Evaluation of the NCPDP Structured and Codified Sig Format for e-prescriptions

Hangsheng Liu,¹ Q Burkhardt,² Douglas S Bell^{2,3}

► Additional material is published online only. To view these files please visit the journal online (www.jamia.org).

¹RAND Corporation, Boston, Massachusetts, USA

²RAND Corporation, Santa Monica, California, USA

³Department of Medicine, David Geffen School of Medicine at UCLA, Los Angeles, CA, USA

Correspondence to

Dr Hangsheng Liu, RAND Corporation, 20 Park Plaza, Suite 720, Boston, MA 02116, USA; hliu@rand.org

Received 4 October 2010

Accepted 17 April 2011

Published Online First

25 May 2011

ABSTRACT

Objective To evaluate the ability of the structure and code sets specified in the National Council for Prescription Drug Programs Structured and Codified Sig Format to represent ambulatory electronic prescriptions.

Design We parsed the Sig strings from a sample of 20 161 de-identified ambulatory e-prescriptions into variables representing the fields of the Structured and Codified Sig Format. A stratified random sample of these representations was then reviewed by a group of experts. For codified Sig fields, we attempted to map the actual words used by prescribers to the equivalent terms in the designated terminology.

Measurements Proportion of prescriptions that the Format could fully represent; proportion of terms used that could be mapped to the designated terminology.

Results The fields defined in the Format could fully represent 95% of Sigs (95% CI 93% to 97%), but ambiguities were identified, particularly in representing multiple-step instructions. The terms used by prescribers could be codified for only 60% of dose delivery methods, 84% of dose forms, 82% of vehicles, 95% of routes, 70% of sites, 33% of administration timings, and 93% of indications.

Limitations The findings are based on a retrospective sample of ambulatory prescriptions derived mostly from primary care physicians.

Conclusion The fields defined in the Format could represent most of the patient instructions in a large prescription sample, but prior to its mandatory adoption, further work is needed to ensure that potential ambiguities are addressed and that a complete set of terms is available for the codified fields.

INTRODUCTION

Patient instructions, for example, 'take 1 tablet once a day,' are a key component of ambulatory prescriptions and are represented in a portion of the prescription called the *signatura*, commonly referred to as the 'Sig.' Based on the premise that standardized electronic Sig information would facilitate communication between prescribers and pharmacists as well as reduce prescribing errors, the National Committee on Vital and Health Statistics asked the National Council for Prescription Drug Programs (NCPDP) to develop a standard for representing Sig information in a standardized, unambiguous format.¹ In 2004, NCPDP convened a group of industry stakeholders to take on this challenge, and released a draft version of the NCPDP Structured and Codified Sig Format in 2006. It is called a 'format' rather than a 'standard' to emphasize its intended use as a component for inclusion in other prescription data standards, including the NCPDP SCRIPT standard for outpatient

prescriptions and the HL7 standard for inpatient prescriptions. In this paper, 'the Format' refers exclusively to the NCPDP Structured and Codified Sig Format.

In 2006, the Centers for Medicare and Medicaid Services (CMS) and the Agency for Healthcare Research and Quality sponsored a pilot testing of several e-prescribing standards, including the 2006 version of the Format. However, the pilot testing with only 42 prescriptions showed that it was not applied consistently by experts.² The Format was not mandated in the current electronic prescribing Final Rule by CMS.³

This study evaluates the revised version of the Format (Version 1.0) released by NCPDP in 2008,¹ as it would be implemented within the NCPDP SCRIPT 10.5 standard for ambulatory prescription transactions. Specifically, we evaluated, for prescriptions from ambulatory care settings, (1) whether the set of fields defined by the Format could fully and faithfully represent the patient instructions, and (2) whether the terms needed for the codified fields existed in the designated terminologies. Our ultimate goal was to assess the Format's ability to unequivocally communicate patient instructions and its readiness for adoption as a component of an outpatient e-prescribing standard.

