palsy

A review

*Kate W. C. Chang, M.A.,[1] Denise Justice, B.S.,[1] Kevin C. Chung, M.D., M.S.,[2] and Lynda J. S. Yang, M.D., Ph.D.[1]

Journal of Pediatric Rehabilitation Medicine: An Interdisciplinary Approach 8 (2015) 167–186
DOI 10.3233/PRM-150335
IOS Press

Review Article

Outcome measures used in clinical studies on neonatal brachial plexus palsy: A systematic literature review using the International Classification of Functioning, Disability and Health

Cigdem Sarac[a,*], Bouke J. Duijnisveld[a], Amber van der Weide[a], Jan W. Schoones[b], Martijn J.A. Malessy[c], Rob G.H.H. Nelissen[a] and Thea P.M. Vliet Vlieland[a]
[a]*Department of Orthopaedic Surgery, Leiden University Medical Center, Leiden, The Netherlands*
[b]*Walaeus Library, Leiden University Medical Center, Leiden, The Netherlands*
[c]*Department of Neurosurgery, Leiden University Medical Center, Leiden, The Netherlands*

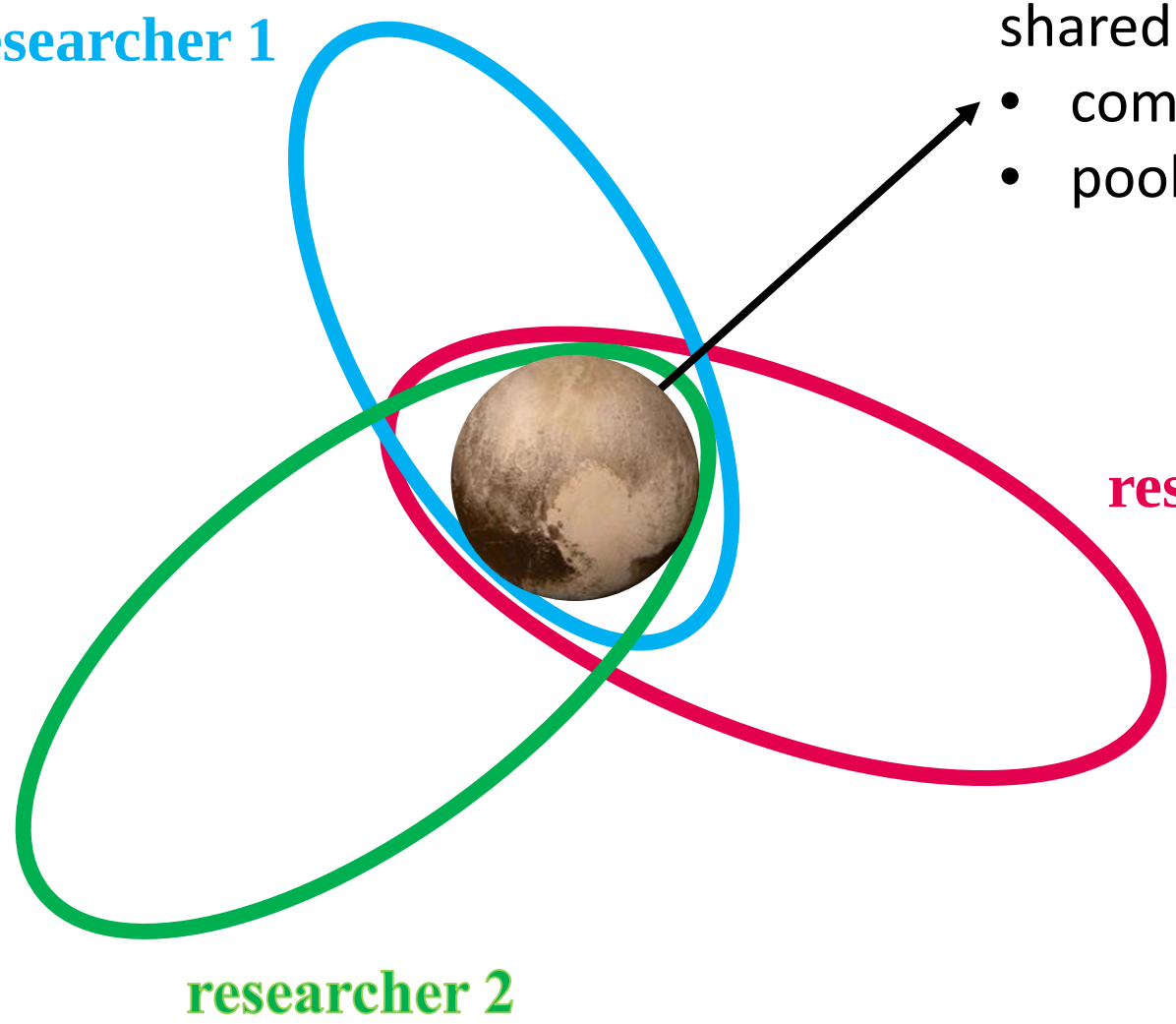
Sarac et al. (2015)

In total 59 different outcome measures were identified in 217 studies.

