6-19-2020

Quantitative measures of health policy implementation determinants and outcomes: A systematic review

Peg Allen

Meagan Pilar

Callie Walsh-Bailey

Cole Hooley

Stephanie Mazzucca

See next page for additional authors

Authors

Peg Allen, Meagan Pilar, Callie Walsh-Bailey, Cole Hooley, Stephanie Mazzucca, Cara C Lewis, Kayne D Mettert, Caitlin N Dorsey, Jonathan Purtle, Maura M Kepper, Ana A Baumann, and Ross C Brownson

Quantitative measures of health policy implementation determinants and outcomes: a systematic review

Peg Allen[1*] , Meagan Pilar[1], Callie Walsh-Bailey[1], Cole Hooley[2], Stephanie Mazzucca[1], Cara C. Lewis[3], Kayne D. Mettert[3], Caitlin N. Dorsey[3], Jonathan Purtle[4], Maura M. Kepper[1], Ana A. Baumann[5] and Ross C. Brownson[1,6]

Abstract

Background: Public policy has tremendous impacts on population health. While policy development has been extensively studied, policy implementation research is newer and relies largely on qualitative methods. Quantitative measures are needed to disentangle differential impacts of policy implementation determinants (i.e., barriers and facilitators) and outcomes to ensure intended benefits are realized. Implementation outcomes include acceptability, adoption, appropriateness, compliance/fidelity, feasibility, penetration, sustainability, and costs. This systematic review identified quantitative measures that are used to assess health policy implementation determinants and outcomes and evaluated the quality of these measures.

Methods: Three frameworks guided the review: Implementation Outcomes Framework (Proctor et al.), Consolidated Framework for Implementation Research (Damschroder et al.), and Policy Implementation Determinants Framework (Bullock et al.). Six databases were searched: Medline, CINAHL Plus, PsycInfo, PAIS, ERIC, and Worldwide Political. Searches were limited to English language, peer-reviewed journal articles published January 1995 to April 2019. Search terms addressed four levels: health, public policy, implementation, and measurement. Empirical studies of public policies addressing physical or behavioral health with quantitative self-report or archival measures of policy implementation with at least two items assessing implementation outcomes or determinants were included. Consensus scoring of the Psychometric and Pragmatic Evidence Rating Scale assessed the quality of measures.

Results: Database searches yielded 8417 non-duplicate studies, with 870 (10.3%) undergoing full-text screening, yielding 66 studies. From the included studies, 70 unique measures were identified to quantitatively assess implementation outcomes and/or determinants. Acceptability, feasibility, appropriateness, and compliance were the most commonly measured implementation outcomes. Common determinants in the identified measures were organizational culture, implementation climate, and readiness for implementation, each aspects of the internal setting. Pragmatic quality ranged from adequate to good, with most measures freely available, brief, and at high school reading level. Few psychometric properties were reported.

(Continued on next page)

---

* Correspondence: pegallen@wustl.edu
[1]Prevention Research Center, Brown School, Washington University in St. Louis, One Brookings Drive, Campus Box 1196, St. Louis, MO 63130, USA
Full list of author information is available at the end of the article

(Continued from previous page)

Conclusions: Well-tested quantitative measures of implementation internal settings were under-utilized in policy studies. Further development and testing of external context measures are warranted. This review is intended to stimulate measure development and high-quality assessment of health policy implementation outcomes and determinants to help practitioners and researchers spread evidence-informed policies to improve population health.

Registration: Not registered

Keywords: Systematic review, Implementation science, Health policy, Policy implementation, Implementation, Public policy, Measures, Psychometric, Pragmatic

Contributions to the literature

- This systematic review identified 70 quantitative measures of implementation outcomes or determinants in health policy studies.
- Readiness to implement and organizational climate and culture were commonly assessed determinants, but fewer studies assessed policy actor relationships or implementation outcomes of acceptability, fidelity/compliance, appropriateness, feasibility, or implementation costs.
- Study team members rated most identified measures' pragmatic properties as good, meaning they are straightforward to use, but few studies documented pilot or psychometric testing of measures.
- Further development and dissemination of valid and reliable measures of policy implementation outcomes and determinants can facilitate identification, use, and spread of effective policy implementation strategies.

Background

Despite major impacts of policy on population health [1–7], there have been relatively few policy studies in dissemination and implementation (D&I) science to inform implementation strategies and evaluate implementation efforts [8]. While health outcomes of policies are commonly studied, fewer policy studies assess implementation processes and outcomes. Of 146 D&I studies funded by the National Institutes of Health (NIH) through D&I funding announcements from 2007 to 2014, 12 (8.2%) were policy studies that assessed policy content, policy development processes, or health outcomes of policies, representing 10.5% of NIH D&I funding [8]. Eight of the 12 studies (66.7%) assessed health outcomes, while only five (41.6%) assessed implementation [8].

Our ability to explore the differential impact of policy implementation determinants and outcomes and disentangle these from health benefits and other societal outcomes requires high quality quantitative measures [9]. While systematic reviews of measures of implementation of evidence-based interventions (in clinical and community settings) have been conducted in recent years [10–13], to our knowledge, no reviews have explored the quality of quantitative measures of determinants and outcomes of policy implementation.

Policy implementation research in political science and the social sciences has been active since at least the 1970s and has much to contribute to the newer field of D&I research [1, 14]. Historically, theoretical frameworks and policy research largely emphasized policy development or analysis of the content of policy documents themselves [15]. For example, Kingdon's Multiple Streams Framework and its expansions have been widely used in political science and the social sciences more broadly to describe how factors related to sociopolitical climate, attributes of a proposed policy, and policy actors (e.g., organizations, sectors, individuals) contribute to policy change [16–18]. Policy frameworks can also inform implementation planning and evaluation in D&I research. Although authors have named policy stages since the 1950s [19, 20], Sabatier and Mazmanian's Policy Implementation Process Framework was one of the first such frameworks that gained widespread use in policy implementation research [21] and later in health promotion [22]. Yet, available implementation frameworks are not often used to guide implementation strategies or inform why a policy worked in one setting but not another [23]. Without explicit focus on implementation, the intended benefits of health policies may go unrealized, and the ability may be lost to move the field forward to understand policy implementation (i.e., our collective knowledge building is dampened) [24].

Differences in perspectives and terminology between D&I and policy research in political science are noteworthy to interpret the present review. For example, Proctor et al. use the term implementation outcomes for what policy researchers call policy outputs [14, 20, 25].

To non-D&I policy researchers, policy implementation outcomes refer to the health outcomes in the target population [20]. D&I science uses the term fidelity [26]; policy researchers write about compliance [20]. While D&I science uses the terms outer setting, outer context, or external context to point to influences outside the implementing organization [26–28], non-D&I policy research refers to policy fields [24] which are networks of agencies that carry out policies and programs.

Identification of valid and reliable quantitative measures of health policy implementation processes is needed. These measures are needed to advance from classifying constructs to understanding causality in policy implementation research [29]. Given limited resources, policy implementers also need to know which aspects of implementation are key to improve policy acceptance, compliance, and sustainability to reap the intended health benefits [30]. Both pragmatic and psychometrically sound measures are needed to accomplish these objectives [10, 11, 31, 32], so the field can explore the influence of nuanced determinants and generate reliable and valid findings.

To fill this void in the literature, this systematic review of health policy implementation measures aimed to (1) identify quantitative measures used to assess health policy implementation outcomes (IOF outcomes commonly called policy outputs in policy research) and inner and outer setting determinants, (2) describe and assess pragmatic quality of policy implementation measures, (3) describe and assess the quality of psychometric properties of identified instruments, and (4) elucidate health policy implementation measurement gaps.

Methods
The study team used systematic review procedures developed by Lewis and colleagues for reviews of D&I research measures and received detailed guidance from the Lewis team coauthors for each step [10, 11]. We followed the PRISMA reporting guidelines as shown in the checklist (Supplemental Table 1). We have also provided a publicly available website of measures identified in this review (https://www.health-policy-measures.org/).

For the purposes of this review, *policy* and *policy implementation* are defined as follows. We deemed public policy to include legislation at the federal, state/province/regional unit, or local levels; and governmental regulations, whether mandated by national, state/province, or local level governmental agencies or boards of elected officials (e.g., state boards of education in the USA) [4, 20]. Here, public policy implementation is defined as the carrying out of a governmental mandate by public or private organizations and groups of organizations [20].

Two widely used frameworks from the D&I field guide the present review, and a third recently developed framework that bridges policy and D&I research. In the

Implementation Outcomes Framework (IOF), Proctor and colleagues identify and define eight implementation outcomes that are differentiated from health outcomes: acceptability, adoption, appropriateness, cost, feasibility, fidelity, penetration, and sustainability [25]. In the Consolidated Framework for Implementation Research (CFIR), Damschroder and colleagues articulate determinants of implementation including the domains of intervention characteristics, outer setting, inner setting of an organization, characteristics of individuals within organizations, and process [33]. Finally, Bullock developed the Policy Implementation Determinants Framework to present a balanced framework that emphasizes both internal setting constructs and external setting constructs including policy actor relationships and networks, political will for implementation, and visibility of policy actors [34]. The constructs identified in these frameworks were used to guide our list of implementation determinants and outcomes.

