Crosswalk August 26, 2010
OMB Comments to modifications made to
Supporting Statement A: Studying the Implementation of
a Chronic Care Toolkit and Practice Coaching
In Practices Serving Vulnerable Populations

	OMB comments	**Action taken**	**Copy of modified text**
1	May want to just add a note for clarification here that this specific ICR is not an evaluation to fulfill a statutory requirement.	Clarifying text added	This ICR is not an evaluation to fulfill a statutory requirement.
2	What measures will indicate patient care quality?	Additional description added	HEDIS or similar care quality indicators that are specific to and appropriate for the change projects selected by the participating practices will be used.
3	Please describe this data collection activity in Supporting Statements A and B, and include, as an IC the data fields that will be collected.	Language added. Data are being collected by task order project staff, not practice staff so status as an IC is unclear. Waiting for clarification from OMB on its status as IC before adding it as an IC.	This data collection will be performed by the task order project staff, not practice personnel, and will not impose a burden on the participating sites.
4	Please remove the term “impact” to distinguish this implementation evaluation from an impact evaluation.		This is a mixed qualitative and quantitative study that is designed to allow us to learn lessons about the feasibility and practicality of implementing a Chronic Care Toolkit and practice coaching in practices that serve vulnerable populations. The focus of this evaluation is on assessing the implementation of the Toolkit and coaching intervention in multiple practices and its potential contributions to improvements in care quality, organizational capacity for quality improvement, and patient and staff satisfaction.

			The qualitative methods and quantitative data collected in the study will be used to document implementation of the intervention, ensuing quality improvement activities, and changes in care quality, organizational capability for change, and patient and staff satisfaction across 20 safety net practice sites that are participating in the study.
5	How will "success" be determined? (Which measure(s) will be used to determine success, and at what threshold will these be deemed as "successful"?)		A successful implementation will be defined as: 1) consistent participation of practice members with the practice coach and participation in quality improvement activities contained in the Toolkit or suggested as a result of the meeting with the coach over the 10 month coaching period; and 2) evidence of improvements in care quality, organizational capacity for engaging in quality improvement, or improvements in staff and patient satisfaction from pre to post test determined through key-informant interviews (reports of perceived improvements across elements of the CCM and organizational capacity for change), staff and patient surveys (increases in incorporation or receipt of care consistent with elements of the CCM), and changes in HEDIS or indicators specific to chronic care activities/conditions the practices select as their area of focus for the project. An unsuccessful implementation will be defined as: 1) the consistent failure (greater than 50%) of the practice to participate in scheduled activities with the practice coach and involving activities contained in the Toolkit or suggested by the coach; and 2) failure to identify areas of perceived improvement care quality, organizational capacity, or patient and staff satisfaction. Note: This study does not employ a control group and so evidence of improvement across the identified domains cannot be directly attributed to the Toolkit and coaching intervention. However, evidence of improvements from pre to post across the identified domains may be suggestive of potential benefit from the intervention and can be used to guide design of future studies as well as future interventions.
6	Please define "success" or revise	Language changed	Specific research questions that will be addressed in evaluating the

	language. Also suggested to clarify that this evaluation is designed to measure <u>implementation process variables</u> after the toolkit was made available (without implying causation), rather than to measure care outcomes resulting from the implementation of the different programs/initiatives that the toolkit encouraged.		implementation of the Toolkit and coaching intervention are:
7	How is capacity measured/quantified?	Language added	Do practices' capacity to manage and sustain change in their organizations increase as measured by Solberg's Change Process Capability Questionnaire which assesses organizational capacity to engage in improvement activities?
8	How is quality of care measured? Are any objective outcomes (rather than patient or doctor perspectives) being recorded?	Clarifying language added	Does the quality of diabetes or other practice-identified chronic care conditions improve after practices participate in the practice coaching intervention as measured by practitioner and staff surveys, key informant interviews, and increased adherence to HEDIS or other appropriate care indicators?
9	Please make it clear here, and in all future reports of results, that if quality of care improves (or changes), no attributions of causation will be drawn to either the programs/processes implemented or to the availability of the toolkit, as there does not seem to be a control group in this	Requested language added	Please note: If changes are detected in quality of care or other variables being assessed in this evaluation, no attribution of causality will be drawn to either the program or processes implemented, or to the Toolkit availability due to lack of a control or comparison group.

