Ubiquitous Computing Laboratory
Kyung Hee University, Korea

PhD Dissertation Presentation

Kyung Hee University

Guideline Enabled Data-driven Knowledge Acquisition and Validation Method for CDSS

Mr. Maqbool Hussain

Department of Computer Science and Engineering

Kyung Hee University, Republic of Korea

Email: maqbool.hussain@oslab.khu.ac.kr

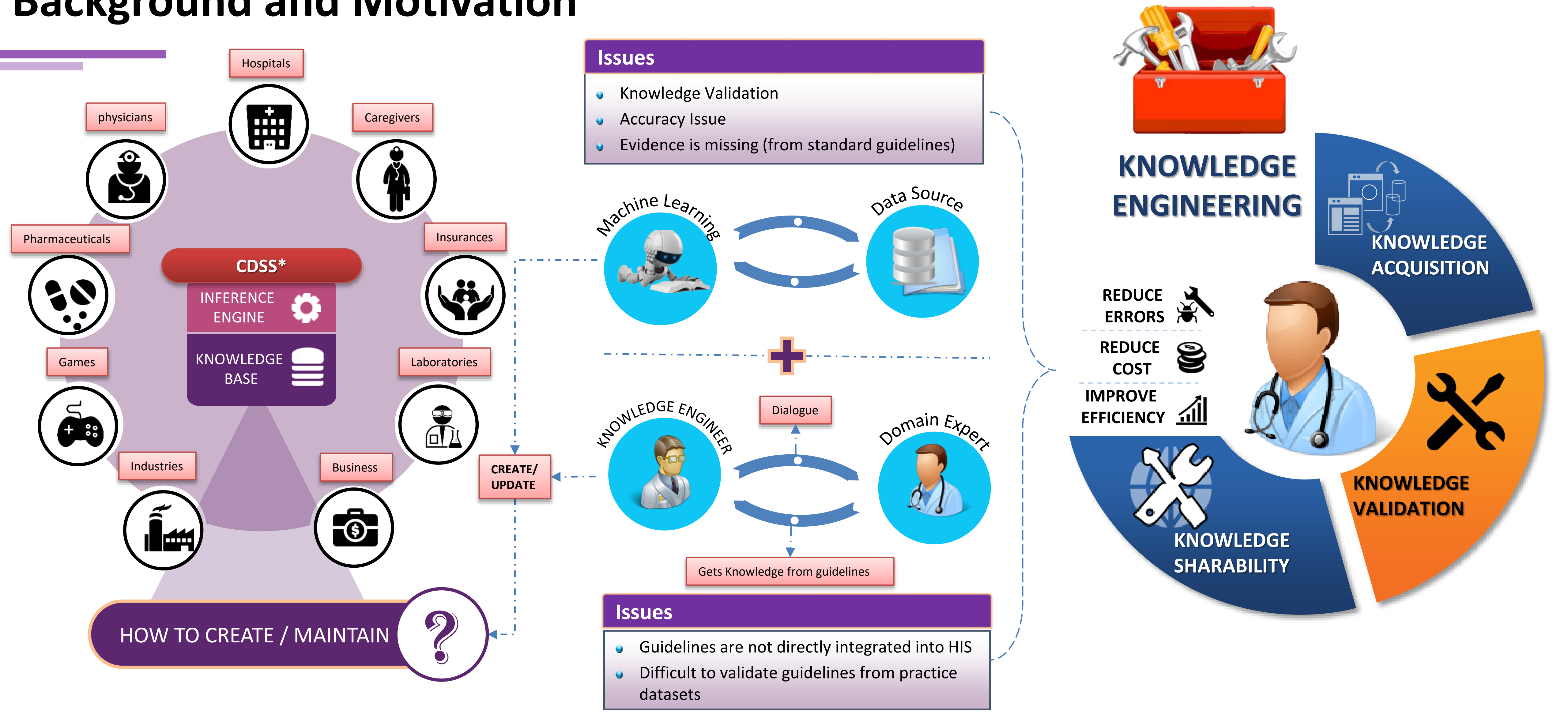
Date: April 25, 2016

Advisor: Prof. Sungyoung Lee, PhD

Presentation Agenda

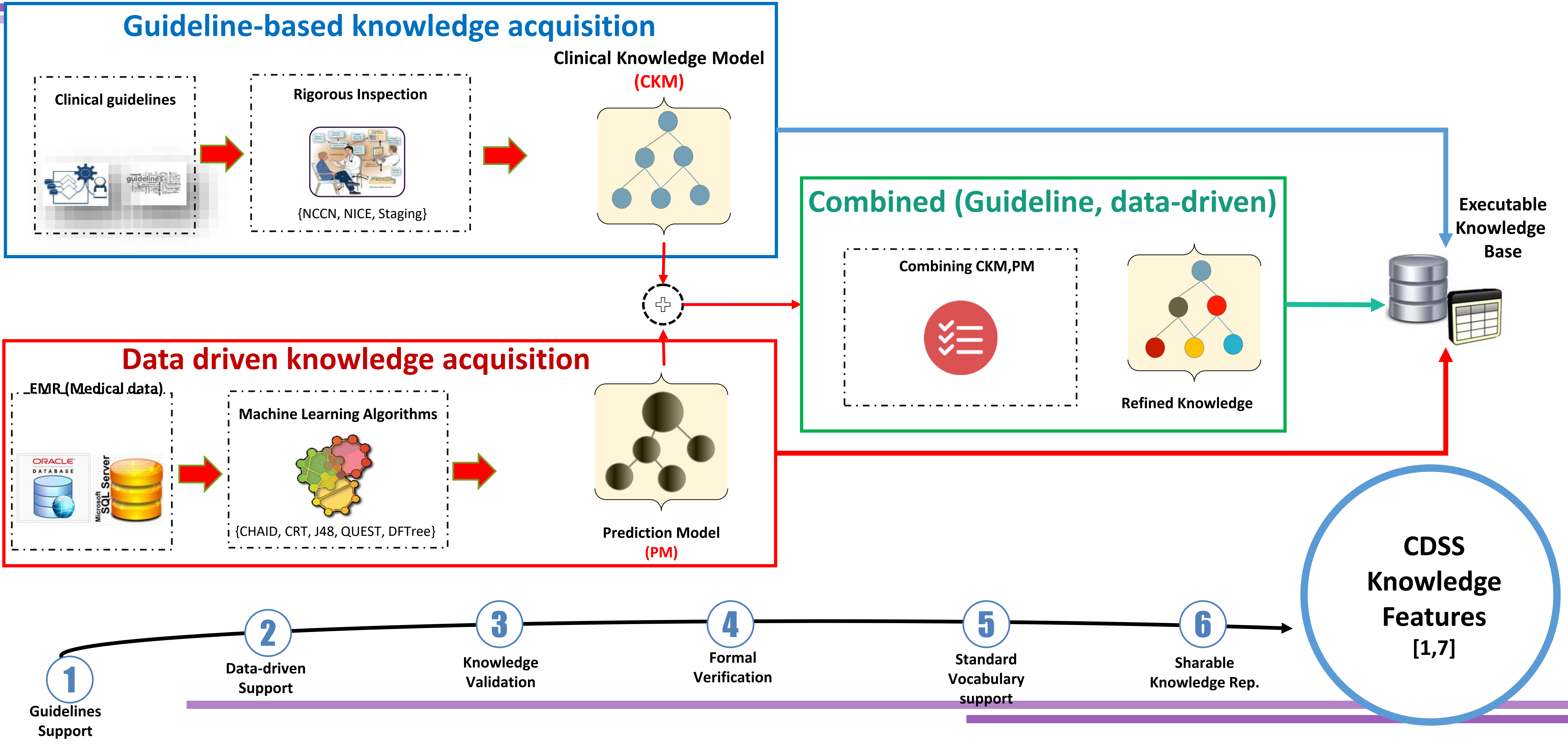
- Introduction
  - Background and Motivation
  - Problem Statement
- Related Work
- Proposed Methodology
- Solutions
- Experiments and Results
- Uniqueness and Contributions
- Conclusion and Future Research

Background and Motivation

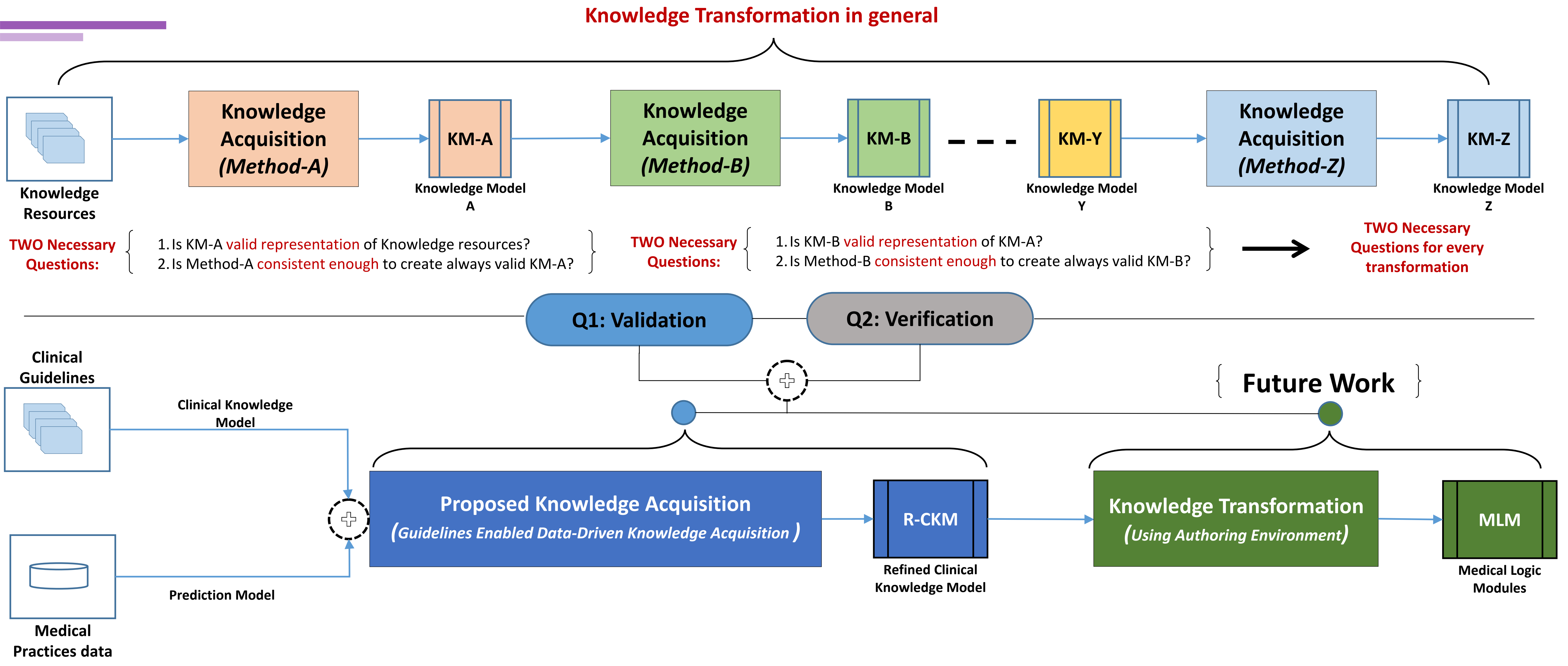


*Clinical Decision Support System

Medical Knowledge acquisition in General



Medical Knowledge acquisition Requirements and Scope of the Proposed work



Problem Statement

- Knowledge acquired from patient data (Data-driven) is considered non-standard and non-validated, Knowledge acquired from CPGs (Guidelines) is generic and non-integrated into real healthcare workflows.

Goal

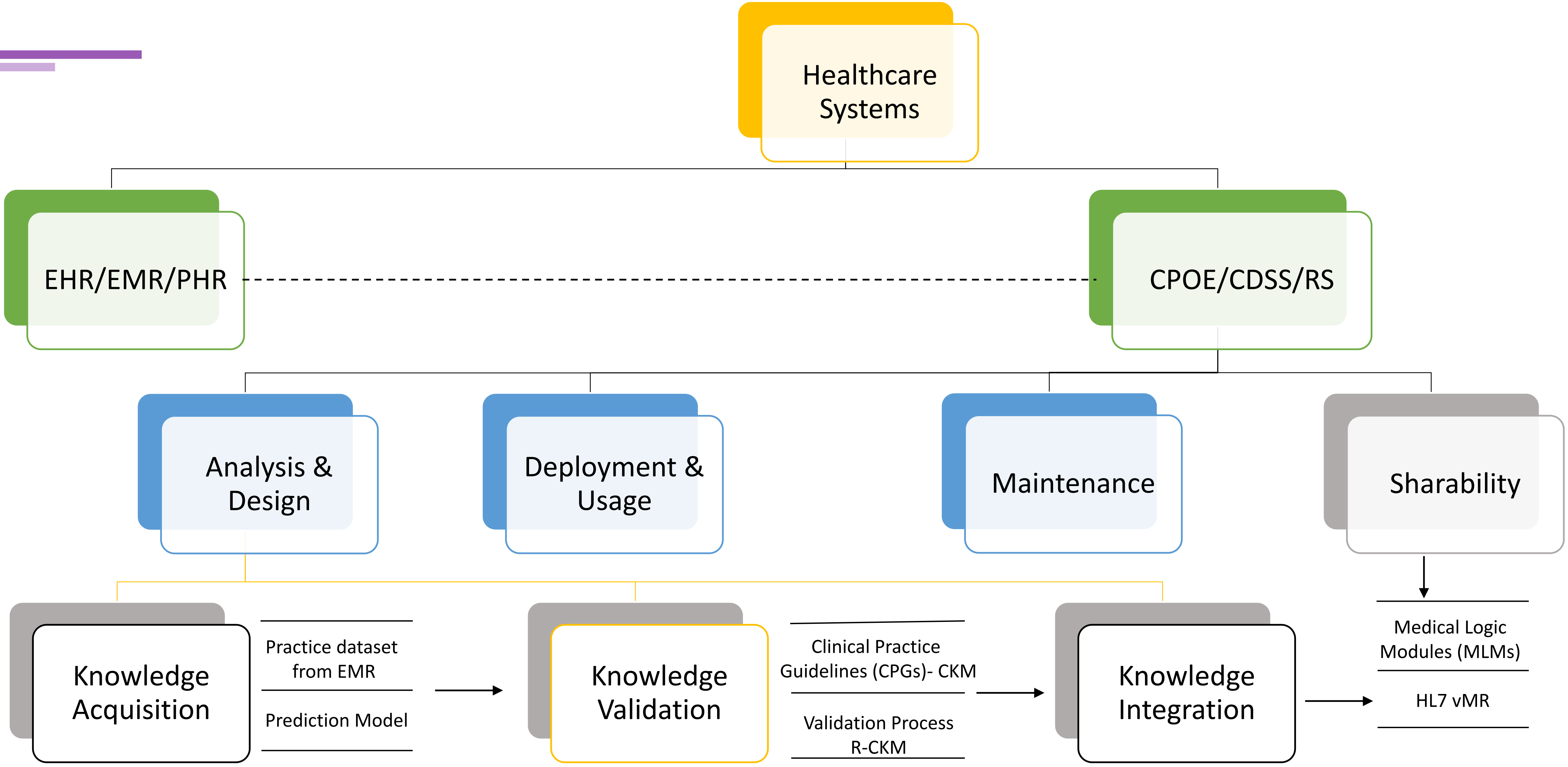
Methodology that exploit real practice dataset (EMRs) for recommendation model and leverage CPGs for validating it for refined standard recommendation model.

Challenges

How to establish validation criteria that align diverse knowledge resources into standardized knowledge acquisition model.

How to verify that validation methodology is consistent and its applicability will result in validated and consistent knowledge acquisition model?

Research Taxonomy



Related Work

Category	Research Work	Guidelines Support	Data-driven Support	Knowledge Validation (patient-cases)	Formal Verification	Standard Vocabulary support	Sharable Knowledge Rep.
Guideline base Knowledge acquisition	Peleg [8]	√	x	x	x	x	√
	Serban [9]	√	x	x	√	√	x
	Shalom [14]	√	x	x	√	x	√
	Miller [15]	√	x	√	x	x	√
Data-driven knowledge acquisition	Perera [11]	x	√	√	x	x	x
	Gomoi [12]	x	√	√	x	x	√
Data-driven and Guideline for knowledge acquisition	Toussi [13]	√	√	x	x	x	x
	Proposed Approach	√	√	√	√	√	√

Limitations: Guideline-base knowledge acquisition

- Knowledge is non-integrated into healthcare workflows
- Not properly validated against patient cases

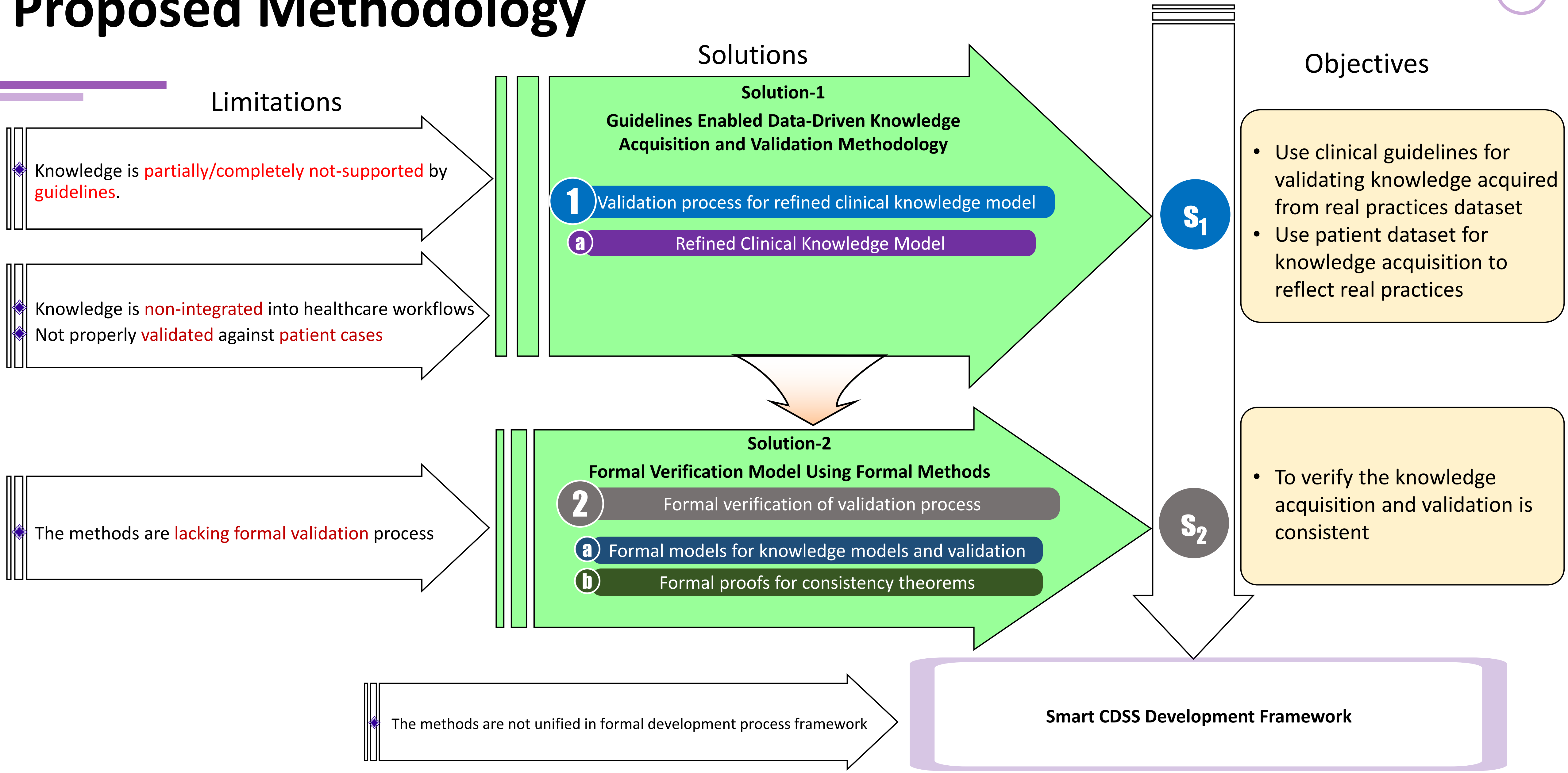
Limitations: Data-driven knowledge acquisition