BACKGROUND

At least 1.5 million medication errors occur annually in the USA, at a cost over \$3.5 billion each year; more than one third of these occur in the ambulatory care setting.⁴ Prior studies have demonstrated that electronic prescribing could reduce medication errors and adverse drug events,⁵ save physician and staff time,^{6–8} and lower medical costs.^{6,9} However, incomplete standards for communicating drug orders have been cited as a key failure in the current system.⁴ Ambulatory e-prescribing standards have now been in place for several years,¹⁰ but in these standards the patient instructions or 'Sig' in electronic prescriptions is represented only as free text, leaving room for misinterpretation and error. A recent study by Singh *et al* highlighted the potential for new kinds of error, showing that nearly 1% of electronic prescriptions contained inconsistent instructions, with the most common error being conflicting drug dosages.¹¹ Palchuk *et al* reported that 16% of electronic prescriptions contained internal discrepancies, a majority of which could potentially lead to adverse drug events.¹² In the absence of a standard for representing patient instructions, which could enable computerized accuracy and dosage checking, the potential benefits of electronic prescribing may not be achieved.¹³

The Structured and Codified Sig Format (Version 1.0) provides a machine-interpretable representation for the patient instructions portion of a prescription to enable more automated safety checking and improve communication between prescribers and pharmacists. To remove sources of misinterpretation and error, the revised 2008 version incorporated the recommendations from the 2006 CMS pilot testing on the 2006 draft version, including improved field naming, field definitions, and examples to illustrate the Format. In addition, the current Format has also been formally incorporated into the NCPDP SCRIPT standard Version 10.5.

The Format includes 13 segments, each of which contains distinct fields to represent potential components of patient instructions (table 1).¹ Some of these fields are intended to take values from designated terminology systems. The Format designates Federal Medication Terminologies (FMT) to codify dose form,¹⁴ while the remaining fields should use the Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT)¹⁵ such as dose delivery method, vehicle, route of administration, site of administration, administration timing, frequency unit, interval unit, duration, and indication precursor. The Format allows repetition of segments (eg, to represent ‘take two tablets once daily for 3 days, then take one tablet once daily for 3 days’) or elements within a segment (eg, to represent ‘take one to two tablets once daily’) to represent complex Sigs, but the entire Format should be repeated in these cases. Additional instructions provided by pharmacists to patients, filling instructions from the prescriber to the pharmacist, and the number of refills are not part of the Format.

METHODS

We expected that most ambulatory Sigs would be relatively simple and easily accommodated by the Format. Thus, we used a large prescription sample in order to isolate and characterize failures of the Format, each of which might represent small but important subsets of prescriptions. To maximize the informational yield from these potentially rare representation failures, we categorized the prescription sample based on a multi-step pre-processing algorithm, followed by over-sampling of categories more likely to contain representation failures. The pre-processing steps were (1) normalization of lexical and other variants, (2) parsing values into each of the applicable fields of the Format, and (3) preliminary identification of representation problems. Each Sig in the final sample was reviewed by two experts to judge whether the Format was capable of fully representing the prescriber’s intent. They also judged whether the Sig should be excluded as inadequately specified (would require a call to the prescriber for clarification). To evaluate the designated terminology systems (separately from the structure of the format), the authors manually searched for an appropriate term to represent each term actually used by prescribers for one of the codified fields. The study was approved by the RAND Institutional Review Board.

Data source

Two e-prescribing system vendors supplied a total sample of 20 161 de-identified (with no individually identifiable information) prescriptions from 82 providers practicing in 48 physician offices in Kansas, Michigan, and Maryland. One vendor’s system only allowed free text for Sigs, while the other had several fields with drop down menus (dose quantity, dose form, and timing) and a free text field. Provider specialties included internal medicine, family practice, pediatrics, and general surgery. The target sample size (20 000) was selected to provide 95% Wald CIs

Table 1 Segments and fields in the structured and codified Sig format

Segment	Field
Repeating Sig segment	Sig sequence position number Multiple Sig modifier
Code system segment	SNOMED version FMT version
Sig free text string segment	Free text string
Dose segment	Dose delivery method Dose delivery method modifier Dose quantity Dose form Dose range modifier
Dose calculation segment*	Dosing basis numeric value* Dosing basis unit of measure* Body metric qualifier* Body metric value* Calculated dose numeric value* Calculated dose unit of measure* Dosing basis range modifier*
Vehicle segment	Vehicle name Vehicle quantity Vehicle unit of measure Multiple vehicle modifier
Route of administration segment	Route Multiple route of administration modifier
Site of administration segment	Site Multiple site of administration modifier
Sig timing segment	Administration timing Multiple administration timing modifier Rate of administration* Rate unit of measure* Time period basis* Frequency numeric value Frequency units Variable frequency modifier Interval numeric value Interval units Variable interval modifier
Duration segment	Duration numeric value Duration text
Maximum dose restriction segment	Maximum dose restriction value Maximum dose restriction units Maximum dose restriction variable numeric value Maximum dose restriction variable units Maximum dose restriction variable duration modifier
Indication segment	Indication precursor Indication text Indication value Indication value unit of measure Indication variable modifier
Stop segment	Stop indicator

*Not evaluated in the study because they are used primarily in the inpatient setting rather than the ambulatory care setting.

of 0.1% around non-accommodated rates as low as 0.5%. The prescriptions had been written from March 2007 to February 2009 (75% being from September 2008 or later). In addition to Sig strings, we also obtained the corresponding drug names, drug strength, and dose form for each prescription.