The most frequently used outcome measures included

- range of motion of the shoulder (*n = 166 studies,* 76%)
- range of motion of the elbow (*n = 87 studies,* 40%)
- the Mallet scale (*n = 66 studies, 30%)*
- Magnetic Resonance Imaging (MRI) (*n = 37 studies, 17%)*
- and the British MRC motor grading scale (*n = 31 studies, 14%)*

Minimal shared dataset



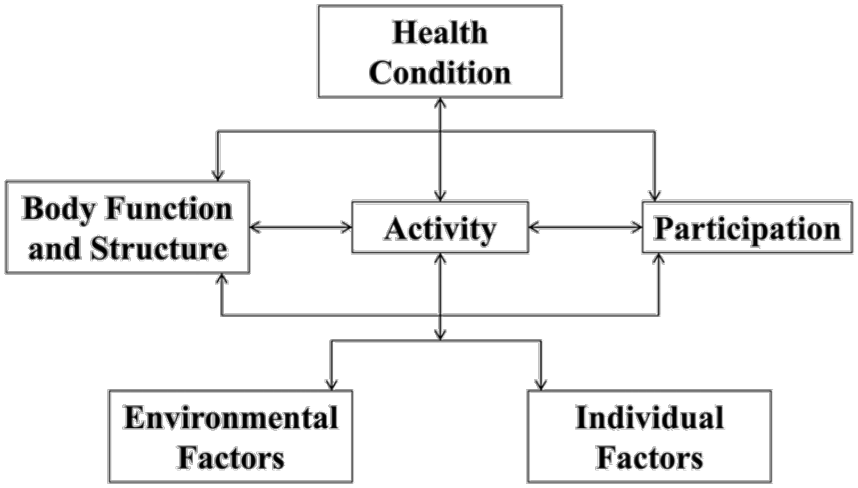
shared items for
- comparison
- pooling

iPluto: Minimal shared dataset

need for a ‘minimal shared dataset’ for pooling and comparing of data

- movement / range of motion
- force
- scoring systems (eg Mallet-score, Raimondi Hand Score, BPOM)
- sensation
- functional outcome (PROMs)

inclusion of all domains of the ICF



collected at standardized time points
e.g. 3 years / 5 years / 7 years

defined by **consensus**

iPluto

Launch of the project:

Narakas-meeting Barcelona, February 2016

“Let’s speak the same language”

narakas
barcelona
2016

Assumption

Consensus on methodology increases participation (?)

How to reach consensus

Methods

- informal
- consensus development conference
- guideline development (GRADE recommendations)
- Nominal group technique (NGT)
- Delphi
- modified NGT (RAND) ~ modified Delphi

Differences

- use of questionnaires
- face-to-face contact
- structure of group interaction

Delphi Survey

Named after the famous oracle at Delphi

Delphi survey:

- a group facilitation technique,
- which is an iterative multistage process,
- designed to transform opinion into group consensus

Assumption: group judgments are more valid than individual judgments.

Delphi Survey

The Delphi method was developed at the beginning of the Cold War to forecast the impact of technology on warfare.

Developed later as a systematic forecasting method which relies on a panel of experts, e.g. long-term trends in science and technology development, economic forecasting.

Use in healthcare: design of Quality Indicators, Outcome Measures.

