

Towards a measurement instrument for determinants of innovations

MARGOT A.H. FLEUREN¹, THEO G.W.M. PAULUSSEN¹, PAULA VAN DOMMELEN¹ AND STEF VAN BUUREN^{1,2}

¹TNO, PO Box 2215, 2301 CE Leiden, The Netherlands, and ²FSS, University of Utrecht, Utrecht, The Netherlands

Address reprint requests to: Margot Fleuren, TNO, PO Box 2215, 2301 CE Leiden, The Netherlands. Tel: +31-88-866-6276; Fax: +31-88-866-0613, E-mail: margot.fleuren@tno.nl

Accepted for publication 15 May 2014

Abstract

Objective. To develop a short instrument to measure determinants of innovations that may affect its implementation.

Design. We pooled the original data from eight empirical studies of the implementation of evidence-based innovations. The studies used a list of 60 potentially relevant determinants based on a systematic review of empirical studies and a Delphi study among implementation experts. Each study used similar methods to measure both the implementation of the innovation and determinants. Missing values in the final data set were replaced by plausible values using multiple imputation. We assessed which determinants predicted completeness of use of the innovation (% of recommendations applied). In addition, 22 implementation experts were consulted about the results and about implications for designing a short instrument.

Setting. Eight innovations introduced in Preventive Child Health Care or schools in the Netherlands.

Participants. Doctors, nurses, doctor's assistants and teachers; 1977 respondents in total.

Results. The initial list of 60 determinants could be reduced to 29. Twenty-one determinants were based on the pooled analysis of the eight studies, seven on the theoretical expectations of the experts consulted and one new determinant was added on the basis of the experts' practical experience.

Conclusions. The instrument is promising and should be further validated. We invite researchers to use and explore the instrument in multiple settings. The instrument describes how each determinant should preferably be measured (questions and response scales). It can be used both before and after the introduction of an innovation to gain an understanding of the critical change objectives.

Keywords: implementation, preventive child healthcare, school-based health promotion

Introduction

The introduction of innovations for improving—public—health outcomes is complex [1–6]. By 'innovations', we mean guidelines, interventions or programmes that are perceived as new by an individual or other unit of adoption [7]. Several models have been proposed that describe similar sequences for the systematic planning of innovations in general terms [1, 4, 8, 9]. Figure 1 shows a generic framework that has been used in the Netherlands and other countries since 1999 for the introduction and evaluation of innovations in healthcare and education [9]. Each of the four main stages in innovation processes (dissemination, adoption, implementation and continuation) can be thought of as critical phases where the desired change may or may not occur. Implementation differs from the preceding phase—adoption—in which people initially acquire and

process information about the innovation and make their decision about using the innovation (behavioural intention). In the implementation phase, the innovation is put into daily practice by the professional (behaviour). The transition from one stage to the next can be affected, positively or negatively, by various determinants associated with characteristics of the innovation (e.g. complexity and clear procedures), the potential user of the innovation (e.g. knowledge and self-efficacy), the organisational context of the user (e.g. staff turnover and financial resources) and the socio-political context (e.g. legislation). A detailed understanding of these determinants can guide the process of designing innovation strategies that will have the potential to produce real change [1, 9, 10].

Although most implementation experts recognize that measuring the determinants of innovation processes is essential, we are not aware of validated instruments for measuring

