Note: Green items reached consensus in the e-Delphi; black items did not reach consensus

Discussion session 1: Introduction

Checklist item		Consensus: % agreement with scores 7-9. **Second round,**(first round)	Item specific comments: First round; **Second round**
3	**Identify the aspect of care that the new service being implemented aims to address (e.g. implementing a guideline recommendation or evidence-based management)**	**90%,**(89%)	
4	**Critically report the evidence underpinning the new service to be implemented: (e.g. phase III randomised controlled trials, systematic review, guideline recommendations)**	**100%,** (90%)	
5	**Include a description of the wider healthcare/policy/commercial context**	**58%,** (55%)	
6	**Describe the rationale for the new service design**	**(95%),** (85%)	• re 6. - this may also be an important aspect of the research i.e. to understand the rationale for the new service design - and how these may differ between different stakeholders
7	**Report the implementation strategy used and its underpinning theory**	**(84%)** (80%)	• Item #7 might be rather part of the methodology section. However, its underpinning theory is rather part of the introduction. So, this might be split: methodology in that section, but implementation theory in the intro. • For # 7 - theory could be theories and might change wording to say and mention it underpinning theory(ies) • implementation strategy might be better placed in methods • **#7 Report the implementation strategy used and its underpinning theory: for me this may also be described in the methods section. It should be somewhere in the paper**
8	**Describe any pilot implementation work and the conclusions from that work**	**(63%),** (60%)	• Item #8 should be included if available. It adds to why the study needed to been done.
9	**Clearly define the aims of the study, differentiating between implementation (process) objectives and**	**(100%),** (90%)	• Re question 9 - depending on the nature of intervention - the outcomes may not be clinical (e.g. if an organisational

	effectiveness (clinical) objectives aims		intervention the summative outcomes may be service design orientated or practitioner behaviour orientated). I am making an assumption the standards that result from this process should be as widely applicable as possible. • re 9. - this assumes as per round 1 that the implementation of the new service is carried out at the same time as the evaluation and by same people - it may not
Wider context	• This part describes the background of any study in which items #3-6 and #9 should be included. Also the (non-)involvement of stakeholders should be addressed. Relates to item #10.		

Discussion sessions 2: Methods part I (Setting, The new service)

Checklist item (Setting)		Consensus: % agreement with scores 7-9. **Second round,**(first round)	**Item specific comments:** First round**; Second round**
10	**Describe the study setting (including health service, personnel involved, patient and public involvement, demography of patients etc)**	**(100%),** (100)	• PPI: describe at what level (tokenism; "subject" = volunteer; advisor). • Ideally, the study design includes the interactive collaboration and dialogue with stakeholders for a priori study improvement, study progress issues and discussion on outcome/valorisation. If not included, this should be explained. • Again, q10. - + these issues may form part of the research
11	**Give year(s) during which the new service was implemented (i.e. planned, initiated and actively developed) and followed up**	**(95%),** (80)	• q.11 - assumes this is straightforward - implementation usually messy and happens over a period of time - that needs to be acknowledged • **As outcomes change over time for many diseases this is essential to my mind** • **the implementation of the new service in which period might be useful given the current changes in healthcare. In the stated period the implementation might have been more (un)successful than in another timeframe.**
	?include an formal assessment made of the context before deciding on an implementation intervention		
	Ultimately, the checklist that results from this delphi process will require specific examples from the implementation science literature to inform authors/editors how to fulfill these recommendations. For example, "Describe the study setting" seems quite vague currently, but is a very important element of transparent reporting in D&I research.		
	OK, I am persuaded...		
Checklist item (The new service)		Consensus: % agreement with scores 7-9. **Second round,**(first round)	**Item specific comments:** First round**; Second round**
12	**Describe the new service (e.g components/content, frequency, duration, intensity, mode of delivery, materials used) with advice on accessing additional detailed information. Use of a standardised checklist**	**(100%),** (100)	

	(e.g. TIDieR) is recommended.		
13	**Describe the professional backgrounds, roles and training requirements of the personnel involved in delivering the intervention with advice on accessing additional detailed information.**	**(84%),** (65)	• part of items (#13, #17) can also be described in addendum. A brief description in the main methodology section, and more details in the addendum. • **Professional training has a wide spread but without this information it is really hard to know if staff in your own context has the skill base required to deliver the intervention**
14	**Define the core components of the intervention, and the processes for assessing fidelity to this core content, and what, if any, local adaptation was allowed.**	**(100%),** (90)	• On item 14, fidelity and adaptation allowed are really two different constructs, not necessarily the mirror image of the other. #14 is duplicative of #12. They should be merged.
15	**Describe the intervention received by control/comparator group not simply stating 'usual care'**	**(95%),** (75)	Describe the intervention received by control/comparator This could be difficult as in a cluster trial may be very different from place to place
16	**What is the relation of components of the intervention to the rationale for the new service design and/or theory underpinning implementation discussed above?**	**(30%),** (21)	I'm not sure whether theory should be given in the methodology section. If briefly described (one-two sentences): yes. Otherwise, embedding in the introduction may be better suited.
17	**Define role of the researchers in design and implementation.**	**(79%),** (60)	• 17 is confusing. Do you mean in program d & I or research d & i? 17 possibly only necessary if they were involved
Change over time	Need to allow for change in intervention over time as well as local adaptability – these q's assume new service is fixed in aspic		
Usual care	**The control intervention needs to be described as "usual care" in one clinical setting may differ from another. Components of the new strategy may be part of the "usual care" given in one centre but not in another. It is important to know whether the intervention works/is efficient/effective, but what makes it precisely working needs to be known.**		

