



Measuring Team Development in Clinical Care Settings

Ronald Stock, MD, MA; Eldon Mahoney, PhD; Patricia A. Carney, PhD

BACKGROUND AND OBJECTIVES: Our objective was to describe the psychometric properties of a measure of team development that can be used to assess and guide team functioning in health care settings.

METHODS: The Team Development Measure (TDM) is a 31-item questionnaire constructed using the Rasch rating scale measurement model. We conducted an Mplus exploratory factor analysis using data collected from 1,194 individuals representing 120 different teams. Team size ranged from three to 39 members from rural and urban inpatient and ambulatory health care settings. Here we characterize the domains of teamness, while taking into account the development of teams over time.

RESULTS: The TDM was found to have good psychometric properties with little measurement error and a Rasch person reliability of 0.95. Overall Cronbach's alpha was 0.97. An Mplus exploratory factor analysis combined with the stochastic nature of the Rasch model suggests a developmental sequence in building teams consisting of four sub-domains with the following mean item difficulty scores: cohesion=40.5, standard deviation (SD)=2.68, communication=49.3 (SD=2.78), roles and goals=52.7 (SD=2.74), and team primacy=53.3 (SD=1.06). This pattern suggests cohesiveness is an initial element for team development, followed by communication, roles and goals clarity, and team primacy.

CONCLUSIONS: We developed and tested a measure of team development that has strong psychometric properties. This tool could be used to study how team functioning affects clinical outcomes and as a quality improvement tool to improve team function.

(Fam Med 2013;45(10):691-700.)

Although working in teams is a central component of the patient-centered medical home,¹ emphasizing teamwork in health care is not new.² Most, if not all, health care systems and practices strive for highly developed teamwork. However, the practical application of team-based care in primary care settings has achieved

only modest penetration into mainstream health care. As the American health care system continues to transform, there is consensus that team-based care is foundational to improving access, quality, and cost of care.³ Team-based care is critical to the creation of a care delivery system in which health professionals are intentionally configured to take

collective responsibility for patient and population health outcomes.^{1,4} Team-based care holds promise as an important strategy to improve patient outcomes and to increase the capacity and sustainability to provide primary care. However, despite this potential, our understanding of the impact of teams and the ways they contribute to key health care outcomes is limited by the lack of robust, validated measures of the effectiveness or quality of team functioning, particularly in primary care.⁵

Decades of research, both in health care^{5,9-13} and outside of health care,⁶⁻⁸ have shown the importance of using groups of people working toward a common goal. While the term "team" is part of the organizational language in health care, work groups are often regarded as teams simply because they are labeled as such. Placing a group of individuals from different disciplines in the same room does not mean they will function as a team. There are numerous definitions of highly effective health care teams, but at the most basic level it requires two or more individuals with specific roles to perform interdependent tasks and share a common goal. Team members must have the skills to share information

From the Department of Family Medicine (Drs Stock and Carney) and the Department of Public Health and Preventive Medicine (Dr Carney), Oregon Health & Science University; and Bellingham, WA (Dr Mahoney).

(communicate) to enable the team to coordinate the care necessary for effective and efficient patient care.^{10,11} Because most medical professionals are not trained in team skills and behaviors, facilitating team development and training is required. A conceptual model of teamwork that includes individual team member competencies, attributes specific to team behaviors, and organizational support for teamwork has been proposed.¹⁰ Each team member brings individual skills, knowledge, attitudes, and necessary behaviors to the practice team, but these are insufficient to fully define a team and generate the necessary outcomes of a team-based care approach. There is widespread agreement that two types of variables influence health care team performance: (1) attributes of the team itself and (2) attributes of the larger organizational context within which the team functions.^{11,12} While tools have been developed to assess individual team member behaviors and skills using standardized observational and self-assessment instruments,¹⁰ few instruments have been developed that use self-reported team member perceptions to evaluate the extent to which known team attributes are present in a team.

A recent review of survey instruments designed to measure teamwork in health care settings¹⁴ identified 35 surveys, nine of which met criteria for psychometric validity. Most of these surveys measured domains related to communication, coordination, and respect, though variability exists in the dimensions of teamwork that were assessed. The number of items from the psychometrically tested surveys ranged from five to 54 items. These existing measures have strengths and weaknesses based on the specific behavioral processes being assessed. An important challenge in studying teamwork is that the items can refer to many different behavioral processes that occur along the course of developing teamwork, which makes

choosing an instrument based on the context of the study very important. The Agency for Healthcare Research and Quality recently developed a measure grounded in the hospital patient safety culture literature that assesses individual perceptions of group level team skills and behaviors.¹⁵ While it has shown utility in evaluation of the Team-STEPPS team development process, it appears to have potential for use in other settings. Further, at the clinic practice level, there is still much to be learned about which measurement tools are best used to evaluate team care and its relationship to health outcomes.