	study.		
10	Please see previous comments about defining "success."	Clarifying language added	…and under what conditions practice coaching is most successful in terms both of a practice's ability to participate in the intervention, and its impact on key variables such as care quality, organizational capacity for sustained improvement, and staff and patient satisfaction.
11	It is our understanding that this is a toolkit implementation evaluation (as opposed to research). Please clarify. (Are the cited studies above potentially duplicative of this ICR? If not, no need to include.)	The two paragraphs describing previous research have been deleted as these are not potentially duplicative of the current ICR, and a summary statement added	No studies have been conducted of the Safety Net Toolkit and coaching intervention in CHCs and so this ICR is not duplicative of other previous research. The questions that will be addressed by this project will support development of improved methods for supporting practices in providing chronic care that can be more widely disseminated and implemented in clinical settings, including those serving vulnerable populations.
12	Please clarify what "the intervention" refers to. Does the "intervention" the dissemination of the toolkit?	Clarifying language added	Key informant interviews will be conducted with practice coaches at mid-point in the intervention and again at the end of the intervention. The intervention is defined as dissemination of the Toolkit combined with support from a practice coach to support quality improvement activities and development of organizational capacity for quality improvement in the practices.
13	For future information (no need to revise in this ICR), specific statistical analysis methods and tests should be described in Supporting Statement Part B (rather than Part A).	Thank you. Information noted	NA
14	How are improvements in care measured?	Additional information added	Because each practice will identify its own specific outcomes and focus for the intervention (e.g. improving diabetes care, improving

			asthma care, improving preventive care services) , these abstraction forms will be developed by the research team for each site, after each site has identified a focus for the intervention at its site. Improvements in care will be determined based on changes in adherence to HEDIS or other appropriate care quality indicators. For example, for a practice that opts to focus on diabetes care, changes in related HEDIS indicators would be tracked such as receipt of annual foot exam, annual eye exam, and appropriately timed tests of HbA1c.
15	Please clarify what "intervention" refers to – dissemination of toolkit materials or implementation of toolkit recommendations? This evaluation was designed to assess what components of, as well as how, the toolkit recommendations are being implemented. It does not test the resulting impact of these suggestions/new programs on care or health outcomes.	Language added and existing language modified	The purpose of this study is to examine the acceptability, feasibility and potential effectiveness of the CCM Practice Coaching Intervention and Toolkit by studying the implementation of practice coaching in 20 safety net practices in California. It is not an experimental study to determine the efficacy of the intervention. The intervention is defined as dissemination of the Toolkit combined with provision of on-site practice coaching support for quality improvement.
16	Please rephrase without the term "effects" or "potential effects" to clarify that this is not an impact evaluation (as stated above). Implementation, not outcome, data is in large part being collected. Any improvements in implementation processes/satisfaction/etc. cannot	Clarifying language added and existing language modified. The term has been changed from effects to benefits	Data analysis will focus on describing the implementation as well as perceived and potential benefits of the intervention at the individual practice level and comparing how implementation and potential benefits varied across the 20 replication sites.

	be attributed causally to the toolkit, or to its recommended programs, through this evaluation (without a control group).		
17	What types of measures specifically are being referred to here?	Clarifying language added	Quantitative data will be collected for this study including staff and provider assessment of CCM implementation (ACIC), patient assessment of the receipt of care consistent with the CCM (PACIC), staff and provider assessment of organizational capacity for quality improvement (CPCQ), staff (PCSSS) and patient satisfaction (CAHPS), care quality (select HEDIS indicators) and provider assessment of the degree to which the practice reflects elements of the patient centered medical home as defined by the NCQA.
18	In what cases would heuristic means be used? If reported at all, please qualify each non-statistically significant finding as such in all reports of results.		These data will be analyzed at the individual practice level using either statistical or heuristic guidelines to describe whether they improved, stayed the same or got worse. If heuristic means are used, the criteria will be determined in advance and based upon clinically important change as determined by the practice in consultation with the research team. In some instances, statistically significant changes may not reflect changes that are clinically significant, meaning they result in real improvements in care or organizational processes. In other instances, small changes that do not achieve statistical significance may be clinically significant. Heuristic criteria developed by each practice will be used to evaluate the import and potential real value of any observed changes in care, organizational capacity or satisfaction as a means of assessing the potential value of changes occurring during the course of the intervention. Any non-statistically significant findings reported from this evaluation will include language clarifying this fact.
19	Please clearly state that the changes from baseline to conclusion of this evaluation reflect increases in	Clarifying language has been added. Please note: data on quality of care are being collected using chart audits	The second step of the analysis is to look at patterns of change across all 20 practices thus treating each practice as a replication. The changes from baseline to conclusion of this evaluation will reflect changes in behaviors/strategies encouraged by the Toolkit