Parsing program

In order to automate the pre-processing steps, we created a processing sequence that normalized common expressions and then parsed words and values from the normalized Sig strings into variables representing the fields of the Format. The

normalization step removed variations due to lower/upper case, spaces, punctuation, and common spelling errors. It also converted numbers expressed in words into numerical values (eg, from 'one' to '1'), expanded abbreviations (eg, converting 'PO BID' to 'by mouth 2 times per day'), and removed extraneous statements (eg, 'thank you'). After normalization, the sample was collapsed into a list of unique Sig strings since many of them were exactly the same. Thus, each unique Sig string represented a number of raw Sig strings. The unique Sig strings were then parsed to populate the structured fields based on regular expression patterns. 'Text' fields were populated with the relevant words actually used in the Sig string. If a Sig string was not fully parsed, any unparsed remainder was output in a separate field. The Perl programming language was used for both normalization and parsing. For example, the program parsed 'Inhale contents of 1 capsule once daily' into several fields: 'inhale' in dose delivery method, '1' in dose quantity, 'capsule' in dose form, '1' in frequency numerical value, 'day' in frequency unit, and 'contents of' in the unparsed field.

Expert review

Twelve prescription technology experts agreed to participate in reviewing the parsed representation of 100 Sig strings each. Because each expert would only be able to review a limited number of representations, we used the parsing results to substantially over-sample Sigs that were more likely to pose representation problems. Figure 1 provides an overview of this strategy. Both the fully parsed strings and not-fully parsed strings were further classified into preliminary problem categories based on the occurrence of string patterns that, on preliminary review by the authors, were likely to pose representation problems (appendix A (available online at www.jamia.org) shows examples of these categories). To create the final expert review sample, we then included (1) one randomly selected example from each of 37 preliminary problem categories, (2) all unique Sig strings that were not fully parsed and not assigned to a preliminary problem

category, and (3) a simple random sample from those unique Sig strings that were fully parsed and not assigned to a preliminary problem category. A final sample size of 600 was targeted to allow for each representation to have two independent reviewers.

For each Sig representation to be reviewed, the experts were presented with the complete original prescription (including the original drug name, strength, dose form, and the raw, unprocessed Sig string), the values parsed into each of Sig Format fields, a re-construction of the Sig based on those values, and the unparsed remainder. They were asked to judge (1) whether the raw Sig string was adequately specified, and (2) whether the Format could fully accommodate the prescriber's intent. 'Adequately specified' meant that a pharmacist could fill the prescription without needing to call the prescriber for clarification. If the two experts reviewing each prescription did not agree with each other, the authors adjudicated the judgments. Five rounds of training among experts on how to make their judgments were conducted prior to the ratings. For each round of training, experts were asked to adjudicate several Sig strings, and then a conference call was convened to reach a consensus on the correct representation of the example Sig strings.

After review by the individual experts, teleconferences with the expert panel were held to review each challenging case, and if the panel agreed upon a judgment that was different from that of the two (or three) expert reviewers, then the panel's consensus decision overrode the individual reviewers' judgments.

Terminology code mapping

To assess how well the terms used in the Sigs could be codified, the authors evaluated whether each term used could be mapped to an appropriate term in the designated vocabulary system. The fields involved were dose delivery method, dose form, route, site, vehicle, indication, and administration timing. Version 1.0 of the Format designates SNOMED CT as the vocabulary system to use for all fields except for dose form. For dose form, the Format specifies use of the relevant FMT vocabulary, which is currently

Figure 1 Sample selection flow chart.

the FDA's 'Structured Product Label' (SPL) subset of the National Cancer Institute Thesaurus (NCIt).¹⁴ We used the January 2009 release of SNOMED and the September 2009 release of the NCIt SPL subset.¹⁶ We report the proportion of terms used in the Sig strings for which a controlled vocabulary term could be found having the same meaning.