- Knowledge is not-supported by guidelines.
- The methods are lacking formal verification process

Limitations: Data-driven and Guideline-based knowledge acquisition

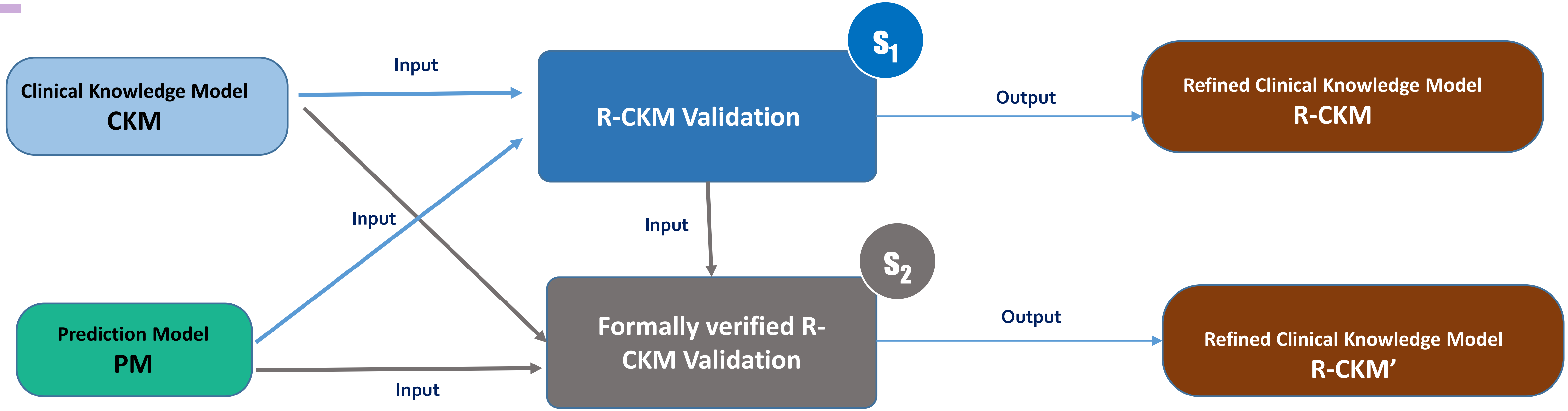
- Knowledge acquired from data which is missing in guidelines is not-supported by guidelines.
- Lack of proper validation

Proposed Methodology



Maqbool Hussain, Muhammad Afzal, Taqdir Ali, Rahman Ali, Wajahat Ali Khan, Arif Jamshed, Sungyoung Lee, Byeong Ho Kang, Khalid Latif; "Data-driven knowledge acquisition, validation, and transformation into HL7 Arden Syntax", Artificial Intelligence in Medicine (Elsevier, SCI, IF:2.03), published (Online Oct, 2015)
Maqbool Hussain, Taqdir Ali, Wajahat Ali Khan, Muhammad Afzal, Sungyoung Lee, Khalid Latif,;"Recommendations service for chronic disease patient in multi-model sensors home environment", Telemedicine and EHealth (SCI, IF:1.6), Vol. 21 Issue 3, pp.185-199, 2015

Conceptual representation of Proposed Methodology



Solution 1

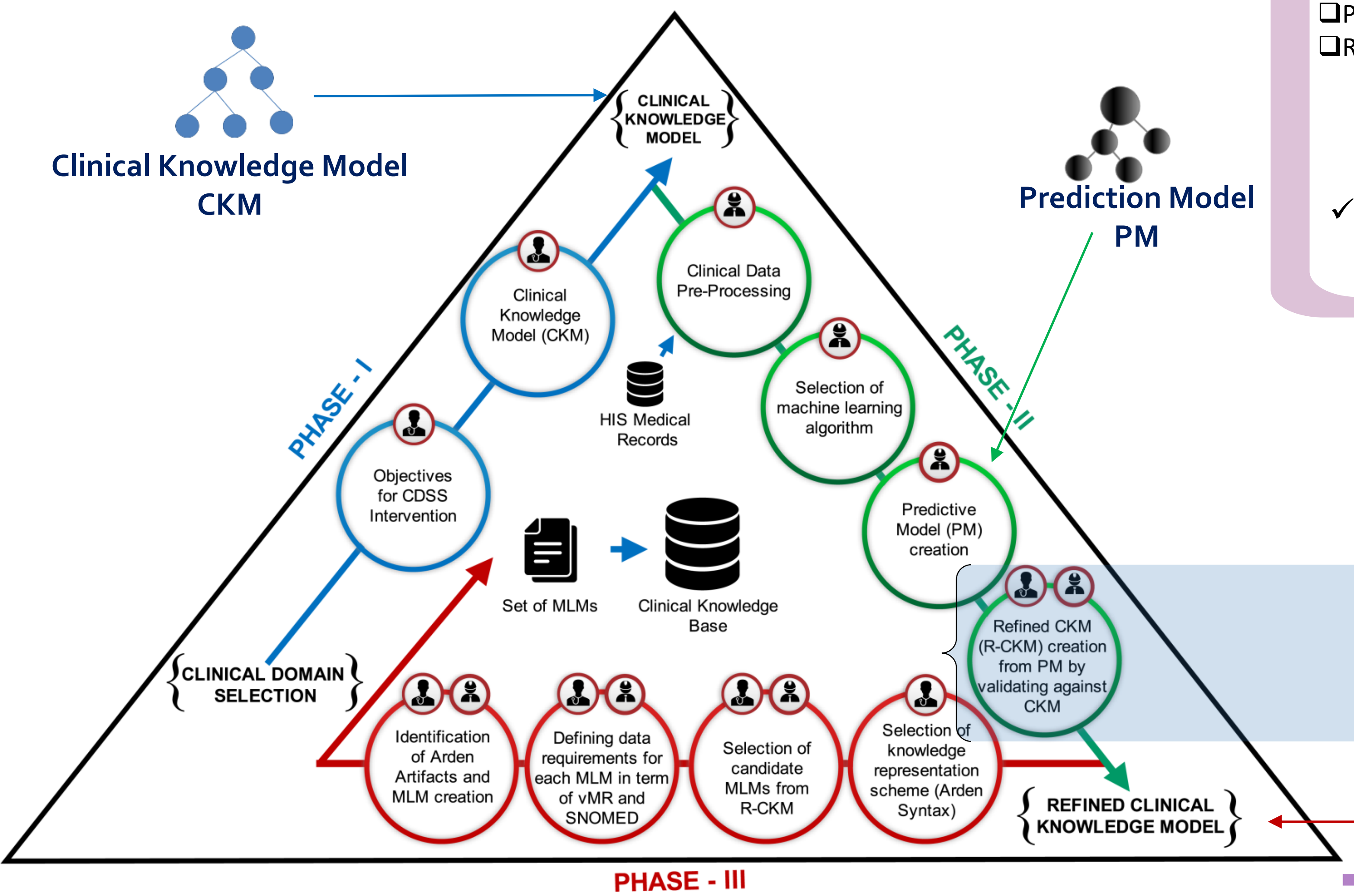
- ❖ Using rigorous validation process to create R-CKM to overcome limitation of
  - ✓ Valid and conformed final model from guidelines - (Data-driven limitation)
  - ✓ The model is evolved from data, so integrable with healthcare system workflow (Guideline-base limitation)
  - ✓ R-CKM pass through validation process (Combined approach limitation)

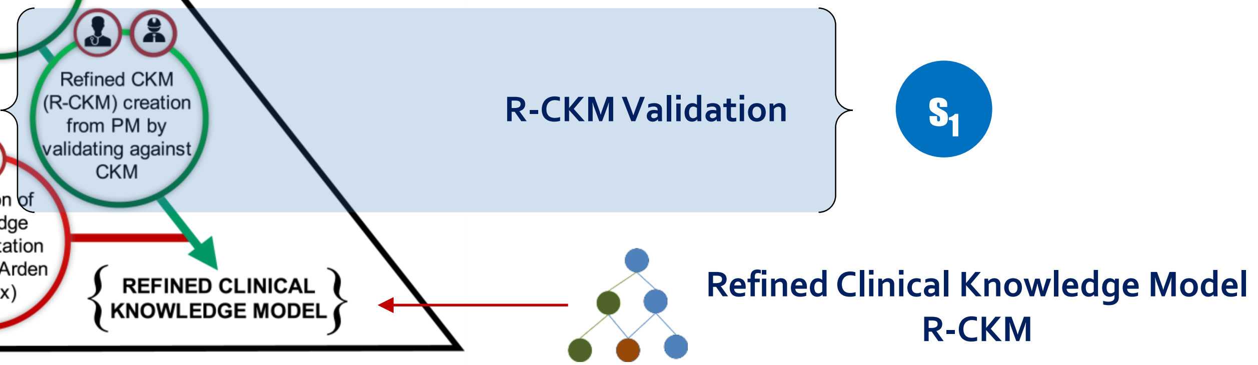
Solution 2

- ❖ Using formal verification to ensure validation process consistency which ultimately;
  - ✓ Enhance validation process for completeness
  - ✓ Produce valid R-CKM

Maqbool Hussain, Muhammad Afzal, Taqdir Ali, Rahman Ali, Wajahat Ali Khan, Arif Jamshed, Sungyoung Lee, Byeong Ho Kang, Khalid Latif; "Data-driven knowledge acquisition, validation, and transformation into HL7 Arden Syntax", Artificial Intelligence in Medicine (Elsevier, SCI, IF:2.03), published (Online Oct, 2015)
Maqbool Hussain, Taqdir Ali, Wajahat Ali Khan, Muhammad Afzal, Sungyoung Lee, Khalid Latif,;"Recommendations service for chronic disease patient in multi-model sensors home environment", Telemedicine and EHealth (SCI, IF:1.6), Vol. 21 Issue 3, pp.185-199, 2015

Solution-1: Process Model





Solution-1

Objectives

- ❑Guideline base validation
- ❑PM from patient data
- ❑Refined sharable clinical knowledge

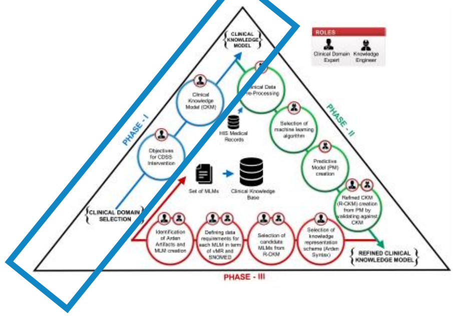
Target Limitations

- ❖Guideline support
- ❖Validated knowledge

Outcome

✓ Guideline enabled clinical knowledge model – R-CKM

Solution-1:
Clinical Knowledge Modeling (Phase-I)

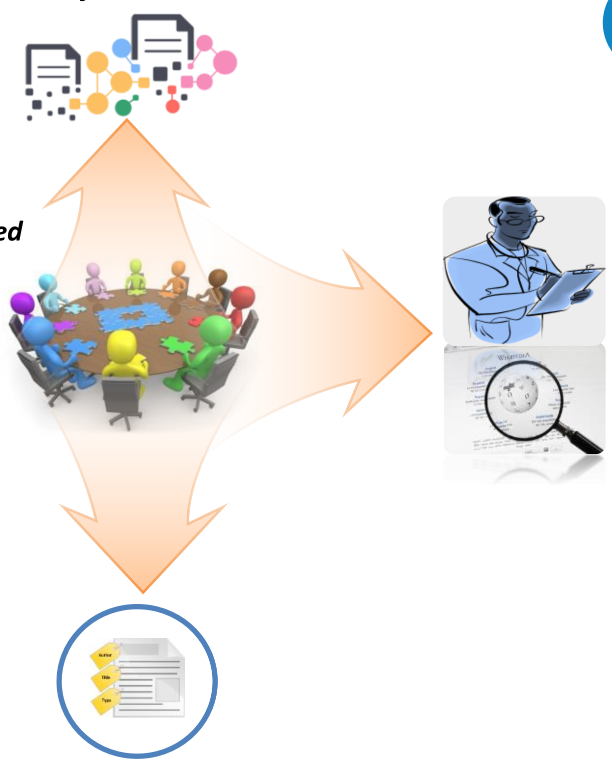


Why CKM?

- **Guidelines** *are semi-structured form (decision tree and description so needs to unify to single representation) and*
- **Guidelines** *are too generic (hard to apply directly to practices)*
- **CKM** *are easy to manage and make knowledge explicit*

1.b Guidelines selection

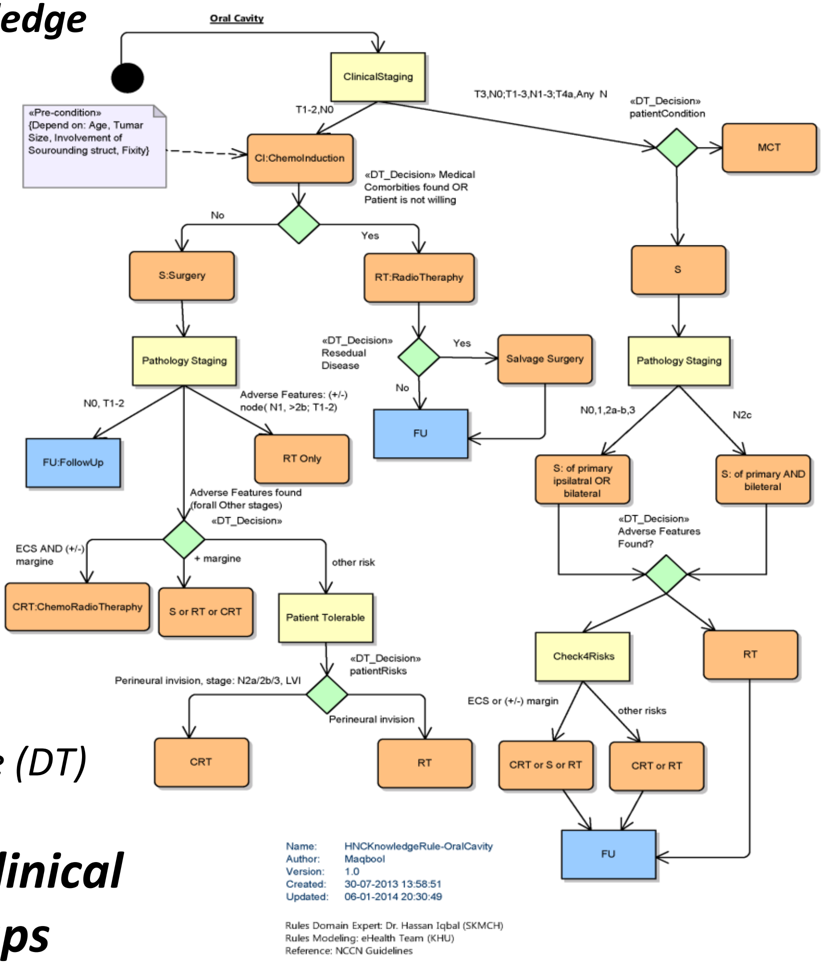
- *NCCN and*
- *TNM for Cancer*



Team Involved
- *Oncologists*
- *Residents*
- *HIS Team*
- *Knowledge Engineers*

2 Clinical Knowledge Model creation

ENTERPRISE ARCHITECT



Contribution

- ***Using Rigorous Inspection process***
  - *Marked the guidelines to explicit the knowledge concepts*
  - *Mapped the concepts into decision tree (DT) formalism*
- ***DT represent more explicitly the clinical knowledge compared to mind-maps***

- *Treatment Plan for oral cavity tumor patients*

1.a Objectives for CDSS Intervention

Solution-1:

Clinical Knowledge Modeling (Phase-I) : Rigorous Inspection process for CKM creation

Defining knowledge representation formalism → Clinical Guidelines Analysis → Guidelines Refinements → Guidelines verification and finalization

Knowledge Representation

- Select the knowledge representation which is more intuitive

Toolset

- Select the toolset which is best fit to the knowledge representation scheme
- Get trained domain experts on the toolset for knowledge representation

Structured

- Mark guidelines that meet clinical objective
- Draw the knowledge model
- Mark all decision paths for further explicit knowledge.

Narrations

- Analyze narrations for marked paths
- Analyze the external evidences
- Revise the marked decision paths.

Local Evidences

- Mark decision paths, candidate for refinement
- Collect local evidences to support refinement
- Refine the paths for local evidences

Concepts Harmonization

- Analyze concepts that are composite or having alternate meanings
- Establish mappings for composite concepts
- Provide translations for alternate meanings

Guideline Modeling

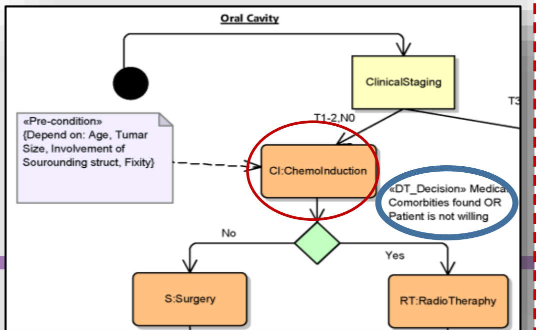
- Model the guidelines in accordance to selected formalism.
- Provides all associated mappings and translations

Verification

- Verify the guideline model from senior domain expert
- Approve or mark improvements
- Finalize the model as CKM after approval

- **Decision Tree**
  - *Suitable for explicit knowledge*
  - *NCCN is semi-structured in DTs*
- **Enterprise Architect**
  - *Support DTs.*
  - *Provide XMI based sharing facilities*

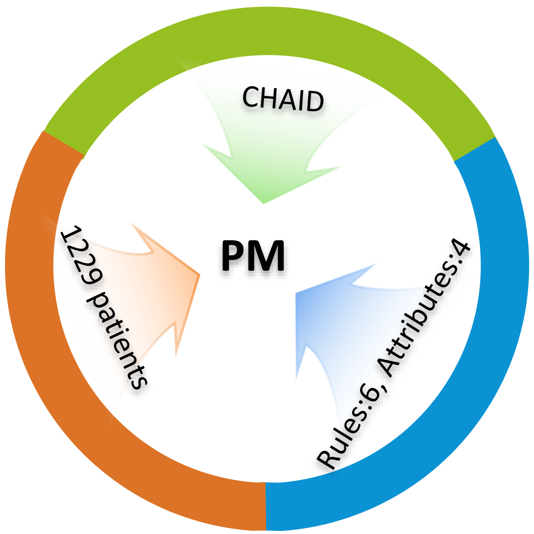
- **Comorbities**
  - *Added for explicit knowledge*



- **ChemoInduction**
  - *Local practices (92%) success for survival*
- **Patient willingness**
  - *Localized decision in addition to comorbities*
- **Palliative patients:**
  - *Stage III or IV*

- **CKM Model**
  - *CKM in decision tree form*
- **Harmonization**
  - *Mapping tables*

Solution-1:
Prediction Model (PM) (Phase-II) (1/2)



$$Ranking^{WSM-score} = \propto \sum_{j=1}^{m} (w_j a_{ij}), for\ i = 1,2,3,\ldots,m$$

Here $\propto$: 0.8 *is scaling constant and* a_{ij} *are attributes with weight* w_j

$$\begin{Bmatrix} a_{ij} & Accuracy(P) & Number\ of\ rules\ (R) & Attributes(A) \\ w_j & 0.8 & -0.1 & -0.1 \end{Bmatrix}$$

1 Data preprocessing

HIS

1. Data from 2181 patients
   - Excluded 204
     - Missing value patients (3)
     - Consulted patients (200)
     - Duplicate patients (1)
2. 1977 patients
   - Clinical stage correction (130)
   - Clinical stage for missing values (118) [Using PART]
3. Clinical stage calculation
   - Excluded 748
     - Patients excluded due to incomplete treatment
     - Increased performance of decision tree algorithm (from 43% to 70%)
4. 1229 patients

1229 patients

2 ML algorithm selection

CRT CHAID DFTree QUEST J48

1. Applying algorithm

$$Ranking^{WSM-score} = \propto \sum_{j=1}^{m} (w_j a_{ij}), for\ i = 1,2,3,\ldots,m$$

Here $\propto$: 0.8 *is scaling constant and* a_{ij} *are attributes with weight* w_j

$$\begin{Bmatrix} a_{ij} & Accuracy(P) & Number\ of\ rules\ (R) & Attributes(A) \\ w_j & 0.8 & -0.1 & -0.1 \end{Bmatrix}$$