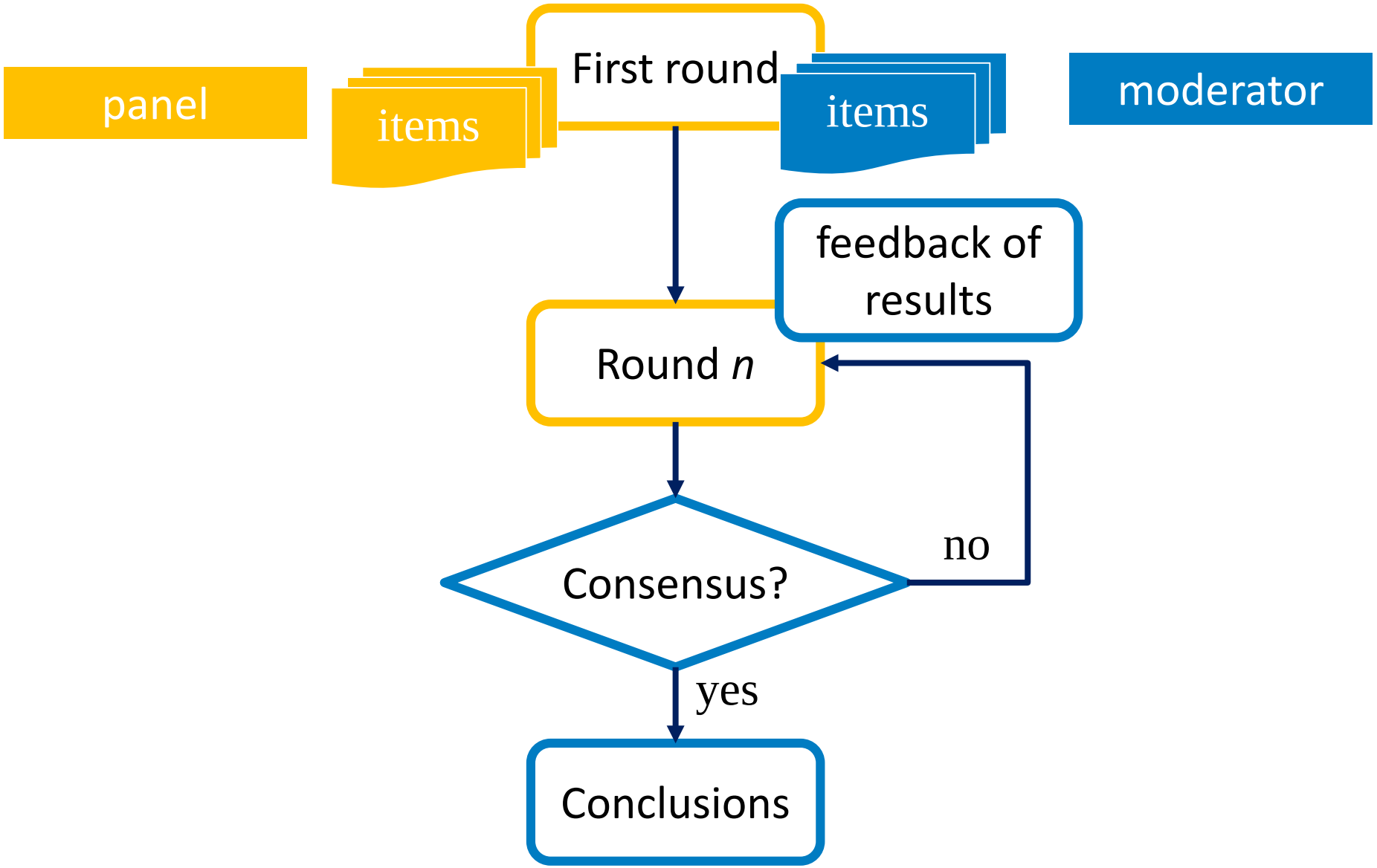
Eminence based Medicine

"All those in favor say 'Aye.'"

"Aye." *"Aye."* *"Aye."*

"Aye." *"Aye."*

Flowchart of Delphi Surveys



Key Points of the Delphi process

- Anonymity of the participants
- Structuring of information flow
- Regular feedback
- Role of the facilitator