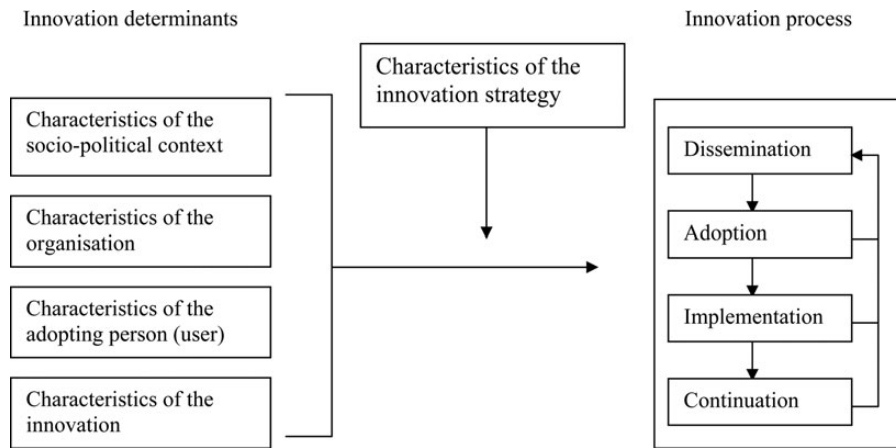


Figure 1 Framework representing the innovation process and related categories of determinants [9].

these determinants and outcomes. In 2004, we published a list of 50 potentially relevant determinants of innovation processes in healthcare based on a systematic review of 57 empirical studies published between 1990 and 2000 and a Delphi study including 44 implementation experts [9]. More recently, ten determinants were added to the original list, mainly based on empirical studies of health promotion programmes [10]. The list is not suitable as a measurement instrument. It is too long, and the impact of specific determinants in the presence of other determinants is not clear.

The list of potentially relevant determinants has now been used in several empirical studies examining the implementation of evidence-based innovations in preventive child healthcare (PCHC) and in schools since 2002. While this is a substantial empirical base to build upon, the studies assessed partially overlapping sets of determinants, so the combined data set is patchy, with no single complete record that includes all the determinants. Conventional techniques are inadequate for the analysis of these missing data. Fortunately, even though the general idea of multiple imputation is already quite old [11], flexible techniques for plausible imputations in multivariate incomplete data have recently become available [12–17].

The aim of the present study is to develop a short instrument to measure determinants of innovations that may affect the stage of implementation (Fig 1). It is based on empirical studies that used the list of potential determinants. We consulted implementation experts, asking them to comment on the results in order to facilitate consensus about the operationalization of each determinant. Our research questions were as follows:

- (i) Which determinants predict the use of innovations in PCHC and schools?
- (ii) Which determinants are seen as relevant by implementation experts and what is their preference for measuring each determinant?

Methods

Databases

Original data were used from five studies in PCHC (this system is open to all children in the Netherlands from birth to

the age of 19 years) and from three curriculum innovation studies in primary and secondary schools. Four of these studies examined the national implementation of PCHC guidelines, and the other four focussed on the implementation of health promotion interventions. Table 1 provides a brief description of the eight empirical studies. Different study designs were used. In order to have comparable data, only cross-sectional data were used, meaning that both determinants and the use of the innovation were measured simultaneously. In case determinants and use were measured in both the pre-test and post-test(s), we only used the pre-test measurements.

Measures

Completeness of use of the innovation is the criterion variable in all eight studies. Completeness refers to the proportion of all activities prescribed by the innovation developers that are put into practice by the user. In all eight studies, the developers of the innovations selected the key activities or recommendations for the assessment of completeness. The range of key activities varied between 8 and 30. In the PCHC studies, the respondents indicated for each key activity on a Likert scale how many children they had exposed to the activity. The possible answers ranged from ‘no child at all’ to ‘all children’. Three studies used six-point Likert scales (categories ranging from 0 to 5) and two used seven-point Likert scales (categories ranging from 0 to 6). We then calculated the mean performance over all key activities and standardized this mean to a range from 0 to 100%. In case of a six-point scale, this was done by multiplying the mean score by a factor of 100/5% and in case of a seven-point scale by a factor of 100/6%. In the school studies, the respondents indicated, for each activity, whether they had applied it yes (1) or no (0). The mean performance of all activities was calculated and multiplied by 100%.