2. Rank algorithm

Algorithm	P	R	A	Ranking
CHAID	71.00%	6	4	0.6975
CRT	71.90%	9	13	0.6915
QUEST	70.90%	6	13	0.6853
DFTree	70.10%	11	8	0.6773
J48	68.90%	55	13	0.6040

CHAID

3 Prediction model creation

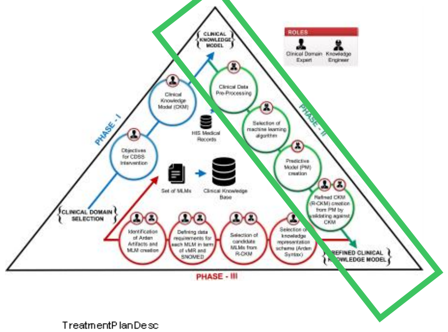
1. CHAID configuration
2. Executing CHAID

Model Summary

Specifications	Growing Method	CHAID
	Dependent Variable	TreatmentPlanDesc
	Independent Variables	Clinical Stage T, Clinical Stage N, Clinical Stage M, Clinical Stage S, TreatmentIntentDesc, PatientStatusDesc, TreatmentStatusDesc, SmokingStatus, PanStatus, NaswarStatus, AlcoholStatus, HistoDesc, GradeDesc
	Validation	Cross Validation
	Maximum Tree Depth	3
	Minimum Cases in Parent Node	100
	Minimum Cases in Child Node	50
Results	Independent Variables Included	TreatmentIntentDesc, Clinical Stage T, Clinical Stage S, HistoDesc
	Number of Nodes	10
	Number of Terminal Nodes	6
	Depth	3

Solution-1:
Prediction Model (PM) (Phase-II) (2/2)

Why PM?
- *Domain experts want knowledge from real practices.*
- *The knowledge from patient data is integrable directly to HIS.*
- *PM represent real practices and easily integrable to HIS.*



Contribution

Comprehensibility and Understandability of algorithm

Algorithm	P	R	A	Ranking
CHAID	71.00%	6	4	0.6975
CRT	71.90%	9	13	0.6915
QUEST	70.90%	6	13	0.6853
DFTree	70.10%	11	8	0.6773
J48	68.90%	55	13	0.6040

RankDecisionTreeAlgo

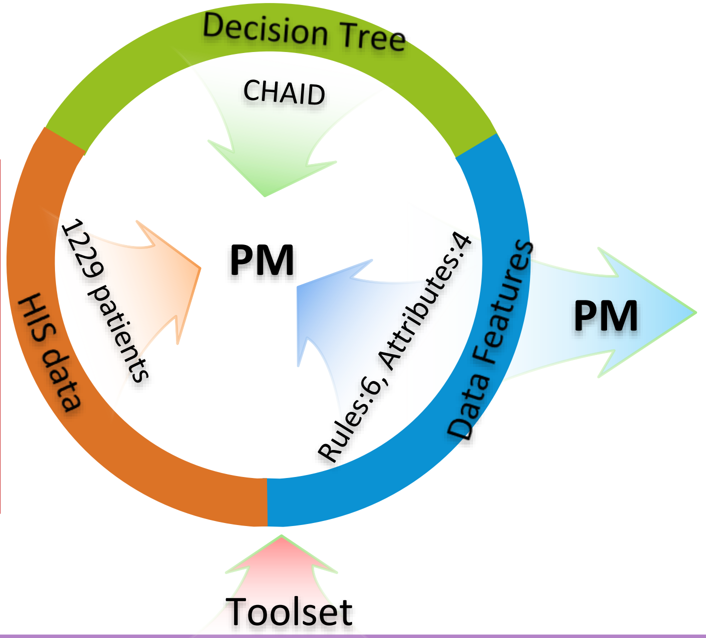
Input: DTAlgos<List>
Output: RankedDTAlogs<List>

1. *Let accuracyP , attributeA, numRulesR, ranking*
2. *ForEach dt in DTAlgos*
3. *Begin*
4. *accuracyP = getDTAccuracy(dt)*
5. *attributeA = getDTNoAttribute(dt)*
6. *numRulesR = getDTNoRules(dt)*
7. ***ranking = Ranking$^{WSM\text{-}score}$ (accuracyP, attributeA , numRulesR)***
8. *RankedDTAlogs.add(dt,ranking)*
9. *End*
10. *Return RankedDTAlogs*

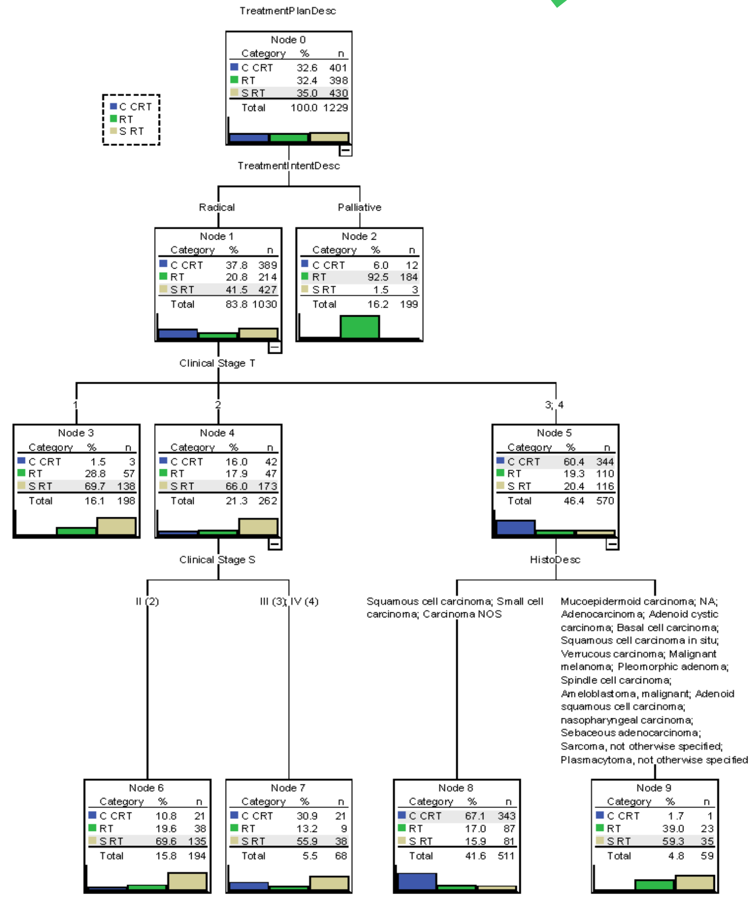
$$Ranking^{WSM-score} = \propto \sum_{j=1}^{m} (w_j a_{ij}), for\ i = 1, 2, 3, \dots, m$$

Here $\propto$: 0.8 *is scaling constant and* a_{ij} *are attributes with weight* w_j

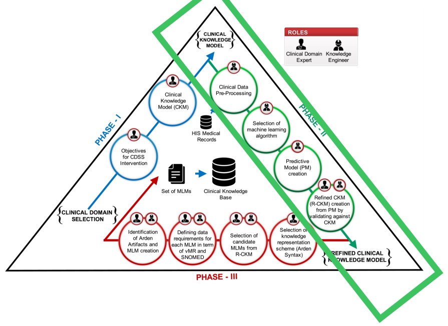
$$\begin{Bmatrix} a_{ij} & Accuracy(P) & Number\ of\ rules\ (R) & Attributes(A) \\ w_j & 0.8 & -0.1 & -0.1 \end{Bmatrix}$$



IBM SPSS

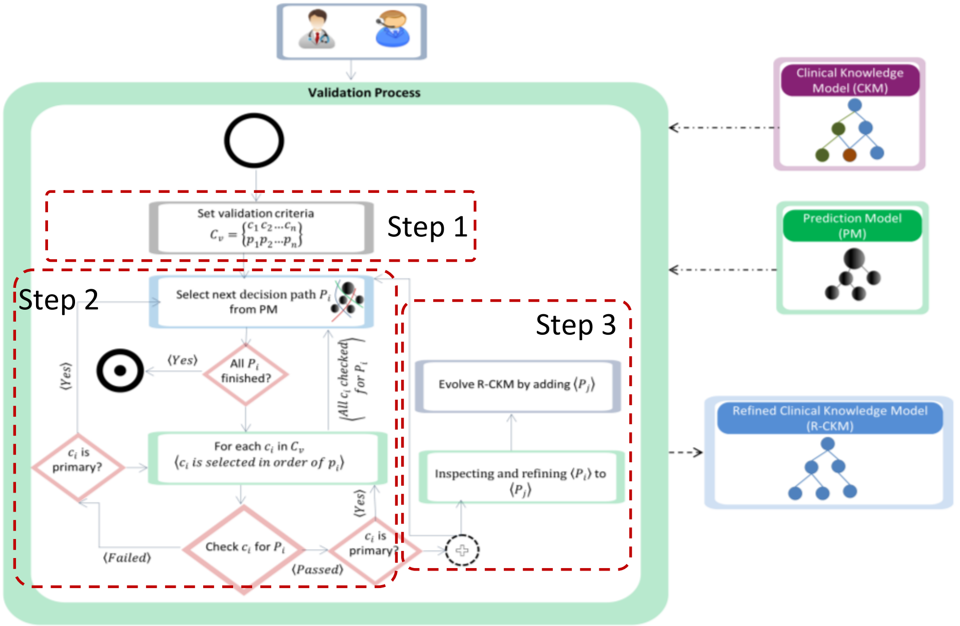


Solution-1:
Refined Clinical Knowledge Model after validation process (Phase-II) (1/5)



Validation Process

Step 1: Validation Criteria setting
Step 2: Conformance of decision path of PM
Step 3: Refine and evolve decision path into R-CKM

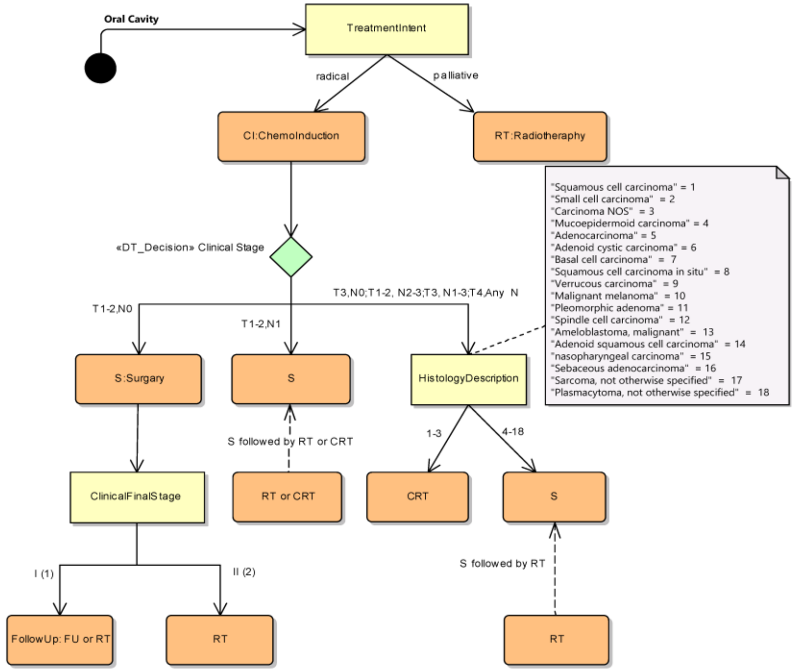


Criteria of validation Process

1) $\{\forall\, P_i \in PM : \textbf{Accuracy}(P_i) > N\%\}$
2) $\{\forall\, P_i \in PM \wedge \forall P_j \in CKM : \textbf{Conflict}(P_i, P_j)\}$
3) $\left\{\forall\, P_i \in PM \wedge \exists P_j \in CKM : \textbf{Conform}(P_i, P_j) \xrightarrow{yields} P_i \in \Delta\, RCKM\right\}$
4) $\{\exists\, P_i \in PM \wedge \forall P_j \in CKM :$

Why R-CKM?

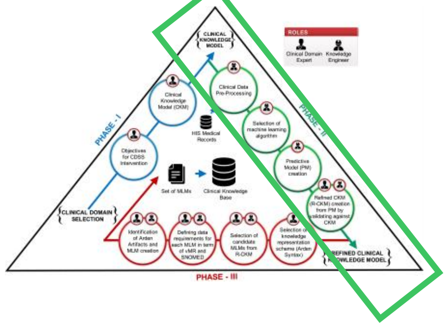
- ***R-CKM** provides refined knowledge which support both Standard guidelines and Real practices data*
- ***Consequently , evidence based knowledge which is directly integrable to HIS***



Contribution

- ***Decision path level conformance***
- ***Conflict resolution***

Solution-1:
Refined Clinical Knowledge Model after validation process (Phase-II) (2/5)



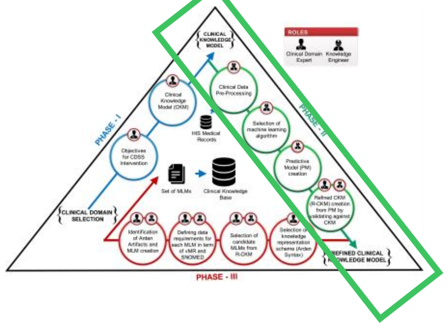
1 Set validation criteria

$$C_v = \begin{Bmatrix} c_1\, c_2 \cdots c_n \\ p_1 p_2 \cdots p_n \end{Bmatrix}$$

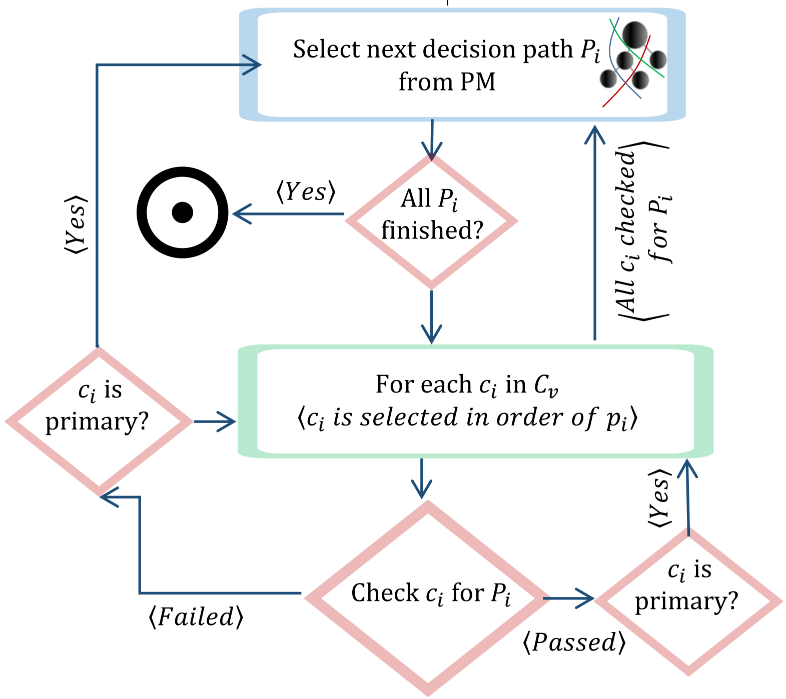
	1 Accuracy	2 Conflicts	3 Conformance	4 Evidence
Rationale	Local evidences (Context/Value)	Sustain Integrity of CPGs (Quality care provision)	Conform to CPGs (Quality care provision)	Quality research result (Quality care provision)
Formal Rep.	$\left\{\begin{matrix} \forall P_i \in PM: \\ \textbf{Accuracy}(P_i) > N\% \end{matrix}\right\}$	$\left\{\begin{matrix} \forall P_i \in PM \wedge \forall P_j \in \\ CKM : \textbf{Conflict}(P_i, P_j) \end{matrix}\right\}$	$\left\{\begin{matrix} \forall P_i \in PM \wedge \exists P_j \in \\ CKM: \textbf{Conform}(P_i, P_j) \\ \xrightarrow{yields} P_i \in \Delta\, RCKM \end{matrix}\right\}$	$\left\{\begin{matrix} \exists P_i \in PM \wedge \forall P_j \in \\ CKM : !Conform(P_i, P_j) \\ \xrightarrow{provide} \textbf{Evidence}(P_i) \wedge \\ Evidence(P_i) \xrightarrow{yields} P_i \\ \in \Delta\, RCKM \end{matrix}\right\}$
Primary/ Priority	Yes/1	Yes/1	No/2	No/3
Example	**Accuracy > 50%** for all decision path	**Chemo-Induction** after **Surgery** has no meanings	**Treatment plans** following same sequence to CPGs are conformed	**Clinical Trails, Systematic reviews, Meta-analysis, Expert opinion**

Proposed criteria support Best practices

Solution-1:
Refined Clinical Knowledge Model after validation process (Phase-II) (3/5)



2 Conformance of decision path of PM



1) **PM Model Processing**
   i. Decision path processing
   ii. Decision path meta-information processing
2) **Decision path selection for evaluation**

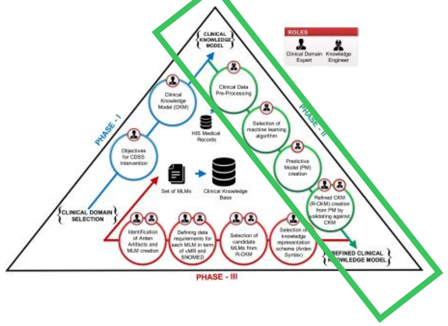
1) **Conformance criteria loading**
2) **Decision path evaluation**
   i. Primary criteria checking
   ii. Non-primary criteria checking
3) **Forwarding passed decision path for refinements**

ConformancePMtoCKM

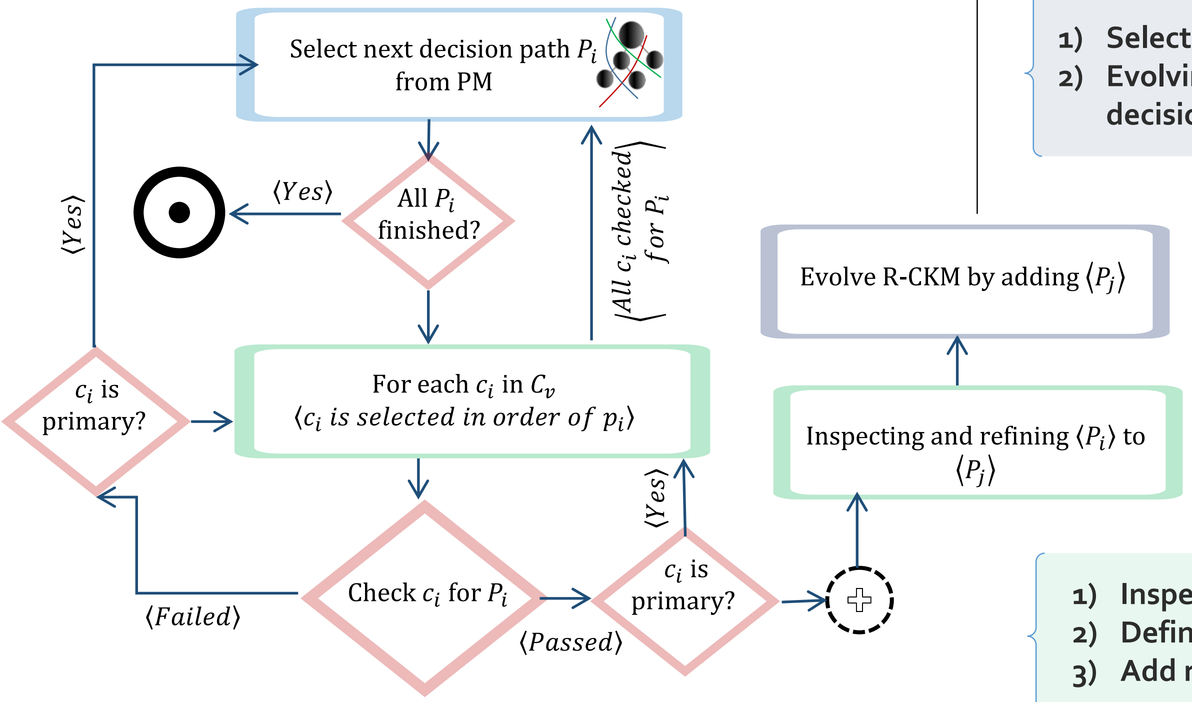
Input: PM, CKM, CriteriaBank: Map<keyCriteria, List<Criteria>>
Output: Conformed:Boolean (True/False)

```
Let decPathPM:List<PMDecisionPath> = PM.getDecPath()

While(CriteriaBank.isNotEmpty())
Begin

    Let criteriaList = CriteriaBank.getNext()

    ForEach dt in criteriaList
    Begin

        Conformed = getConformance(dt,criteriaList)
        If(!Conformed & isPrimary(criteriaList))
            breakAll;
        ElseIf( isNotPrimary(criteriaList) )
            breakAll;

    End

End
Return Conformed
```