We used existing literature¹⁰⁻¹³ and extensive experience working with teams to develop an instrument designed to measure team functioning from the perspective of individual members within the team. Here our intent is to describe the psychometric testing we conducted to evaluate the properties of a 31-item questionnaire used to measure team attributes and, more specifically, team development. The purpose of the Team Development Measure (TDM) is to measure the degree to which a team has components needed for highly effective teamwork in place.

Methods

Survey Instrument Design and Development

The initial questionnaire contained 39 items derived from literature on health care teams where attributes important to the building and maintenance of teams were identified.¹¹⁻¹³ These variables were supplemented with the authors' experience conducting field work with teams providing clinical care. The tool was designed to include a 4-point Likert-like scale response format (strongly disagree=1, disagree=2, agree=3, strongly agree=4). These items were operationalized into a questionnaire that was tested for the degree to which they measured team development. Four of the items were reverse scored. The 39 items were initially

tested with five teams that included a total of 73 members. Based on the initial analysis, two items were dropped due to poor item fit (where fit values were above the upper control limit of the fit statistic range of 0.5–1.50¹⁶). Both dropped items were designed to assess leadership. A qualitative investigation at the time suggested that the term “leader” was interpreted by many to mean formal organization leaders, rather than rotating, task-specific, leaders within the team. A subsequent item was developed to change the wording to “lead” (“On this team, the person who takes the lead differs depending on who is best suited for the task”). To reduce response burden, eight items that calibrated similarly were dropped, which did not significantly decrease the precision of the measure. The final instrument includes 30 items from the original set plus a new item on leadership resulting in a 31-item questionnaire (Table 1).

Setting for Psychometric Assessment

Participants in the full psychometric analysis included 1,194 team members representing 120 different working health care teams. These teams practiced within a non-profit health care system that includes urban and rural critical access hospitals and outpatient clinics across Oregon, Washington, and Alaska. Teams ranged in size from three to 39 members, with a mean of 9.95 and median of 8.0. Team composition included physicians, nurses, receptionists, administrators, and ancillary therapists. The response rate for completion of the test instruments was high, with 85% of participants answering all 31 questions and 95% answering 29 or more questions. No correlation was found between team size and number of questions answered ($r=-0.046$, $P>.05$). All TDM data collected were anonymous. Oregon Health & Science University's Institutional Review Board reviewed this study and granted an exemption under 45 CFR 46.102(f).

Table 1: The Team Development Measure

Instructions: This questionnaire is a measure of team characteristics. Please indicate how much you strongly disagree, disagree, agree, or strongly agree to each statement as it applies to your team at the present time. There are no right or wrong answers, just your perceptions. Before beginning please write the name of your team on this line:

	Strongly Disagree	Disagree	Agree	Strongly Agree
1. Team members say what they really mean.	1	2	3	4
2. Team members say what they really think.	1	2	3	4
3. Team members talk about other team members behind their back.	1	2	3	4
4. All team members participate in making decisions about the work of the team.	1	2	3	4
5. All team members feel free to share their ideas with the team.	1	2	3	4
6. All team members feel free to express their feelings with the team.	1	2	3	4
7. The team practices tolerance, flexibility, and appreciation of the unique differences between team members.	1	2	3	4
8. The team handles conflicts in a calm, caring, and healing manner.	1	2	3	4
9. Regardless of the topic, communication between the people on this team is direct, truthful, respectful, and positive.	1	2	3	4
10. The team openly discusses decisions that affect the work of the team before they are made.	1	2	3	4
11. In this team, members support, nurture, and care for each other.	1	2	3	4
12. The team has agreed upon clear criteria for evaluating the outcomes of the team's effort.	1	2	3	4
13. As a team we come up with creative solutions to problems.	1	2	3	4
14. In the team there is more of a WE feeling than a ME feeling.	1	2	3	4
15. There is confusion about what the work is that the team should be doing.	1	2	3	4
16. There is confusion about how to accomplish the work of the team.	1	2	3	4
17. Roles and responsibilities of individual team members are clearly understood by all members of the team.	1	2	3	4
18. All team members place the accomplishments of the team ahead of their own individual accomplishments.	1	2	3	4
19. The goals of the team are clearly understood by all team members.	1	2	3	4
20. All team members define the goals of the team as more important than their own personal goals.	1	2	3	4
21. I am happy with the outcomes of the team's work so far.	1	2	3	4
22. I enjoy being in the company of the other members of the team.	1	2	3	4
23. This team is a personally meaningful experience for me.	1	2	3	4
24. I have a clear understanding of what other team members expect of me as a team member.	1	2	3	4
25. The work I do on this team is valued by the other team members.	1	2	3	4
26. I am allowed to use my unique personal skills and abilities for the benefit of the team.	1	2	3	4
27. Some members of this team resist being led.	1	2	3	4
28. Information that is important for the team to have is openly shared by and with all team members.	1	2	3	4
29. All individuals on this team feel free to suggest ways to improve how the team functions.	1	2	3	4
30. When team problems arise the team openly explores options to solve them.	1	2	3	4
31. On this team, the person who takes the lead differs depending on who is best suited for the task.	1	2	3	4