Table 2 contains the full list of 60 determinants and those determinants that were measured in the eight studies (first and second columns). The selection of determinants measured differed in the eight studies depending on the context. Sometimes there was good reason to exclude a specific determinant

Table 1 Description of the eight empirical studies in the meta-analysis

Type of study	Setting	Design	Respondents	N
Guideline for visual disorders: levels of use [18]	PCHC	Cross sectional	Doctors, nurses, doctor's assistant ^a	311
Guideline for congenital heart disorders: levels of use [19]	PCHC	Cross sectional	Doctors, nurses ^a	210
Guideline for prevention of child abuse: effect of planned innovation strategy on levels of use [20]	PCHC	Pre-test and post-test ^c ; experimental vs. control group	Doctors, nurses ^b	302
Guideline for congenital heart disorders: effect of e-learning vs. traditional learning on levels of use [21]	PCHC	Pre-test and post-test ^c ; experimental vs. control group	Doctors, nurses ^b	317
Education programme for prevention of passive smoking in infants: levels of continuation of use [22]	PCHC	Cross sectional	Doctors, nurses ^a	465
School-wide programme for prevention of bullying: effect of planned innovation strategy on levels of use ^d	Primary schools	Pre-test and post-test ^c	Teachers ^b	125
Mental health promotion programme: effect of planned innovation strategy on levels of use ^d	Primary schools	Pre-test and two post-tests ^c	Teachers ^b	188
Sex education programme: effect of planned innovation strategy on levels of use [23]	Secondary schools	Pre-test and post-test ^c ; experimental vs. control group	Teachers ^b	59
Total				1977

PCHC, Preventive Child Health Care.

^aAll PCHC organizations.

^bSelected sample of PCHC organizations/schools.

^cWe only used pre-test measurements.

^dPublication in preparation.

Table 2 List of the 60 original determinants (first column), determinants included in the analyses (second and third column), results of meta-analysis and experts' opinions (fourth and fifth column) and determinants in the final measurement instrument (last column)

Original list of determinants ^a	Number of empirical studies in which the determinant was measured	Determinants included (yes), excluded (no) or combined (yes, →) in meta-analysis ^b	Significantly associated with completeness of use in univariate analysis	Determinants including (yes), excluding (no) or combining (→) in final instrument according to the experts ^b	Name of determinant in final measurement instrument
Socio-political context					
1 Patient cooperation (measured)	4	Yes → 29	–	Combining → 29	–
2 Patient awareness benefits	–	–	–	Combining → 30	–
3 Patient doubts user's expertise	–	–	–	Combining → 30	–
4 Financial burden on patient	–	–	–	Combining → 30	–
5 Patient discomfort	–	–	–	Combining → 30	–
6 Legislation and regulations	–	–	–	Yes	Legislation and regulations
Organisation					
7 Decision-making process	1	No, only one study	–	Combining → 9	–
8 Hierarchical structure	–	–	–	No, only in adoption stage	–
9 Reinforcement management	7	Yes	Yes	Yes	Formal ratification by management
10 Organisational size	4	No, inconsistency	–	Combining → 45	–
11 Functional or product structure	–	–	–	Combining → 19, 56	–
12 Relationship with other organisations	1	No, only one study	–	No, only in adoption stage	–
13 Collaboration between departments	–	–	–	Combining → 19	–
14 Staff turnover	3	Yes	Yes	Yes	Replacement when staff leave
15 Staff capacity	1	No, only one study	–	Yes	Staff capacity
16 Available expertise	–	–	–	Combining → 15, 26	–
17 Logistical procedures	–	–	–	Combining → 45	–
18 Number of potential users	–	–	–	Combining → 19, 23	–
43 Financial resources	5	Yes → 47	–	Yes	Financial resources
44 Reimbursement	–	–	–	Combining → 43	–
45 Material resources	2	Yes	–	Yes	Material resources and facilities
46 Administrative support	1	Yes → 47	–	Combining → 45	–
47 Time available	8	Yes	Yes	Yes	Time available
48 Coordinator	5	Yes	Yes	Yes	Coordinator
49 Users involved in development	–	–	–	No, only in adoption stage	–
50 Opinion leader	–	–	–	No, only in adoption stage	–