Searches
Through EBSCO, we searched MEDLINE, PsycInfo, and CINAHL Plus. Through ProQuest, we searched PAIS, Worldwide Political, and ERIC. Due to limited time and staff in the 12-month study, we did not search the grey literature. We used multiple search terms in each of four required levels: health, public policy, implementation, and measurement (Table 1). Table 1 shows search terms for each string. Supplemental Tables 2 and 3 show the final search syntax applied in EBSCO and ProQuest.

The authors developed the search strings and terms based on policy implementation framework reviews [34, 35], additional policy implementation frameworks [21, 22], labels and definitions of the eight implementation outcomes identified by Proctor et al. [25], CFIR construct labels and definitions [9, 33], and additional D&I research and search term sources [28, 36–38] (Table 1). The full study team provided three rounds of feedback on draft terms, and a library scientist provided additional synonyms and search terms. For each test search, we calculated the percentage of 18 benchmark articles the search captured. We determined a priori 80% as an acceptable level of precision.

Inclusion and exclusion criteria
This review addressed only measures of implementation by organizations mandated to act by governmental units or legislation. Measures of behavior changes by individuals in target populations as a result of legislation or governmental regulations and health status changes were outside the realm of this review.

There were several inclusion criteria: (1) empirical studies of the implementation of public policies already passed or approved that addressed physical or behavioral

Table 1 Search terms and strings

String	Search terms
Health	"health" OR "healthcare" OR "healthy" OR "healthier" OR "wellness"
Public policy	"policy" OR "policies" OR "law" OR "laws" OR "legislation" OR "legislative" OR "statute" OR "statutes" OR "regulation" OR "regulations" OR "regulatory" OR "executive order" OR "executive orders" OR "congress" OR "congresses" OR "congressional" OR "city council" OR "city councils" OR "county council" OR "county councils" OR mandat* OR "ordinance" OR "ordinances" OR "rule" OR "rules"
Implementation	"implement*" OR disseminat* OR "institutionalization" OR "institutionalisation" OR "integrate" OR "integrates" OR "integrated" OR "integrating" OR "integration" OR "integrations" OR "knowledge transfer" OR "knowledge exchange" OR "knowledge translation" OR "knowledge diffusion" OR "knowledge utilization" OR "research utilization" OR "innovation"
Measurement	"measure" OR "measures" OR "measurement" OR "measurements" OR "instrument" OR "instruments" OR "survey" OR "surveys" OR "questionnaire" OR "questionnaires" OR "scale" OR "scales" OR "self-report" OR "self-reports" OR "self-reported" OR "archived data" OR "archival data" OR "quantitative" OR "quantitatively" OR "inventory" OR "inventories" OR "rating" OR "ratings" OR "assessment form" OR "assessment forms" OR "evaluation form" OR "evaluation forms" OR "tool" OR "tools" OR "index" OR "indexes" OR "indices"

health, (2) quantitative self-report or archival measurement methods utilized, (3) published in peer-reviewed journals from January 1995 through April 2019, (4) published in the English language, (5) public policy implementation studies from any continent or international governing body, and (6) at least two transferable quantitative self-report or archival items that assessed implementation determinants [33, 34] and/or IOF implementation outcomes [25]. This study sought to identify transferable measures that could be used to assess multiple policies and contexts. Here, a transferable item is defined as one that needed no wording changes or only a change in the referent (e.g., policy title or topic such as tobacco or malaria) to make the item applicable to other policies or settings [11]. The year 1995 was chosen as a starting year because that is about when web-based quantitative surveying began [39]. Table 2 provides definitions of the IOF implementation outcomes and the selected determinants of implementation. Broader constructs, such as readiness for implementation, contained multiple categories.

Exclusion criteria in the searches included (1) non-empiric health policy journal articles (e.g., conceptual articles, editorials); (2) narrative and systematic reviews; (3) studies with only qualitative assessment of health policy implementation; (4) empiric studies reported in theses and books; (5) health policy studies that only assessed health outcomes (i.e., target population changes in health behavior or status); (6) bill analyses, stakeholder perceptions assessed to inform policy development, and policy content analyses without implementation assessment; (7) studies of changes made in a private business not encouraged by public policy; and (8) countries with authoritarian regimes. We electronically programmed the searches to exclude policy implementation studies from countries that are not democratically governed due to vast differences in policy environments and implementation factors.

Screening procedures
Citations were downloaded into EndNote version 7.8 and de-duplicated electronically. We conducted dual independent screening of titles and abstracts after two group pilot screening sessions in which we clarified inclusion and exclusion criteria and screening procedures. Abstract screeners used Covidence systematic review software [40] to code inclusion as yes or no. Articles were included in full-text review if one screener coded it as meeting the inclusion criteria. Full-text screening via dual independent screening was coded in Covidence [40], with weekly meetings to reach consensus on inclusion/exclusion discrepancies. Screeners also coded one of the pre-identified reasons for exclusion.

Data extraction strategy
Extraction elements included information about (1) measure meta-data (e.g., measure name, total number of items, number of transferable items) and studies (e.g., policy topic, country, setting), (2) development and testing of the measure, (3) implementation outcomes and determinants assessed (Table 2), (4) pragmatic characteristics, and (5) psychometric properties. Where needed, authors were emailed to obtain the full measure and measure development information. Two coauthors (MP, CWB) reached consensus on extraction elements. For each included measure, a primary extractor conducted initial entries and coding. Due to time and staff limitations in the 12-month study, we did not search for each empirical use of the measure. A secondary extractor checked the entries, noting any discrepancies for discussion in consensus meetings. Multiple measures in a study were extracted separately.

Quality assessment of measures
To assess the quality of measures, we applied the Psychometric and Pragmatic Evidence Rating Scales (PAPERS) developed by Lewis et al. [10, 11, 41, 42]. PAPERS includes assessment of five pragmatic instrument characteristics that affect the level of ease or difficulty to use the instrument: brevity (number of items), simplicity of language (readability level), cost (whether it is freely available), training burden (extent of data collection training needed), and analysis burden (ease or difficulty

Allen et al. Implementation Science   (2020) 15:47

Page 5 of 17

Table 2 Health policy implementation outcomes and determinants assessed in included measures (N = 70 unique measures in 66 health policy implementation studies)

Domain	Construct	Included measures (N = 70) n (%)	Definition	Source
Implementation outcomes	Acceptability	17 (24%)	Perceptions by staff in organizations mandated to implement the policy, or perceptions of other stakeholders, that the policy mandate is agreeable, palatable, or satisfactory	Proctor et al. 2011 [25]
	Adoption*	8 (11%)	Intention and initial actions of mandated organizations to revise their organizational policies to address policy mandates (not policy development or passage of bills into law).	Proctor et al. 2011 [25]
	Appropriateness	12 (17%)	"Perceived fit, relevance, or compatibility of the [policy] for a given practice setting, provider, or consumer; and/or perceived fit of the [policy] to address a particular issue or problem"; context fit	Proctor et al. 2011, pg. 69 [25]
	Costs	10 (14%)	"Cost impact of an implementation effort"	Proctor et al. 2011, pg. 69 [25]
	Feasibility	12 (17%)	"Extent to which a new [policy] can be successfully used or carried out within a given agency or setting" Level of administration required to implement a policy, often called policy automaticity	Proctor et al. 2011, pg. 69 [25] Howlett et al. 2015 [19]
	Fidelity/compliance	18 (26%)	"Degree to which a [policy] was implemented as it was prescribed" [mandated]	Proctor et al. 2011, pg. 69 [25]
	Penetration	8 (11%)	"Integration of a [policy] within a service setting and its subsystems"	Proctor et al. 2011, pg. 70 [25]
	Sustainability	1 (1%)	"Extent [new policy] is maintained or institutionalized within a service setting's ongoing, stable operations"	Proctor et al. 2011, pg. 70 [25]
Determinants of implementation assessed	Adaptability	7 (10%)	"Degree to which an intervention can be adapted, tailored, refined, or reinvented to meet local needs"	Damschroder et al. 2009, pg. 6 [33]
	Complexity	4 (6%)	"Perceived difficulty of implementation, reflected by duration, scope, radicalness, disruptiveness, centrality, and intricacy and number of steps required to implement"	Damschroder 2009, pg. 6 [33]
	Presence of champions	3 (4%)	Field or practice leaders, people who can facilitate, and support practice change among professionals	Bullock 2019 [34], Damschroder et al. 2009 [33]
	Organizational culture and climate (general)	27 (39%)	Culture: "Norms, values, and basic assumptions of a given organization"; or Climate: "Absorptive capacity for change", extent policy compliance will be "rewarded, supported, and expected within their organization"	Damschroder et al. 2009, pg. 8 [33] Damschroder et al. 2009, pg.8 [33]
	Policy implementation climate	16 (23%)		
	a. Goals and feedback	6 (9%)	"Degree [the policy-mandate] goals are clearly communicated, acted upon, and fed back to staff and alignment of that feedback with goals"	Damschroder et al. 2009, pg. 9 [33]
	b. Relative priority	8 (11%)	"Individuals' shared perception of importance of the [policy] implementation within the organization", competing priorities	Damschroder et al. 2009, pg. 8 [33]

Table 2 Health policy implementation outcomes and determinants assessed in included measures (*N* = 70 unique measures in 66 health policy implementation studies) *(Continued)*