Data analysis

Descriptive statistics were generated for the original raw Sig string sample. Both unweighted and weighted statistics were produced based on the parsing results and the judgments made by the panelists. Each raw Sig string in the final sample for review was assigned two sampling weights from the two-stage sampling process shown in figure 1 (stage 1 collapsed raw Sig strings into unique Sig strings; stage 2 was sampling within each preliminary problem category). The product of these two sampling weights (after normalization) estimates the number of original (raw Sig) prescriptions that each item in the final sample represents.

RESULTS

Sample description

Thirty of the 20 161 prescriptions had no Sig information, leaving a final sample of 20 131 raw Sig strings transmitted between vendors and pharmacy systems. The most frequent raw Sig string—'1 tablet once a day'—represented 2810 (13.96%) raw Sig strings, followed by 'Take 1 tablet daily' (4.69%) and '1 tablet twice a day' (4.40%). The 20 most frequent raw Sig strings represented 42.74% of all raw Sig strings. A total of 143 (0.71%) raw Sig strings included notes to pharmacists on medication dispensing (in addition to instructions for the patient), and 193 (0.91%) raw Sig strings included information for patients not related to medication use. This extraneous information was deleted, either through the normalization program or manually. After normalization, the raw Sig strings collapsed to 3060 unique Sig strings.

Parsing results and sample selection for review

The parsing program fully parsed 2573 (84.1%) unique normalized Sig strings, corresponding to 95.3% of the raw Sig strings. That is, all the words or terms used in these 2573 normalized Sig strings were populated into specific Structured and Codified Sig data fields, with no unparsed words remaining.

Our further stratification of the Sig sample prior to sampling for expert review (figure 1) resulted in 3.4% (88 out of 2573) of fully parsed unique Sig strings being classified into preliminary problem categories, and 34.9% (170 out of 487) of unique Sig strings that were not fully parsed being assigned to preliminary problem categories. The final sample for expert review (as shown in the dash line boxes in figure 1, as well as in the first column of table 2) included all 317 non-fully parsed unique Sig strings that

were not assigned to a known preliminary problem category, 23 out of 170 that were not fully parsed but were assigned to a preliminary problem category, and 14 out of 88 unique Sig strings that were fully parsed but nonetheless contained a known preliminary problem. An additional 247 Sigs were then randomly selected from the 2485 fully parsed unique Sig strings that were not assigned to a preliminary problem category, to create a final sample of 601 Sig strings for expert review.

Expert review results

The review process resulted in 75 of 601 Sigs being categorized for exclusion as inadequately specified, representing 7.2% of the original 20 131 raw Sig strings after weighting (table 2). The most common reasons for Sigs being inadequately specified (ie, requiring a call-back to the prescriber) were conflicting information within the same Sig, and missing information that pharmacists would consider necessary (appendix B (available online at www.jamia.org) shows examples). For example, according to the Sig '1 tablet once a day take 1 tablet by mouth twice daily as needed,' a patient would not know whether s/he should take a tablet once or twice a day. A pharmacist would not be able to fill a prescription based on a Sig string such as 'BID. Up to QID for severe pain' because dose quantity is missing.

Among the 18 683 adequately specified Sig strings, 5.0% (940) were classified as not accommodated by the current Structured and Codified Sig Format (table 2), with a 95% CI from 2.7% to 7.4%. Within the category of fully parsed Sig strings not preliminarily classified as having a representation problem, 99% were accommodated by the current Format, versus 47.0% for Sig strings not fully parsed and not so categorized. Approximately 20% of Sig strings assigned to preliminary problem categories were judged to be representable.

Based on the reviewers' judgments and additional consensus meetings among the expert panel, we identified 23 distinct types of patient instructions that were used in the sample but are not accommodated by the Format. As shown in table 3, nine of these non-representable Sig concepts were related to the dose segment, including instructions for preparing a dose, multiple-step dose delivery, durations of action specified for the dose delivery method, etc. 'Until current supply gone' was the most prevalent non-representable concept, occurring in 1.2% of the total Sig string sample. Several other timing concepts are also not accommodated, including explicit start or stop dates, alternating timing, repeating cycles, and the timing of one medication being related to other medications. Finally, there were many miscellaneous additional patient instructions that are not representable, for instance, 'Use ice cold. Do not hold tear duct.'