Discussion session 3: Methods part II (Population, randomisation, data, analysis)

Checklist item (Population)		Consensus: % agreement with scores 7-9. **Second round,**(first round)	Item specific comments: First round; **Second round**
18	**Describe sites invited/excluded with reasons**	**(100%),** (95)	item #18: in- or exclusion criteria of sites should be given, as well as which sites were included. This may also be part of an addendum to the methodology section
19	**Describe the population targeted by the intervention and any eligibility criteria**	**(100%),** (95)	
20	**Report method by which patients are referred to, or access the new service.**	**(100%),** (90)	• **#21 confusing as writtenâ€¦..what if consent for data collection, but not intervention?**
21	**If applicable, describe any consent required (which should be to the new service and not to research)**	**(53%),** (60)	
22	**Describe recruitment of any sub-groups recruited for additional research-tasks (e.g. questionnaire completion, physiological measures, detailed record analysis)**	**(47%),** (55)	• 22 seems off topic to me if this is in relation to programming. • I don't understand the wording of #22. • **#22: for any new research the sub-groups should be described in detail. Otherwise the study and related analysis is scientifically not sound/repeatable.**
	These items get at but do not precisely describe the characteristics of those who end up participating- and contrasting those who participate with those who decline- at both the setting and the individual patient level		

Checklist item (Randomisation)		Consensus: % agreement with scores 7-9. **Second round,**(first round)	Item specific comments: First round; **Second round**
23	**Description of randomisation (or if not randomised how comparator group was selected)**	**(95%),** (85)	• part of methodology, needed for repetition of the study elsewhere. However, an extensive description fits an addendum to the methodology section, not the main core of this section • Important, but covered in other guidelines

Checklist item (Data)	Consensus: % agreement with scores 7-9. **Second**	Item specific comments: First round; **Second round**

		round,(first round)	
24	**Describe outcome measurements (specifically describing any that are at population level) distinguishing between process and clinical outcomes, health economic data**	**(100%),** (100)	
25	**Describe data collection processes (specifically including methods of extracting routine data).**	**(100%),** (90)	
26	**Describe any processes for quality assurance (especially for use of routine data)**	**(84%),** (65)	• Item #26: to be described in brief terms in main core of this section. Can be described more extensively in addendum. Or may be referred to if described in another scientific or openly available publication. • **#26: may also be given in an addendum to the paper.**
Checklist item (Analysis)		**Consensus: % agreement with scores 7-9. Second round,**(first round)	**Item specific comments:** First round**; Second round**
27	**Describe power calculation and rationale for sample size**	**(100%),** (90)	
28	**Describe methods of statistical analysis (with reasons for that choice) including approach to clustering, handling of missing data, intention to treat analysis, and adjustment for confounders etc)**	**(100%),** (90)	
29	**Specify *a priori* sub-group analyses (e.g. between different sites in a multicentre study, different clinical or demographic populations)**	**(95%),** (85)	• Agree with item #29, but may be brief in main core of this section and extended in addendum. • **#29: any analysis regarding primary and secondary outcome should be defined on beforehand. Other analyses may be derived from the (unexpected?) results. Therefore, I rate this question a bit higher**
	• Above are all important, but covered in other guidelines- should be integrated with others • One of the struggles to develop unique D&I reporting guidelines highlighted by the United States' NIH efforts to do what StaRI is undertaking is pertinent here. There are over 25 guidelines archived on the EQUATOR network already. Some of them have components applicable to D&I research (for example the methods section of the CONSORT Pragmatic trial guidelines). Our NIH group argued whether it would be better to refer D&I investigators to existing publication guidelines sometimes		

Discussion session 4: Results

Checklist item (Population)		Consensus: % agreement with scores 7-9. **Second round,**(first round)	**Item specific comments:** First round**; Second round**
30	**Report the number of sites approached, reasons for non-participation, and characteristics of participating sites**	**(89%),** (75)	• #30 is duplicative of #18. These should be merged.
31	**Report the total eligible population (e.g. number of people with the relevant condition registered with the practice, or eligible for a service), number approached and any exclusions**	**(100%),** (90)	• It is important to try and report the eligible population though this might sometimes be tricky to know (e.g.. when recruiting in the community rather than through the health service). Still, an attempt would be appreciated.
32	**Report participation rate among the eligible population, compare characteristics with the eligible population as a whole, and describe any known reasons for non-participation.**	**(95%),** (75)	• 32 can be difficult data may not be available
33	**Report compliance with/attrition from the service as a process outcome**	**(95%),** (85)	• Item #33 deals with why people were not compliant with the study. This might complicate the research ethics procedure (people may withdraw at any moment w/o giving a reason). • 33. not sure this is essential • q.33 - don't think 'compliance' is appropriate term here - participation would be better
34	**Report details of any subgroups recruited to specific research tasks (e.g. questionnaire completion, physiological testing) as opposed to the clinical service. Compare characteristics of any sub-groups to the whole eligible population**	**(74%),** (70)	
35	**Include a CONSORT diagram (modified as necessary) to illustrate the recruitment of sites, provision of service to patients, and any sub-groups**	**(84%),** (80)	• Item #35: would be profitable. But not necessary per se
	In the ideal situation we like to see all of the data stated above.		