Psychometric Analysis

Psychometric analyses involved a two-step process. In the first step, an MPlus promax exploratory factor analysis¹⁹ with items treated as ordinal variables was conducted to determine the number of underlying dimensions contained in the item set and to identify the subset of items that corresponds to each of the underlying dimensions. In the second step, a Rasch rating scale measurement model¹⁶⁻¹⁸ was applied with Winsteps¹⁶ software to transform ordinal numeric results to an interval score reflecting a range from 0–100. To assure that the created measurement scale is in fact equal interval (like a “yardstick”), the metric for all Rasch analysis is in logits (log odds units). To convert logit values to a user-friendly 0–100 scale (yardstick) from the summated 1–4 response category data, two components of conversion information are supplied by Winsteps: (1) the mean or center of the scale in 0–100 units, which in this study was 49.02 and (2) the size of one logit in 0–100 units, which in this study was 6.02. While, by definition, the center of a logit scale is zero and the size of a logit is one logit, the use of this conversion transforms logits into a 0–100 range where zero is the lowest possible score and 100 is the highest possible score, and the 0–100 remains an equal interval score.

For a comprehensive discussion of Rasch measurement models compared to other approaches, see Massof.¹⁷ Briefly, the Rasch model is a stochastic (represents random observations) mathematical statement of measurement that specifies that the probability of a respondent giving a specific response to a measurement item is a function of the ability of the respondent to agree or disagree with the item minus the difficulty of the item minus the difficulty of the response category.¹⁸ When a set of observations fits the model, the result is a true linear measure constituting a real number line representing the quantity of the construct present (eg, the amount of “teamness”).

The location of an item on the scale represents the difficulty of the item or, in this case, how much team development a team member perceives as present to be able to endorse the item. The location of the individual team members on the scale is referred to as their “ability” or how much team development is present.

For the measure to be useful, all items must fit a single real number line (be uni-dimensional) as assessed by mean square error fit statistics.¹⁷ Item fit is the degree to which responses to an item fit expectations, given the ability of the respondent and the difficulty of the item. When this expectation is not met, the item is not contributing useful information in construct measurement. The overall precision of the measure is assessed by Rasch person reliability. One minus person reliability is the amount of error in measuring the construct.¹⁷ Mean square error fit statistics estimate the degree to which responses match what is expected given the calibrated item difficulty and person ability. A fit coefficient of 1.0 indicates perfect fit to Rasch model expectations. Fit values greater than 1.50 indicate the item is measuring something other than the intended construct. Fit values less than 0.5 indicate a lack of expected variability in responses (persons of very different ability are giving the same answer).¹⁶ Typically, few items are expected to have a perfect fit.

Winsteps software provides two estimates of measurement error.^{16,17} Real person reliability assumes that all measurement error is due to a misfit of the data to the model (worst case reliability). Model person reliability assumes that all errors of measurement are due to chance (best case reliability). The “true” reliability of a measure is somewhere between real and model. Since both persons and items have a location on the measurement scale, two item fit values were calculated. Infit is most sensitive to unexpected responses when the item scale location is close to the person’s scale location.¹⁷ Outfit is most sensitive to unexpected

responses when item scale location is far from the person’s scale location.¹⁷ In addition, items are ideally spread across the scale with no large gaps. Very few items were left incomplete by respondents. However, since Rasch is a stochastic model, missing data is irrelevant. Thus, for the psychometric analysis, missing raw data were recoded using a “9” in the dataset, per Winstep protocol.