Solution-1:
Refined Clinical Knowledge Model after validation process (Phase-II) (4/5)



3 Refine and evolve decision path into R-CKM



1) **Selection of candidate path in R-CKM**
2) **Evolving R-CKM with newly refined decision path**

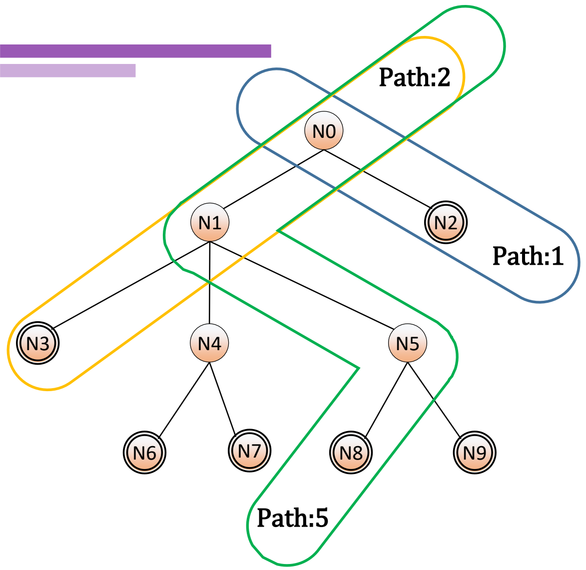
1) **Inspecting decision path**
2) **Defining refinements from CKM**
3) **Add refinements to decision path**

EvolveRCKM

Input: P_i : PMDecisionPath, listRefinements: List<Refinement>
Output: R-CKM

```
Let rckm: R − CKM = ∅, Pj : RCKMDecPath
Pj = Pi

ForEach ref in listRefinements
Begin

    Pj .addrefinements( ref )
    If(rckm = ∅)
        rckm.addFirstPath( Pj )
    Else
        rckm.updatePath( Pj )

End
Return rckm
```

Solution-1:
Refined Clinical Knowledge Model after validation process (Phase-II) (5/5)



Path:1 { If(TreatmentIntentDesc = Palliative) TreatmentPlan = RT }

Path:2 { If(TreatmentIntentDesc = Radical and ClinicalStage T = 1) TreatmentPlan = S RT }

Path:5 { If(TreatmentIntentDesc = Radical and ClinicalStage T = 3 or 4 and histology = 1 or 2 or 3) TreatmentPlan = C CRT }

1 Set validation criteria

(1) Accuracy >50%

(2) Conflicts

(3) Conformance

(4) Evidence

2 Conformance of decision path of PM

Path:1

1 — Accuracy: 92%

2 — No conflict

3 —
- Palliative patients are Stage III/IV (90%)
- RT is secondary level treatment for stage III/IV patient

Path:2

1 — Accuracy: 69.7%

2 — No conflict

3 —
- S RT is recommended treatment for all stage patients
- Radical patients may have any of Stage (I,II,III, and IV)
- So S RT is conformed.

Path:5

1 — Accuracy: 67.1%

2 — No conflict

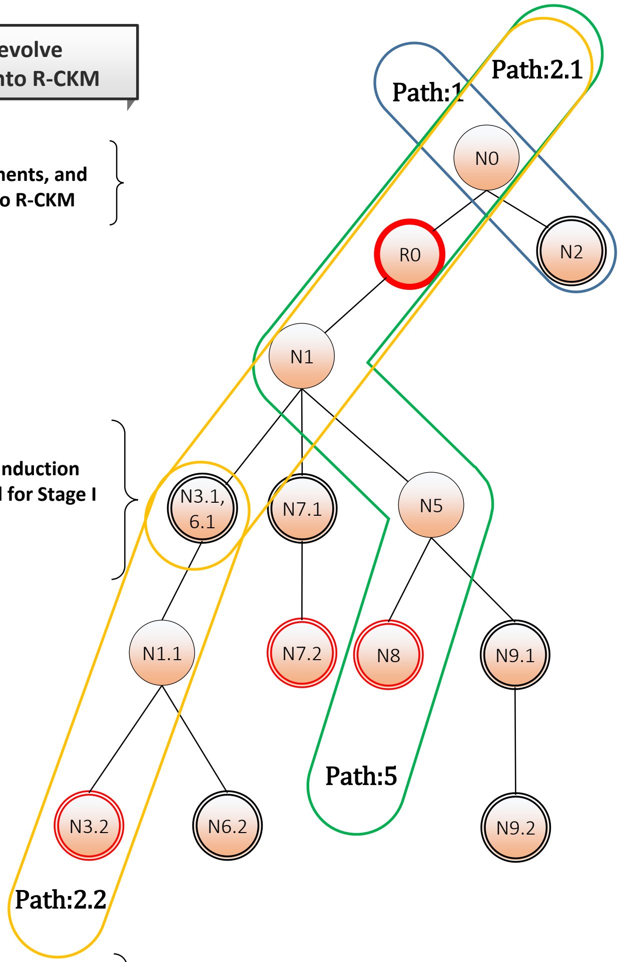
4 —
- Path 5 is not conformed to CKM
- Physicians provided local evidences and published expert opinion for effectiveness.
- So C CRT is kept in R-CKM.

3 Refine and evolve decision path into R-CKM

Refinements: { • No refinements, and • Add path to R-CKM }

Refinements: { • R0: ChemoInduction • FU is added for Stage I patients }

Refinements: { • R0: ChemoInduction • C is already provided in R0, so C CRT is refined to CRT. }



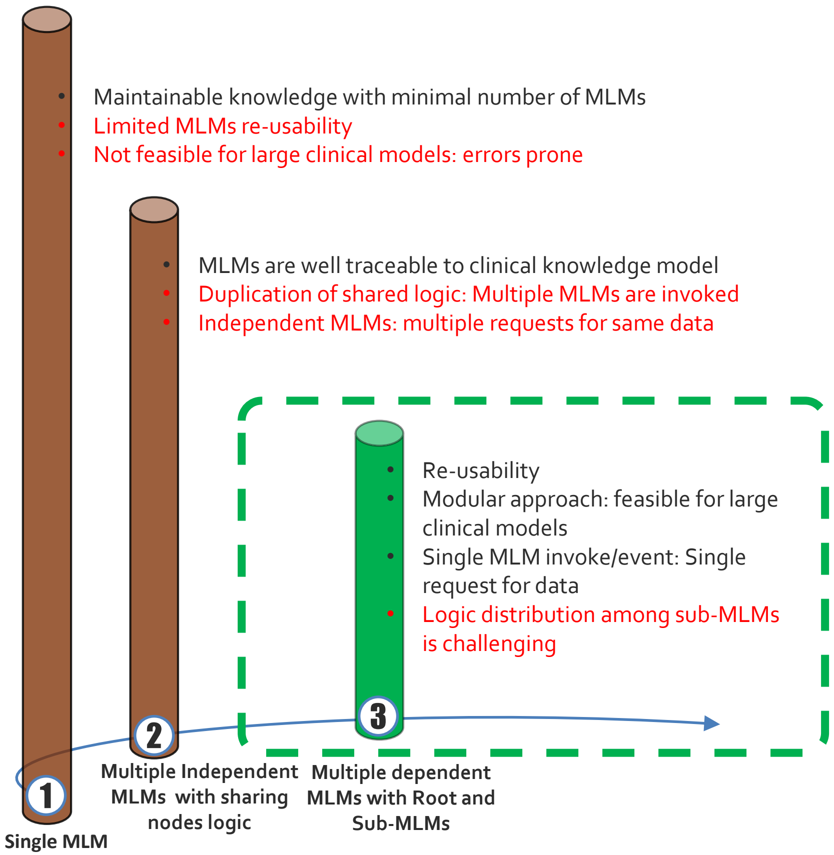
Solution-1:
R-CKM as Sharable Knowledge: MLMs (Phase-III)

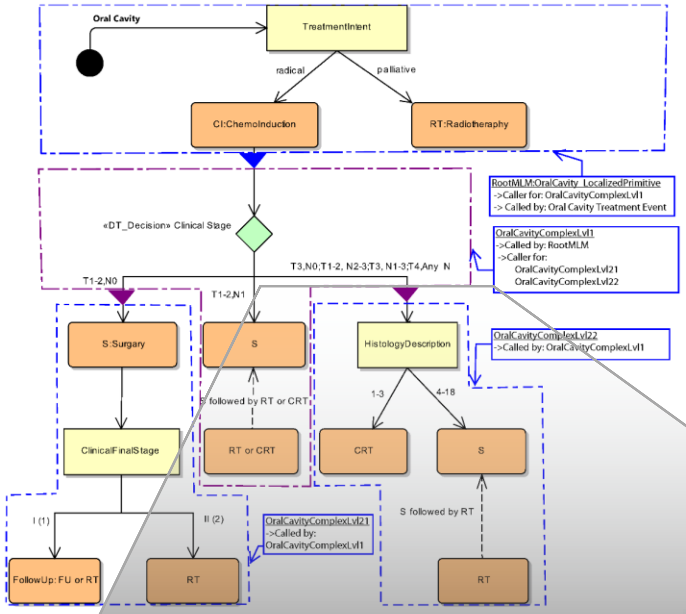
Why Arden Syntax?

- ***R-CKM is only knowledge representation – so called CPG (clinical practice guideline)***
- ***So Computer Interpretable Guideline(CIG) representation is required.***
- ***Arden syntax is HL7 standard and commercially used CIG scheme***

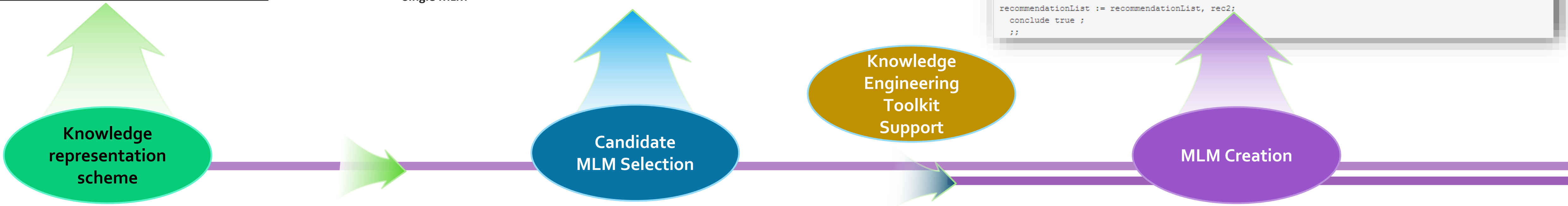
HL7 Arden Syntax

- ✓ Medical Logic Module (MLM)
- ✓ Vocabulary (SNOMED)
- ✓ Data Model (HL7 vMR)






```
IF( (ProcedureEvent1.procedureCode = "395077000" And ProcedureEvent1.procedureMethod = "27762005")  AND(
(ObservationResult1.observationFocus = "385356007" And ObservationResult1.observationValue = "23351008")  OR
(ObservationResult2.observationFocus = "385356007" And ObservationResult2.observationValue = "67673008"))  )
THEN


recPart1 := new ProcedureEvent with "387713003";
rec1 := new Recommendation1 with recPart1;
recommendationList := recommendationList, rec1;
recPart2 := new ProcedureEvent with "108290001";
rec2 := new Recommendation2 with recPart2;
recommendationList := recommendationList, rec2;
  conclude true ;
  ;;
```



Results:

Comparison with Data driven approach

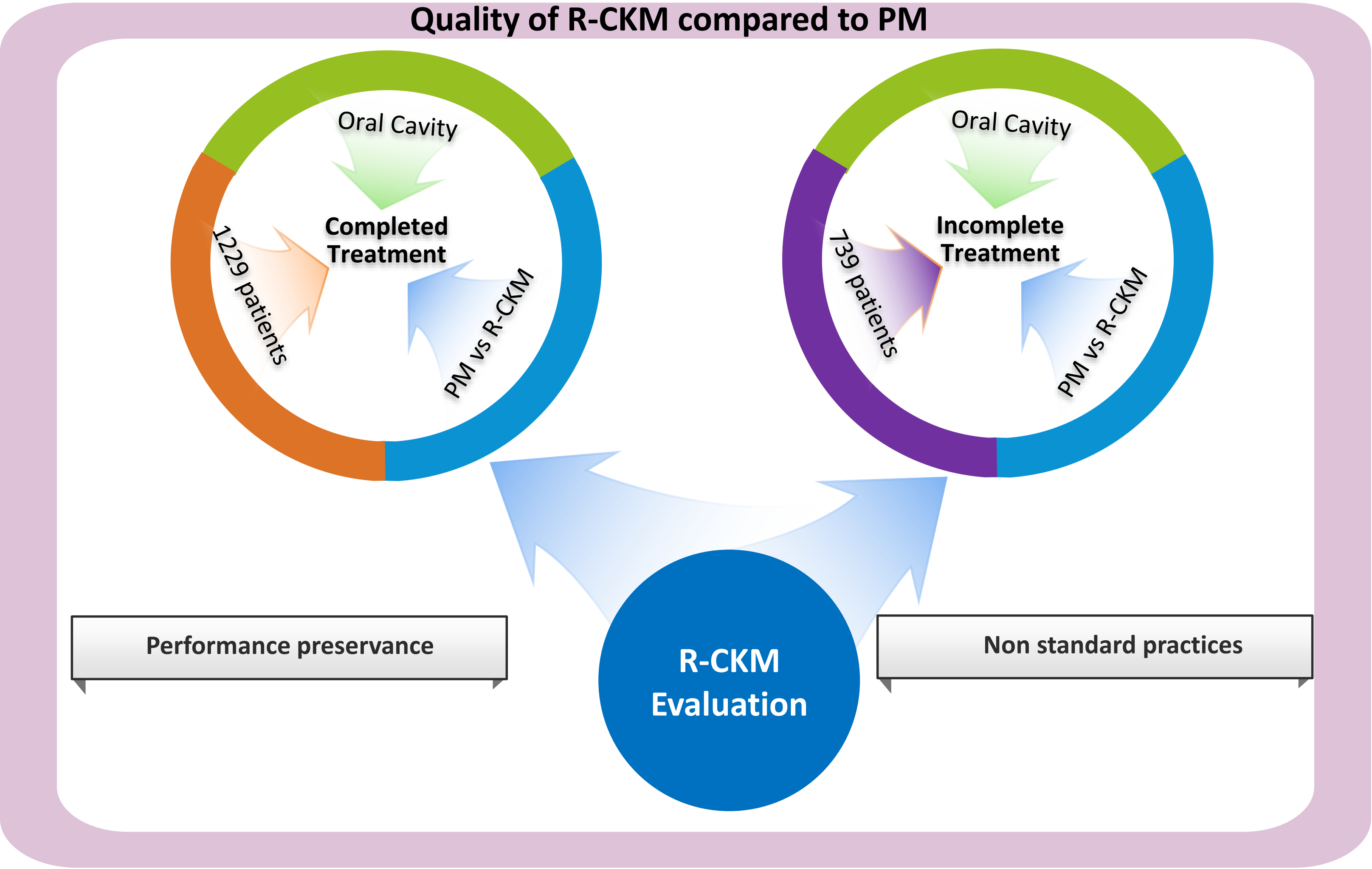
Medical Data
***SKMCH**

Patient with completed treatments

- ✓ Cancer site: Oral cavity
- ✓ Patients: 1229

Patient with incomplete treatments

- ✓ Cancer site: Oral cavity
- ✓ Patients: 739



*SKMCH: Shukat Khanum Memorial Cancer Hospital, Lahore, Pakistan

Maqbool Hussain, Muhammad Afzal, Taqdir Ali, Rahman Ali, Wajahat Ali Khan, Arif Jamshed, Sungyoung Lee, Byeong Ho Kang, Khalid Latif; "Data-driven knowledge acquisition, validation, and transformation into HL7 Arden Syntax", Artificial Intelligence in Medicine (Elsevier, SCI, IF:2.03), published (Online Oct, 2015)

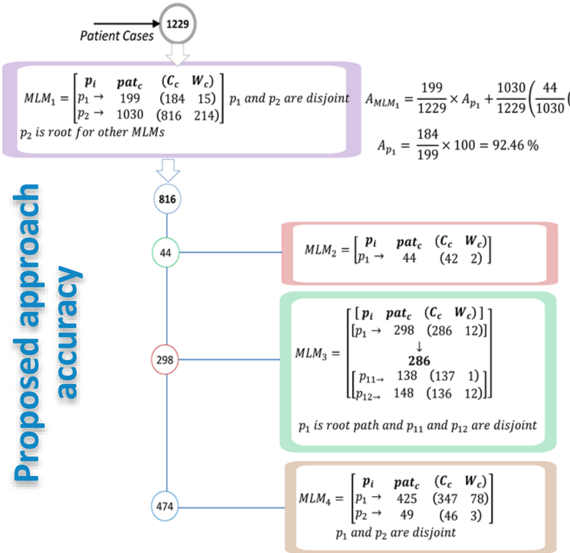
Results & Evaluations:

Comparison with Data driven: [Performance preserve and standard knowledge]

PERFORMANCE PRESERVANCE

Completed Treatment(SKMCH*): Oral Cavity patients (1229)

Proposed approach accuracy



$$A_{MLM_1} = \frac{199}{1229} \times A_{p_1} + \frac{1030}{1229}\left(\frac{44}{1030}(A_{MLM_2}) + \frac{298}{1030}(A_{MLM_3}) + \frac{474}{1030}(A_{MLM_4})\right) = 72.57\,\%$$

$$A_{p_1} = \frac{184}{199} \times 100 = 92.46\,\%$$

$$A_{MLM_2} = \frac{42}{44} \times 100 = 95.45\,\%$$

$$A_{MLM_3} = \frac{(136+137)}{298} \times 100 = 91.61\,\%$$

$$A_{MLM_4} = \frac{(347+46)}{474} \times 100 = 82.91\%$$

Comparison

- *Accuracy perseverance*

Remarks

- *R-CKM preserves the accuracy of PM*
- *At same time, it is standardized*

R-CKM vs Existing model

- ✓ R-CKM: 72.57%
- ✓ Existing: 71.0%

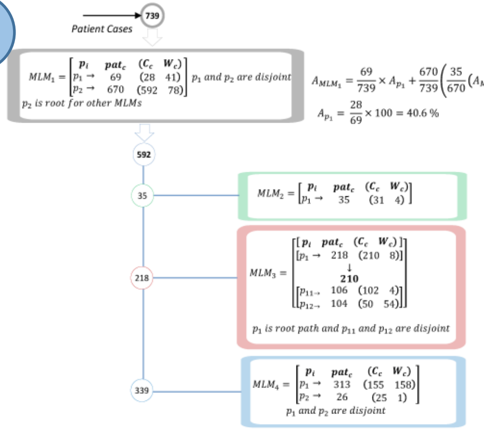
Data-driven approach Accuracy [11]

Observed	Predicted				
	C CRT	RT	S RT	Total Cases	Accuracy
C CRT	343	12	46	401	85.5%
RT	87	184	127	398	46.2%
S RT	81	3	346	430	80.5%
Total	511	199	519	1229	
Overall Percentage	41.6%	16.2%	42.2%		71.0%

NON-STANDARD PRACTICES

Incomplete Treatment(SKMCH*): Oral Cavity patients (739)

Proposed approach accuracy



$$A_{MLM_1} = \frac{69}{739} \times A_{p_1} + \frac{670}{739}\left(\frac{35}{670}(A_{MLM_2}) + \frac{218}{670}(A_{MLM_3}) + \frac{339}{670}(A_{MLM_4})\right) = 53\,\%$$

$$A_{p_1} = \frac{28}{69} \times 100 = 40.6\,\%$$

$$A_{MLM_2} = \frac{31}{35} \times 100 = 88.6\,\%$$

$$A_{MLM_3} = \frac{(102+50)}{218} \times 100 = 69.7\,\%$$

$$A_{MLM_4} = \frac{(155+25)}{339} \times 100 = 53.1\,\%$$

Comparison

- *Non-guideline based treatment*

Remarks

- *R-CKM is guideline enabled, so incomplete treatment plans (non-standard) are discouraged*