Checklist item (Fidelity)	Consensus: % agreement with scores 7-9. **Second round,**(first round)	**Item specific comments:** First round**; Second round**

36	**Report fidelity to the core components of the planned intervention (including, in multicentre studies, in the different settings)**	**(100%),** (85)	• Important, but fidelity is a very complex construct and probably needs reporting standards of its own • Again, good to report fidelity but if measuring it might affect it (ie. the act of observation changes fidelity) then perhaps it's less meaningful than we might think. Important to try if possible though.
37	**Report any modifications or adaptations to the new service during the course of the study**	**(100%),** (95)	• #37 - would add significant before modifications need to provide an analyis of why as well as what modifications/adaptations made
	Overlap with earlier items- very important to report both fidelity and adaptations/ variations separately and non judgmentally		
Checklist item (Outcomes)		**Consensus: % agreement with scores 7-9. Second round,**(first round)	**Item specific comments:** First round**; Second round**
38	**Report outcomes for the whole eligible population, before an analysis of any sub-groups**	**(100%),** (85)	
39	**Report process and clinical outcomes**	**(100%),** (95)	
40	**If relevant, report impact on use of health service resources (and ideally cost of the intervention)**	**(84%),** (70)	• Item #40: involves HTA. This can be included in the study design, but needs to be stated in e.g. methodology. If HTA is taken along as stakeholder (not partner), this may rather be part of the discussion.
41	**Report any unintended consequences, or adverse effects**	**(100%),** (95)	

Discussion session 5: Discussion, Abstract, General

Checklist item (Population)		Consensus: % agreement with scores 7-9. **Second round,**(first round)	Comments
42	**Include a structured abstract (for example including summary of findings, strengths and limitations, comparison with other studies, conclusions and implications)**	**(58%),** (55)	• I didn't understand 'structured abstract' in question 42. I assume this means structured summary • Not sure how I feel about structured abstracts for the Discussion. I could be pursuaded although maybe a 'Results in context' box would be better. • #42 is duplicative of #2.
43	**Reflect on the processes of implementing the service, barriers or facilitators, and lessons learned**	**(79%),** (75)	
44	**How did the setting enable or hinder the implementation of the new service**	**(79%),** (60)	• To me item #44 relates to item #43. It seems to be part of item #43. As such, I rated #44 lower (not unnecessary). • 44. Can be quite subjective especially for people bought into the idea of a new service
45	**How was the new service ~~was~~ implemented highlighting (if relevant) variations between sites and over time and the impact on treatment outcomes and unintended consequences**	**(74%),** (60)	• 45. may be in results • Item 45 seems redundant with the others- if report on them, not sure need it
46	**Interpret findings in the light of the general body of literature, and consider implications for healthcare services (including issues of generalizability, transferability, strategies for facilitating and normalising into routine care)**	**(100%),** (80)	
	• Include interpretation in relation to theory - reflecting back on the theory underpinning the intervention reported/stated earlier		
	• q's 43, 44, 45 - relevant data on these should be in findings, not just reflected on in discussion		
Checklist item (Title and abstract)		Consensus: % agreement with scores 7-9. **Second round,**(first round)	**Item specific comments:** First round**; Second round**
1	**The title (or abstract if word count of title precludes) should include a description of the methodology (e.g.**	**79%,** (80%)	

	phase IV implementation study, cluster randomised implementation trial, interrupted time series, before and after, stepped wedge study)		
2	**There should be a structured abstract which clearly states aim, study design, setting, population, intervention, outcomes, conclusion and implications.**	**95%,** (95%)	
	• This seems a very general standard, nothing specific to these studies		
Checklist item (General)		**Consensus: % agreement with scores 7-9. Second round,**(first round)	**Item specific comments:** First round**; Second round**
47	**Include statement(s) on regulatory approvals (including, as appropriate, ethical approval, confidential use of routine data, governance approval), trial/study registration, funding and conflicts of interest.**	**(89%),** (75)	**#47 mixes up many different topics- some like registration and COI are critical; others much less so**
General comments			
	I would also like to know if they had any stakeholder advisory group and how they engaged with them - for what purposes, with what frequency.		
	• These are covered under other guidelines		
	Sorry, lots of essentials in my response. Hard to say much shouldn't be there really, good suggestions for a reporting standard.		