Overall, the adequately specified prescriptions made use of some fields more than others (table 4), with the most frequently used fields being dose quantity and dose form (used in 96% of prescriptions), followed by frequency numeric value and

Table 2 Expert review results

Sampling stratum	Total sample		Inadequately specified		Adequately specified			
	Unweighted count	Weighted to full sample*	Unweighted count	Weighted to full sample*	Format inadequate		Format accommodates	
					Unweighted	Weighted†	Unweighted	Weighted†
Fully parsed	247	18 862	14	1307	6	158	227	17 397
Fully parsed with preliminary problem	14	318	1	39	7	226	6	53
Not fully parsed	317	490	58	93	121	210	138	187
Not fully parsed with preliminary problem	23	461	2	10	9	346	12	105
Total	601	20 131	75	1449 (7.2%)*	143	940 (5.0%)†	383	17 742 (95.0%)†

*Weighted to the full sample of 20 131 prescriptions.

†Weighted to the 18 682 adequately specified Sigs.

Table 3 Sig concepts not accommodated by the Structured and Codified Sig Format

Sig concept*	Complete Sig example	N (%)†
Dose segment	1. Dose preparation	Place capsule in device as directed and inhale contents once daily 118 (0.63)
	2. Multi-step dose delivery	Apply sparingly and rub in well to affected area(s) twice daily. Not for face, breast or groin. 138 (0.74)
	3. Dose delivery method action duration	Apply to nares and rub well for one minute. Apply twice daily for 7 days 7 (0.04)
	4. May take additional actions	2 teaspoons four times a day use as a gargle or mouthwash. May swallow. 1 (0.01)
	5. Inexact dose quantity for ointment/cream	Apply small amount of ointment every 3–4 hours and at bedtime for 7–10 days 74 (0.39)
	6. Inexact dose quantity depending on patient body size	Use two or three drops to fill the ear canal as needed for pain. 7 (0.04)
	7. Increase dose quantity or frequency if tolerated	Begin with 4 tabs once weekly and increase to 6 tabs weekly if tolerated 12 (0.06)
	8. Undefined dose form	Inhale contents of 1 capsule once daily 15 (0.08)
	9. Remove the remaining medication	Apply patch for 12 to 14 h daily, then remove 35 (0.19)
Vehicle segment	10. Vehicle preparation	Place 17 g (1 heaping tablespoonful) in beverage of choice and drink at bedtime 20 (0.10)
Site of administration segment	11. Site preparation	Cleanse with saline, then apply and cover wound. When practicable, daily or twice daily changes are preferred 2 (0.01)
Sig timing segment	12. Irregular intervals within specified frequency	Instil 1 drop into right eye twice daily at 6–8 h intervals. 70 (0.37)
Duration segment	13. Until current supply is gone	1 teaspoon (5 ml) every 12 h until finished 220 (1.18)
	14. Duration of trigger event	A small amount twice a day use until rash is gone. 22 (0.12)
	15. Maximum duration	A small amount twice a day use for only 10 days at a time 6 (0.03)
Other timing concepts	16. Explicit start or stop date	3 tablets twice daily ×14 days. Start on 3 December 2008 81 (0.43)
	17. Alternating timing	Alternate with 2 tablets every other day with 1 and 1/2 tablets every other day 31 (0.16)
	18. Repeating cycle prescription	Take 1 tablet daily for 21 days, then 7 tablet-free days; repeat. 95 (0.51)
Stop segment	19. Timing related to other medications	Take 1 tablet at HS, 30 min. After ibuprofen 4 (0.02)
	20. Stop indication	Take one tablet daily. Hold for diarrhea 2 (0.01)
Other concepts	21. Additional patient instructions	1 GT four times a day OU ×10 days. Use ice cold. Do not hold tear duct 55 (0.29)
	22. Take or not take along with other medications	1 tablet once a day use daily along with isosorbide 60 mg 18 (0.10)
	23. Switching to another medication or dose form	1 capsule once a week then start over the counter Vitamin D 1000 units every day 1 (0.01)

*A Sig string can have multiple non-representable concepts.

†Weighted to the 18 682 adequately specified Sigs.

Table 4 Frequency of use for individual structured fields

Segment	Field	N (%)*
Repeating segment	Sequence position number	NA
	Multiple Sig modifier	614 (3.5)
Dose segment	Dose delivery method text	10 840 (61.1)
	Dose quantity	16 958 (95.6)
	Dose form text	16 958 (95.6)
	Vehicle text	523 (3.0)
Route of administration segment	Route of administration text	2906 (16.4)
Site of administration segment	Site of administration text	1522 (8.6)
Sig timing segment	Administration timing text	1991 (11.2)
	Frequency numeric value	14 346 (80.9)
	Frequency units text	14 328 (80.8)
	Interval numeric value	1923 (10.8)
	Interval units text	1923 (10.8)
Duration segment	Duration numeric value	1632 (9.2)
	Duration text	1632 (9.2)
Maximum dose restriction segment	Maximum dose restriction numeric value	211 (1.2)
	Maximum dose restriction units text	211 (1.2)
	Maximum dose restriction variable numeric value	195 (1.1)
	Maximum dose restriction variable units text	211 (1.2)
Indication segment	Indication precursor text	2264 (12.8)
	Indication text	889 (5.0)
Stop segment	Stop indicator	0 (0.0)†

*Weighted to the 17 742 adequately specified and accommodated Sig strings.