R-CKM vs Existing model

- ✓ R-CKM: 53%
- ✓ Existing: 59%

Data-driven approach Accuracy [11]

Path#	PM Decision Path	Candidate Patient Cases	PM Path Accuracy
Path-1	Node 0->RT	Palliative patients: **69**	40.58% (C:28 W:41)
Path-2	Node 0 -> Node 1 -> Node 3 -> S RT	Patient with radical and CS: T1: **139**	95.68%(C:133, W:6)
Path-3	Node 0 -> Node 1 -> Node 4 -> Node 6 -> S RT	Patients with radical and FCS II: **123**	73.98%(C:91, W:32)
Path-4	Node 0 -> Node 1 -> Node 4 -> Node 7 -> S RT	Patients with radical and FCS III;IV: **56**	67.86%(C:38, W:18)
Path-5	Node 0 -> Node 1 -> Node 5 -> Node 8 -> C CRT	Patients with radical, CS:T3-4 and HistoDescription 1,2,3: **324**	38.58% (C:125, W:199)
Path-6	Node 0 -> Node 1 -> Node 5 -> Node 9 -> S RT	Patients with radical, CS:T3-4 and HistoDescription other than 1,2,3: **28**	85.71%(C:24, W:4)
Overall PM Accuracy: PM_{acc}			**59.0%**

- CS: Clinical Stage
- FCS: Final Clinical Stage
- HistoDescription 1: Squamous cell carcinoma
- HistoDescription 2: Small cell carcinoma
- HistoDescription 3: Carcinoma NOS
- C: Correctly classified patient cases
- W: Wrongly classified patient cases

*SKMCH: Shukat Khanum Memorial Cancer Hospital, Lahore, Pakistan

Results & Evaluations: [Approach-I]

Comparison with Combined approach [Non-validation VS Validation]

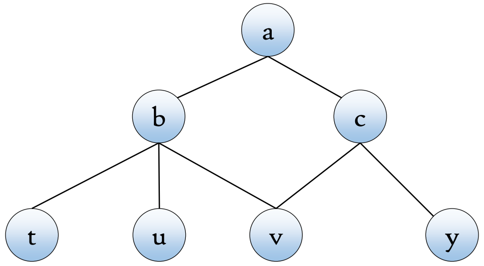
Approach-1

CKM attribute

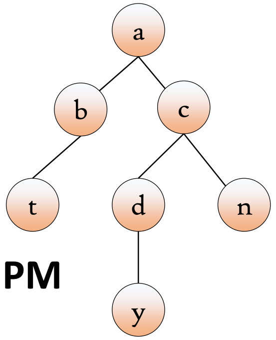
a, b, c, t, u, v, y

Patient data attribute

a, b, c, t, d, n, y, u

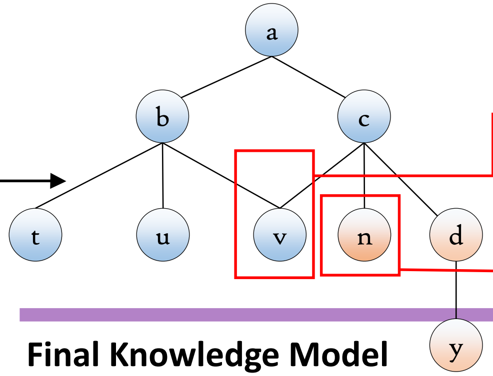


CKM



PM

- ❖ Investigate decision path in PM which is not in CKM
- ❖ Extend CKM for missing decision path



Final Knowledge Model

Limitations of existing approach

- ❑ No validation to extension of path in CKM
- ❑ Final Knowledge Model is not integrable

- ▪ **"v" is not added to R-CKM while only "u" is added as refinements.**
- ▪ **Presence of "v" make R-CKM non-integrable as** ***v is not existed in patietn data***

❖ **"n" is not properly validated against CKM, so in R-CKM it is removed**

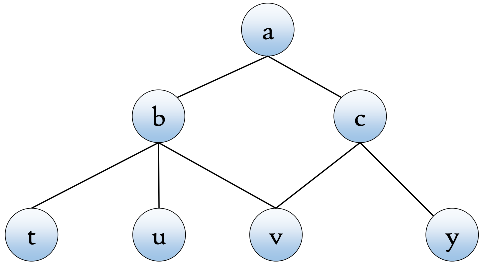
Proposed

CKM attribute

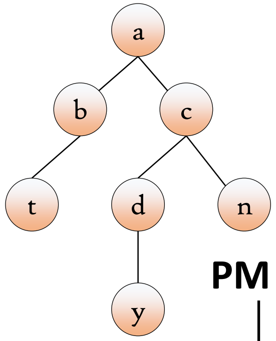
a, b, c, t, u, v, y

Patient data attribute

a, b, c, t, d, n, y, u

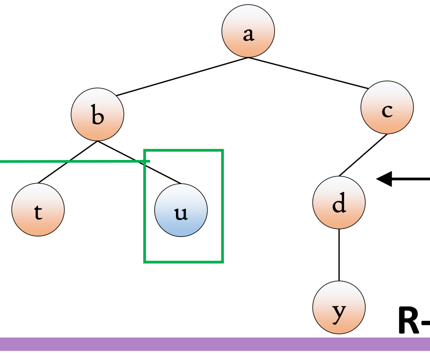


CKM



PM

- ❖ Investigate each decision path in PM and apply conformance
- ❖ Extend PM to R-CKM after conformance and refinements

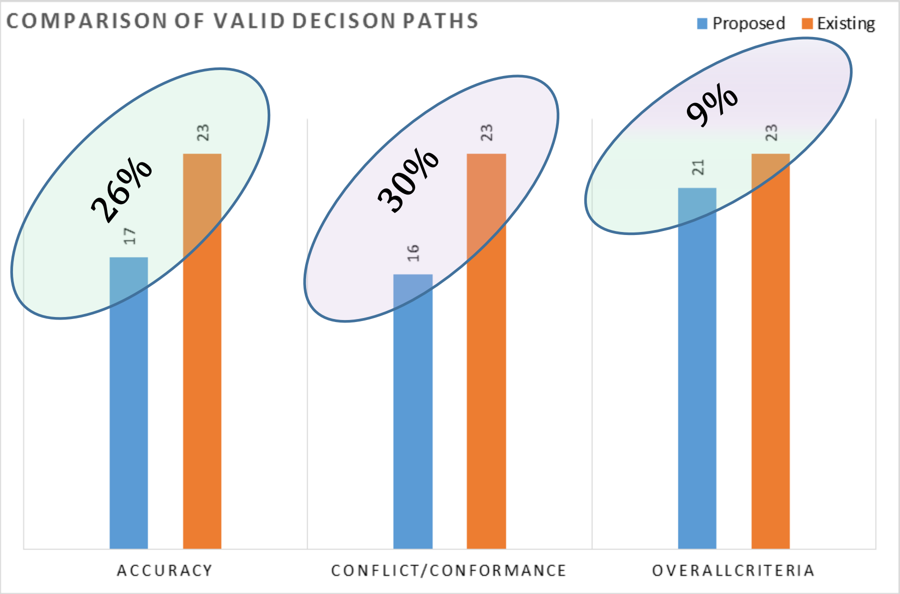


R-CKM

Results & Evaluations: [Approach-I]

Comparison with Combine approach [Non-validation vs Validation]

Algorithm: C4.5 (with accuracy 69.7%)
Dataset: 1229 (H&N cancer dataset of SKMCH)
Guideline: NCCN
Total decision paths: 23



Six (6) decision paths have lower accuracy than the targeted

Seven (7) decision paths are not conformed

Two(2) decision paths are not conformed and having lower accuracy

```
C4.5 pruned tree
------------------

TreatmentIntentDescription = Radical
|   ClinicalStageS = I (1)
|   |   GrdDescription = NA: RT (26.0/7.0)
|   |   GrdDescription != NA: S RT (95.0/17.0)
|   ClinicalStageS != I (1)
|   |   ClinicalStageS = II (2): S RT (132.0/40.0)
|   |   ClinicalStageS != II (2)
|   |   |   HistologyDescription = Adenoid cystic carcinoma: S RT (15.0/7.0)
|   |   |   HistologyDescription != Adenoid cystic carcinoma
|   |   |   |   ClinicalStageT = 2: S RT (44.0/18.0)
|   |   |   |   ClinicalStageT != 2
|   |   |   |   |   HistologyDescription = Mucoepidermoid carcinoma: S RT (5.0/1.0)
|   |   |   |   |   HistologyDescription != Mucoepidermoid carcinoma
|   |   |   |   |   |   HistologyDescription = NA: S RT (4.0/1.0)
|   |   |   |   |   |   HistologyDescription != NA
|   |   |   |   |   |   |   HistologyDescription = Adenocarcinoma: RT (3.0/1.0)
|   |   |   |   |   |   |   HistologyDescription != Adenocarcinoma
Accuracy: 50%  |   |   |   |   |   |   |   |   HistologyDescription = Malignant melanoma: RT (4.0/2.0)
|   |   |   |   |   |   |   |   HistologyDescription != Malignant melanoma
|   |   |   |   |   |   |   |   |   ClinicalStageT = 1: S RT (11.0/4.0)
|   |   |   |   |   |   |   |   |   ClinicalStageT != 1
|   |   |   |   |   |   |   |   |   |   HistologyDescription = Squamous cell carcinoma
Conformance Issue  |   |   |   |   |   |   |   |   |   |   |   Alcohol = Yes: C CRT (6.0)
|   |   |   |   |   |   |   |   |   |   |   Alcohol != Yes
Conformance Issue and Accuracy: 47%  |   |   |   |   |   |   |   |   |   |   |   |   GrdDescription = NA: C CRT (19.0/10.0)
|   |   |   |   |   |   |   |   |   |   |   |   GrdDescription != NA
Conformance Issue  |   |   |   |   |   |   |   |   |   |   |   |   |   TreatStatusDesc = Incomplete: C CRT (29.0/4.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   TreatStatusDesc != Incomplete
|   |   |   |   |   |   |   |   |   |   |   |   |   |   ClinicalStageN = 0
Conformance Issue and Accuracy: 45%  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   GrdDescription = Poor: C CRT (11.0/6.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   GrdDescription != Poor
Conformance Issue  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   PatStatusDesc = On treatment: C CRT
(14.0/3.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   PatStatusDesc != On treatment
Conformance Issue  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   TreatStatusDesc = Complete: C CRT
(128.0/51.0)
Accuracy: 40%  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   TreatStatusDesc != Complete: RT
(10.0/6.0)
Conformance Issue  |   |   |   |   |   |   |   |   |   |   |   |   |   |   ClinicalStageN != 0: C CRT (120.0/29.0)
|   |   |   |   |   |   |   |   |   |   HistologyDescription != Squamous cell carcinoma
|   |   |   |   |   |   |   |   |   |   |   Smoking = No (None)
|   |   |   |   |   |   |   |   |   |   |   |   PatStatusDesc = Alive with disease: RT (2.0)
|   |   |   |   |   |   |   |   |   |   |   |   PatStatusDesc != Alive with disease
|   |   |   |   |   |   |   |   |   |   |   |   |   PatStatusDesc = Alive: S RT (3.0/1.0)
Accuracy: 50%  |   |   |   |   |   |   |   |   |   |   |   |   |   PatStatusDesc != Alive: RT (2.0/1.0)
Accuracy: 50%  |   |   |   |   |   |   |   |   |   |   |   Smoking != No (None): C CRT (4.0/2.0)
TreatmentIntentDescription != Radical: RT (133.0/10.0)

Number of Leaves  :     23
Size of the tree :  45
```

Results & Evaluations:[Approach-II]

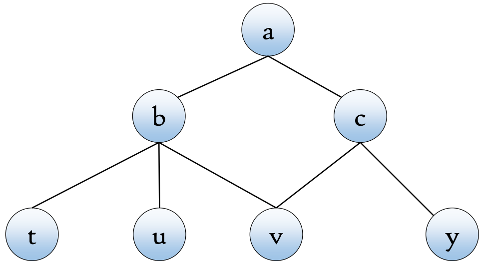
Comparison with Combined approach [Non-validation VS Validation]

CKM_1 attribute

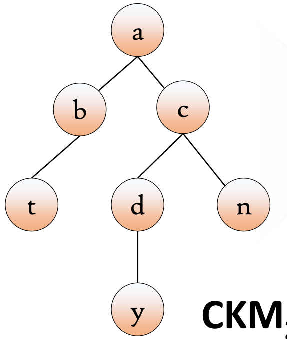
{ a, b, c, t, u, **v**, y }

CKM_2 attribute

{ a, b, c, t, d, n, y, **u** }



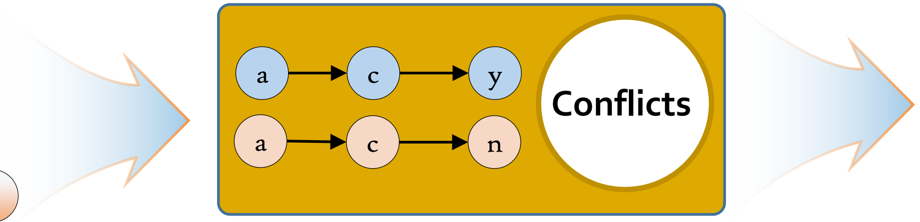
CKM_1

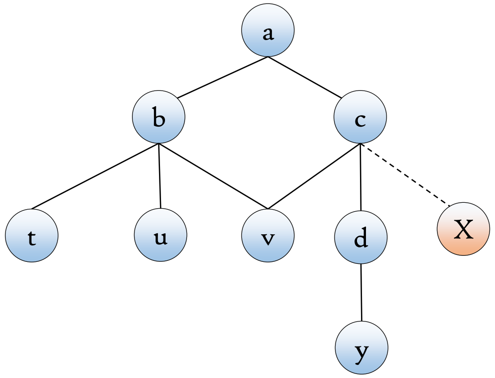


CKM_2

Key Validation Process

- ❖Combine all decision paths (Union)
- ❖Revise (if resolution exist) decision paths with conflicting decision
- ❖Conformance = Conflict





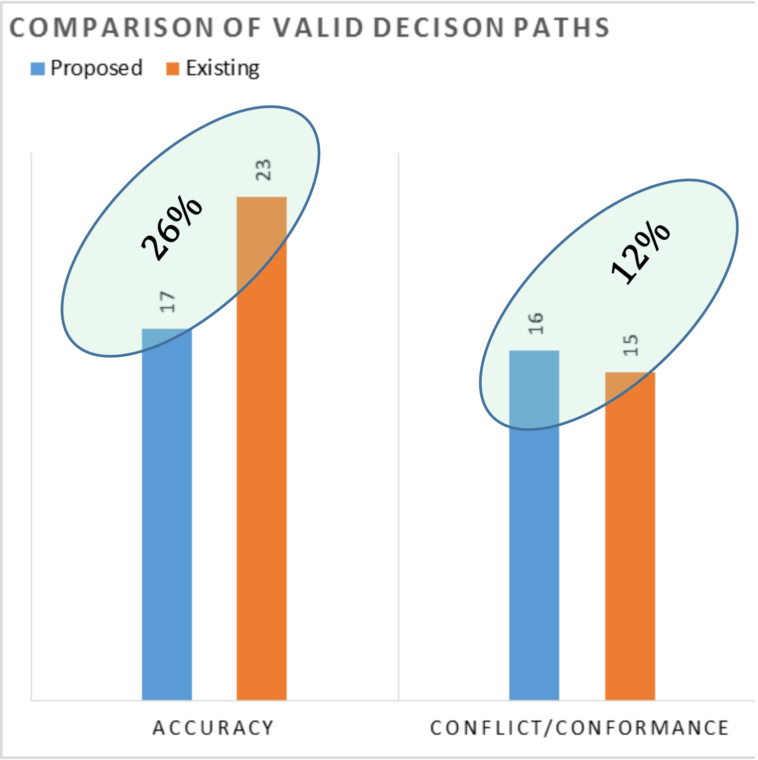
Limitations of existing approach

- ❑ Final Knowledge Model is not integrable
- ❑ Conflict resolution only depends on limited evidence.

Results & Evaluations: [Approach-II]

Comparison with Combine approach [Non-validation vs Validation]

Algorithm: C4.5 (with accuracy 69.7%)
Dataset: 1229 (H&N cancer dataset of SKMCH)
Guideline: NCCN
Total decision paths: 23



6/23 decision paths have lower accuracy than the targeted

1/8 non-conformed paths was having published proved evidence to consider

Conflict Issue: Revised with CRT

```
C4.5 pruned tree
------------------

TreatmentIntentDescription = Radical
|   ClinicalStageS = I (1)
|   |   GrdDescription = NA: RT (26.0/7.0)
|   |   GrdDescription != NA: S RT (95.0/17.0)
|   ClinicalStageS != I (1)
|   |   ClinicalStageS = II (2): S RT (132.0/40.0)
|   |   ClinicalStageS != II (2)
|   |   |   HistologyDescription = Adenoid cystic carcinoma: S RT (15.0/7.0)
|   |   |   HistologyDescription != Adenoid cystic carcinoma
|   |   |   |   ClinicalStageT = 2: S RT (44.0/18.0)
|   |   |   |   ClinicalStageT != 2
|   |   |   |   |   HistologyDescription = Mucoepidermoid carcinoma: S RT (5.0/1.0)
|   |   |   |   |   HistologyDescription != Mucoepidermoid carcinoma
|   |   |   |   |   |   HistologyDescription = NA: S RT (4.0/1.0)
|   |   |   |   |   |   HistologyDescription != NA
|   |   |   |   |   |   |   HistologyDescription = Adenocarcinoma: RT (3.0/1.0)
|   |   |   |   |   |   |   HistologyDescription != Adenocarcinoma
|   |   |   |   |   |   |   |   HistologyDescription = Malignant melanoma: RT (4.0/2.0)
|   |   |   |   |   |   |   |   HistologyDescription != Malignant melanoma
|   |   |   |   |   |   |   |   |   ClinicalStageT = 1: S RT (11.0/4.0)
|   |   |   |   |   |   |   |   |   ClinicalStageT != 1
|   |   |   |   |   |   |   |   |   |   HistologyDescription = Squamous cell carcinoma
|   |   |   |   |   |   |   |   |   |   |   Alcohol = Yes: C CRT (6.0)
|   |   |   |   |   |   |   |   |   |   |   Alcohol != Yes
|   |   |   |   |   |   |   |   |   |   |   |   GrdDescription = NA: C CRT (19.0/10.0)
|   |   |   |   |   |   |   |   |   |   |   |   GrdDescription != NA
|   |   |   |   |   |   |   |   |   |   |   |   |   TreatStatusDesc = Incomplete: C CRT (29.0/4.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   TreatStatusDesc != Incomplete
|   |   |   |   |   |   |   |   |   |   |   |   |   |   ClinicalStageN = 0
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   GrdDescription = Poor: C CRT (11.0/6.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   GrdDescription != Poor
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   PatStatusDesc = On treatment: C CRT
(14.0/3.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   PatStatusDesc != On treatment
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   TreatStatusDesc = Complete: C CRT
(128.0/51.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   TreatStatusDesc != Complete: RT
(10.0/6.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   ClinicalStageN != 0: C CRT (120.0/29.0)
|   |   |   |   |   |   |   |   |   |   HistologyDescription != Squamous cell carcinoma
|   |   |   |   |   |   |   |   |   |   |   Smoking = No (None)
|   |   |   |   |   |   |   |   |   |   |   |   PatStatusDesc = Alive with disease: RT (2.0)
|   |   |   |   |   |   |   |   |   |   |   |   PatStatusDesc != Alive with disease
|   |   |   |   |   |   |   |   |   |   |   |   |   PatStatusDesc = Alive: S RT (3.0/1.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   PatStatusDesc != Alive: RT (2.0/1.0)
|   |   |   |   |   |   |   |   |   |   |   Smoking != No (None): C CRT (4.0/2.0)
TreatmentIntentDescription != Radical: RT (133.0/10.0)

Number of Leaves  :     23
Size of the tree :  45
```