Domain	Construct	Included measures (*N* = 70) *n* (%)	Definition	Source
	Readiness for implementation	43 (61%)		Damschroder et al. 2009 [33]
	a. Communication of policy	22 (31%)	Actions taken to disseminate policy requirements and guidelines to implementers.	Identified in screening [33]
	b. Policy awareness and knowledge	18 (26%)	Implementing staff/provider awareness the policy mandate exists, or knowledge of policy content	Identified in screening [33]
	c. Leadership for implementation	13 (19%)	"Commitment, involvement, and accountability of leaders and managers with the implementation"	Damschroder et al. 2009, pg. 9 [33]
	d. Training	14 (20%)	Training of staff/providers on how to implement the policy-mandated practices	Identified in screening [33]
	e. Non-training resources	19 (27%)	"Level of resources dedicated for implementation and on-going operations including money…physical space, and time" other than training resources	Damschroder et al. 2009, pg. 9 [33]
	Structure of organization	2 (3%)	"The social architecture, age, maturity, and size of an organization"	Damschroder et al. 2009, pg. 7 [33]
	Actor relationships and networks	12 (17%)	Presence and characteristics of relationships between parallel organizations that must collaborate for policy implementation to be effective	Bullock 2019 [34]
	Visibility of policy role/policy actors	7 (10%)	Perceived presence and importance of different actors pertinent to implementation of the policy	Bullock 2019 [34]
	Political will for policy implementation	8 (11%)	Societal desire and commitment to generate resources to carry out policies	Bullock 2019 [34]
	Target population characteristics	3 (4%)	Demographics, norms, neighborhood environments of population groups that affect implementation	Bullock 2019 [34]

of interpretation of scoring and results). Lewis and colleagues developed the pragmatic domains and rating scales with stakeholder and D&I researchers input [11, 41, 42] and developed the psychometric rating scales in collaboration with D&I researchers [10, 11, 43]. The psychometric rating scale has nine properties (Table 3): internal consistency; norms; responsiveness; convergent, discriminant, and known-groups construct validity; predictive and concurrent criterion validity; and structural validity. In both the pragmatic and psychometric scales, reported evidence for each domain is scored from poor (− 1), none/not reported (0), minimal/emerging (1), adequate (2), good (3), or excellent (4). Higher values are indicative of more desirable pragmatic characteristics (e.g., fewer items, freely available, scoring instructions, and interpretations provided) and stronger evidence of psychometric properties (e.g., adequate to excellent reliability and validity) (Supplemental Tables 4 and 5).

Data synthesis and presentation
This section describes the synthesis of measure transferability, empiric use study settings and policy topics, and

PAPERS scoring. Two coauthors (MP, CWB) consensus coded measures into three categories of item transferability based on quartile item transferability percentages: mostly transferable (≥ 75% of items deemed transferable), partially transferable (25–74% of items deemed transferable), and setting-specific (< 25% of items deemed transferable). Items were deemed transferable if no wording changes or only a change in the referent (e.g., policy title or topic) was needed to make the item applicable to the implementation of other policies or in other settings. Abstractors coded study settings into one of five categories: hospital or outpatient clinics; mental or behavioral health facilities; healthcare cost, access, or quality; schools; community; and multiple. Abstractors also coded policy topics to healthcare cost, access, or quality; mental or behavioral health; infectious or chronic diseases; and other, while retaining documentation of subtopics such as tobacco, physical activity, and nutrition. Pragmatic scores were totaled for the five properties, with possible total scores of − 5 to 20, with higher values indicating greater ease to use the instrument. Psychometric property total scores for the nine properties were also calculated, with possible scores

Table 3 Psychometric and Pragmatic Evidence Rating Scale (PAPERS) domains and definitions

Scale	Domain	Definition
Pragmatic criteria	Brevity	Number of items; excellent < 10 items
	Language simplicity	Readability of items, ranging from accessible only to experts (poor) to readable at or below an 8th grade level (excellent)
	Cost to use instrument	Monetary amount researchers pay to use the instrument; excellent = freely available in the public domain
	Training ease	Extent of assessor burden due to required trainings versus manualized self-training; excellent = no training required by instrument developer
	Analysis ease	Extent of assessor burden due to complexity of scoring interpretation; excellent = cutoff scores with value labels and automated calculations
Psychometric properties	Norms	A measure of generalizability based on sample size and means and standard deviations of item values
	Internal consistency	Reliability
	Convergent construct validity	Observed association in data of two theoretically related constructs, assessed through effect sizes and correlations
	Discriminant construct validity	Observed differentiation (lack of association) of two theoretically distinct constructs, assessed through effect sizes and correlations
	Known-groups validity	Extent to which groups known to have different characteristics can be differentiated by the measure
	Predictive criterion validity	Extent to which a measure can predict or be associated with an outcome measured at a future time
	Concurrent criterion validity	Correlation of a measure's observed scores with scores from a previously established measure of the construct
	Responsiveness	Extent to which a measure can detect changes over time, i.e., clinically important not just statistically significant changes over time
	Structural validity	Structure of test covariance, i.e., extent to which groups of items increase or decrease together versus a different pattern, assessed by goodness of fit of factor analyses or principal component analyses

Lewis et al. [11], Stanick et al. [42]
Each domain is scored from poor (− 1), none/not reported (0), minimal/emerging (1), adequate (2), good (3), or excellent (4). Specific rating scales for each domain are provided in Supplemental Tables 4 and 5

of − 9 to 36, with higher values indicating evidence of multiple types of validity.

Results

Yields

The database searches yielded 11,684 articles, of which 3267 were duplicates (Fig. 1). Titles and abstracts of the 8417 articles were independently screened by two team members; 870 (10.3%) were selected for full-text screening by at least one screener. Of the 870 studies, 804 were excluded at full-text screening or during extraction attempts with the consensus of two coauthors; 66 studies were included. Two coauthors (MP, CWB) reached consensus on extraction and coding of information on 70 unique quantitative eligible measures identified in the 66 included studies plus measure development articles where obtained. Nine measures were used in more than one included study. Detailed information on identified measures is publicly available at https://www.health-policy-measures.org/.

The most common exclusion reason was lack of transferable items in quantitative measures of policy implementation (*n* = 597) (Fig. 1). While this review focused on transferable measures across any health issue or setting,

researchers addressing specific health policies or settings may find the excluded studies of interest. The frequencies of the remaining exclusion reasons are listed in Fig. 1.

A variety of health policy topics and settings from over two dozen countries were found in the database searches. For example, the searches identified quantitative and mixed methods implementation studies of legislation (such as tobacco smoking bans), regulations (such as food/menu labeling requirements), governmental policies that mandated specific clinical practices (such as vaccination or access to HIV antiretroviral treatment), school-based interventions (such as government-mandated nutritional content and physical activity), and other public policies.

Among the 70 unique quantitative implementation measures, 15 measures were deemed mostly transferable (at least 75% transferable, Table 4). Twenty-three measures were categorized as partially transferable (25 to 74% of items deemed transferable, Table 5); 32 measures were setting-specific (< 25% of items deemed transferable, data not shown).

Implementation outcomes

Among the 70 measures, the most commonly assessed implementation outcomes were fidelity/compliance of
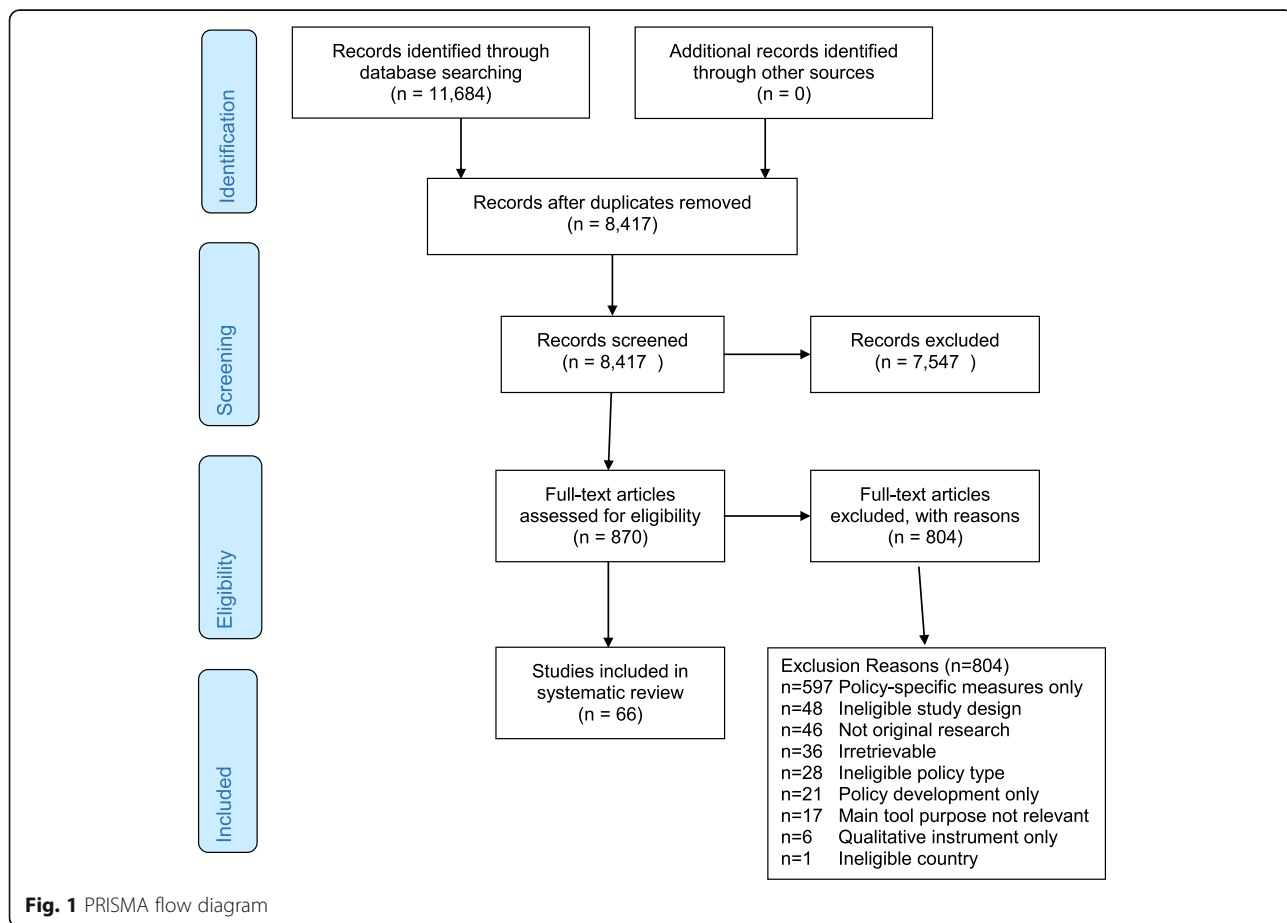
Fig. 1 PRISMA flow diagram