Results

Factor Structure

Results of the Mplus19 promax exploratory factor analysis best solution indicated four factors (standardized root mean square residual=0.045) (Table 2). The item content of the four factors suggests that factor one is communication (n=14 variables), factor two is roles and goals clarity (n=4 variables), factor three is cohesion (n=4 variables), and factor four is primacy of the team (n=2 variables). Twenty-four of the 31 items loaded to a factor with five items not loading to any factor, though these were used in the Rasch Model. Alpha coefficients for each domain ranged from 0.76–0.94.

Person Reliability

Table 3 shows the number of teams and the within-team person reliability coefficients. For the 120 teams, neither real ($r=0.136$, $P>.05$) nor model ($r=0.171$, $P>.05$) person reliability was related to the number of persons within a given team. Real person reliability of the TDM in this psychometric analysis was 0.95, and model reliability was 0.96. Cronbach’s alpha for the scale overall was 0.97.

Item Fit and Difficulty

Table 4 shows item difficulties (measure) and fit statistics. The TDM measurement is scaled as 0–100, where zero is the lowest possible score and 100 the highest possible score. As demonstrated in Table 4, both infit and outfit for all items were found to be acceptable based on a fit statistics range of 0.5–1.5.¹⁶ The majority of items had fit values

Table 2: M-plus Factor Loadings of Team Development Measure Items*

Team Development Items	Communi- cation	Roles and Goals	Cohesion	Team Primacy
All team members feel free to share their ideas with the team.	0.886	-0.017	0.011	-0.055
All team members participate in making decisions re: work of the team.	0.803	0.195	-0.114	-0.074
The team openly discusses decisions that affect their work before they are made.	0.766	0.287	-0.045	-0.194
All team members feel free to express their feelings with the team.	0.766	-0.011	0.090	-0.039
Team members say what they really think.	0.732	-0.104	0.107	0.095
Communication between the people on this team is direct, truthful, positive.	0.702	-0.053	0.110	0.182
All individuals feel free to suggest how to improve team function.	0.681	0.133	0.093	0.008
When team problems arise we openly explore options to solve them.	0.673	0.186	0.041	0.052
Team members say what they really mean.	0.664	-0.071	0.165	0.136
The team handles conflicts in a calm, caring, and healing manner.	0.636	0.002	0.125	0.153
Information important to have is openly shared by and with all.	0.631	0.240	0.057	-0.089
Team members talk about other team members behind their back.	0.531	-0.046	-0.021	0.249
The team practices tolerance, flexibility, appreciation of differences.	0.533	-0.028	0.275	0.119
There is more of a “we” feeling than a “me” feeling.	0.548	0.056	0.207	0.213
There is confusion about what the work of the team should be.	0.004	0.772	0.103	-0.021
There is confusion about how to accomplish the work of the team.	0.068	0.786	0.056	0.000
Roles and responsibilities of team members are clearly understood.	0.042	0.710	-0.024	0.160
The goals of the team are clearly understood by all team members.	-0.030	0.776	-0.047	0.275
I enjoy being in the company of other members of the team.	0.192	-0.172	0.774	0.085
This team is a personally meaningful experience for me.	0.047	0.034	0.789	0.059
The work I do on this team is valued by the other team members.	0.111	0.131	0.752	-0.040
I am allowed to use my unique skills and abilities.	0.129	0.166	0.724	-0.105
Members place the accomplishments of the team ahead of their own.	0.101	0.143	0.025	0.730
Members define the goals of the team as more important than their own.	0.073	0.142	-0.028	0.799
I am happy with the outcomes of the team’s work so far.	0.177	0.254	0.222	0.389
I have a clear understanding of what other team members expect of me.	-0.095	0.533	0.588	-0.106
In this team, members support, nurture, and care for each other.	0.385	-0.046	0.402	0.207
The team has agreed upon clear criteria for evaluating team outcomes.	0.405	0.480	-0.157	0.067
As a team, we come up with creative solutions to problems.	0.452	0.229	0.253	0.017
Cronbach’s Alpha Coefficients	0.94	0.87	0.86	0.76

* This table includes 29 of the 31 items.

between 0.8 and 1.2, representing very good fit and contributing to the high person reliability. Items 3 and 27 have marginal fit, but deleting either item decreases person reliability.