	implementation of behaviors/strategies encouraged by the toolkit. "Improvements" in care or outcomes are not measured.	conducted by Task Order project staff, not practice staff, and are being collected for exploratory and descriptive purposes only. They will not be used to determine effectiveness or to suggest a causal relationship between the intervention and any observed improvements (or declines) due to limitations of the current study design. All reporting of findings will include language clarifying this.	and practice coach intervention. Improvements in patient outcomes are not being assessed. Data on quality of care indicators selected by each project will be collected but will be treated as descriptive and exploratory data. These data will be not be analyzed in order to determine statistically significant improvements in care nor used to ascribe causality given the lack of a control group. The changes being measured will include changes in processes of care (increased implementation of components of the CCM) and organizational capacity for engaging in quality improvement and on criteria reflecting quality of care defined by the practices.
20	What are some examples of what these criteria might be? (Is there a list for sites to draw from?) Are they all categorical variables? Why would a chi square test be used here? How would thresholds for each criterion (forming the three cells) be determined? Why three categories?		When possible, these criteria will be based on HEDIS quality indicators or those of another well recognized source. An example of these criteria might be HEDIS indicators for diabetes care (annual foot exam, annual eye exam, appropriately timed visits and lab tests). Care issues that practices are expected to be interested in include areas such as management of asthma, diabetes, preventive services, cardiovascular disease, depression, hypertension, and obesity among others. The majority of these variables that would be used to explore possible improvements in these areas are expected to be categorical. Improvements in patient outcomes are not being assessed. Each practice will be rated as either improved, stayed the same or got worse on their individually identified care criterion. The use of 3 categories was chosen to reflect the level of precision that we thought was reasonable given the nature of the observations involved and the fact that different practices would be choosing different criteria to evaluate. We were looking for a level of precision which

			would give us a common metric among the various possible change criteria. The results of these individual analyses will be assessed using a Chi Square test of Goodness of Fit where the null hypothesis would be that a third of the sample would fall into each cell if the pattern was truly random. If the universe of possible change states is defined as "stayed the same", "got better" or "got worse", then a truly random outcome would specify that each of the conditions would have equal probability of occurrence. The chi square test of goodness of fit takes this condition as the null against which to test the distribution of sample outcomes. In the present case, we would hope that the distribution of the sample outcomes would shift toward the "improved" category and this would result in a distribution of outcome states that is no longer .33, .33, .33. For example with 21 practices the null would hypothesize a distribution of 7 practices in each outcome state. If the sample data had a distribution of 3 practices in the 'got worse' category, 6 practices in the stayed the same category and 12 practices in the got better category, the null could be rejected at the .05 level f significance with a chi square of 6.0 and 2 df. Other distributions of sample data would also be possible that would also lead to a rejection of the null hypothesis of equal probability. Thresholds will be individually defined and specified at the practice level and will include careful operational definitions so that each outcome state can be reliably determined. For example, the criterion may be that a practice has to have more than a 20% increase in # of specific types of evaluations to be considered 'improved' and that a reduction of 10% in the number of evaluations would be categorized as 'got worse', while staying within the -10% to +20% band would result in a category of 'stayed the same'.
21	Please remove this statement. This study is an implementation	Statement: "evaluate the effectiveness of the	With 20 practices this approach does not have high power, but the patterns of change across independent practices at different times