†No stop indicators were found among the 601 Sigs sampled for expert review. In the original 20 131 raw Sigs, about 0.02% of Sigs have stop indicators assuming the parsing program correctly parsed all the Sig strings.

frequency units text (used in 81%). Among the 4.4% of Sigs for which the dose was not specified, 1.5% were PRNs (either 'as needed' or 'as directed') and a majority of the remaining 2.9% were instructions on applying cream, gel, or ointment. Only 61% of Sig strings specified a dose delivery method, 13% an indication precursor, 16% a route of administration, and 11% an administration timing. Interval numeric value and interval units text were used in only 11% of Sigs.

Mapping to SNOMED and FMT terminology codes

Of 35 distinct dose delivery method terms used in the sample, a SNOMED term with the same meaning could be found for 21 (60%). Forty eight (84%) out of the 57 dose form terms could be similarly matched to an NCI term. A majority of the vehicles used (82%, 36 out of 44) were represented in SNOMED and 19 out of 20 route of administration terms were represented. SNOMED could represent about 70% of the 76 site terms. Among the 194 administration timing terms, only 64 (33%) were represented in SNOMED. About 93% (170 out of 182) of indication terms were fully represented in SNOMED. Of note, in some cases, each portion of a term was representable, but the whole term was not representable due to lack of a mechanism for representing composite terms, for example, 'nausea' with modifier 'severe'. (See appendix C (available online at www.jamia.org) for more details on the terms that were not representable.)

Unresolved problems with the Format

Our experts did not reach agreement on the proper use of the Format's free text field. This field is mandatory in the current Format to ensure that the pharmacist can see language that

represents the prescriber's complete intent. The Format allows three types of values in the free text field: 'pure free text,' with no values in the structured fields, 'reconstructed from structured Sig,' or 'capture what the MD ordered.' The instructions for using the Format imply that the latter type of free text value can be used to represent information that would add meaning to the values captured in the structured fields (see p 41 in the *Structured and Codified Sig Format Implementation Guide*¹). This use of free text was controversial among panelists, however, because some expected that physicians would not reconcile free text edits with structured values and would therefore frequently transmit contradictions, undermining the certainty of all the structured data.

The panel also did not reach consensus on the use of different conjunctions within Sigs containing multiple repeating segments. A Sig that consisted of '[segment] AND [segment] OR [segment] AND [segment]' could be interpreted in multiple ways depending on the operator precedence assumed. The Format does not include a mechanism for expressing what would be parentheses in logical expressions. The use of multiple conjunctions would arise in particular because numeric ranges are represented using the same 'repeating segment' structure used for multi-part Sigs. For example, 'take 1 or 2 tablets upon awakening and in the evening' could be represented by the Format using four repeating segments, but it would be difficult to reconstruct the original Sig in the absence of grouping logical expressions.

DISCUSSION

Our evaluation of the current Structured and Codified Sig Format (Version 1.0, 2008) suggests that the content areas defined in the current Format cover the majority of information domains needed for patient instructions in ambulatory care. The NCPDP has achieved its initial target that the Format 'must support 80% of the Sigs being written today.'¹ However, several significant problems would need resolution for it to reach NCPDP's long-term goal of representing 99% of Sigs and to function well as a federally mandated component of an outpatient e-prescribing standard. These include the absence of many needed terms in the designated terminologies (SNOMED CT and FMT), the lack of a mechanism for representing composite terms, the uncertain status of the free text field, the need to simplify the representation of numeric ranges, and the lack of a mechanism analogous to parentheses for grouping logical expressions. In order to accommodate the non-representable information, new structured fields could be added to the Format, but this would increase the Format's complexity. With this caveat, the following discussion proposes some possibly useful modifications.