Since the Rasch model is stochastic, the difficulty order of items suggests the developmental character of team development. The easiest

item, with a mean calibrated value of 36.6, is item 22 (“I enjoy being in the company of the other members of the team”). Agreeing to this item requires the least amount of team development. The most difficult item is 27 (“Some members of this team resist being led”). Agreeing to this item requires the most amount of team development.

Table 5 outlines the distribution of transformed scores using the Rasch approach. The mean 0–100 difficulty of items loading within each of the four factors is: cohesion=40.5 (standard deviation [SD]=2.68), communication=49.3 (SD=2.78), roles and goals=52.7 (SD=2.74), and team primacy=53.3 (SD=1.06). This pattern suggests that cohesiveness is

Table 3: Relationship of Actual and Model Person Reliability Within Study Teams*

Real Person Reliability	Number of Teams	Percent of Teams		Model Person Reliability	Number of Teams	Percent of Teams
0.52	1	0.83		0.60	1	0.83
0.66	1	0.83		0.71	1	0.83
0.67	1	0.83		0.74	1	0.83
0.73	1	0.83		0.76	1	0.83
0.75	1	0.83		0.79	2	1.67
0.76	1	0.83		0.80	2	1.67
0.78	2	1.67		0.81	1	0.83
0.79	2	1.67		0.82	1	0.83
0.8	1	0.83		0.84	2	1.67
0.82	4	3.33		0.85	2	1.67
0.83	2	1.67		0.86	6	5.00
0.84	4	3.33		0.87	2	1.67
0.85	3	2.50		0.88	6	5.00
0.86	3	2.50		0.89	3	2.50
0.87	3	2.50		0.90	7	5.83
0.88	2	1.67		0.91	3	2.50
0.89	3	2.50		0.92	5	4.17
0.9	6	5.00		0.93	12	10.00
0.91	6	5.00		0.94	8	6.67
0.92	10	8.33		0.95	13	10.83
0.93	9	7.50		0.96	17	14.17
0.94	13	10.83		0.97	12	10.00
0.95	12	10.00		0.98	9	7.50
0.96	11	9.17		0.99	3	2.50
0.97	9	7.50		Total	120	100.00
0.98	7	5.83				
0.99	2	1.67				
Total	120	100.00				

* Each line represents a team or teams of a different size.

an initial element for team development, followed by communication, then roles and goals clarity, and team primacy. The Table 5 graphical depiction readily shows that some item-response combinations are more difficult than others even though the item domain may, in general, be easier. The eight stages were identified by naturally occurring break points of the item-response category combinations on the 0–100 scale. Since the concept is development, it is reasonable to consider that a team putting

components in place and answering “agree” is clearly less “in place” than a team in which team members are more likely to respond to an item with “agree strongly.”

Discussion

This study found that the TDM has strong psychometric properties and is valid, based on its development using existing literature and extensive field experience, and is a reliable instrument for measuring perceptions of team development. It is potentially

a tool for assisting health care team members understand the extent that attributes of teamwork are present within their group. We used two rounds of testing with two independent groups of teams to assess its psychometric parameters. The initial test was used to conduct the data reduction step and then a second test was conducted to complete a full psychometric analysis of the tool. Importantly, it holds its psychometric strength in groups as small as three members and as large as 30. It was