Contributions of Solution-1

Solution-1

Objectives

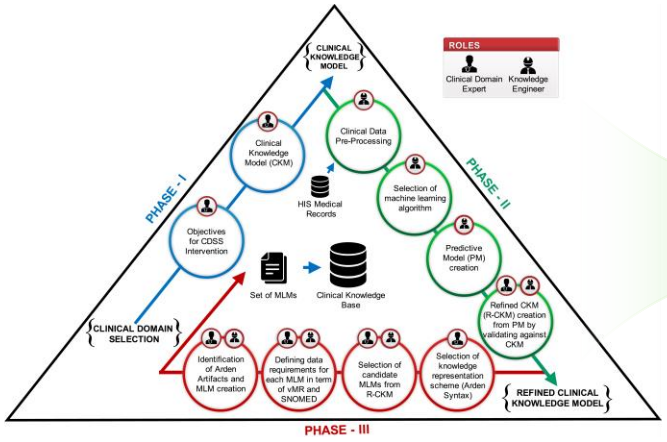
- ❑ Guideline base validation
- ❑ PM from patient data
- ❑ Refined sharable clinical knowledge

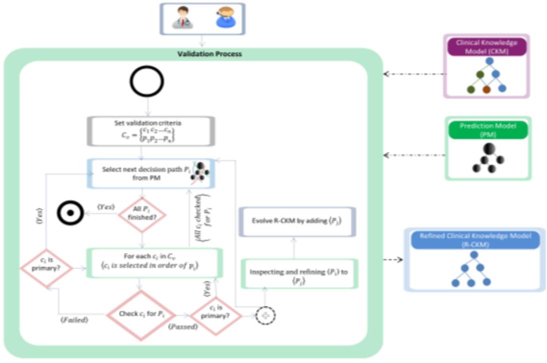
Target Limitations

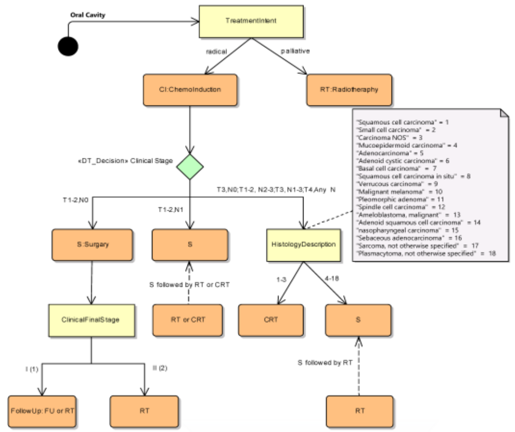
- ❖ Guideline support
- ❖ Validated knowledge

Outcome

✓ Guideline enabled clinical knowledge model – R-CKM







How to verify that:
Validation process is consistent to produce consistent knowledge model

Contributions

- ✓ Validation process for clinical knowledge model
- ✓ Guideline enabled knowledge model – R-CKM

Solution-2: Process model

Solution-2

Objectives

- ❑ Formal modeling
- ❑ Formal verification of validation process and R-CKM

Target Limitations

- ❖ Inconsistency in knowledge and validation process

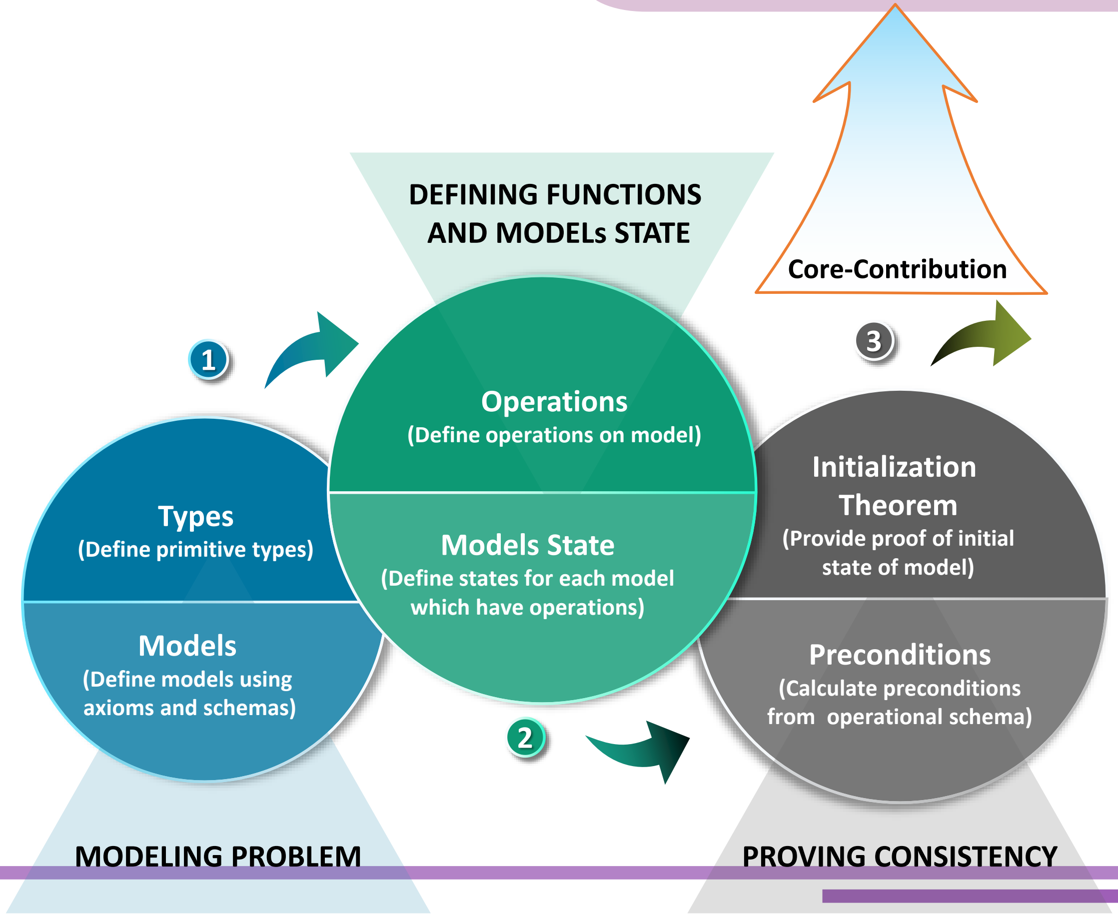
Outcome

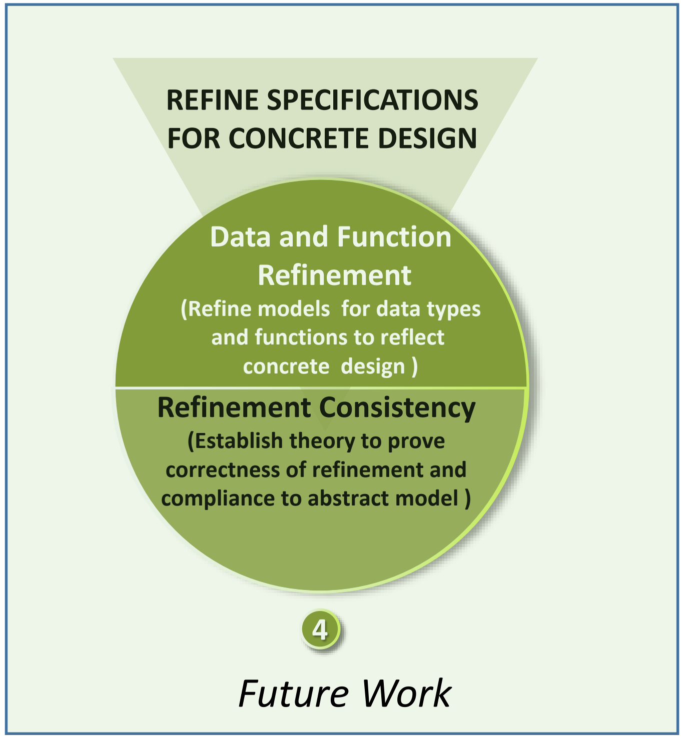
- ✓ Proving consistency of validation processes and R-CKM by eliminating inconsistencies using formal method

Z Formal Notations

Why Z for KB verifications?

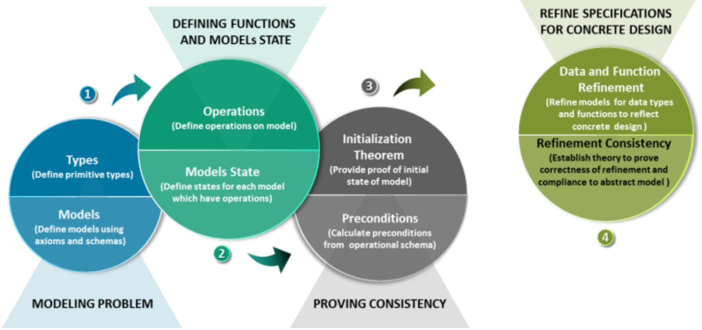
- ✓ Easy Knowledge Modeling
  - ❖ Knowledge decomposition
  - ❖ Static feature modeling
  - ❖ Dynamic feature modeling
- ✓ Data-rich formalism
  - ❖ Types: primitive contents
  - ❖ Free Types: Complex semantics
  - ❖ Axioms and Schema: Complex semantic with constraints
- ✓ Tool support
  - ✓ Concrete design
  - ✓ Auto code generation



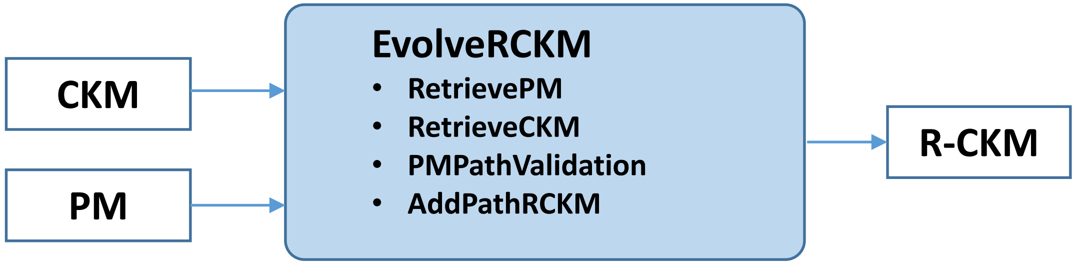


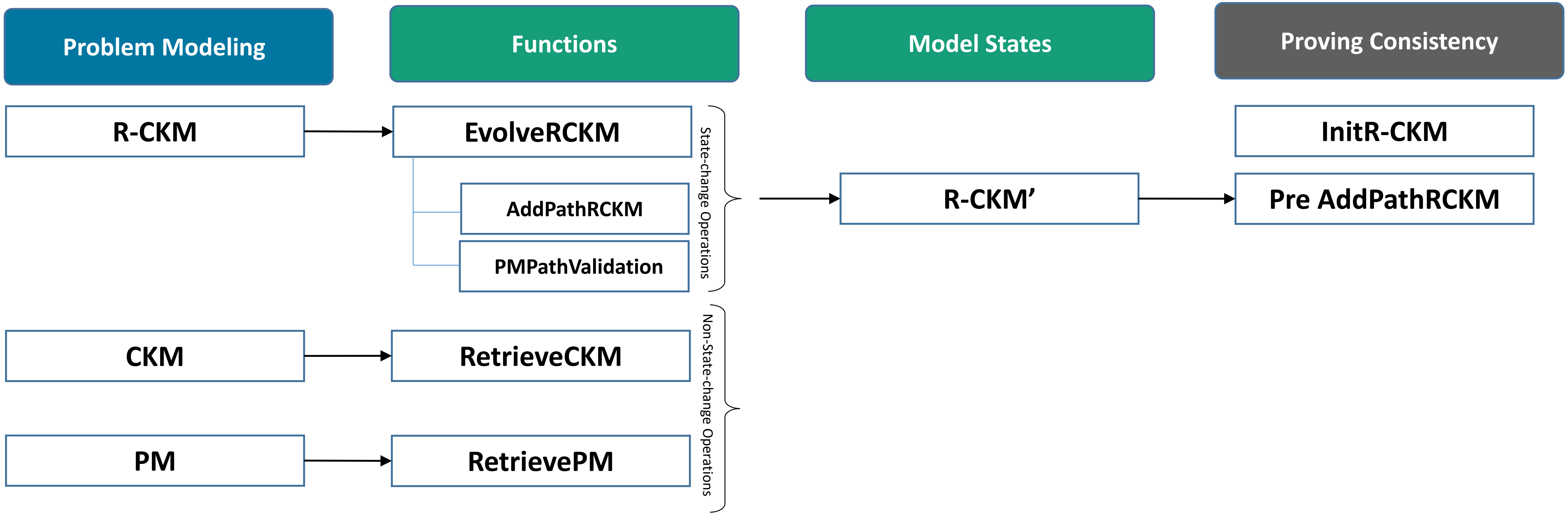
Solution-2: Process model

Formal Verification of Validation process

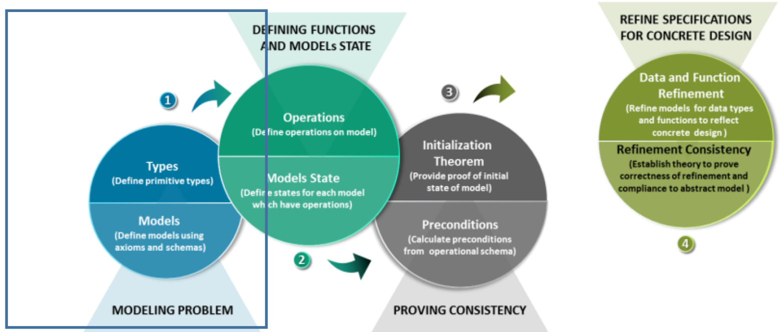


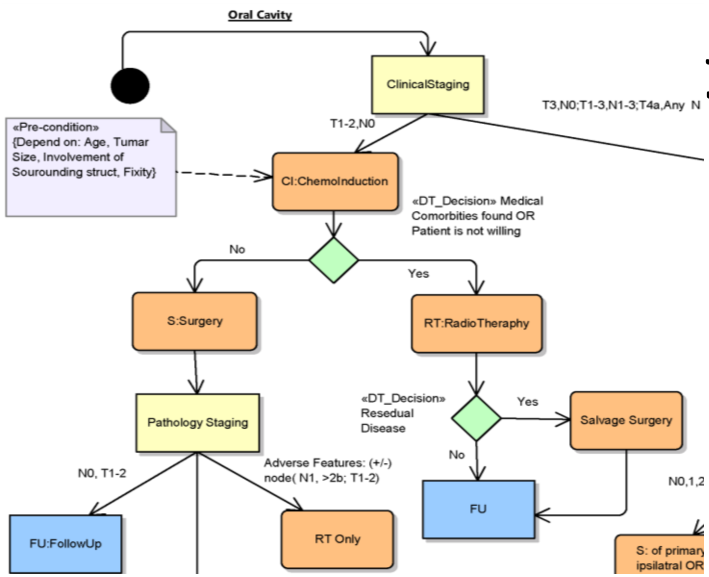
Validation Process





Solution-2:
Modeling problem: Models (CKM)





Axiom

Declarations

$$decisionPathConditionCKM : \mathbb{F}\ ConditionKMs$$
$$ConclusionCKM : \mathbb{F}\ TreatmentPlan$$
$$decisionPathCKM : ConditionKMs \nrightarrow TreatmentPlan$$

Predicates

$$decisionPathConditionCKM = \mathrm{dom}\ decisionPathCKM$$
$$ConclusionCKM = \mathrm{ran}\ decisionPathCKM$$
$$(\mathrm{ran}\ ConclusionCKM \cap \mathrm{ran}(\mathrm{dom}\ decisionPathCKM)) \subset \mathrm{ran}\ decisionPathConditionCKM$$
$$head\ (\mathrm{dom}\ decisionPathCKM) \notin \mathrm{ran}\ ConclusionCKM \cap \mathrm{ran}\ decisionPathConditionCKM$$

Condition part of one decision-path can include conclusion as condition

Root condition of decision-path must be condition

Schema

$ClinicalKnowledgeModel$

Declarations

$$CKM : \mathbb{P}\ decisionPathCKM$$
$$guidelines : decisionPathCKM \rightarrow ClinicalObjectives$$
$$rootCKM : \mathrm{seq}\ Condition$$

Predicates

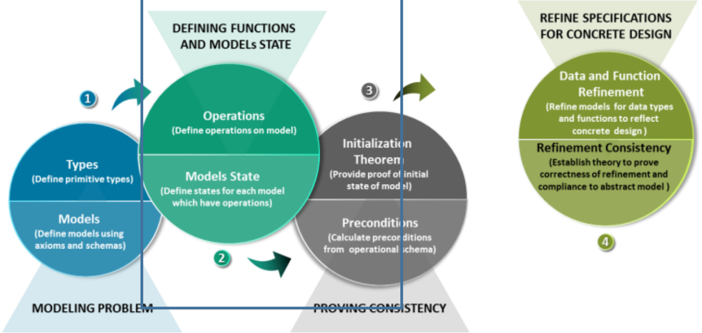
$$CKM = \mathrm{dom}\ guidelines$$
$$\forall\ dp : decisionPathCKM \mid dp \in CKM \bullet$$
$$head\ (\mathrm{dom}\ dp) \notin \mathrm{ran}\ ConclusionCKM \cap \mathrm{ran}\ decisionPathConditionCKM$$
$$\exists\ dp : decisionPathCKM\ ;\ dp_1 : decisionPathCKM \mid dp, dp_1 \in CKM \bullet$$
$$last\ (\mathrm{dom}\ dp) = \mathrm{ran}\ dp_1 \Leftrightarrow \mathrm{dom}\ dp_1 = \mathrm{dom}\ dp \setminus last\ (\mathrm{dom}\ dp)$$
$$rootCKM = \forall\ dp : decisionPathCKM \mid dp \in CKM \bullet \langle \bigcap(\mathrm{ran}(\mathrm{dom}\ dp)) \rangle$$

All decision paths must have no conclusion in first condition

Conclusion of one decision-path can be condition for another decision path.