Table 4 Mostly transferable measures identified in studies of health policy implementation (n = 15)

Tool name	Number of items	Development Author, year	Empirical use Author, year Setting, country	Implementation outcomes and determinants assessed	Pragmatic PAPERS score[a]	Psychometric properties assessed
Adaptations of Evidence-Based Practices	9	Stirman et al. 2013 [44]	Lau and Brookman-Frazee 2016 [45] Mental health, USA	Fidelity/compliance, adaptability	12	Norms
Creative Climate Questionnaire	10	Ekvall 1996 [46]	Lövgren 2002 [47] Healthcare, Sweden	Organizational culture and climate	13	Norms[b]
Job Control Scale	22	Dwyer and Ganster 1991 [48]	Condon-Paoloni 2015 [49] Nutrition, Australia	Organizational culture/climate	12	Norms, internal consistency
Organizational Climate Measure	82	Patterson et al. 2005 [50]	Lau and Brookman-Frazee 2016 [45] Mental health, USA	Organizational culture/climate	10	Norms[b]
Organizational Social Context Measurement System	105	Glisson et al. 2012 [51]	Beidas et al. 2013 [52] Mental or behavioral health, USA	Organizational culture/climate, communication of policy	5	Norms, structural validity
Perceived Organizational Support Survey	8	Eisenberger et al. 1997 [53]	Eby et al. 2013 [54] Tobacco, USA	Organizational culture/climate	12	Norms, structural validity, internal consistency
Pharmaceutical Policies Survey	17	Vogler et al. 2016 [55]	Vogler et al. 2016 [55] Healthcare, Europe	Costs of implementation	11	Norms
Planning for Change Survey	4	Wanberg 2000 [56]	Eby et al. 2013 [54] Tobacco, USA	Organizational culture/climate	12	Norms, structural validity, internal consistency
Policy Coalition Evaluation Tool	15	Hardy et al. 2013 [57]	Hardy et al. 2013 [57] Community nutrition, USA	Fidelity/compliance, sustainability, readiness, organizational culture/climate, actor relationships/networks	9	Not reported
Policy Empowerment Index	12	Gavriilidis and Östergren 2012 [58]	Gavriilidis and Östergren 2012 [58] Hospitals/clinics, traditional medicine policy, South Africa	Adaptability, readiness, actor relationships, political will for implementation, target population characteristics affecting implementation	16	Not reported
Policy Implementation Barometer	10	Hongoro et al. 2018 [59]	Hongoro et al. 2018 [59] Access to care, Uganda	Appropriateness, readiness to implement	11	Norms
Policy Readiness and Stage of Change Needs Assessment	130	Roeseler et al. 2016 [60]	Roeseler et al. 2016 [60] Tobacco, USA	Adoption, fidelity/compliance	13	Norms
Rehabilitation Policy Questionnaire	7	Brämberg et al. 2015 [61]	Brämberg et al. 2015 [61] Hospitals/clinics, Sweden	Acceptability, adoption, fidelity/compliance, penetration, readiness	11	Norms
Rütten's Health Policy Questionnaire	24	Rütten et al. 2003 [62]	Rütten et al. 2003 [62] Cancer, tobacco, physical activity, Europe (6 countries)	Acceptability, cost, org culture/climate, readiness to implement, political will implementation	15	Norms[b]
Veteran's Administration All Employee Survey	14	Smith et al. 2017 [63]	Smith et al. 2017 [63] Mental health, USA	Organizational culture/climate	11	Norms[b]

Mostly transferable measures are defined here as those in which ≥ 75% of items can readily be used in multiple settings without change or by changing only the referent (i.e., policy name, setting)

[a]Pragmatic PAPERS score—Psychometric and Pragmatic Evidence Rating Scale [11, 41, 42], five domains assessed: brevity (score based on number of items), language simplicity, burden/ease of interpretation of scoring, and training burden, total possible score 20, higher numbers indicate greater ease to use the measure

[b]Additional subscale level psychometric properties were reported

Allen *et al. Implementation Science* (2020) 15:47

Page 10 of 17

Table 5 Partially transferable measures identified in studies of health policy implementation (*n* = 23)

Tool name	Number of items	Development Author, year	Empirical use Author, year Setting/topic, country	Implementation outcomes and determinants assessed	Pragmatic PAPERS score[a]	Psychometric properties assessed
Carasso User Fee Removal Questionnaire	18	Carasso et al. 2012 [64]	Carasso et al. 2012 [64] Healthcare, Zambia	Organizational culture/climate, readiness to implement	10	Norms[b]
Domain-Specific Innovativeness	6	Adapted from Goldsmith 1991 [65]	Webster et al. 2013 [66] Schools, physical activity, USA	Adoption	10	Norms, internal consistency
Evidence-Based Practice Attitude Scale	15	Aarons et al. 2010 [67]	Gill et al. 2014 [68], Beidas et al. 2013 [52] Mental health, USA, Canada	Acceptability, feasibility	12	Norms, internal consistency, structural validity[b]
Environmental Assessment Instrument	133	Lavinghouze et al. 2009 [69]	Lavinghouze et al. 2009 [69] Oral health, USA	Organizational culture/climate, champions, readiness to implement, structure of organization, actor relationships/networks, visibility of policy role/actors, political will for implementation	16	Norms
Health Enhancing Physical Activity Policy Audit Tool	75	Bull et al. 2014 [70]	Bull et al. 2015 [71] Physical activity, Europe	Readiness to implement, actor relationships/networks, political will for implementation, target population characteristics affecting implementation	12	Norms
Fall Prevention Coalition Survey	203	Schneider et al. 2016 [72]	Schneider et al. 2016 [72] Community, injury prevention, USA	Organizational culture/climate, champions, readiness to implement, actor relationships/network, visibility policy actors	7	Norms
Health Disparities Collaborative Staff Survey	21	Helfrich et al. 2007 [73]	Helfrich et al. 2007 [73] Healthcare, chronic disease, USA	Appropriateness, feasibility, adaptability, organizational climate/culture	8	Not reported
Healthy Cities Questionnaire	125	Donchin et al. 2006 [74]	Donchin et al. 2006 [74] Community, health promotion, Israel	Communication of policy, leadership for implementation, resources (non-training), actor relationships/networks, visibility of policy role/actors, political will for implementation	10	Norms[b]
Konduri Disease Registry Survey	12	Were et al. 2010 [75]	Konduri et al. 2017 [76] Hospital/clinics, tuberculosis, Ukraine	Acceptability, feasibility, readiness to implement	11	Norms, internal consistency
Local Wellness Policy Survey	39	McDonnell and Probart 2008 [77]	McDonnell and Probart 2008 [77] Schools—nutrition, physical activity, USA	Acceptability, readiness to implement, actor relationships/networks	10	Norms
Logical Assessment Matrix	9	Mersini et al. 2017 [78]	Mersini et al. 2017 [78] Nutrition, Albania	Adoption, costs of implementation, penetration, target population characteristics affecting implementation	13	Not reported
Maternal Child and Newborn Health Indicators	13	Cavagnero et al. 2008 [79]	Cavagnero et al. 2008 [79] Healthcare, global	Penetration, cost	7	Not reported
Organizational Readiness for Change	125	Lehman et al. 2002 [80]	Lau and Brookman-Frazee 2016 [45] Gill et al. 2014 [68] Mental health, USA	Organizational culture/climate	14	Norms
Perceived Attributes of Physical Activity Promotion in the Academic Classroom (PAPAC)	18	Adapted from Pankratz et al. 2002 [81]	Webster et al. 2018 [66] Schools, physical activity, USA	Appropriateness, feasibility, complexity, relative advantage	10	Norms[b]
Perceived Characteristics of Intervention Scale	20	Cook et al. 2015 [82]	Lau and Brookman-Frazee 2016 [45] Mental health, USA	Appropriateness, feasibility, adaptability, readiness to implement, relative advantage	13	Norms, structural validity[b]

Table 5 Partially transferable measures identified in studies of health policy implementation (*n* = 23) (*Continued*)