iPluto

panel selection

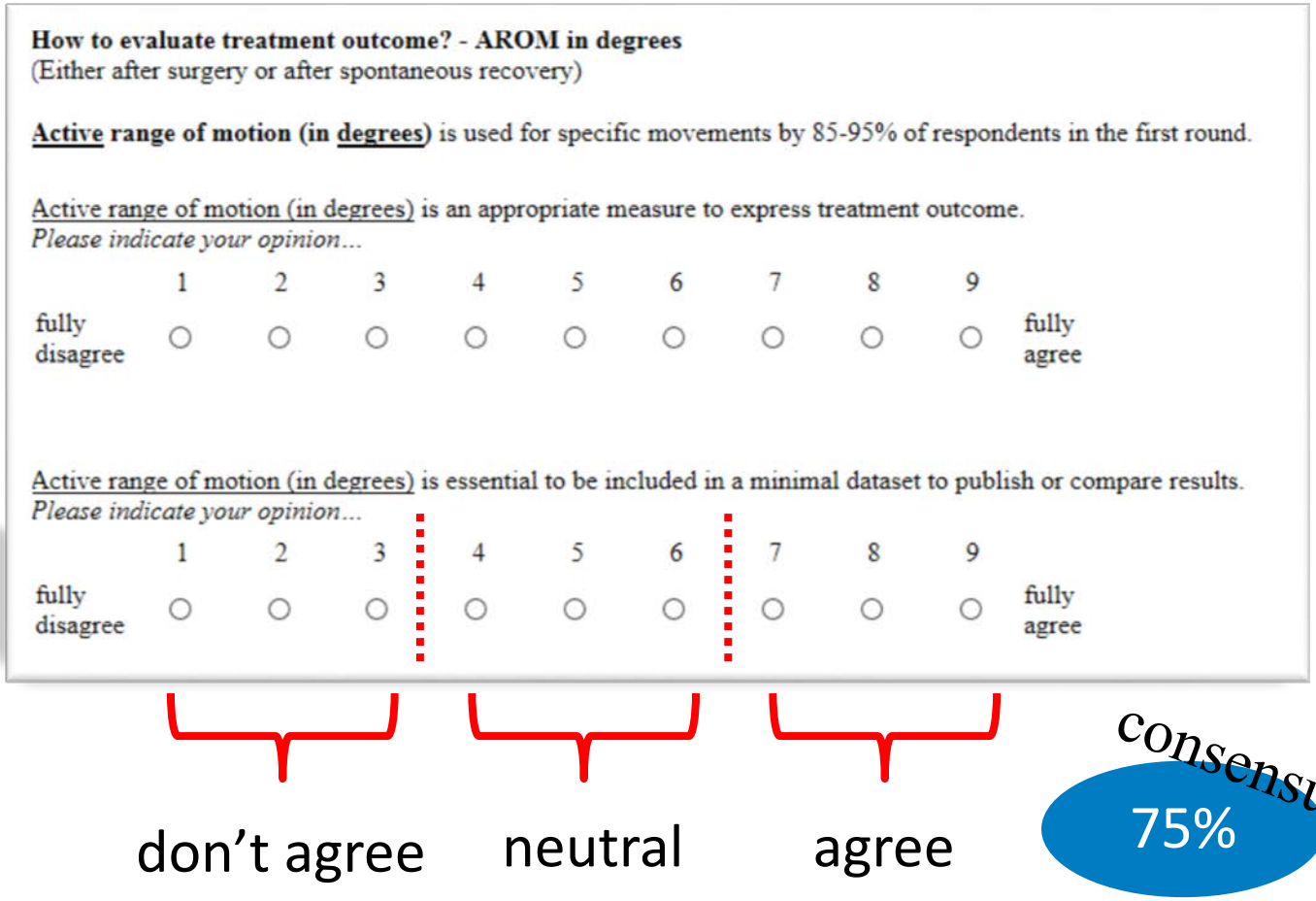
- email addresses of participants of
  - Narakas meetings 2011 / 2016
  - Toronto Obstetrical Brachial Plexus Palsy Workshop 2014
- call out for participation during Narakas meeting 2016
- literature / own mail address book
- first email encouraged forwarding to interested colleagues