Root of CKM

Solution-2:
Functions and Model States: (RCKM)



RCKM evolution function:

- *Pass validation criteria for each PM path*
- *Add to RCKM path after refinements*

$$EvolveRCKM \mathrel{\widehat=} PMPathValidation \land AddPathRCKM$$

Input:

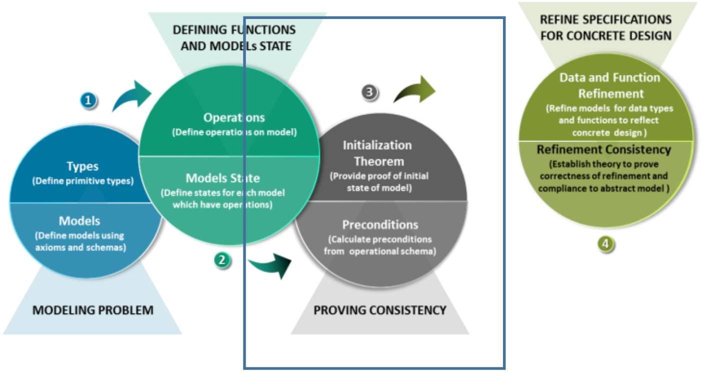
- *Decision-path of PM*
- *Qualified accuracy for decision-path by domain expert*

Output:

- *True: passed / False: Failed*

$$
\begin{array}{l}
\textit{PMPathValidation} \\
\hline
\textit{RefinedClinicalKnowledgeModel} \\
dp_{pm}? : decisionPath \\
qualifiedAcc? : \mathbb{Z} \\
\hline
\text{Criteria-1:}\quad dp_{pm}? \in PM \land decisionPathAccuracy(dp_{pm}?) \geq qualifiedAcc? \\
\text{Criteria-2:}\quad \forall\, t_1, t_2 : treatmentSet \mid t_1, t_2 \in \mathrm{ran}(\mathrm{ran}(dp_{pm}?)) \land TreatmentPlan^{\sim}(t_1) > TreatmentPlan^{\sim}(t_2) \bullet \\
\qquad \exists\, dp_{ckm} : decisionPathCKM;\ t_3, t_4 : treatmentSet \mid dp_{ckm} \in CKM, \\
\qquad t_3, t_4 \in (\mathrm{ran}(\mathrm{dom}(dp_{ckm})) \cap \mathrm{ran}(ConclusionCKM)) \cup \mathrm{ran}(\mathrm{ran}(dp_{ckm})) \bullet \\
\qquad (t_3 = t_1 \land t_4 = t_2) \Rightarrow TreatmentPlan^{\sim}(t_3) > TreatmentPlan^{\sim}(t_4) \\
\text{Criteria-3,4:}\quad decPathEvidences(dp_{pm}?) \neq \emptyset \lor \\
\qquad \exists\, dp_{ckm} : decisionPathCKM \mid dp_{ckm} \in CKM \bullet \\
\qquad\qquad (\mathrm{ran}(\mathrm{dom}(dp_{pm}?)) \subseteq \mathrm{ran}(\mathrm{dom}(dp_{ckm})) \Rightarrow \\
\qquad\qquad \mathrm{ran}(\mathrm{ran}(dp_{pm}?)) \subseteq \\
\qquad\qquad (\mathrm{ran}(\mathrm{dom}(dp_{ckm})) \cap \mathrm{ran}(ConclusionCKM)) \cup \mathrm{ran}(\mathrm{ran}(dp_{ckm})))
\end{array}
$$

Solution-2:
Proving Consistency: (Initialization and Precondition Theorem)



$$\exists\ RefinedClinicalKnowledgeModel' \bullet InitRCKM$$

Initialization Theorem states:

There exist initial state of schema for which the model must be consistent

$InitRCKM$

$RefinedClinicalKnowledgeModel'$

$$
\begin{aligned}
&accuracyRCKM' = 0\\
&RCKM' = \emptyset\\
&refinedCKM' = \emptyset\\
&rootRCKM' = \emptyset\\
&refinedCKMsAccuracy' = \emptyset\\
&decisionPathRCKM' = \emptyset\\
&decisionPathConditionRCKM' = \emptyset\\
&ConclusionRCKM' = \emptyset\\
&decPathRCKMAccuracy' = \emptyset\\
&evidences' = \emptyset\\
&decPathRCKMEvidences' = \emptyset
\end{aligned}
$$

$$\mathrm{pre}\ AddPathRCKM = \exists\ RefinedClinicalKnowledgeModel' \bullet AddPathRCKM$$

Preconditions Theorem:

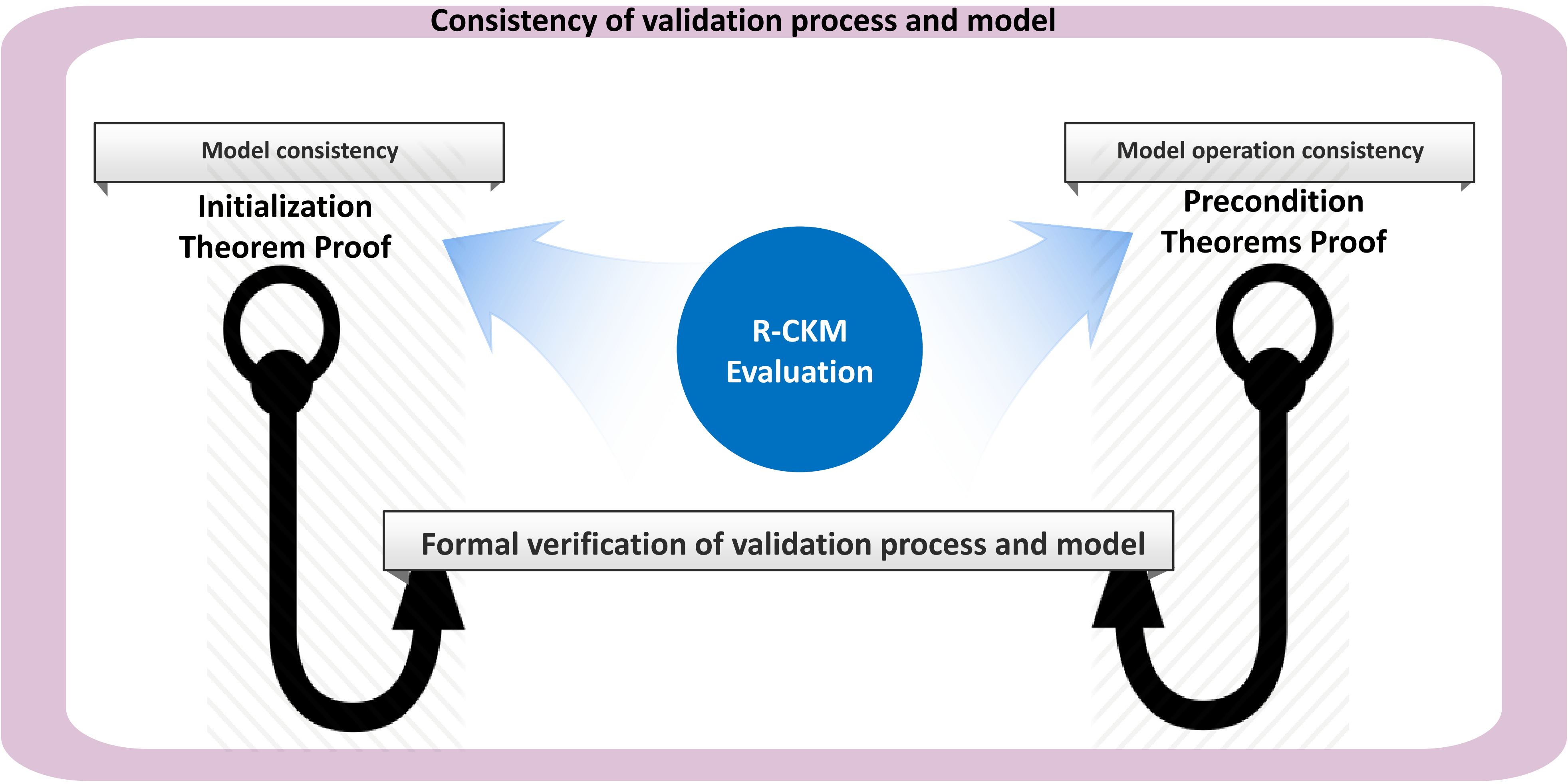
- *Proofs for simplifying AddPathRCKM operation for pre-condition*

$\mathrm{pre}\ AddPathRCKM$

$$
\begin{aligned}
&RefinedClinicalKnowledgeModel\\
&dp_{pm}? : decisionPath\\
&qualifiedAcc? : \mathbb{Z}\\
&refinements? : \mathbb{F}\ RefinedTreatmentPlan
\end{aligned}
$$

$$\exists\ RefinedClinicalKnowledgeModel';\ rckmPath! : decisionPathRCKM \bullet AddPathRCKM$$

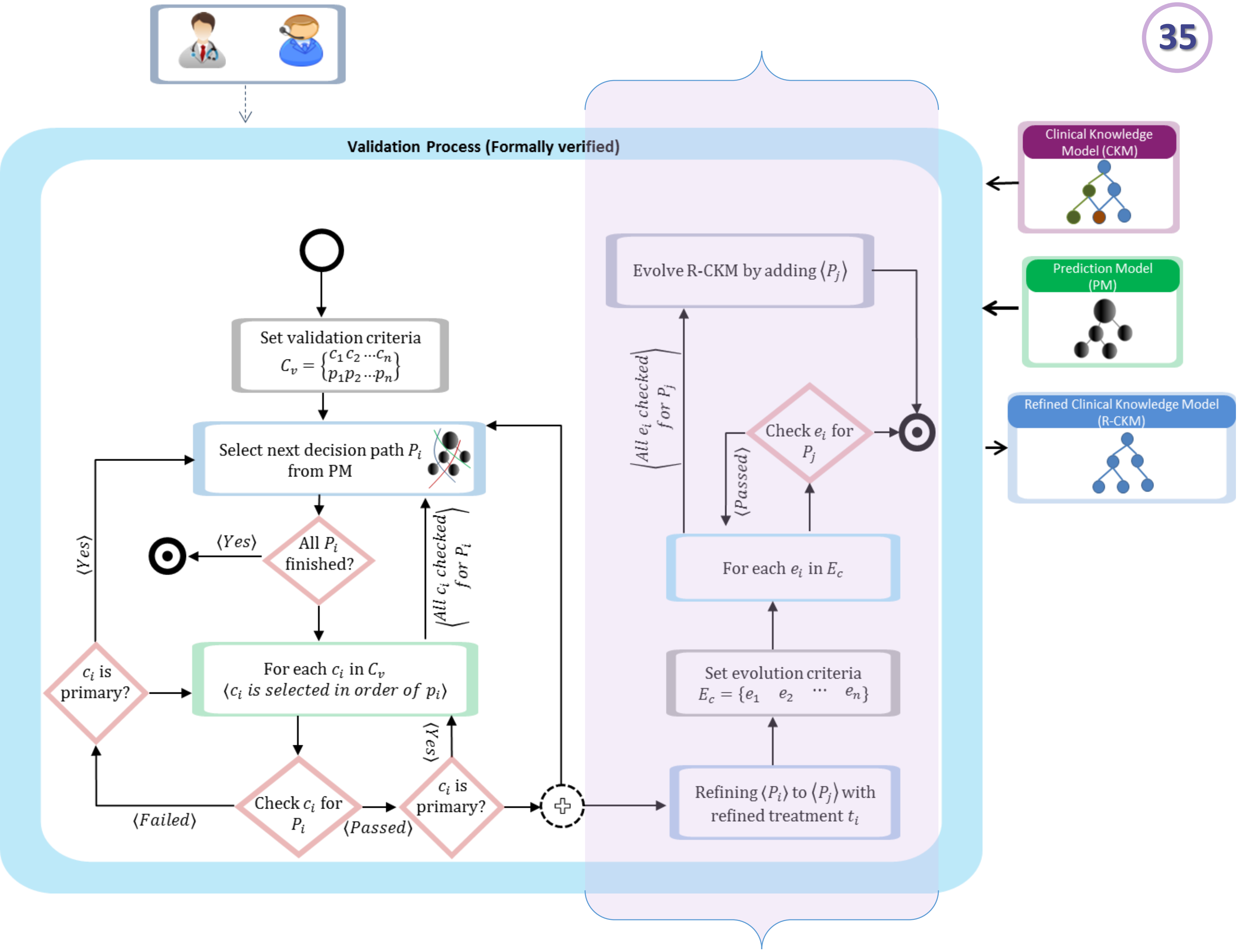
Results


Consistency of validation process and model
Model consistency
Initialization
Theorem Proof
Model operation consistency
Precondition
Theorems Proof
R-CKM
Evaluation
Formal verification of validation process and model

Results:
Formally verified Validation Process

Additional criteria of validation Process

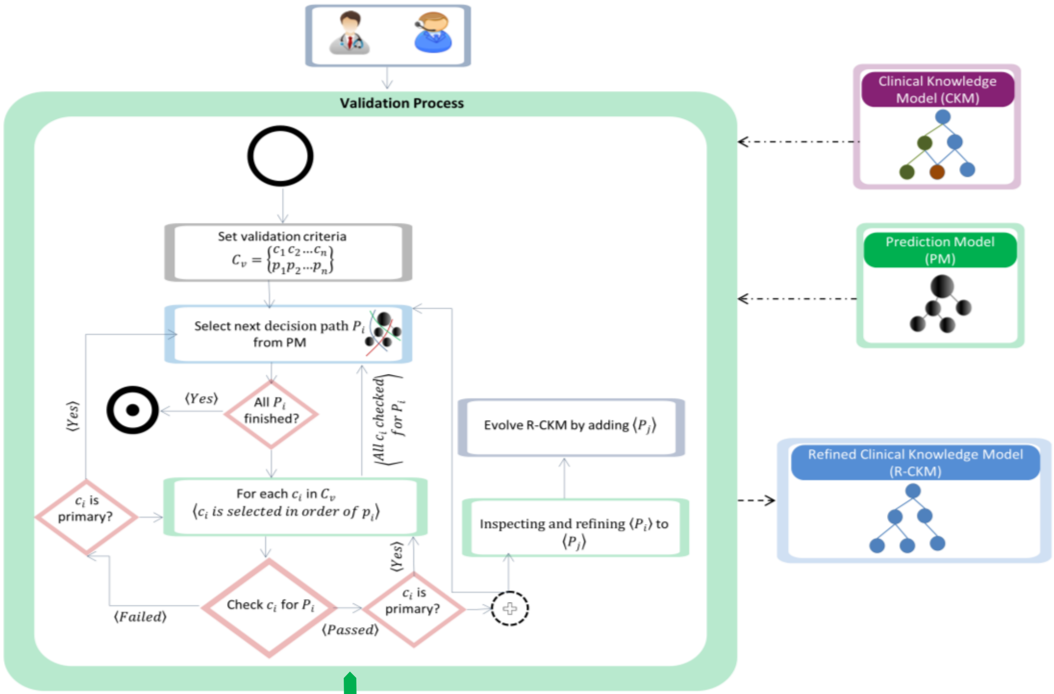
$$RCKM \neq \emptyset \Rightarrow head\,(\mathrm{dom}\; dp_{pm}?) = rootRCKM$$
$$\forall\, pos : \mathbb{N} \mid pos \in \mathrm{dom}\; refinements? \bullet pos > 1 \land$$
$$\quad pos \leq (\#(\mathrm{dom}\; dp_{pm}?) + \#(\mathrm{ran}\; dp_{pm}?))$$
$$\mathrm{ran}(\mathrm{dom}\; rckmPath!) \subset \mathrm{ran}\; decisionPathConditionRCKM$$
$$\mathrm{ran}(\mathrm{ran}\; rckmPath!) \subset \mathrm{ran}\; ConclusionRCKM$$
$$(\mathrm{ran}(\mathrm{ran}\; rckmPath!) \cap \mathrm{ran}\; decisionPathConditionRCKM) \subset$$
$$\quad \mathrm{ran}\; decisionPathConditionRCKM$$
$$0 \leq decPathRCKMAccuracy(rckmPath!) \leq 100$$
$$head\,(\mathrm{dom}\; rckmPath!) \notin \mathrm{ran}\; ConclusionRCKM \cap \mathrm{ran}\; decisionPathConditionRCKM$$
$$\exists\, dp : decisionPathRCKM \mid dp \in RCKM \bullet$$
$$\quad \mathrm{dom}\; rckmPath! = \mathrm{dom}\; dp \setminus last\,(\mathrm{dom}\; dp) \Rightarrow last\,(\mathrm{dom}\; dp) = \mathrm{ran}\; rckmPath!$$
$$\mathrm{dom}\; rckmPath! = \exists\, p_{ckm} : decisionPathCKM \mid p_{ckm} \in CKM \bullet$$
$$\quad \mathrm{dom}(p_{pm}?) \cup \mathrm{dom}\; p_{ckm}$$



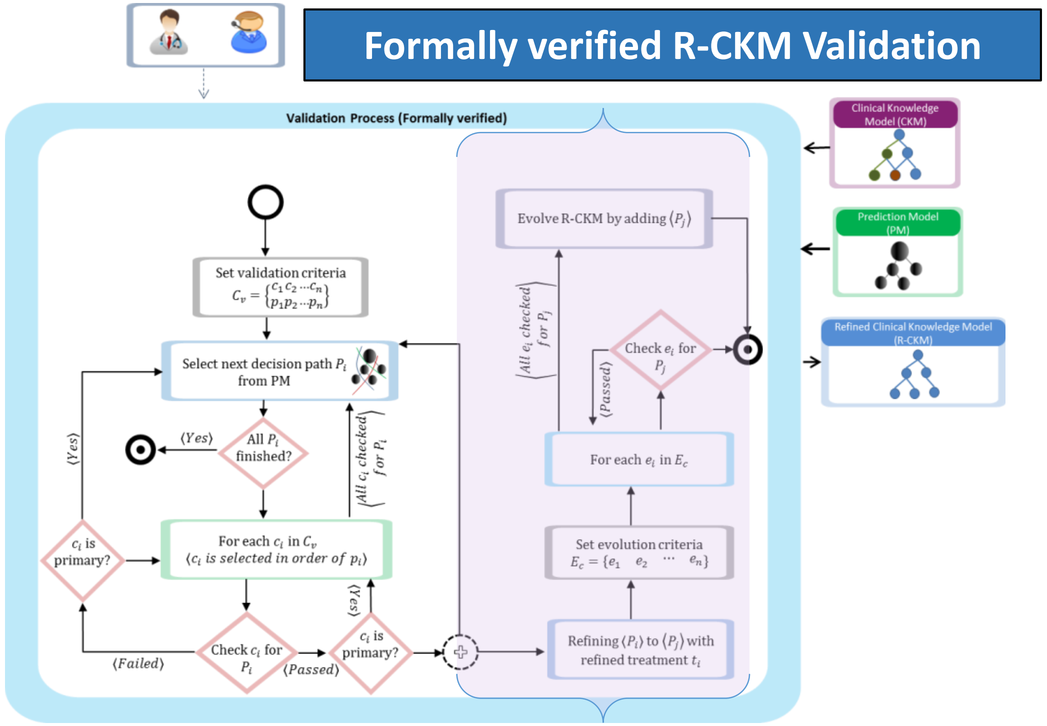
Nine(9) criteria after refinements

Results:
Comparison of formally verified validation process

R-CKM Validation



Formally verified R-CKM Validation



Criteria of validation Process

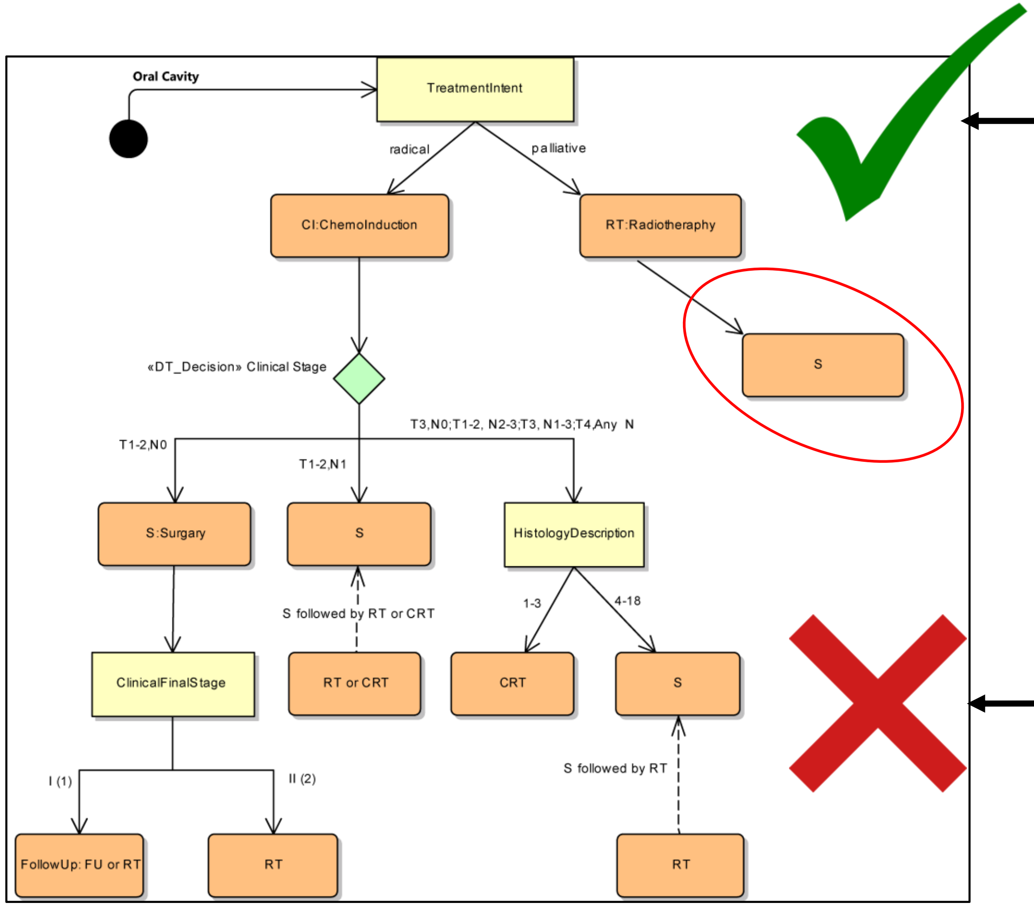
- ✓ $\{\forall P_i \in PM: \textbf{Accuracy}(P_i) > N\%\}$
- ✓ $\{\forall P_i \in PM \wedge \forall P_j \in CKM : \textbf{Conflict}(P_i, P_j)\}$
- ✓ $\left\{\forall P_i \in PM \wedge \exists P_j \in CKM: \textbf{Conform}(P_i, P_j) \xrightarrow{yields} P_i \in \Delta\, RCKM\right\}$
- ✓ $\left\{\exists P_i \in PM \wedge \forall P_j \in CKM : \; !\, Conform(P_i, P_j) \xrightarrow{provide} \textbf{Evidence}(P_i) \wedge Evidence(P_i) \xrightarrow{yields} P_i \in \Delta\, RCKM\right\}$