Tool name	Number of items	Development Author, year	Empirical use Author, year Setting/topic, country	Implementation outcomes and determinants assessed	Pragmatic PAPERS score[a]	Psychometric properties assessed
Probart School Wellness Survey	39	Probart et al 2010 [83]; Probart et al 2008 [84]; McDonnell and Probart 2008 [77]	Probart et al. 2010 Schools, nutrition, physical activity, USA	Adoption, cost, fidelity/compliance, adaptability, organizational climate/culture	9	Norms, internal consistency
Rakic Quality and Safety Survey	50	Rakic et al. 2018 [85]	Rakic et al. 2018 [85] Healthcare QI, Bosnia and Herzegovina	Acceptability, appropriateness, feasibility, complexity, organizational culture/climate, readiness to implement, actor relationships/networks	10	Norms
Rozema Outdoor Smoking Ban Survey	14	Rozema et al. 2018 [86]	Rozema et al. 2018 [86] Schools, tobacco, Netherlands	Fidelity/compliance, organizational culture/climate, readiness to implement	14	Norms, internal consistency
School Tobacco Policy Index	40	Barbero et al. 2013 [87]	Barbero et al. 2013 [87] Schools, tobacco, USA	Fidelity/compliance, communication of policy, resources (non-training), visibility of policy role/actors	17	Norms
Specialty Care Transformation Survey	26	Williams et al. 2017 [88]	Williams et al. 2017 [88] Healthcare, access to care, USA	Appropriateness, organizational culture/climate, readiness to implement, leadership for implementation	10	Norms
Spencer Quality Improvement Survey	120	Spencer and Walshe 2009 [89]	Spencer and Walshe 2009 [89] Healthcare, quality improvement, European Union	Readiness to implement, leadership for implementation, actor relationships/networks	8	Norms
Tobacco Industry Interference Index	20	Assunta and Dorotheo 2016 [90]	Assunta and Dorotheo 2016 [90] Tobacco, Southeast Asia	Policy implementation climate, visibility of policy role/actors, political will for implementation	13	Not reported
Tummers' Diagnosis Related Group Policy Survey 2	21	Tummers 2012 [91]	Tummers and Bekkers 2014 [92] Mental or behavioral health, Netherlands	Acceptability, adoption, appropriateness, feasibility, adaptability, champions, organizational culture/climate, relative priority, readiness to implement	11	Norms[b]

Partially transferable measures are defined here as those in which 25 to < 75% of items can readily be used in multiple settings without change or by changing only the referent (i.e., policy name, setting)

QI quality improvement

[a]Pragmatic PAPERS score—Psychometric and Pragmatic Evidence Rating Scale [11, 41, 42], five domains assessed: brevity (score based on number of items), language simplicity, burden/ease of interpretation of scoring, and training burden, total possible score 20, higher numbers indicate greater ease to use the measure

[b]Additional subscale level psychometric properties were reported

the policy implementation to the government mandate (26%), acceptability of the policy to implementers (24%), perceived appropriateness of the policy (17%), and feasibility of implementation (17%) (Table 2). Fidelity/compliance was sometimes assessed by asking implementers the extent to which they had modified a mandated practice [45]. Sometimes, detailed checklists were used to assess the extent of compliance with the many mandated policy components, such as school nutrition policies [83]. Acceptability was assessed by asking staff or healthcare providers in implementing agencies their level of agreement with the provided statements about the policy mandate, scored in Likert scales. Only eight (11%) of the included measures used multiple transferable items to assess adoption, and only eight (11%) assessed penetration.

Twenty-six measures of implementation costs were found during full-text screening (10 in included studies and 14 in excluded studies, data not shown). The cost time horizon varied from 12 months to 21 years, with most cost measures assessed at multiple time points. Ten of the 26 measures addressed direct implementation costs. Nine studies reported cost modeling findings. The implementation cost survey developed by Vogler et al. was extensive [53]. It asked implementing organizations to note policy impacts in medication pricing, margins, reimbursement rates, and insurance co-pays.

Determinants of implementation

Within the 70 included measures, the most commonly assessed implementation determinants were readiness for implementation (61% assessed any readiness component) and the general organizational culture and climate (39%), followed by the specific policy implementation climate within the implementation organization/s (23%), actor relationships and networks (17%), political will for policy implementation (11%), and visibility of the policy role and policy actors (10%) (Table 2). Each component of readiness for implementation was commonly assessed: communication of the policy (31%, 22 of 70 measures), policy awareness and knowledge (26%), resources for policy implementation (non-training resources 27%, training 20%), and leadership commitment to implement the policy (19%).

Only two studies assessed organizational structure as a determinant of health policy implementation. Lavinghouze and colleagues assessed the stability of the organization, defined as whether re-organization happens often or not, within a set of 9-point Likert items on multiple implementation determinants designed for use with state-level public health practitioners, and assessed whether public health departments were stand-alone agencies or embedded within agencies addressing additional services, such as social services [69]. Schneider and colleagues assessed coalition structure as an implementation determinant, including

items on the number of organizations and individuals on the coalition roster, number that regularly attend coalition meetings, and so forth [72].
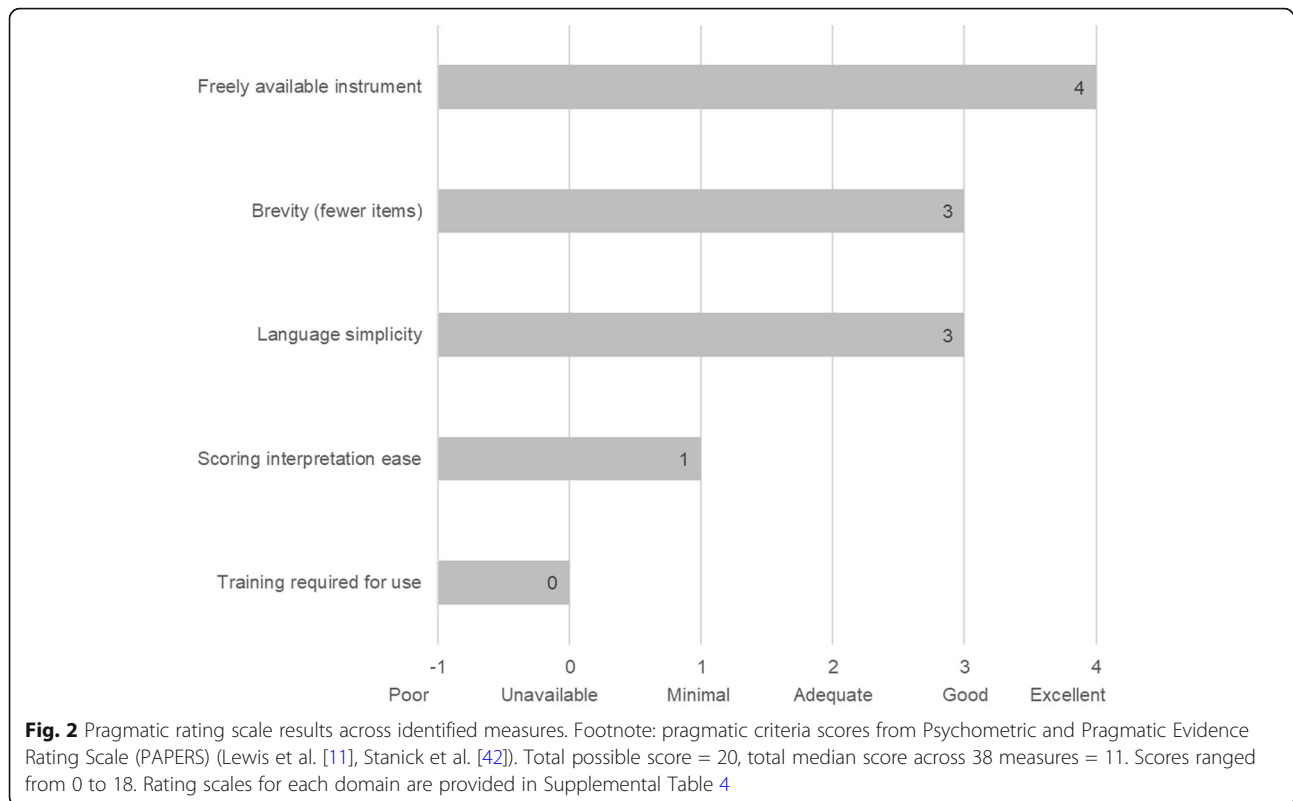
Tables of measures

Tables 4 and 5 present the 38 measures of implementation outcomes and/or determinants identified out of the 70 included measures with at least 25% of items transferable (useable in other studies without wording changes or by changing only the policy name or other referent). Table 4 shows 15 mostly transferable measures (at least 75% transferable). Table 5 shows 23 partially transferable measures (25–74% of items deemed transferable). Separate measure development articles were found for 20 of the 38 measures; the remaining measures seemed to be developed for one-time, study-specific use by the empirical study authors cited in the tables. Studies listed in Tables 4 and 5 were conducted most commonly in the USA ($n = 19$) or Europe ($n = 11$). A few measures were used elsewhere: Africa ($n = 3$), Australia ($n = 1$), Canada ($n = 1$), Middle East ($n = 1$), Southeast Asia ($n = 1$), or across multiple continents ($n = 1$).