~ 300 email addresses

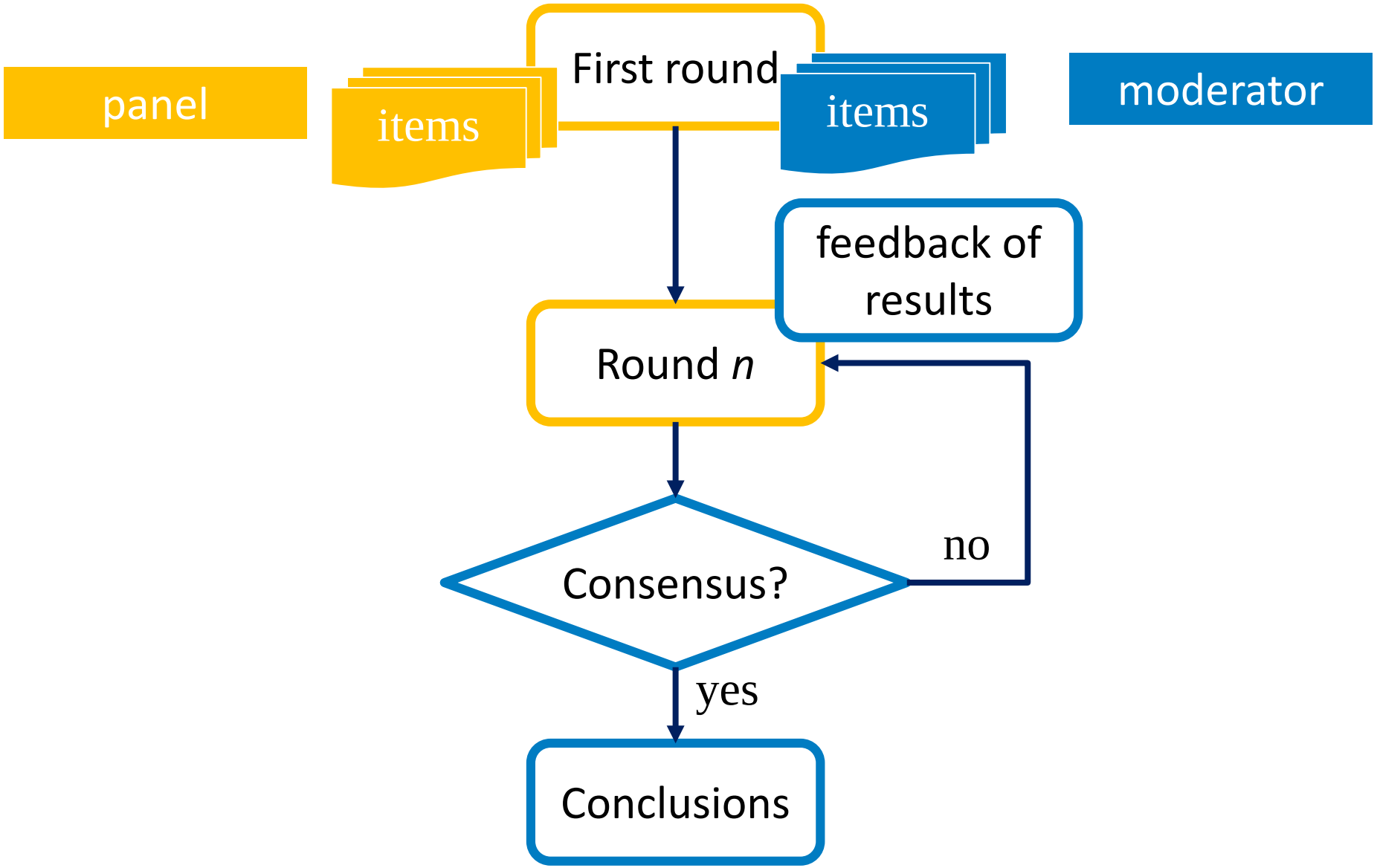
Methods

First Round: inventory and open end questions for suggestion of items

subsequent rounds: evaluation of items using a 9 point Likert-scale



Flowchart of Delphi Surveys



Response rate

First round:

300 invitations sent

20 email addresses proved to be outdated or false

107 responders

- 27 completely empty (just clicked the link)
- 3 very incomplete
- 9 double from the same center
- 68 analyzed

Participants

Round 1: June 2016 – September 2016

Round 2: September 2016 – November 2016

Round 3: December 2016 – February 2017

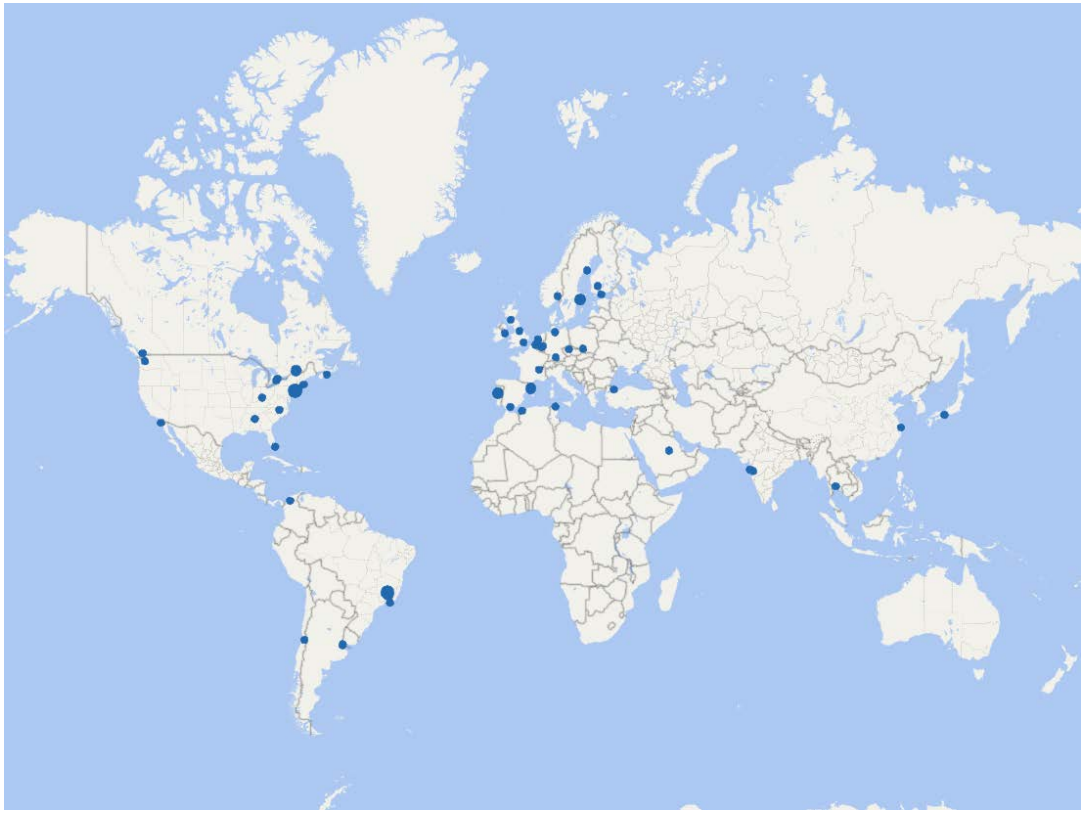
n	Round1	Round2	Round3
13	+	-	-
5	+	+	-
2	+	-	+
49	+	+	+
10	-	+	+
4	-	+	-
	69	68	61