Table 4: Item Fit and Difficulty

Questionnaire Item # (see Table 1)	Item Difficulty Score	Infit	Outfit
9. Regardless of the topic, communication between the people on this team is direct, truthful, respectful, and positive.	50.0	0.73	0.71
13. As a team we come up with creative solutions to problems.	45.4	0.75	0.70
30. When team problems arise the team openly explores options to solve them.	48.6	0.77	0.75
8. The team handles conflicts in a calm, caring, and healing manner.	48.2	0.77	0.75
21. I am happy with the outcomes of the team's work so far.	47.7	0.79	0.75
1. Team members say what they really mean.	46.0	0.84	0.81
25. The work I do on this team is valued by the other team members.	42.8	0.84	0.80
14. In the team there is more of a "we" feeling than a "me" feeling.	46.8	0.85	0.81
29. All individuals on this team feel free to suggest ways to improve how the team functions.	47.1	0.86	0.82
2. Team members say what they really think.	49.2	0.89	0.89
11. In this team, members support, nurture, and care for each other.	44.8	0.89	0.84
20. All team members define the goals of the team as more important than their own personal goals.	54.1	0.90	0.92
19. The goals of the team are clearly understood by all team members.	52.8	0.90	0.90
18. All team members place the accomplishments of the team ahead of their own individual accomplishments.	52.6	0.91	0.91
26. I am allowed to use my unique personal skills and abilities for the benefit of the team.	41.3	0.93	0.90
7. The team practices tolerance, flexibility, and appreciation of the unique differences between team members.	45.6	0.95	0.91
22. I enjoy being in the company of the other members of the team.	36.6	0.98	0.92
10. The team openly discusses decisions that affect the work of the team before they are made.	51.6	1.00	0.99
28. Information that is important for the team to have is openly shared by and with all team members.	48.8	1.03	1.03
5. All team members feel free to share their ideas with the team.	47.9	1.03	1.01
12. The team has agreed upon clear criteria for evaluating the outcomes of the team's effort.	56.9	1.04	1.09
4. All team members participate in making decisions about the work of the team.	52.5	1.07	1.07
6. All team members feel free to express their feelings with the team.	52.5	1.07	1.09
17. Roles and responsibilities of individual team members are clearly understood by all members of the team.	53.4	1.10	1.12
16. There is confusion about how to accomplish the work of the team.	53.1	1.13	1.13
23. This team is a personally meaningful experience for me.	41.2	1.15	1.11
31. On this team, the person who takes the lead differs depending on who is best suited for the task.	50.6	1.21	1.30
24. I have a clear understanding of what other team members expect of me as a team member.	46.6	1.21	1.17
15. There is confusion about what the work is that the team should be doing.	51.4	1.30	1.30
27. Some members of this team resist being led.	58.6	1.43	1.59
3. Team members talk about other team members behind their back.	54.7	1.49	1.60

Table 5: Stages of Team Development

Stage	Score Range	Components Present *	Solidification
Pre-Team	0–36	None to building cohesiveness	Initial development
1	37–46	Cohesiveness	In place
2	47–54	Communication	
3	55–57	Role and goal clarity	
4	58–63	Team primacy	
5	64–69	Cohesiveness	Firmly in place
6	70–77	Communication	
7	78–80	Role and goal clarity	
8	81–86	Team primacy	
Fully developed	87–100	Everything	

developed and tested in a variety of health care settings including ambulatory, hospital based, and administrative health care teams in both rural and urban environments. The TDM measures four domains, two of which are similar to domains identified in the recent systematic review,¹⁴ such as communication and roles/goals. However, the TDM was designed to measure teamwork development and, thus, it also assesses cohesion and team primacy, both of which are crucial to the development of teams. With the increased focus of team-based care as part of the patient-centered medical home,²² and health care reform initiatives in general, the ability to accurately measure team development is vitally important.

This paper presents detailed methods on factor analysis and use of the Rasch approach for score conversion. While the topic of measuring team development is clearly important for family medicine research and practice, the methodological detail provided here may be difficult to understand and interpret without training in psychometrics. We have provided this detail because instrument design, development, and testing contributes to reliability and validity of measures, which are important for accuracy. Family physicians often work with teams

of investigators, including biostatisticians, as part of their scholarly work. We hope this paper will foster the use of a common language to make it easier to dialogue about instrument testing. We were unable to provide a tutorial on factor analysis and Rasch modeling in this paper due to space constraints, but we do provide citations that can be used for further study.^{20,21}

The development of accurate tools to measure team-based care, as well as other aspects of health care reform, are important to academic family medicine for a number of reasons. First of all, educators in family medicine lead the development of training programs that promote team-based care, which includes teamwork as a critical component, especially in the development of primary care medical homes. Assisting our student and resident learners in understanding and improving the domains of effective practice teamwork is critical to our educational mission to prepare them to be leaders in their health care practice and community. Secondly, research in team-based care needs to be core to the mission of family medicine's academic institutions emphasizing accurate measurement of team development and creating best methods for developing teams. Finally, the use of factor analysis and Rasch measurement

methods in this study will hopefully encourage others in the academic family medicine research fields to consider these types of survey development and scoring methods.