Quality of identified measures

Figure 2 shows the median pragmatic quality ratings across the 38 measures with at least 25% transferable items shown in Tables 4 and 5. Higher scores are desirable and indicate the measures are easier to use (Table 3). Overall, the measures were freely available in the public domain (median score = 4), brief with a median of 11–50 items (median score = 3), and had good readability, with a median reading level between 8th and 12th grade (median score = 3). However, instructions on how to score and interpret item scores were lacking, with a median score of 1, indicating the measures did not include suggestions for interpreting score ranges, clear cutoff scores, and instructions for handling missing data. In general, information on training requirements or availability of self-training manuals on how to use the measures was not reported in the included study or measure development article/s (median score = 0, not reported). Total pragmatic rating scores among the 38 measures with at least 25% of items transferable ranged from 7 to 17 (Tables 4 and 5), with a median total score of 12 out of a possible total score of 20. Median scores for each pragmatic characteristic were the same across all measures as for the 38 mostly or partially transferable measures, with a median total score of 11 across all measures.

Few psychometric properties were reported. The study team found few reports of pilot testing and measure refinement as well. Among the 38 measures with at least 25% transferable items, the psychometric properties from the PAPERS rating scale total scores ranged from − 1 to 17 (Tables 4 and 5), with a median total score of

Fig. 2 Pragmatic rating scale results across identified measures. Footnote: pragmatic criteria scores from Psychometric and Pragmatic Evidence Rating Scale (PAPERS) (Lewis et al. [11], Stanick et al. [42]). Total possible score = 20, total median score across 38 measures = 11. Scores ranged from 0 to 18. Rating scales for each domain are provided in Supplemental Table 4

5 out of a possible total score of 36. Higher scores indicate more types of validity and reliability were reported with high quality. The 32 measures with calculable norms had a median norms PAPERS score of 3 (good), indicating appropriate sample size and distribution. The nine measures with reported internal consistency mostly showed Cronbach's alphas in the adequate (0.70 to 0.79) to excellent (≥ 90) range, with a median of 0.78 (PAPERS score of 2, adequate) indicating adequate internal consistency. The five measures with reported structural validity had a median PAPERS score of 2, adequate (range 1 to 3, poor to good), indicating the sample size was sufficient and the factor analysis goodness of fit was reasonable. Among the 38 measures, no reports were found for responsiveness, convergent validity, discriminant validity, known-groups construct validity, or predictive or concurrent criterion validity.

Discussion

In this systematic review, we sought to identify quantitative measures used to assess health policy implementation outcomes and determinants, rate the pragmatic and psychometric quality of identified measures, and point to future directions to address measurement gaps. In general, the identified measures are easy to use and freely available, but we found little data on validity and reliability. We found more quantitative measures of intra-organizational determinants of policy implementation than measures of

the relationships and interactions between organizations that influence policy implementation. We found a limited number of measures that had been developed for or used to assess one of the eight IOF policy implementation outcomes that can be applied to other policies or settings, which may speak more to differences in terms used by policy researchers and D&I researchers than to differences in conceptualizations of policy implementation. Authors used a variety of terms and rarely provided definitions of the constructs the items assessed. Input from experts in policy implementation is needed to better understand and define policy implementation constructs for use across multiple fields involved in policy-related research.

We found several researchers had used well-tested measures of implementation determinants from D&I research or from organizational behavior and management literature (Tables 4 and 5). For internal setting of implementing organizations, whether mandated through public policy or not, well-developed and tested measures are available. However, a number of authors crafted their own items, with or without pilot testing, and used a variety of terms to describe what the items assessed. Further dissemination of the availability of well-tested measures to policy researchers is warranted [9, 13].

What appears to be a larger gap involves the availability of well-developed and tested quantitative measures of the external context affecting policy implementation that can be used across multiple policy settings and topics

Allen *et al. Implementation Science*   (2020) 15:47

Page 14 of 17

[9]. Lack of attention to how a policy initiative fits with the external implementation context during policy-making and lack of policymaker commitment of adequate resources for implementation contribute to this gap [23, 93]. Recent calls and initiatives to integrate health policies during policymaking and implementation planning will bring more attention to external contexts affecting not only policy development but implementation as well [93–99]. At the present time, it is not well-known which internal and external determinants are most essential to guide and achieve sustainable policy implementation [100]. Identification and dissemination of measures that assess factors that facilitate the spread of evidence-informed policy implementation (e.g., relative advantage, flexibility) will also help move policy implementation research forward [1, 9].

Given the high potential population health impact of evidence-informed policies, much more attention to policy implementation is needed in D&I research. Few studies from non-D&I researchers reported policy implementation measure development procedures, pilot testing, scoring procedures and interpretation, training of data collectors, or data analysis procedures. Policy implementation research could benefit from the rigor of D&I quantitative research methods. And D&I researchers have much to learn about the contexts and practical aspects of policy implementation and can look to the rich depth of information in qualitative and mixed methods studies from other fields to inform quantitative measure development and testing [101–103].

Limitations
This systematic review has several limitations. First, the four levels of the search string and multiple search terms in each level were applied only to the title, abstract, and subject headings, due to limitations of the search engines, so we likely missed pertinent studies. Second, a systematic approach with stakeholder input is needed to expand the definitions of IOF implementation outcomes for policy implementation. Third, although the authors value intra-organizational policymaking and implementation, the study team restricted the search to governmental policies due to limited time and staffing in the 12-month study. Fourth, by excluding tools with only policy-specific implementation measures, we excluded some well-developed and tested instruments in abstract and full-text screening. Since only 12 measures had 100% transferable items, researchers may need to pilot test wording modifications of other items. And finally, due to limited time and staffing, we only searched online for measures and measures development articles and may have missed separately developed pragmatic information, such as training and scoring materials not reported in a manuscript.

Despite the limitations, several recommendations for measure development follow from the findings and related literature [1, 11, 20, 35, 41, 104], including the need to (1) conduct systematic, mixed-methods procedures (concept mapping, expert panels) to refine policy implementation outcomes, (2) expand and more fully specify external context domains for policy implementation research and evaluation, (3) identify and disseminate well-developed measures for specific policy topics and settings, (4) ensure that policy implementation improves equity rather than exacerbating disparities [105], and (5) develop evidence-informed policy implementation guidelines.

Conclusions
Easy-to-use, reliable, and valid quantitative measures of policy implementation can further our understanding of policy implementation processes, determinants, and outcomes. Due to the wide array of health policy topics and implementation settings, sound quantitative measures that can be applied across topics and settings will help speed learnings from individual studies and aid in the transfer from research to practice. Quantitative measures can inform the implementation of evidence-informed policies to further the spread and effective implementation of policies to ultimately reap greater population health benefit. This systematic review of measures is intended to stimulate measure development and high-quality assessment of health policy implementation outcomes and predictors to help practitioners and researchers spread evidence-informed policies to improve population health and reduce inequities.

Author details
[1]Prevention Research Center, Brown School, Washington University in St. Louis, One Brookings Drive, Campus Box 1196, St. Louis, MO 63130, USA. [2]School of Social Work, Brigham Young University, 2190 FJSB, Provo, UT 84602, USA. [3]Kaiser Permanente Washington Health Research Institute, 1730 Minor Ave, Seattle, WA 98101, USA. [4]Department of Health Management & Policy, Drexel University Dornsife School of Public Health, Nesbitt Hall, 3215 Market St, Philadelphia, PA 19104, USA. [5]Brown School, Washington University in St. Louis, One Brookings Drive, Campus Box 1196, St. Louis, MO 63130, USA. [6]Department of Surgery (Division of Public Health Sciences) and Alvin J. Siteman Cancer Center, Washington University School of Medicine, 4921 Parkview Place, Saint Louis, MO 63110, USA.