analysis n=59

Participants Round 2 & 3

n=59

Europe	25
North-America	15
South-America	10
Asia	7
Africa	2



Methods

How to evaluate treatment outcome? - AROM in degrees
(Either after surgery or after spontaneous recovery)

Active range of motion (in degrees) is used for specific movements by 85-95% of respondents in the first round.

Active range of motion (in degrees) is an appropriate measure to express treatment outcome.
Please indicate your opinion…

	1	2	3	4	5	6	7	8	9	
fully disagree	○	○	○	○	○	○	○	○	○	fully agree

Active range of motion (in degrees) is essential to be included in a minimal dataset to publish or compare results.
Please indicate your opinion…

	1	2	3	4	5	6	7	8	9	
fully disagree	○	○	○	○	○	○	○	○	○	fully agree

don't agree neutral agree

consensus
75%

Results (examples)

	Round2			Round3		
	mean	7-9	7-9 (%)	mean	7-9	7-9 (%)
Active Range of Motion in degrees						
…is an appropriate outcome measure.	8,76	58	98%			
…is essential to be included in a minimal dataset.	8,44	56	95%			

consensus

	Round2			Round3		
	mean	7-9	7-9 (%)	mean	7-9	7-9 (%)
The Gilbert Shoulder Score						
…is an appropriate outcome measure.	4,20		19%	3,44	9	15%
…is essential to be included in a minimal dataset.	3,73			3,12	9	15%

no consensus

Results (examples)

	Round2			Round3		
	mean	7-9	7-9 (%)	mean	7-9	7-9 (%)
Active Range of Motion (in degrees) **Items**						
ExtRot_abd	7,61	44	75%			
ExtRot_add	8,44	56	95%			
Abd	8,61	58	98%			
IntRot	7,25	39	66%	5,25	28	47%
ElbFlex	8,68	58	98%			
ElbExt	7,98	50	85%			
Sup	7,49	43	73%	6,46	37	63%
Pron	7,19	40	68%	5,95	32	54%
WriFlex	7,07	39	66%	5,19	26	44%
WriExt	7,98	55	93%			
FiFlex	7,90	49	83%			
FiExt	7,92	52	88%			
ThFlex	7,29	41	69%	4,97	24	41%
ThExt	7,46	44	75%	5,41	28	47%

How to measure Treatment Outcome

… is an appropriate measure to express treatment outcome. (1-9)	7/8/9 (%)
Passive range of motion (in degrees)	68%
Active range of motion (in degrees)	**98%**
Active range of motion (in AMS)	61%
Muscle Force	69%
Mallet Score	**83%**
Gilbert Elbow Score	15%
Raimondi Hand Score	41%
BPOM Brachial Plexus Outcome Measure	39%
AHA Assisting Hand Assessment	29%
Nine hole peg test	8%
Testing sensibility with Semmes Weinstein filaments	31%
Testing sensibility with two point discrimination	29%
Using Pain Questionnaires	42%

PROMs

	Round2		
	mean	7-9	7-9 (%)
Patient Reported Outcome Measures (PROMs)			
…is an appropriate outcome measure.	5,46	25	42%
…is essential to be included in a minimal dataset.	5,15	21	36%
I have sufficient knowledge to judge different PROMS	3,56	11	19%

Timing of evaluation

Standardized time points should be used to compare results.

iPluto proposed to use the age of the infant, and not the follow-up time after a specific intervention.

The first proposal was to evaluate at the age of 1 / 3 / 5 / 7 years.