	evaluation, not an impact (or effectiveness) evaluation.	intervention" removed and replaced with language on implementation	linked to similar interventions, accompanied by clear descriptions of each intervention should allow us to describe the implementation of the intervention and variations in its implementation across the 20 sites, and to explore the potential benefit (or lack of benefit) of the intervention to the different sites.
22	Please define.	Explanatory text added	Baseline levels of compliance (present/not present) with HEDIS (or other relevant care indicator) will be calculated at the practice level and will be compared to follow-up rates using the Z test for a single proportion where the baseline rate will define the null condition and the follow-up rate will be the alternate condition. For example, an assessment of compliance with HEDIS indicators for diabetes might be assessed as follows: annual foot exam (present), annual eye exam (not present), planned visit (not present), etc.
23	Please add that data will be checked for such an occurrence (worsening outcomes)	Analysis changed from one-sided to two-sided tests to allow detection of both improvements and the possibility of worsening performance	Two-sided tests will be used to allow us to assess for potential improvements as well as the possibility of worsening outcomes.
24	What might these reports look like? What findings specifically may be reported? Please state that a thorough discussion of limitations will be included in all reporting of results, including the fact that causation of any outcomes cannot be attributed causally to the presence	Additional description added and requested content of limitations of study added	A report describing evaluation findings will be developed that will describe: (1) The course of the intervention across the 20 sites and observed variations in its implementation across these sites (for example, variations in degree of practice participation in the intervention (hours allocated, buy-in, content, activities engaged in, etc), variations in ways practices engaged with or used the Toolkit and practice coach, variations in focus for the various quality improvement activities, and variations in approach by practices in undertaking the activities – for example use of a teamlette vs. a full quality improvement team); (2) Degree of observed changes in implementation of the CCM, organizational capacity for change, and staff/patient satisfaction based on interviews and surveys and variations across sites in degree and nature of change; (3) Barriers and facilitators to implementation and use of the Toolkit and practice

			coach and variations in these across sites; (4) Perceived value and feasibility of use of the Toolkit and practice coaching intervention and variations in this across sites; (5) Lessons learned in implementing the intervention that may have relevance to future efforts. A thorough discussion of the limitation of the study will be included in all reports.
25	Are there privacy concerns with this method?	We have utilized electronic data sources such as registries, billing records and EHRs at these practices in past studies and will ensure that all data collection methods, whether using paper records or existing electronic data, are HIPAA compliant, adhere to human subjects regulations regarding confidentiality and secure data transmission and storage. Clarifying language has been added	For practices with electronic data, we will work with IT personnel at the site to obtain these same data through their electronic health record or registries. All data abstraction will be HIPAA compliant and utilize appropriate security to protect confidentiality of data. Electronic or hand abstracted data will then be entered into a project database for analysis.
24	These examples have to do with how frequently tests/exams are performed – presumably more is better. My question is, does higher frequency testing in these areas count as an improvement in care per se? Is higher frequency testing always associated positively with improved outcomes?	Language added to page 17 of statement	Higher frequency of testing does not equate directly with improvements in care, is not always associated positively with improved outcomes, and is not in and of itself an effective method for assessing improvements in care. Because of this, we will use HEDIS measures to assess quality, not frequency of testing. Frequency of testing will be considered only to the degree to which it is included as a part of the HEDIS measures. HEDIS is a tool developed by the National Committee for Quality Assurance (NCQA) to measure performance on important dimensions of care and service. It consists of 71 measures across 8 domains of care and is used by 90 percent of health plans in America

			to measure quality, and frequently used by researchers to assess improvements in care. More information on HEDIS measures can be accessed online at: http://www.ncqa.org/tabid/187/default.aspx In the diabetes example provided on page 17 of the supporting statement, HEDIS measures for quality of diabetes care include (but are not limited to): a) percent of diabetic patients at practice receiving an annual foot exam, b) percent of patients at practice receiving annual eye exam, and c) percent of patients receiving annual HbA1c test. So for practices in this project that opt to focus on improving chronic care for diabetes, we will use these measures to assess improvements or declines in quality. Practices that show an increase in percentage of patients receiving care consistent with these measures will be counted as improved and visa versa. Our focus will be only on consistency with HEDIS measures. Care that extends beyond that recommended in HEDIS will not be considered in this project. For example, a patient who receives the annual foot exam as recommended in HEDIS and in addition to this then receives an additional 3 foot exams in the year, will be considered to have received care consistent with the HEDIS standard of an annual foot exam, but will not be considered to have received incrementally better care than another patient who received only the single annual foot exam per HEDIS standards.