References
1. Purtle J, Dodson EA, Brownson RC. Policy dissemination research. In: Brownson RC, Colditz GA, Proctor EK, editors. Dissemination and Implementation Research in Health: Translating Science to Practice, Second Edition. New York: Oxford University Press; 2018.
2. Brownson RC, Baker EA, Deshpande AD, Gillespie KN. Evidence-based public health. Third ed. New York, NY: Oxford University Press; 2018.
3. Guide to Community Preventive Services. About the community guide.: community preventive services task force; 2020 [updated October 03, 2019; cited 2020. Available from: https://www.thecommunityguide.org/.
4. Eyler AA, Chriqui JF, Moreland-Russell S, Brownson RC, editors. Prevention, policy, and public health, first edition. New York, NY: Oxford University Press; 2016.
5. Andre FE, Booy R, Bock HL, Clemens J, Datta SK, John TJ, et al. Vaccination greatly reduces disease, disability, death, and inequity worldwide. Geneva, Switzerland: World Health Organization; 2008 February 2008. Contract No.: 07-040089.
6. Cheng JJ, Schuster-Wallace CJ, Watt S, Newbold BK, Mente A. An ecological quantification of the relationships between water, sanitation and infant, child, and maternal mortality. Environ Health. 2012;11:4.
7. Levy DT, Li Y, Yuan Z. Impact of nations meeting the MPOWER targets between 2014 and 2016: an update. Tob Control. 2019.
8. Purtle J, Peters R, Brownson RC. A review of policy dissemination and implementation research funded by the National Institutes of Health, 2007-2014. Implement Sci. 2016;11:1.
9. Lewis CC, Proctor EK, Brownson RC. Measurement issues in dissemination and implementation research. In: Brownson RC, Ga C, Proctor EK, editors. Disssemination and Implementation Research in Health: Translating Science to Practice, Second Edition. New York: Oxford University Press; 2018.
10. Lewis CC, Fischer S, Weiner BJ, Stanick C, Kim M, Martinez RG. Outcomes for implementation science: an enhanced systematic review of instruments using evidence-based rating criteria. Implement Sci. 2015;10:155.
11. Lewis CC, Mettert KD, Dorsey CN, Martinez RG, Weiner BJ, Nolen E, et al. An updated protocol for a systematic review of implementation-related measures. Syst Rev. 2018;7(1):66.
12. Chaudoir SR, Dugan AG, Barr CH. Measuring factors affecting implementation of health innovations: a systematic review of structural, organizational, provider, patient, and innovation level measures. Implement Sci. 2013;8:22.
13. Rabin BA, Lewis CC, Norton WE, Neta G, Chambers D, Tobin JN, et al. Measurement resources for dissemination and implementation research in health. Implement Sci. 2016;11:42.
14. Nilsen P, Stahl C, Roback K, Cairney P. Never the twain shall meet?--a comparison of implementation science and policy implementation research. Implement Sci. 2013;8:63.
15. Sabatier PA, editor. Theories of the Policy Process. New York, NY: Routledge; 2019.
16. Kingdon J. Agendas, alternatives, and public policies, second edition. Second ed. New York: Longman; 1995.
17. Jones MD, Peterson HL, Pierce JJ, Herweg N, Bernal A, Lamberta Raney H, et al. A river runs through it: a multiple streams meta-review. Policy Stud J. 2016;44(1):13–36.
18. Fowler L. Using the multiple streams framework to connect policy adoption to implementation. Policy Studies Journal. 2020 (11 Feb).
19. Howlett M, Mukherjee I, Woo JJ. From tools to toolkits in policy design studies: the new design orientation towards policy formulation research. Policy Polit. 2015;43(2):291–311.
20. Natesan SD, Marathe RR. Literature review of public policy implementation. Int J Public Policy. 2015;11(4):219–38.
21. Sabatier PA, Mazmanian. Implementation of public policy: a framework of analysis. Policy Studies Journal. 1980 (January).
22. Sabatier PA. Theories of the Policy Process. Westview; 2007.
23. Tomm-Bonde L, Schreiber RS, Allan DE, MacDonald M, Pauly B, Hancock T, et al. Fading vision: knowledge translation in the implementation of a public health policy intervention. Implement Sci. 2013;8:59.
24. Roll S, Moulton S, Sandfort J. A comparative analysis of two streams of implementation research. Journal of Public and Nonprofit Affairs. 2017;3(1):3–22.
25. Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Admin Pol Ment Health. 2011;38(2):65–76.
26. Brownson RC, Colditz GA, Proctor EK, editors. Dissemination and implementation research in health: translating science to practice, second edition. New York: Oxford University Press; 2018.
27. Tabak RG, Khoong EC, Chambers DA, Brownson RC. Bridging research and practice: models for dissemination and implementation research. Am J Prev Med. 2012;43(3):337–50.
28. Rabin BA, Brownson RC, Haire-Joshu D, Kreuter MW, Weaver NL. A glossary for dissemination and implementation research in health. J Public Health Manag Pract. 2008;14(2):117–23.

29. Lewis CC, Klasnja P, Powell BJ, Lyon AR, Tuzzio L, Jones S, et al. From classification to causality: advancing understanding of mechanisms of change in implementation science. Front Public Health. 2018;6:136.

30. Boyd MR, Powell BJ, Endicott D, Lewis CC. A method for tracking implementation strategies: an exemplar implementing measurement-based care in community behavioral health clinics. Behav Ther. 2018;49(4):525–37.

31. Glasgow RE. What does it mean to be pragmatic? Pragmatic methods, measures, and models to facilitate research translation. Health Educ Behav. 2013;40(3):257–65.

32. Glasgow RE, Riley WT. Pragmatic measures: what they are and why we need them. Am J Prev Med. 2013;45(2):237–43.

33. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci. 2009;4:50.

34. Bullock HL. Understanding the implementation of evidence-informed policies and practices from a policy perspective: a critical interpretive synthesis in: How do systems achieve their goals? the role of implementation in mental health systems improvement [Dissertation]. Hamilton, Ontario: McMaster University; 2019.

35. Watson DP, Adams EL, Shue S, Coates H, McGuire A, Chesher J, et al. Defining the external implementation context: an integrative systematic literature review. BMC Health Serv Res. 2018;18(1):209.

36. McKibbon KA, Lokker C, Wilczynski NL, Ciliska D, Dobbins M, Davis DA, et al. A cross-sectional study of the number and frequency of terms used to refer to knowledge translation in a body of health literature in 2006: a Tower of Babel? Implement Sci. 2010;5:16.

37. Terwee CB, Jansma EP, Riphagen II, de Vet HC. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. Qual Life Res. 2009;18(8):1115–23.

38. Egan M, Maclean A, Sweeting H, Hunt K. Comparing the effectiveness of using generic and specific search terms in electronic databases to identify health outcomes for a systematic review: a prospective comparative study of literature search method. BMJ Open. 2012;2:3.

39. Dillman DA, Smyth JD, Christian LM. Internet, mail, and mixed-mode surveys: the tailored design method. Hoboken, NJ: John Wiley & Sons; 2009.

40. Covidence systematic review software. Melbourne, Australia: Veritas Health Innovation. https://www.covidence.org. Accessed Mar 2019.

41. Powell BJ, Stanick CF, Halko HM, Dorsey CN, Weiner BJ, Barwick MA, et al. Toward criteria for pragmatic measurement in implementation research and practice: a stakeholder-driven approach using concept mapping. Implement Sci. 2017;12(1):118.

42. Stanick CF, Halko HM, Nolen EA, Powell BJ, Dorsey CN, Mettert KD, et al. Pragmatic measures for implementation research: development of the Psychometric and Pragmatic Evidence Rating Scale (PAPERS). Transl Behav Med. 2019.

43. Henrikson NB, Blasi PR, Dorsey CN, Mettert KD, Nguyen MB, Walsh-Bailey C, et al. Psychometric and pragmatic properties of social risk screening tools: a systematic review. Am J Prev Med. 2019;57(6S1):S13–24.

44. Stirman SW, Miller CJ, Toder K, Calloway A. Development of a framework and coding system for modifications and adaptations of evidence-based interventions. Implement Sci. 2013;8:65.

45. Lau AS, Brookman-Frazee L. The 4KEEPS study: identifying predictors of sustainment of multiple practices fiscally mandated in children's mental health services. Implement Sci. 2016;11:1–8.

46. Ekvall G. Organizational climate for creativity and innovation. European J Work Organizational Psychology. 1996;5(1):105–23.

47. Lövgren G, Eriksson S, Sandman PO. Effects of an implemented care policy on patient and personnel experiences of care. Scand J Caring Sci. 2002; 16(1):3–11.

48. Dwyer DJ, Ganster DC. The effects of job demands and control on employee attendance and satisfaction. J Organ Behav. 1991;12:595–608.

49. Condon-Paoloni D, Yeatman HR, Grigonis-Deane E. Health-related claims on food labels in Australia: understanding environmental health officers' roles and implications for policy. Public Health Nutr. 2015;18(1):81–8.

50. Patterson MG, West MA, Shackleton VJ, Dawson JF, Lawthom R, Maitlis S, et al. Validating the organizational climate measure: links to managerial practices, productivity and innovation. J Organ Behav. 2005;26:279–408.

51. Glisson C, Green P, Williams NJ. Assessing the Organizational Social Context (OSC) of child welfare systems: implications for research and practice. Child Abuse Negl. 2012;36(9):621–32.

52. Beidas RS, Aarons G, Barg F, Evans A, Hadley T, Hoagwood K, et al. Policy to implementation: evidence-based practice in community mental health-- study protocol. Implement Sci. 2013;8(1):38.

53. Eisenberger R, Cummings J, Armeli S, Lynch P. Perceived organizational support, discretionary treatment, and job satisfaction. J Appl Psychol. 1997;82:812–20.

54. Eby L, George K, Brown BL. Going tobacco-free: predictors of clinician reactions and outcomes of the NY state office of alcoholism and substance abuse services tobacco-free regulation. J Subst Abus Treat. 2013;44(3):280–7.

55. Vogler S, Zimmermann N, de Joncheere K. Policy interventions related to medicines: survey of measures taken in European countries during 2010- 2015. Health Policy. 2016;120(12):1363–77.

56. Wanberg CRB, Banas JT. Predictors and outcomes of openness to change in a reorganizing workplace. J Applied Psychology. 2000;85:132–42.

57. Hardy LJ, Wertheim P, Bohan K, Quezada JC, Henley E. A model for evaluating the activities of a coalition-based policy action group: the case of Hermosa Vida. Health Promot Pract. 2013;14(4):514–23.

58. Gavriilidis G, Östergren P-O. Evaluating a traditional medicine policy in South Africa: phase 1 development of a policy assessment tool. Glob Health Action. 2012;5:17271.

59. Hongoro C, Rutebemberwa E, Twalo T, Mwendera C, Douglas M, Mukuru M, et al. Analysis of selected policies towards universal health coverage in Uganda: the policy implementation barometer protocol. Archives Public Health. 2018;76:12.