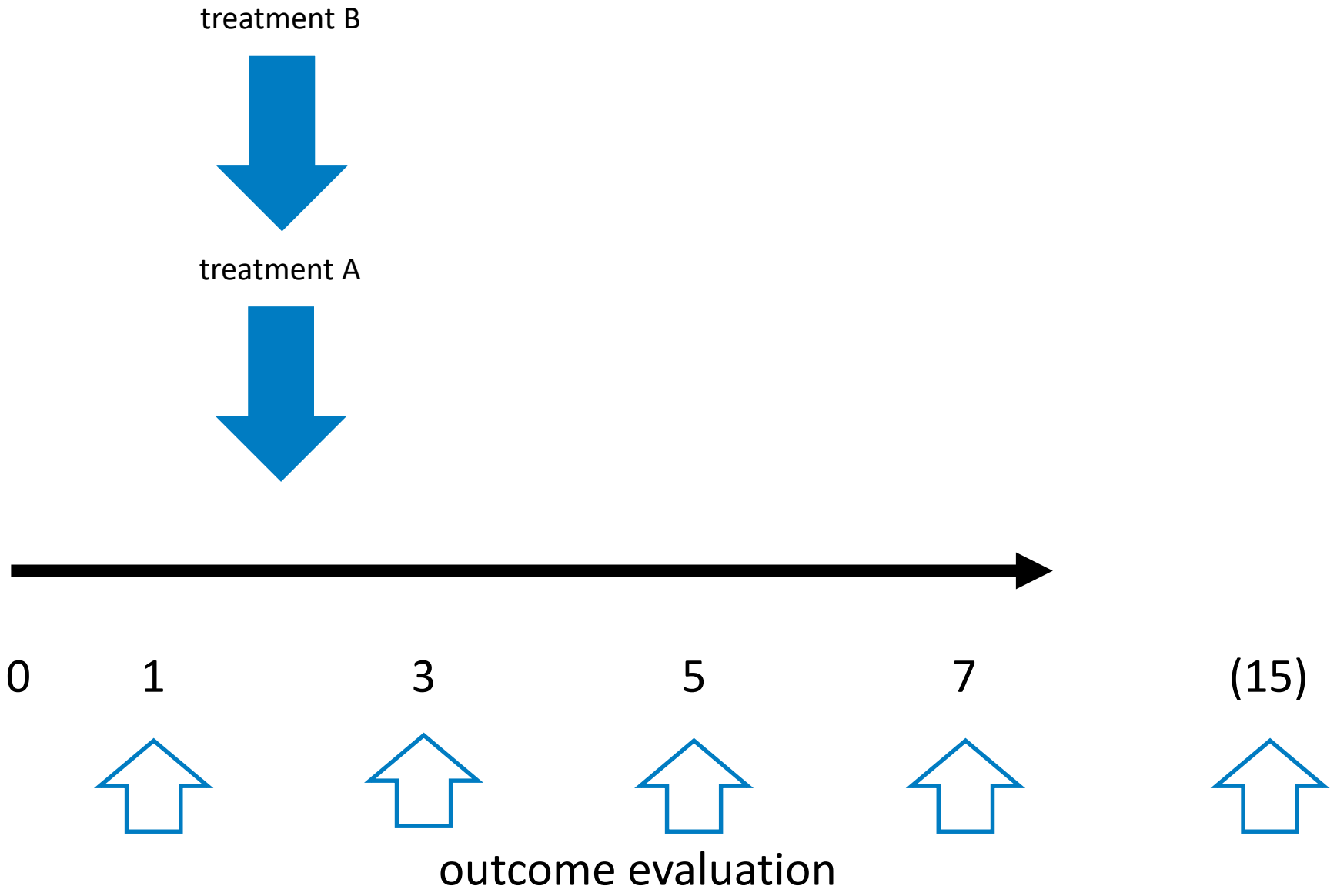
In the first round 63/68 (93%) supported this concept.

Many participants suggested to add a time point at 2 years of age, and one as teenager, e.g. at 15 years of age.