51 Feedback to user about innovation process	7	Yes	Yes	Yes	Feedback to user about innovation process
54 Information available about use of innovation	3	Yes	Yes	Yes	Information accessible about use of innovation
61 Unrest in organisation	–	–	–	Yes	Unrest in organisation
User / health professional					
19 Social support colleagues	5	Yes	Yes	Yes	Social support
20 Social support other professionals	2	Yes → 19	–	Combining → 19	–
21 Social support supervisors	5	Yes → 19	–	Combining → 19	–
22 Social support higher management	–	–	–	Combining → 19	–
23 Modelling	3	Yes	Yes	Yes	Descriptive norm
24 Skills	3	Yes → 26	–	Combining → 26	–
25 Knowledge	1	No, only one study	–	Yes	Knowledge
26 Self-efficacy	8	Yes	Yes	Yes	Self-efficacy
27 Ownership	–	–	–	Combining → 56	–
28 Task orientation	3	Yes	–	Yes	Job perception
29 Expects patient cooperation	4	Yes	–	Yes	Client/patient cooperation
30 Expects patient satisfaction	4	Yes	Yes	Yes	Client/patient satisfaction
31 Work-related stress	1	No, only one study	–	Combining → 56	–
32 Contradictory goals	1	No, only one study	–	Combining → 56	–
33 Ethical problems	–	–	–	Combining → 56	–
52 Outcome expectations	8	Yes	Yes	Yes	Outcome expectations
53 Subjective norm	7	Yes	Yes	Yes	Subjective norm
55 Extent innovation is read	5	Yes	Yes	Yes	Awareness of content of innovation
56 Personal benefits/drawbacks	7	Yes	Yes	Yes	Personal benefits / drawbacks
Innovation					
34 Clear procedures	8	Yes	Yes	Yes	Procedural clarity
35 Compatibility	5	Yes	Yes	Yes	Compatibility
36 Trialability	1	No, only one study	–	No, relates to level of use	–
37 Relative advantage	–	–	–	Combining → 52, 56	–
38 Observability	3	Yes	Yes	Yes	Observability
39 Appealing	–	–	–	Combining → 56	–
40 Relevance for patient	2	Yes	Yes	Yes	Relevance for client/patient
41 Risks for patient	–	–	–	Combining → 40	–
42 Frequency in the use	–	–	–	Combining → 56	–
57 Perceived prevalence of health problem	2	Yes	–	Combining → 52, 56	–
58 Correctness	5	Yes	Yes	Yes	Correctness

(continued)

Table 2 Continued

Original list of determinants ^a	Number of empirical studies in which the determinant was measured	Determinants included (yes), excluded (no) or combined (yes, →) in meta-analysis ^b	Significantly associated with completeness of use in univariate analysis	Determinants including (yes), excluding (no) or combining (→) in final instrument according to the experts ^b	Name of determinant in final measurement instrument
59 Completeness	7	Yes	Yes	Yes	Completeness
60 Complexity	5	Yes	Yes	Yes	Complexity

^a Numbers correspond with original list [9]; 51–60 new determinants since 2002; 61 added by experts.

^b Combining because determinants belong to a single underlying construct; → refers to the number(s) of the similar determinant(s).

beforehand, for example because the determinant was considered irrelevant for that particular innovation. Furthermore, some determinants relating to the socio-political context or the organisation were excluded since these characteristics, such as formal financial arrangements or legislation, were assumed to be the same for all organisations.

Concept of multiple imputation

Multiple imputation is nowadays accepted as a practical and valid way to handle incomplete data. Statistical inference through multiple imputation involves three steps: imputation, analysis and pooling. All major software packages now support multiple imputation.

Starting from the observed, incomplete data set, imputation finds replacement values for each missing entry in the data set. The special feature of multiple imputation is that several (instead of just one) replacement values are sought. The variability between the replacement values reflects our uncertainty about what value to impute. If we are almost sure what the missing value should have been, then the replacements will be very similar to each other. On the other hand, if we have little to no idea about what to impute, the variability between the replacement values will be large. In general, the replacement values are drawn from a distribution specifically modelled for each missing entry and will take all relevant information into account. In this article, we have created $m = 50$ completed data sets using the multiple imputation by chained equations (MICE) approach [15]. These 50 imputed sets are identical for the non-missing data entries but differ in the imputed values.