Recommendations for the format structure

The need to use repeating Sig segments for compound expressions greatly expands the information needed to represent instructions that require relatively few words in English. Allowing quantity fields, such as the dose quantity, to take on ranges, such as '1 to 2,' rather than repeating the entire Sig segment with the dose quantity field taking only the individual values in the range, would avoid ambiguities that might arise from combining numeric range repeats with other kinds of repeats, such as 'AND' or 'THEN' multi-part instructions. Newer data formats such as XML can accommodate such complex data structures. A similar approach might also be used to efficiently accommodate multiple instances of vehicle, route of administration, site of administration, administration timing, indication, and maximum dose restriction, for which there is currently no method of representation.

Because missing information was one of the more frequent causes for prescriptions to be inadequately specified, it might improve the current Format to make more segments or fields mandatory. Dose quantity, for instance, may be made mandatory as it is a critical component of patient instructions. A caveat is that mandatory use of the dose fields might create problems for Sigs of topical medications, which are not accommodated by the current Format. In addition, timing-related fields and the indication segment, another key element of patient instructions, are optional. Some inadequately specified instructions might be avoided by making it mandatory that at least one set of the following fields should be used: frequency numeric value and frequency unit, interval numeric value and interval unit, administration timing, or indication precursor. This would make timing information optional for PRN prescriptions but otherwise required for patient safety.

Resolution is also needed on the question of whether prescriber-edited free text instructions should be allowed to coexist with values in the structured and codified fields versus only permitting 'pure free text' instructions (with no values in the structured and codified fields) if any editing or supplemental instructions are added, to prevent potential inconsistencies. One possible solution would be to create two text fields, one for reconstructed text from the structured fields, which the prescriber could not edit, and a separate 'additional instructions' field where they enter additional information. If prescribers were shown the reconstruction of what is already in the structured fields adjacent to where they are typing additional instructions, this could potentially reduce the possibility of entering conflicting information.

Recommendations for terminology mapping

The Format does not support compositionality of terms, for example, combining a main code with modifiers like 'severe' or 'right.' This type of expression, however, is available in SNOMED, and it is increasingly assumed that SNOMED will be used in a context that can support such compositionality (such as the HL7 Clinical Document Architecture). The nested, complex data structures that would be required to support compositionality are not compatible with the relatively flat, EDIFACT-based structure of the current Format. If compositionality in the Format were supported, substantially more Sig terms could be codified.

Study limitations

Our findings are based on ambulatory prescriptions that were generated mostly by primary care physicians. Thus, our findings may not generalize to prescriptions written by specialists or apply to the inpatient setting. In addition, our study is a laboratory evaluation using historical Sig strings rather than being a prospective pilot test. As such, it may overestimate the proportion of Sigs that are not representable using the Format because providers may change their prescribing patterns using a well-designed e-prescribing interface that steers them toward standardized expressions.

Safety implications

Improving the Format could have important safety benefits. A recent study found that nearly one in five primary care patients misinterpreted their prescription instructions,¹⁷ half of which were dosage errors that might be averted through better explanations of amounts, frequencies, and acceptable ranges. Clarifying the use of the free text field in combination with the

structured fields would reduce the potential for internal inconsistencies within the Sig, which currently appear in 1%~16% of electronic prescriptions containing a free text field, and about 20% of these inconsistencies could lead to moderate or severe adverse drug events.^{11 12} Finally, the Format could even help pharmacists to generate more consistent labels, which is often a challenge.¹⁸

CONCLUSION

The Structured and Codified Sig Format is not ready for immediate use as a federally mandated component of an outpatient e-prescribing standard, because substantial proportions of the terms needed for codified fields could not be mapped in the designated terminology systems and because there are no precedence rules for combining different types of repeating Sig segments. Further work is also needed to achieve consensus on the semantics of each field in the Format and to resolve the proper use of the free text field, among other ambiguities.

Acknowledgments We are indebted to many people for providing expert advice during the course of this project, including Laura Topor, who represented the NCPDP Structured and Codified Sig task group, and our expert panel, including Peter Kaufman and Michelle Soble-Lerner (DrFirst), Ajit Dhavle, PharmD (Surescripts), Dan Makowski, RPh (Allscripts), George Robinson, RPh and Shobha Phansalkar, RPh, PhD (Partners Healthcare System), Scott Robertson, PharmD (Kaiser Permanente), Alan Zuckerman, MD (Georgetown), Rick Peters, MD (OpenHealth Consulting), Casey Kozlowski, RPh, Miranda Rochol, CPhT, and Michelle Davidson, RPh (Walgreens), Jim Hancock, RPh (QS/1), Rob Franz, RPh (Medco), and Connie Sinclair, RPh (Point of Care Partners). We also thank Scot Hickey for professional database management and programming support and Diane Schoeff for extraordinary project management. We are also indebted to Jon White of the Agency for Healthcare Research and Quality for providing seminal feedback on our initial study design. Finally, we thank our project officer, Andrew Morgan, at the Centers for Medicare and Medicaid Services (CMS), for providing consistent guidance, support, and encouragement.