The evaluation of the TDM has a number of strengths. The intention in the developmental phase of the instrument was to improve patient care through teamwork across a variety of health care settings, so the relevance to the teams and the care they provide was high. The data were collected with individual anonymity and, although it does prevent analyzing the data by team member profession or role, it ensures that response and group bias was minimized. We collected data from health care professionals performing a spectrum of roles with all individuals identifying themselves as part of a team. Use in health care settings, including inpatient and outpatient, occurring in moderate-sized communities in three states is beneficial as the tool is likely generalizable to similar communities and health care settings in the United States. Methodologically strong, the tool's items were constructed using a two-phase approach, and large numbers of individuals and teams were surveyed. The MPlus factor analysis identified initial domains, and the use of the Rasch rating scale measurement model allowed us to apply a powerful

method of developing a concise, accurate linear measurement by converting individual ordinal response data (Likert scale responses) into true interval data (like a “yardstick”) allowing statistical analysis across time and populations.

Some of the weaknesses to the current analysis include the lack of data to explore the relationship between perceived team attributes and the professional role of the team member and lack of information about the influence of the organizational context in which the teams performed. Future research is needed to conduct this step. What is known, however, is that all the teams were employed by the same health system, albeit in different geographic regions and different institutional settings. Further examination of data collected by professional role and clinical care setting and the association of observed team behaviors to the measured attributes would augment the analysis of this instrument. Typical limitations of the Rasch model include the need for a large number of observations or replications to estimate the parameters of the model, which we were able to obtain in this study.²³ In addition, the Rasch model holds strong assumptions, which can be challenging to meet by the data collected. In fact, Rasch specifications are often not perfectly met but can be usefully met by careful data collection.²¹

In our experience in the clinical setting, information from the TDM is shared with team members to collectively discuss how to improve their teamwork. Survey data is presented to teams in three ways. First, an overall score of team development is calculated, and the team is shown where that score falls on the developmental scale as shown in Table 5. This provides the team with a “yardstick” measure of where they are currently in their stage of development and the potential for growth. It also assesses whether a team attribute, or domain, is “in place” (eg, the attribute is present, but the team

is still working on it), or “firmly in place” (eg, the team attribute is such that external and/or internal stresses to the team are unlikely to impact that attribute). Secondly, the team is shown a graph that plots the frequency of individual team member scores on the 0–100 developmental scale. This allows team members to begin to understand that a range of perceptions of teamness exists within the clinic. In the early development of teams, the variation of team member perceptions can be quite broad. Often, the work of the team is to reduce this variability of perceptions (getting on the “same page”), thus allowing the group to develop together over time. Using the data in this way can help teams determine what strategies can improve their team functioning. Finally, individual items from the TDM, usually three to four, are presented to the team as potential items for the team to discuss and create strategies to improve. These items are specifically chosen from the item set based on evidence that some members agreed to the item and others disagreed, thus indicative of team attribute perception misalignment. Importantly, serial measures of the same team over time allows the team to observe changes in team functioning that demonstrates improvement in the four domains being measured (cohesiveness, communication, roles and goals, and team primacy). Areas of future research include assessment of the impact of highly developed teams on patient outcomes in which surprisingly little research has been performed. These studies could provide a better understanding of the relationship between the team development measure and clinical outcome measures, team productivity, effectiveness, and the impact of interventions that promote the development of team-based care. Further validation of the tool could be accomplished through correlation of direct observation to begin to develop observable behaviors that are closely associated with development of team

attributes and vice versa. More inquiry would answer questions about potential differences in perceptions of “teamness” based on professional role, age, gender, race, urban versus rural, and hospital versus clinic teams. Some team attribute sub-domains noted in this analysis would benefit from testing of more items related to that domain, such as team primacy and leadership, and development of a shorter instrument would mitigate any item response burden. In conclusion, we developed and tested a measure of team development that has very strong psychometric properties. This tool could be used as an evaluative tool that can be linked to clinical outcomes and as a quality improvement tool to improve team function.

ACKNOWLEDGMENTS: PeaceHealth Oregon Region, Eugene, OR, and the Family Medicine Research Program at Oregon Health & Science University supported this research. The John A. Hartford Foundation Geriatrics Interdisciplinary Teams in Practice Initiative provided support for one of the authors (RS).

The authors would like to acknowledge the major contributions made to the development of the instrument by Bill Mahoney, PhD, and Carolyn Turkovich, as well as the clinicians and staff who contributed their time to complete the instrument. We also wish to thank LeNeva Spire, Publications Manager, OHSU Department of Family Medicine, for her editorial assistance.