The second step is to estimate the statistic of scientific interest from each imputed data set. Typically, we apply the method we would have used had the data been complete to each imputed data set. This is now easy since all data are complete. Note that we now have 50 estimates (and not just one), which moreover differ from each other because their imputed data sets are different. It is important to realize that these differences are caused by our uncertainty about what value to impute.

In practice, we want 1 result, not 50. The last step is to pool the 50 estimates into 1 combined estimate and calculate its variance. For quantities that are approximately normally distributed, we can calculate the mean over 50 estimates and sum the within- and between-imputation variance according to the method developed by Rubin, called Rubin's rules [11]. The final estimate and its variance can be used to calculate correct *P*-values and 95% confidence intervals.

Data analysis

The goal of the analysis is to infer how well we can predict completeness of use on the basis of a set of determinants. The eight studies generated data about a total of 39 determinants in the list, with a mean of 20 determinants per study. Seven determinants were removed since they were measured in one study only. One determinant (size of the organisation) was removed because the different studies produced data that were found to be inconsistent. The data of ten determinants were combined to form four determinants because of their

resemblance. The criteria were (i) determinants should belong to the same underlying theoretical construct as described in literature and (ii) additional factor analysis in case of doubts about the same underlying constructs. For example, four determinants related to perceived support from respectively colleagues, other healthcare professionals, supervisors and higher management. All these determinants belong to the construct of ‘social support’ and were therefore combined. The number of determinants remaining in the meta-analysis was 25. Each determinant was, on average, measured in five out of eight studies. Table 2 shows the outcomes of the reduction process (second and third columns).

The same determinant was sometimes operationalized in different ways. For example, two variants of perceived social support from colleagues were (i) ‘The support from my colleagues is...’ (four-point Likert scale ranging from ‘very insufficient’ to ‘very sufficient’) and (ii) ‘I can always count on my colleagues when needed’ (seven-point Likert scale ranging from ‘totally disagree’ to ‘totally agree’). The 25 unique determinants were operationalized as 62 variants; 11 determinant were operationalized as 1 variant and 14 determinants consisted of more variants with a maximum of 4. There were 11 variants that had <100 (of 1977) observations; they were discarded before the analysis. In addition, data from 152 participants were discarded before the analysis because the outcome measure (completeness of use) was lacking. The final data set consisted of 1825 participants and 51 variants representing 25 unique determinants.

Multiple imputation by chained equations [14, 15] was used to produce 50 multiply-imputed data sets under the missing at random (MAR) assumption. We chose MICE since it is flexible, produces imputed data that are close to the measured data and has held up very well against alternatives in simulations [17]. The MAR assumption is plausible since nearly all missing data are created by design (i.e. the investigator did not collect the determinant) and is robust to substantial violations [24].

The imputation model was specified for each variant. Predictors were selected that had an absolute correlation with the primary outcome (‘completeness of use’) of at least 0.1 and a minimum proportion of usable observations of 0.3 [15]. Alongside the primary outcome, study identification (eight classes) was always included in the model.

Each of the 50 completed data sets was analysed by linear regression analysis, with determinants and study identification as independent variables and completeness of use as the dependent variable. The estimated regression weights for different imputations were combined using Rubin’s rules [11]. Completeness of use was regressed for each potential determinant (univariate analysis). Since we wanted to avoid including different variants of the same determinant in the analysis, the variant that had the strongest association with completeness of use, as determined by the Wald statistic of the regression coefficient, was selected for further analysis. *P*-values of <0.05 (two-sided) were considered statistically significant.

Consultation of experts

In the next step, 22 implementation researchers, implementation advisors/consultants and policymakers commented on the

results of the analyses. The majority of experts were purposely chosen from outside the field of PCHC and schools (e.g. working in mental healthcare or hospitals). Although, in the previous study [9], implementation experts reached consensus about the operationalization of each determinant, we first asked the experts whether they still agreed with the operationalization and response scales for each determinant that was significantly related to completeness of use in the meta-analysis. Second, for each determinant that was not included in the final model—because the data were insufficient or because the determinant had not been measured—or that was not significantly related to completeness of use, the experts indicated whether theoretical expectations suggested that a determinant should be retained in the final instrument and why. They also provided an operationalization when they thought the determinant should be retained. Third, they were also given the opportunity to add, on empirical or theoretical grounds, new determinants that might have been overseen in the original list.