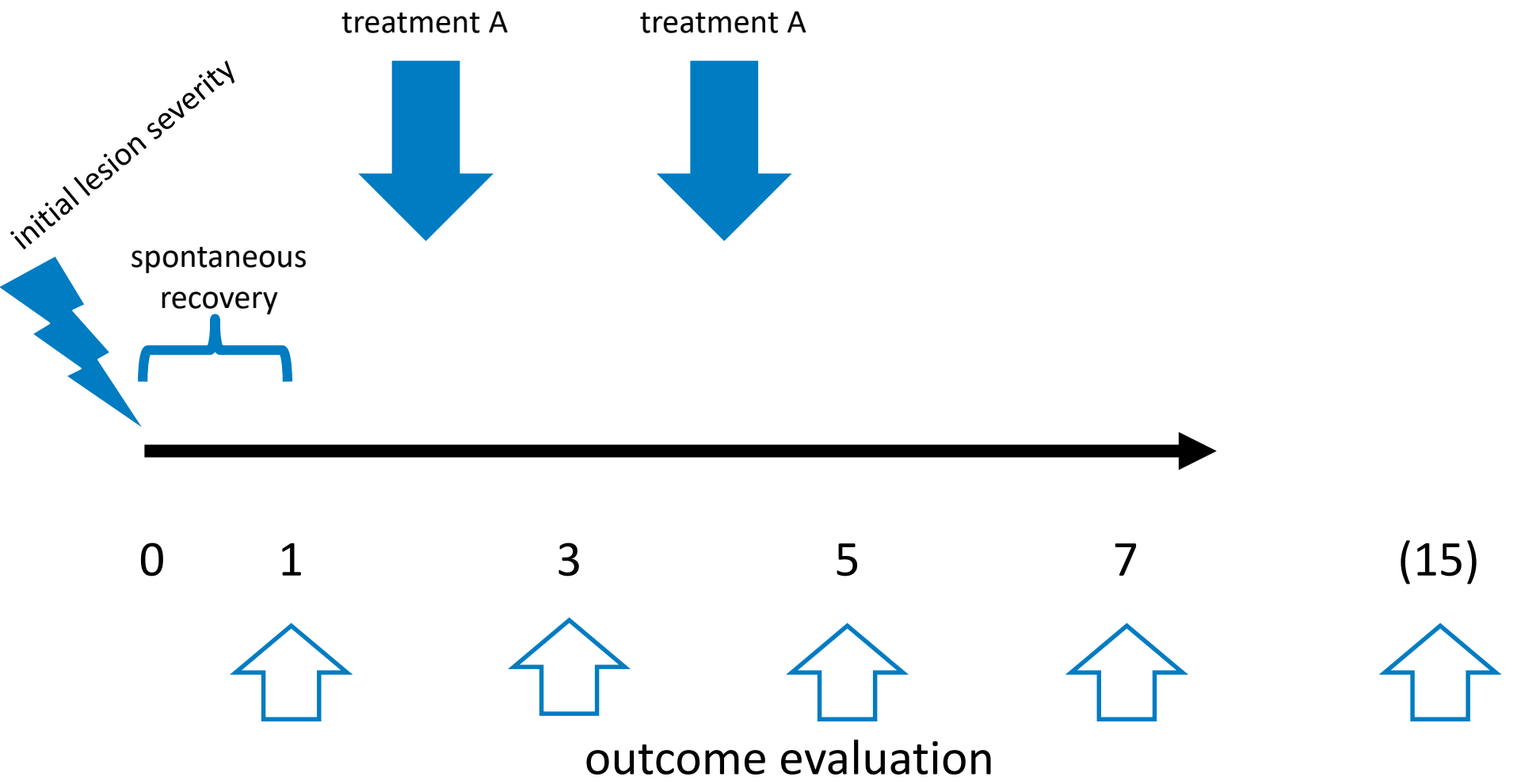
In Round 2/3: >75% consensus for evaluation at the age of 1 / 3 / 5 / 7 years.

For evaluation at 15 years, 74.6% agreed, nearly consensus.

Stratify and correct



Stratify and correct



Stratify and correct

To stratify or correct outcomes for pooling or comparison of series from different centers, it is necessary to

- assess initial lesion severity
- spontaneous recovery

Initial lesion severity

"… is suitable to express lesion severity"

Consensus on

- the Narakas Classification (which should be assessed at 1 month)
- recovery of elbow flexion
- time to recovery of elbow flexion

No consensus on

- elbow strength (MRC)
- Toronto Test Score
- Cookie Test

No consensus that MRI/CTM is essential to assess root avulsions, but participants agreed that the number of root avulsions is appropriate to express lesion severity.

Spontaneous recovery

“Which key-movements should be serially investigated at 1-3-6-9 months..”

Serial evaluation in the first year of life to record spontaneous recovery should include, consensus was reached:

- External rotation (measured in adduction)
- Abduction
- Elbow flexion
- Wrist extension
- Finger flexion
- Finger extension

Conclusion

There is consensus among participants to iPluto

- how to express initial lesion severity
  - Narakas classification at 1 month
  - recovery of elbow flexion and timing of recovery of elbow flexion
- how to document spontaneous recovery
  - serial evaluation of key movements at 1-3-6-9 months
- when to assess outcome
  - at the age of 1-3-5-7(-15) years
- which items should be minimally assessed
  - movement in degrees: abduction, external rotation, elbow flexion, wrist extension, finger flexion, finger extension
  - Mallet scores (as subscores)

Beware…

absence of consensus to include

≠

consensus not to include

Next steps

- no new rounds, as responses did not change between Rounds 2 and 3

- publication of the results (preferably OpenAccess)
- promote the use of the minimal dataset defined by consensus

- check your own protocol for missing items of the minimal dataset
- use the minimal dataset, keep using any other outcome measures as you wish
- start using standardized time points

- specific survey on PROMs – to follow soon

consensus conference: Narakas meeting 2019 in Leiden (?)

- report your experiences
- further refining

https://ipluto.org/
✉ *ipluto@lumc.nl*