60. Roeseler A, Solomon M, Beatty C, Sipler AM. The tobacco control network's policy readiness and stage of change assessment: what the results suggest for moving tobacco control efforts forward at the state and territorial levels. J Public Health Manag Pract. 2016;22(1):9–19.

61. Brämberg EB, Klinga C, Jensen I, Busch H, Bergström G, Brommels M, et al. Implementation of evidence-based rehabilitation for non-specific back pain and common mental health problems: a process evaluation of a nationwide initiative. BMC Health Serv Res. 2015;15(1):79.

62. Rütten A, Lüschen G, von Lengerke T, Abel T, Kannas L, Rodríguez Diaz JA, et al. Determinants of health policy impact: comparative results of a European policymaker study. Sozial-Und Praventivmedizin. 2003;48(6):379–91.

63. Smith SN, Lai Z, Almirall D, Goodrich DE, Abraham KM, Nord KM, et al. Implementing effective policy in a national mental health reengagement program for veterans. J Nerv Ment Dis. 2017;205(2):161–70.

64. Carasso BS, Lagarde M, Cheelo C, Chansa C, Palmer N. Health worker perspectives on user fee removal in Zambia. Hum Resour Health. 2012;10:40.

65. Goldsmith REH, C.F. Measuring consumer innovativeness. J Acad Mark Sci. 1991;19(3):209–21.

66. Webster CA, Caputi P, Perreault M, Doan R, Doutis P, Weaver RG. Elementary classroom teachers' adoption of physical activity promotion in the context of a statewide policy: an innovation diffusion and socio-ecologic perspective. J Teach Phys Educ. 2013;32(4):419–40.

67. Aarons GA, Glisson C, Hoagwood K, Kelleher K, Landsverk J, Cafri G. Psychometric properties and U.S. National norms of the Evidence-Based Practice Attitude Scale (EBPAS). Psychol Assess. 2010;22(2):356–65.

68. Gill KJ, Campbell E, Gauthier G, Xenocostas S, Charney D, Macaulay AC. From policy to practice: implementing frontline community health services for substance dependence--study protocol. Implement Sci. 2014;9:108.

69. Lavinghouze SR, Price AW, Parsons B. The environmental assessment instrument: harnessing the environment for programmatic success. Health Promot Pract. 2009;10(2):176–85.

70. Bull FC, Milton K, Kahlmeier S. National policy on physical activity: the development of a policy audit tool. J Phys Act Health. 2014;11(2):233–40.

71. Bull F, Milton K, Kahlmeier S, Arlotti A, Juričan AB, Belander O, et al. Turning the tide: national policy approaches to increasing physical activity in seven European countries. British J Sports Med. 2015;49(11):749–56.

72. Schneider EC, Smith ML, Ory MG, Altpeter M, Beattie BL, Scheirer MA, et al. State fall prevention coalitions as systems change agents: an emphasis on policy. Health Promot Pract. 2016;17(2):244–53.

73. Helfrich CD, Savitz LA, Swiger KD, Weiner BJ. Adoption and implementation of mandated diabetes registries by community health centers. Am J Prev Med. 2007;33(1,Suppl):S50-S65.

74. Donchin M, Shemesh AA, Horowitz P, Daoud N. Implementation of the Healthy Cities' principles and strategies: an evaluation of the Israel Healthy Cities network. Health Promot Int. 2006;21(4):266–73.

75. Were MC, Emenyonu N, Achieng M, Shen C, Ssali J, Masaba JP, et al. Evaluating a scalable model for implementing electronic health records in resource-limited settings. J Am Med Inform Assoc. 2010;17(3):237–44.

76. Konduri N, Sawyer K, Nizova N. User experience analysis of e-TB Manager, a nationwide electronic tuberculosis recording and reporting system in Ukraine. ERJ Open Research. 2017;3:2.
77. McDonnell E, Probart C. School wellness policies: employee participation in the development process and perceptions of the policies. J Child Nutr Manag. 2008;32:1.
78. Mersini E, Hyska J, Burazeri G. Evaluation of national food and nutrition policy in Albania. Zdravstveno Varstvo. 2017;56(2):115–23.
79. Cavagnero E, Daelmans B, Gupta N, Scherpbier R, Shankar A. Assessment of the health system and policy environment as a critical complement to tracking intervention coverage for maternal, newborn, and child health. Lancet. 2008;371 North American Edition(9620):1284-93.
80. Lehman WE, Greener JM, Simpson DD. Assessing organizational readiness for change. J Subst Abus Treat. 2002;22(4):197–209.
81. Pankratz M, Hallfors D, Cho H. Measuring perceptions of innovation adoption: the diffusion of a federal drug prevention policy. Health Educ Res. 2002;17(3):315–26.
82. Cook JM, Thompson R, Schnurr PP. Perceived characteristics of intervention scale: development and psychometric properties. Assessment. 2015;22(6):704–14.
83. Probart C, McDonnell ET, Jomaa L, Fekete V. Lessons from Pennsylvania's mixed response to federal school wellness law. Health Aff. 2010;29(3):447–53.
84. Probart C, McDonnell E, Weirich JE, Schilling L, Fekete V. Statewide assessment of local wellness policies in Pennsylvania public school districts. J Am Diet Assoc. 2008;108(9):1497–502.
85. Rakic S, Novakovic B, Stevic S, Niskanovic J. Introduction of safety and quality standards for private health care providers: a case-study from the Republic of Srpska, Bosnia and Herzegovina. Int J Equity Health. 2018;17(1):92.
86. Rozema AD, Mathijssen JJP, Jansen MWJ, van Oers JAM. Sustainability of outdoor school ground smoking bans at secondary schools: a mixed-method study. Eur J Pub Health. 2018;28(1):43–9.
87. Barbero C, Moreland-Russell S, Bach LE, Cyr J. An evaluation of public school district tobacco policies in St. Louis County, Missouri. J Sch Health. 2013;83(8):525–32.
88. Williams KM, Kirsh S, Aron D, Au D, Helfrich C, Lambert-Kerzner A, et al. Evaluation of the Veterans Health Administration's specialty care transformational initiatives to promote patient-centered delivery of specialty care: a mixed-methods approach. Telemed J E-Health. 2017;23(7):577–89.
89. Spencer E, Walshe K. National quality improvement policies and strategies in European healthcare systems. Quality Safety Health Care. 2009;18(Suppl 1):i22–i7.
90. Assunta M, Dorotheo EU. SEATCA Tobacco Industry Interference Index: a tool for measuring implementation of WHO Framework Convention on Tobacco Control Article 5.3. Tob Control. 2016;25(3):313–8.
91. Tummers L. Policy alienation of public professionals: the construct and its measurement. Public Adm Rev. 2012;72(4):516–25.
92. Tummers L, Bekkers V. Policy implementation, street-level bureaucracy, and the importance of discretion. Public Manag Rev. 2014;16(4):527–47.
93. Raghavan R, Bright CL, Shadoin AL. Toward a policy ecology of implementation of evidence-based practices in public mental health settings. Implement Sci. 2008;3:26.
94. Peters D, Harting J, van Oers H, Schuit J, de Vries N, Stronks K. Manifestations of integrated public health policy in Dutch municipalities. Health Promot Int. 2016;31(2):290–302.
95. Tosun J, Lang A. Policy integration: mapping the different concepts. Policy Studies. 2017;38(6):553–70.
96. Tubbing L, Harting J, Stronks K. Unravelling the concept of integrated public health policy: concept mapping with Dutch experts from science, policy, and practice. Health Policy. 2015;119(6):749–59.
97. Donkin A, Goldblatt P, Allen J, Nathanson V, Marmot M. Global action on the social determinants of health. BMJ Glob Health. 2017;3(Suppl 1):e000603-e.
98. Baum F, Friel S. Politics, policies and processes: a multidisciplinary and multimethods research programme on policies on the social determinants of health inequity in Australia. BMJ Open. 2017;7(12):e017772-e.
99. Delany T, Lawless A, Baum F, Popay J, Jones L, McDermott D, et al. Health in All Policies in South Australia: what has supported early implementation? Health Promot Int. 2016;31(4):888–98.
100. Valaitis R, MacDonald M, Kothari A, O'Mara L, Regan S, Garcia J, et al. Moving towards a new vision: implementation of a public health policy intervention. BMC Public Health. 2016;16:412.
101. Bennett LM, Gadlin H, Marchand, C. Collaboration team science: a field guide. Bethesda, MD: National Cancer Institute, National Institutes of Health; 2018. Contract No.: NIH Publication No. 18-7660.
102. Mazumdar M, Messinger S, Finkelstein DM, Goldberg JD, Lindsell CJ, Morton SC, et al. Evaluating academic scientists collaborating in team-based research: a proposed framework. Acad Med. 2015;90(10):1302–8.
103. Brownson RC, Fielding JE, Green LW. Building capacity for evidence-based public health: reconciling the pulls of practice and the push of research. Annu Rev Public Health. 2018;39:27–53.
104. Brownson RC, Colditz GA, Proctor EK. Future issues in dissemination and implementation research. In: Brownson RC, Colditz GA, Proctor EK, editors. Dissemination and Implementation Research in Health: Translating Science to Practice. Second Edition ed. New York: Oxford University Press; 2018.
105. Thomson K, Hillier-Brown F, Todd A, McNamara C, Huijts T, Bambra C. The effects of public health policies on health inequalities in high-income countries: an umbrella review. BMC Public Health. 2018;18(1):869.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.