The first two authors evaluated the answers. Consensus was considered to have been achieved if 75% of the experts agreed about whether determinants not included in the meta-analysis should be maintained or not in the final list. In all other cases, the researchers made a decision based on theoretical grounds.

Results

Pooled analysis

The model with only study identification as a predictor explained 35% of variance in completeness of use. Incorporating all 25 determinants raised this to 51%, and so the determinants explained an additional 16% of variance in completeness of use. Table 3 shows the 21 determinants that were significantly associated with completeness of use in the univariate analysis.

Consultation of experts

According to the experts, seven determinants from the original list that were not included in the analysis or were not significantly related to completeness of use should be included in a measurement instrument on the basis of theoretical expectations. These determinants are as follows: legislation and regulations, staff capacity, financial resources, material resources and facilities, knowledge, job perception and client/patient cooperation. The others could be excluded mainly because of their resemblance to other determinants (Table 2, fifth column).

The experts found that only minor changes were required to the operationalization of the ‘original’ determinants, and seven determinants were rephrased: ‘staff turnover’ as ‘replacement when staff leave’, ‘material resources’ as ‘material resources and facilities’, ‘information available’ as ‘information accessible’, ‘modelling’ as ‘descriptive norm’, ‘task orientation’ as ‘job perception’, ‘extent innovation is read’ as ‘awareness of content of innovation’ and ‘clear procedures’ as ‘procedural clarity’ (Table 2, last column).

Table 3 Results of the univariate regression models for predicting completeness of use

Determinant ^a	Beta	SE	P-value	Number of observations
Formal ratification by management (9)	9.61	1.08	<0.001	1764
Replacement when staff leave (14)	3.33	0.65	<0.001	914
Social support (19)	2.80	0.87	0.001	1362
Descriptive norm (23)	2.06	0.40	<0.001	671
Self-efficacy (26)	19.26	1.14	<0.001	1810
Client/patient satisfaction (30)	4.19	0.68	<0.001	464
Procedural clarity (34)	6.01	1.01	<0.001	1583
Compatibility (35)	4.48	0.68	<0.001	1452
Observability (38)	3.50	0.66	<0.001	464
Relevance for client/patient (40)	2.82	0.80	0.001	464
Time available (47)	3.02	0.63	<0.001	464
Coordinator (48)	6.19	1.19	<0.001	1396
Feedback to user about innovation process (51)	3.91	0.89	<0.001	826
Outcome expectations (52)	0.98	0.24	<0.001	693
Subjective norm (53)	0.92	0.09	<0.001	1481
Information accessible about use of innovation (54)	7.11	2.99	0.02	637
Awareness of content of innovation (55)	7.06	1.01	<0.001	1368
Personal benefits/drawbacks (56)	5.37	0.74	<0.001	464
Correctness (58)	6.44	1.86	0.001	923
Completeness (59)	5.94	1.13	<0.001	464
Complexity (60)	2.29	0.80	0.005	761

^aNumber corresponds to original list (Table 2, first column); name corresponds to determinant in final measuring instrument (Table 2, last column).

Nine experts suggested adding a total of twelve new determinants to the list. Finally, only ‘unrest in organization’ was added since the others closely resembled determinants already on the list.

Final measurement instrument

The list of 60 potentially relevant determinants was reduced to 29. Twenty-one determinants were based on the pooled analysis of the empirical data and seven on the theoretical expectations of the implementation experts consulted; one new determinant was added on the basis of the implementation experts’ practical experience. These 29 determinants have been included and described in the measurement instrument for determinants of innovations (MIDI). It describes how each determinant should preferably be measured (questions and response scales) and the syntax for analyses (Supplementary Appendix A).

Discussion

We developed the MIDI to improve our understanding of the critical determinants that may affect implementation and to target the innovation strategy better. The instrument can be used both before and after the introduction of an innovation. The instrument is intended primarily for implementation researchers, but it can also be used by implementation

consultants/advisors. We expect that the instrument will help to collect information as a basis for an empirically grounded judgement of the relative importance of determinants.