CORRESPONDING AUTHOR: Address correspondence to Dr Stock, Oregon Health & Science University, Department of Family Medicine, Mail Code: FM, 3181 SW Sam Jackson Park Road, Portland, OR 97239-3098. 503-494-6810. stockr@ohsu.edu.

References

1. Future of Family Medicine Leadership Committee. The Future of Family Medicine: a collaborative project of the family medicine community. *Ann Fam Med* 2004;2:S3-S32.
2. Carney PA, Dietrich AJ, Keller A, Landgraf JM, O'Connor GT. Tools, teamwork, and tenacity: an office system for cancer prevention. *J Fam Pract* 1992;35(4):388-94.
3. Institute of Medicine. Crossing the quality chasm: a new health system for the 21st century. Washington, DC: National Academies Press, 2001.
4. David AK. Preparing the Personal Physician for Practice (P4): residency training in family medicine for the future. *J Am Board Fam Med* 2007;20:332-41.

5. Naylor M, Coburn K, Kurtzman E, et al. Inter-professional team-based primary care for chronically ill adults: state of the science 2010. Manuscript commissioned for ABIM Foundation and American Academy of Nursing March 24-25, 2010 meeting on advancing team-based care for the chronically ill.
6. Salas E, Bowers CA, Cannon-Bowers JA. Military team research: 10 years of progress. *Military Psychology* 1995;7(2):55-75.
7. Helmreich RL, Merritt AC, Wilhelm JA. The evolution of Crew Resource Management training in commercial aviation. *Int J Aviat Psychol* 1999;9(1):19-32.
8. McCarthy A, Garavan TN. Team learning and metacognition: a neglected area of HRD research and practice. *Advances in Developing Human Resources* 2008;10(4):509-24.
9. Boulton C, Green A, Boulton L, et al. Successful models of comprehensive health care for multi-morbid older persons: a review of effects on health and health care. Washington, DC: Institute of Medicine, National Academy of Sciences, 2008. Report commissioned for committee on Retooling for an Aging America: Building the Health Care Workforce.
10. Baker DP, Salas E, King H, Battles J, Barach P. The role of teamwork in the professional education of physicians. *J Qual Pat Safety* 2005;31(4):185-202.
11. Drinka TJK, Clark PG. Health care teamwork: interdisciplinary practice and teaching. Westport, CT: Auburn Press, 2000.
12. Fried BJ, Topping S, Rundall TG. Groups and teams in health service organizations. In: Shortell S, Kaluzny AD. Health care management: organization and design. Albany, NY: Delmar Thomson Learning Publishing, 2000; Chapter 6:154-90.
13. Rush Geriatric Interdisciplinary Team Training Project. Principles of successful team work and team competencies, Version 2.0. Rush University, June 1999.
14. Valentine MA, Nembhard IM, Edmondson AC. Measuring teamwork in health care settings, April 12, 2012. www.hbs.edu/research/pdf/11-116.pdf. Accessed May 14, 2012.
15. Battles J, King HB. TeamSTEPPS Teamwork Perceptions Questionnaire (T-TPC) Manual. Washington, DC: American Institute for Research, 2010.
16. Linacre JM. WINSTEPS® Rasch measurement computer program. 2009. Beaverton, OR. www.Winsteps.com.
17. Massof RW. The measurement of vision disability. *Optom Vis Sci* 2002;79(8):516-52.
18. Bond T, Fox C. Applying the Rasch model: fundamental measurement in the human sciences. Mahwah, NJ: Lawrence Erlbaum Associates, 2001.
19. Muthen LK, Muthen BO. Mplus: the comprehensive modeling program for applied researchers guide. Los Angeles: Muthen & Muthen, 2001.
20. Furr RM, Bacharach VR. Psychometrics: an introduction. Thousand Oaks, CA: Sage Publishers, 2008.
21. Smith RM. Introduction to Rasch Measurement: theory, models and applications. Maple Grove, MN: JAM Press, 2004.
22. Robert Graham Center for Policy Studies in Family Medicine and Primary Care. The patient-centered medical home: history, seven core features, evidence, and transformational change. November 2007. www.aafp.org/online/etc/medialib/aafp_org/documents/about/pcmh.Par.0001.File.dat/PCMH.pdf.
23. Linacre JM. Likert or Rasch? *Rasch Measurement Transactions* 1994;8(2):35.