Funding This project was funded by the Centers for Medicare and Medicaid Services Office of E-Health Standards and Services (contract number: HHSM-500-2005-000281).

Competing interests None.

Ethics approval This study was approved by the RAND Institutional Review Board.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. **National Council for Prescription Drug Programs.** *Structured and Codified Sig Format Implementation Guide*. Scottsdale, AZ: National Council for Prescription Drug Programs, 2008.
2. **Bell DS, Schueth A, Crosson JC, et al.** *Pilot Testing of Electronic Prescribing Standards*, 2007. http://www.rand.org/health/projects/erx/documents/NJ-EPAC_Pilot_Testing_of_Electronic_Prescribing_Standards.pdf (accessed 9 Feb 2010).
3. **Centers for Medicare and Medicaid Services (CMS), HHS.** Medicare program; standards for e-prescribing under Medicare Part D and identification of backward compatible version of adopted standard for e-prescribing and the Medicare prescription drug program (version 8.1). Final rule. *Fed Regist* 2008;**73**:18917–42.
4. **Aspden D, Wolcott J, Bootman J, et al.** *Preventing Medication Errors*. Washington, DC: The National Academies Press, 2007.
5. **Ammenwerth E, Schnell-Inderst P, Machan C, et al.** The effect of electronic prescribing on medication errors and adverse drug events: a systematic review. *J Am Med Inform Assoc* 2008;**15**:585–600.
6. **Johnston D, Pan E, Middleton B, et al.** *The Value of Computerized Order Entry in Ambulatory Settings*. Boston, MA: Center for Information Technology Leadership, 2004.
7. **Mandel R, Boulter P.** *E-Rx Collaborative, E-Prescribing: A Bridge to the 21st Century*. 2004. <http://www.ncvhs.hhs.gov/040525p4.pdf> (accessed 9 Feb 2010).
8. **Wang SJ, Middleton B, Prosser LA, et al.** A cost-benefit analysis of electronic medical records in primary care. *Am J Med* 2003;**114**:397–403.
9. **Department of Health and Human Services.** *Pilot Testing of Initial Electronic Prescribing Standards—Cooperative Agreements*, 2007. http://healthit.ahrq.gov/portal/server.pt/gateway/PTARGS_0_227312_0_0_18/eRxReport_041607.pdf (accessed 9 Feb 2010).
10. **Centers for Medicare and Medicaid Services.** *CMS Legislative Summary: Medicare Prescription Drug, Improvement, and Modernization Act of 2003*. 2004. <http://www.cms.hhs.gov/MMAUpdate/downloads/PL108-173summary.pdf> (accessed 8 Feb 2010).
11. **Singh H, Mani S, Espadas D, et al.** Prescription errors and outcomes related to inconsistent information transmitted through computerized order entry: a prospective study. *Arch Intern Med* 2009;**169**:982–9.
12. **Palchuk MB, Fang EA, Cygielnik JM, et al.** An unintended consequence of electronic prescriptions: prevalence and impact of internal discrepancies. *J Am Med Inform Assoc* 2010;**17**:472–6.
13. **Bell DS, Friedman MA.** E-prescribing and the medicare modernization act of 2003. *Health Aff (Millwood)* 2005;**24**:1159–69.
14. **National Cancer Institute.** *Federal Medication Terminologies*, 2009. <http://www.cancer.gov/cancertopics/cancerlibrary/terminologyresources/fmt> (accessed Feb 2010).
15. **U.S. National Library of Medicine.** *SNOMED Clinical Terms® (SNOMED CT®)*, 2011. http://www.nlm.nih.gov/research/umls/Snomed/snomed_main.html (accessed Apr 2011).
16. **National Cancer Institute.** *Directory Listing of /evs/ftp1/FDA/SPL*, 2009. <http://evs.nci.nih.gov/ftp1/FDA/SPL> (accessed Dec 2009).
17. **Davis TC, Wolf MS, Bass PF 3rd, et al.** Literacy and misunderstanding prescription drug labels. *Ann Intern Med* 2006;**145**:887–94.
18. **Wolf MS, Shekelle P, Choudhry NK, et al.** Variability in pharmacy interpretations of physician prescriptions. *Med Care* 2009;**47**:370–3.