In practice, the user first needs to decide which determinants should be measured. The main criterion is—given the nature of the innovation and/or the context—the anticipated impact of the determinant on possible variations in completeness of use of the innovation.

Characteristics of the socio-political context do not often result in differentiation, particularly when a study is limited to a single country. For example, legislation and regulations relating to child healthcare in the Dutch context will differentiate little, if at all, between the different organizations. However, in international research, this is likely to be a relevant factor for explaining variations in the implementation of particular innovations.

For a sound assessment of an innovation strategy, it is advisable to measure as many determinants as possible since they may all be of practical relevance for designing that strategy. Yet, it is sometimes not feasible to measure particular determinants before an innovation is implemented because the user will not yet have a clear picture of what the innovation entails. For example, a subjective assessment of certain characteristics of an innovation will not be possible if people do not have knowledge awareness of the innovation.

Our study was limited in some respects. Though MIDI is promising, it is not yet a validated instrument and further development is needed. The results in the current version are

specific to preventive innovations in the context of PCHC and schools. It is not known to what extent the results can be generalized to other settings or to more technology-based innovations. On the other hand, there are some reasons suggesting that the instrument may be applicable to a broader range of settings. First, the original list of determinants was developed and used in many different healthcare settings and school settings. Second, the experts in the present study, as well as in a Delphi study performed previously [9], indicated that most determinants were generic. It should also be noted that we did not have empirical data to test all the determinants from the original list. We used self-reported adherence measures. There is sometimes a large discrepancy between self-report and objective measurements [25]. This may have affected our measures of completeness of use. As the same methods were applied in all eight studies, we expect that this will have resulted in a systematic bias (if at all). However, underestimation of completeness of use will not necessarily affect the associations with the determinants, which was the primary objective of our study.

Since our publication of the original list of determinants in 2004, several other frameworks and checklist have been published. Recently, Flottorp *et al.* published a nice systematic review and synthesis of frameworks and taxonomy of determinants, resulting in checklist of 57 determinants grouped in 7 domains [26]. This checklist is consensus based, and the 57 determinants closely resemble the ones in our original list. As we were able to verify determinants empirically, the present study represents a next logical step of development.

We envisage the development of MIDI into a validated instrument with sensibly chosen cut-offs for the scores for each determinant. To achieve this goal, we invite implementation researchers to use and explore MIDI in applied settings and to report and share their results. Empirical data from a broader range of innovations and settings will help to substantiate the sensitivity of the instrument in practice. As measurement in implementation research is still in its infancy, we hope that MIDI will be of interest to both implementation researchers and advisors.

Supplementary material

Supplementary Material is available at *INTQHC* online.

Acknowledgements

The authors thank Ab Rijpstra for his assistance with data management.

Funding

This project is supported by the Netherlands Organisation for Health Research and Development (ZonMw) [number 20040008]. Funding to pay the Open Access publication

charges for this article was provided by The Netherlands Organisation for Scientific Research (NWO).

References

- Greenhalgh T, Robert G, Macfarlane F *et al.* Diffusion of innovations in service organizations: systematic review and recommendations. *Milbank Q* 2004;**82**:581–629.
- Fixsen DL, Noam SF, Blase KA *et al.* *Implementation Research: A Synthesis of the Literature*. Florida: University of South Florida, 2005.
- Grimshaw J, Eccles M, Thomas R *et al.* Toward evidence-based quality improvement. Evidence (and its limitations) of the effectiveness of guideline dissemination and implementation strategies 1966–1998. *J Gen Intern Med* 2006;**21**(Suppl 2):S14–20.
- Grol R, Wensing M, Eccles M. *Improving Patient Care: The Implementation of Change in Clinical Practice*. Edingburgh: Elsevier, 2005.
- Prior M, Guerin M, Grimmer-Somers K. The effectiveness of clinical guideline implementation strategies—a synthesis of systematic review findings. *J Eval Clin Pract* 2008;**14**:888–97.
- Gulbrandsson K. *From News to Everyday Use: The Difficult Art of Implementation*. Ostersund: Swedish National Institute of Public Health, 2008.
- Rogers EM. *Diffusion of Innovations*. New York: Free press, 2003.
- Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. *Am J Public Health* 1999;**89**:1322–7.
- Fleuren M, Wiefferink K, Paulussen T. Determinants of innovation within health care organizations: literature review and Delphi study. *Int J Qual Health Care* 2004;**16**:107–23.
- Bartholomew LK, Parcel GS, Kok G *et al.* *Planning Health Promotion Programs: an Intervention Mapping Approach*. San Francisco: Jossey-Bass, 2006.
- Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. New York: Wiley, 1987.
- Buuren van S, Boshuizen HC, Knook DL. Multiple imputation of missing blood pressure covariates in survival analysis. *Stat Med* 1999;**18**:681–94.
- Buuren van S. Multiple imputation of discrete and continuous data by fully conditional specification. *Stat Methods Med Res* 2007;**16**:219–42.
- Buuren van S, Groothuis-Oudshoorn K. MICE: multivariate imputation by chained equations in R. *J Stat Software* 2011;**45**:1–67.
- Buuren van S. *Flexible Imputation of Missing Data*. Boca Raton: Chapman & Hall/CRC Press, 2012.
- Sterne JA, White IR, Carlin JB *et al.* Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009;**338**:b2393.
- Lee KJ, Carlin JB. Multiple imputation for missing data: fully conditional specification versus multivariate normal imputation. *Am J Epidemiol* 2010;**171**:624–63.

18. Fleuren MAH, Verlaan ML, Velzen-Mol van HWM *et al.* Zicht op het gebruik van de JGZ-standaard Opsporing van Visuele Stoornissen 0–19 jaar: een landelijk implementatieproject. [Looking closely at adherence to the Preventive Child Health Care guideline on early detection of visual disorders]. Leiden: TNO Kwaliteit van Leven, 2006.
19. Fleuren MAH, Dommelen van P, Kamphuis M *et al.* Landelijke Implementatie JGZ-standaard Vroegtijdige Opsporing van Aangeboren Hartafwijkingen 0–19 jaar [National implementation of the Preventive Child Health Care guideline on early detection of congenital health disorders]. Leiden: TNO Kwaliteit van Leven, 2007.
20. Broerse A, Fleuren MAH, Kamphuis M *et al.* Effectonderzoek proefimplementatie JGZ-richtlijn secundaire preventie kindermishandeling [Evaluation of the pilot implementation of the Preventive Child Health Care guideline on secondary prevention of child abuse]. Leiden: TNO Kwaliteit van Leven, 2009.
21. Galindo Garre F, Kamphuis M, Verheijden MW *et al.* Onderzoek naar de mogelijkheden en haalbaarheid van E-learning bij de implementatie van richtlijnen in de jeugdgezondheidszorg [Feasibility and effect of e-learning on the implementation of Preventive Child Health Care guidelines]. Leiden: TNO Kwaliteit van Leven, 2010.
22. Crone MR, Verlaan M, Willemsen MC *et al.* Sustainability of the prevention of passive infant smoking within well-baby clinics. *Health Educ Behav* 2006;**33**:178–96.
23. Wiefferink CH, Poelman J, Linthorst M *et al.* Outcomes of a systematically designed strategy for the implementation of sex education in Dutch secondary schools. *Health Educ Res* 2005;**20**:323–33.
24. Collins LM, Schafer JL, Kam CM. A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychol Methods* 2001;**6**:330–51.
25. Adams AS, Soumerai SB, Lomas J *et al.* Evidence of self-report bias in assessing adherence to guidelines. *Int J Qual Health Care* 1999;**11**:187–92.
26. Flottorp SA, Oxman AD, Krause J *et al.* A checklist for identifying determinants of practice: a systematic review and synthesis of frameworks and taxonomies of factors that prevent or enable improvements in healthcare professional practice. *Implement Sci* 2013;**8**:35.