Additional criteria of validation Process

$$
\begin{aligned}
& RCKM \neq \emptyset \Rightarrow head\,(\mathrm{dom}\ dp_{pm}?) = rootRCKM \\
& \forall\, pos : \mathbb{N} \mid pos \in \mathrm{dom}\ refinements? \bullet pos > 1 \wedge \\
& \quad pos \leq (\#(\mathrm{dom}\ dp_{pm}?) + \#(\mathrm{ran}\ dp_{pm}?)) \\
& \mathrm{ran}(\mathrm{dom}\ rckmPath!) \subset \mathrm{ran}\ decisionPathConditionRCKM \\
& \mathrm{ran}(\mathrm{ran}\ rckmPath!) \subset \mathrm{ran}\ ConclusionRCKM \\
& (\mathrm{ran}(\mathrm{ran}\ rckmPath!) \cap \mathrm{ran}\ decisionPathConditionRCKM) \subset \\
& \quad \mathrm{ran}\ decisionPathConditionRCKM \\
& 0 \leq decPathRCKMAccuracy(rckmPath!) \leq 100 \\
& head\,(\mathrm{dom}\ rckmPath!) \notin \mathrm{ran}\ ConclusionRCKM \cap \mathrm{ran}\ decisionPathConditionRCKM \\
& \exists\, dp : decisionPathRCKM \mid dp \in RCKM \bullet \\
& \quad \mathrm{dom}\ rckmPath! = \mathrm{dom}\ dp \setminus last\,(\mathrm{dom}\ dp) \Rightarrow last\,(\mathrm{dom}\ dp) = \mathrm{ran}\ rckmPath! \\
& \mathrm{dom}\ rckmPath! = \exists\, p_{ckm} : decisionPathCKM \mid p_{ckm} \in CKM \bullet \\
& \quad \mathrm{dom}(p_{pm}?) \cup \mathrm{dom}\ p_{ckm}
\end{aligned}
$$

Maqbool Hussain et. al "Development Framework for evolutionary clinical decision support system", Artificial Intelligence (Elsevier, SCI, IF:3.37), under review (2016)
Maqbool Hussain, Taqdir Ali, Wajahat Ali Khan, Muhammad Afzal, Sungyoung Lee, Khalid Latif,;"Recommendations service for chronic disease patient in multi-model sensors home environment", Telemedicine and EHealth (SCI, IF:1.6), Vol. 21 Issue 3, pp.185-199, 2015

Results:
R-CKM Validation vs Formally verified R-CKM Validation



S1 — R-CKM Validation

Surgery (S) followed by RT is passed

Reason – It is refinement

S2 — Formally verified R-CKM Validation

Surgery (S) followed by RT is inconsistent

Reason – It is refinement, but it conflicts with guideline

$$\mathrm{dom}\ rckmPath! = \exists\, p_{ckm} : decisionPathCKM \mid p_{ckm} \in CKM \bullet \mathrm{dom}(p_{pm}?) \cup \mathrm{dom}\ p_{ckm}$$

According to guidelines:

For palliative patient followUp is recommended after RT

Contributions of Solution-2

Solution-2

Objectives

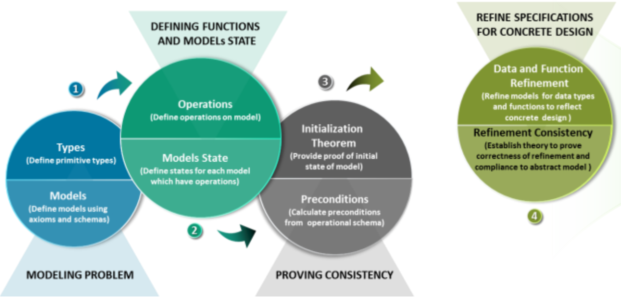
- ❑Formal modeling
- ❑Formal verification of validation process and R-CKM

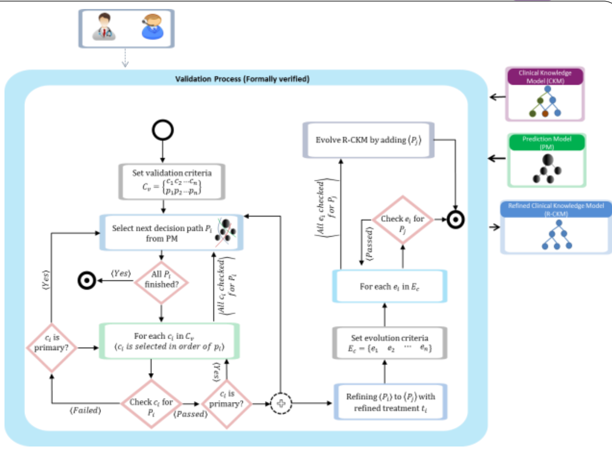
Target Limitations

❖Inconsistency in knowledge and validation process

Outcome

✓ Proving consistency of validation processes and R-CKM by eliminating inconsistencies using formal method





Additional criteria of validation Process

$$RCKM \neq \emptyset \Rightarrow head\,(\mathrm{dom}\ dp_{pm}?) = rootRCKM$$
$$\forall\, pos : \mathbb{N} \mid pos \in \mathrm{dom}\ refinements? \bullet pos > 1 \wedge$$
$$pos \leq (\#(\mathrm{dom}\ dp_{pm}?) + \#(\mathrm{ran}\ dp_{pm}?))$$
$$\mathrm{ran}(\mathrm{dom}\ rckmPath!) \subset \mathrm{ran}\ decisionPathConditionRCKM$$
$$\mathrm{ran}(\mathrm{ran}\ rckmPath!) \subset \mathrm{ran}\ ConclusionRCKM$$
$$(\mathrm{ran}(\mathrm{ran}\ rckmPath!) \cap \mathrm{ran}\ decisionPathConditionRCKM) \subset$$
$$\mathrm{ran}\ decisionPathConditionRCKM$$
$$0 \leq decPathRCKMAccuracy(rckmPath!) \leq 100$$
$$head\,(\mathrm{dom}\ rckmPath!) \notin \mathrm{ran}\ ConclusionRCKM \cap \mathrm{ran}\ decisionPathConditionRCKM$$
$$\exists\, dp : decisionPathRCKM \mid dp \in RCKM \bullet$$
$$\mathrm{dom}\ rckmPath! = \mathrm{dom}\ dp \setminus last\,(\mathrm{dom}\ dp) \Rightarrow last\,(\mathrm{dom}\ dp) = \mathrm{ran}\ rckmPath!$$
$$\mathrm{dom}\ rckmPath! = \exists\, p_{ckm} : decisionPathCKM \mid p_{ckm} \in CKM \bullet$$
$$\mathrm{dom}(p_{pm}?) \cup \mathrm{dom}\ p_{ckm}$$

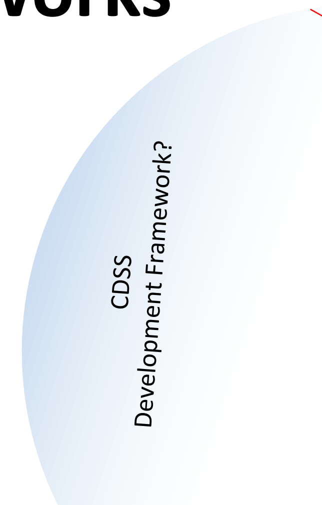
After all:
Formally verified knowledge validation processes and knowledge model.

Motivation:
Can solutions (1 and 2) be represent in a unified process model

Contributions

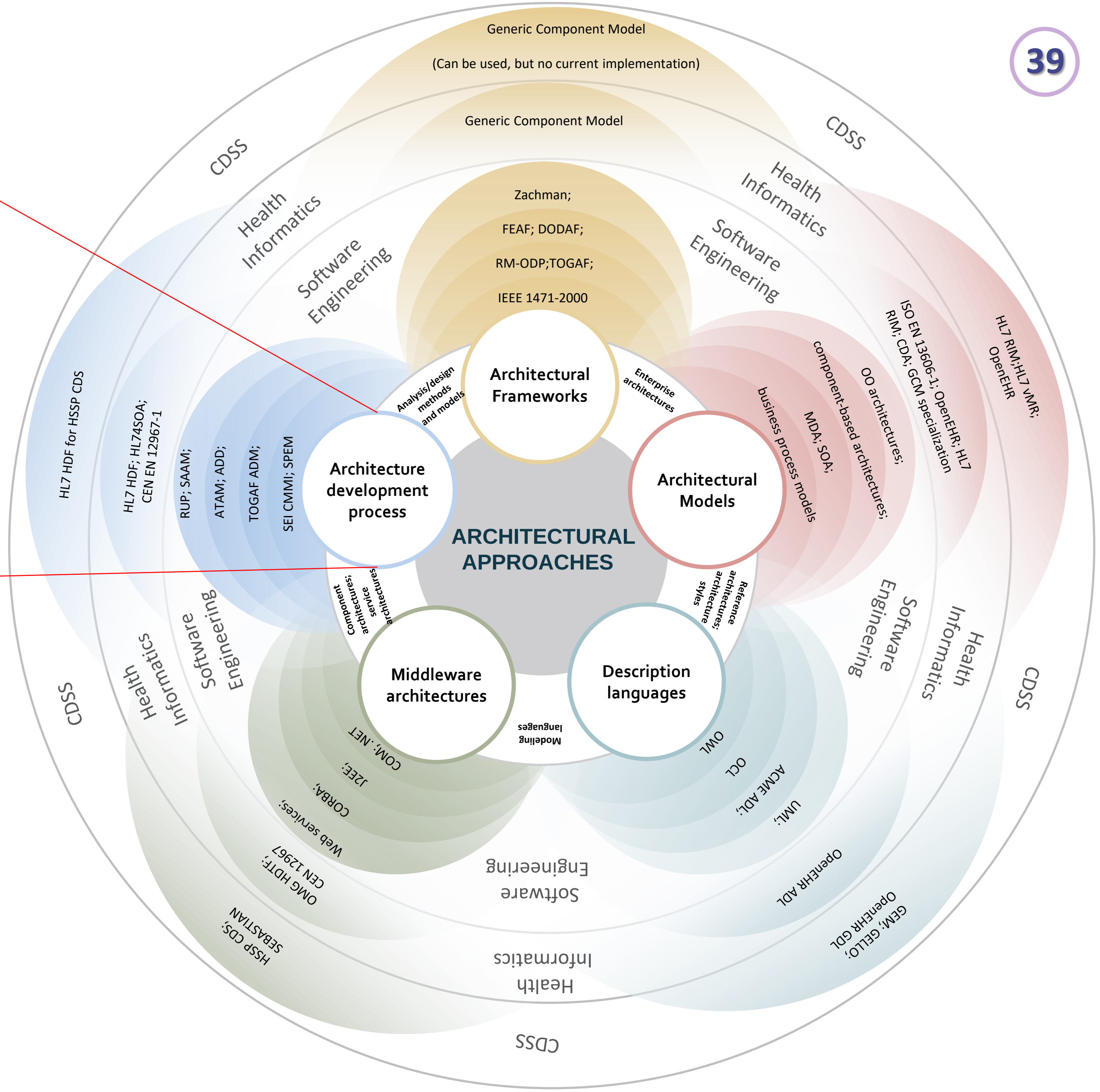
- ✓ Formally verified validation process
- ✓ Additional consistency criteria

Development Frameworks



Existing architectural approaches [24]

- Architectural approaches focuses on single aspect for CDSS:
  - Reference Models, Knowledge representation, Recommendation delivery mode
- Sophisticated knowledge acquisition methods are available, but they lack:
  - Process guidelines to plug into a complete CDSS development process



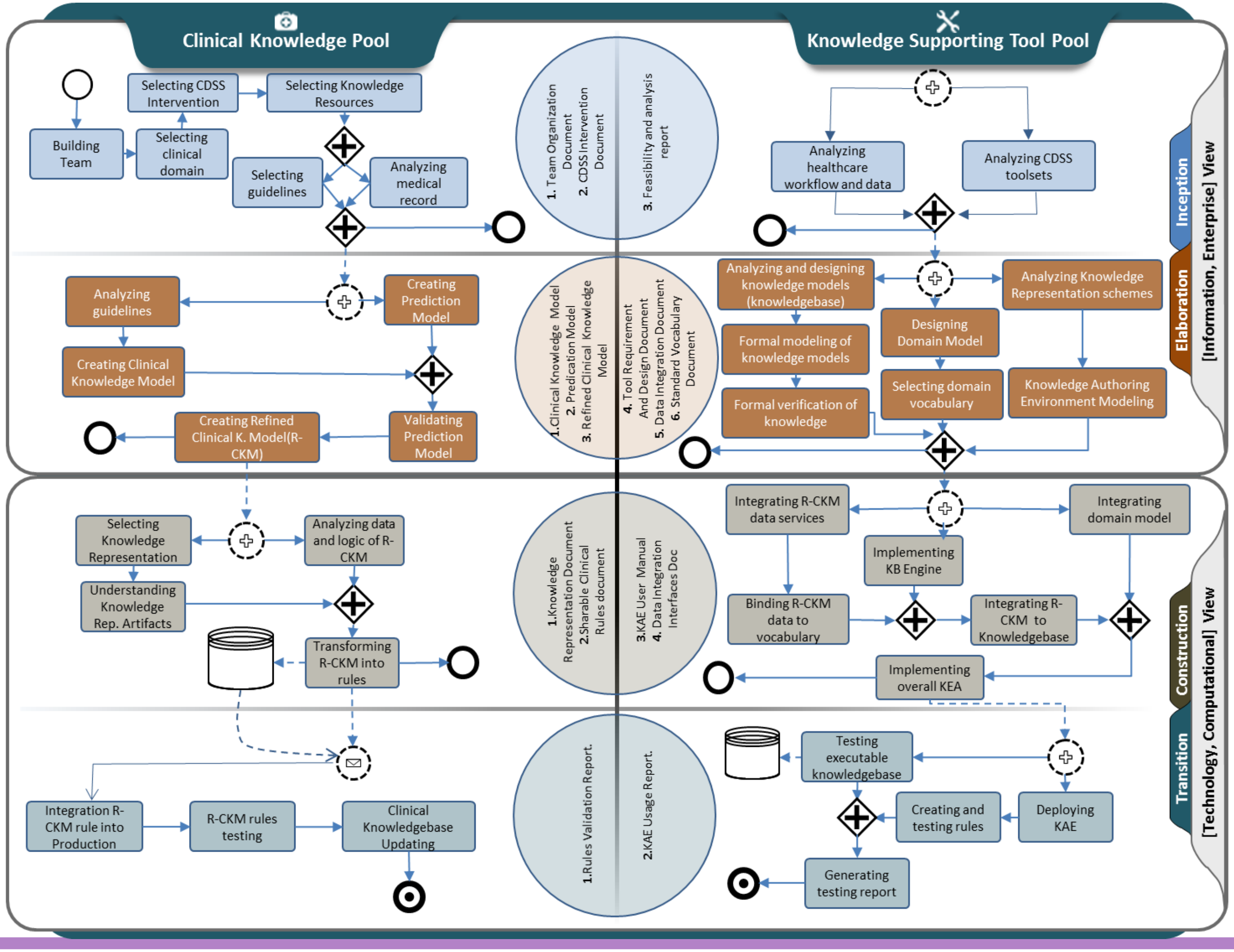
Smart CDSS Development Framework

Clinical Knowledge Pool

- ✓ Three phase model
  - ❖ Main Role: Domain Experts
  - ❖ Facilitator: Developers
- ✓ RUP phases
  - ❖ Inception: clinical objectives
  - ❖ Elaboration: PM, CKM, R-CKM (CPG building)
  - ❖ Construction: MLMs (CIG building)
  - ❖ Transition: MLM test base validation

Knowledge Supporting Tool Pool

- ✓ Authoring tools and formalism
  - ❖ Main Role: Developers
  - ❖ Facilitator: Domain Experts
- ✓ RUP phases
  - ❖ Inception: Tools analysis
  - ❖ Elaboration: Domain modeling, Tools modeling, Formal modeling
  - ❖ Construction: Tools implmentation
  - ❖ Transition: Tools deployments



Motivations for RUP

- ✓ Iterative process model
  - ❖ Best fit for evolutionary knowledge
  - ❖ Distribute development of system in coherent phases with concrete milestone
- ✓ Configurable processes
  - ❖ Tailoring mechanism for update or new processes
  - ❖ Allow to configure processes for CDSS specialized requirements

Motivations for RM-ODP

- ✓ Consistent viewpoints
  - ❖ Separate concerns of different stakeholders
  - ❖ Consistent viewpoints of the system
- ✓ Wide adoption
  - ❖ Widely adoption for industrial applications
  - ❖ Most healthcare frameworks adoption: ISO HISA, NEHTA-IF, HL7 SAIF

Uniqueness and Contributions

Knowledge acquisition and validation method

- ❑ The knowledge acquired is integratable with healthcare workflows
- ❑ The knowledge acquired is influenced from guidelines – supporting evidence

Quality of sharable knowledge

- ❑ Performance is preserved (accuracy:72.5%) over pure data-driven approach (accuracy:71.0%) [for completed treatments]

Formal verification of knowledge acquisition and validation method

- ❑ The method is formally verified for completeness and consistencies

Conclusion and Future Work

- This thesis contributes to:

Providing knowledge acquisition using validation methodology

- Incorporate guidelines (CKM) as source to validate knowledge (PM) acquired from practice data and refined as guideline supported knowledge (R-CKM).
- The refined knowledge (R-CKM) is converted into sharable knowledge (MLMs).

Formally verified the validation methodology

- Using formal methods (Z notation) to create formal model for knowledge acquisition and validation process
- The model is formally verified for completeness and internal consistency by applying proofs for initialization and precondition theorem on the model

- Future Research
  - Incorporating validation at sharable knowledge format.

Publications

- Patents (2)
  - International
- SCI/ SCIE Journals (10)
  - First Author- TWO Published
  - Co-Author- Eight Published
- Non SCI Journals (1)
  - First Author- One Published
- Conferences (24)
  - First Author- Four Publications
  - Co-Author- Twenty Publications

Total Publications = 35

THANK YOU!

Any questions or comments?

Selected References

[1] A. Wright, D. F. Sittig, A framework and model for evaluating clinical decision support architectures, Journal of biomedical informatics 41 (6) (2008) 982{990.
[2] A. Wright, D. F. Sittig, J. S. Ash, S. Sharma, J. E. Pang, B. Middleton,Clinical decision support capabilities of commercially-available clinical information systems, Journal of the American Medical Informatics Association 16 (5) (2009) 637-644.
[3] National comprehensive cancer network, http://www.nccn.org/, (Accessed: 24 Arpil, 2015) (2014).
[4] M. Samwald, K. Fehre, J. De Bruin, K.-P. Adlassnig, The arden syntax standard for clinical decision support: Experiences and directions, Journal of biomedical informatics 45 (4) (2012) 711-718.
[5] HL7CDS WG, Hl7 version 3 standard: Virtual medical record for clinical decision support (vmr-cds) logical model, Ann Arbor, MI:Health Level Seven, Inc, (Release 2) (2014).
[6] IHTSDO, Snomed ct, http://www.ihtsdo.org/snomed-ct/, (Accessed: 26 December, 2014) (2014).
[7] M. Peleg, Computer-interpretable clinical guidelines: A methodological review, Journal of biomedical informatics 46 (4) (2013) 744-763.
[8] M. Peleg, L. A. Gutnik, V. Snow, V. L. Patel, Interpreting procedures from descriptive guidelines, Journal of biomedical informatics 39 (2) (2006) 184{195.
[9] R. Serban, A. Ten Teije, Exploiting thesauri knowledge in medical guideline formalization, Methods of information in medicine 48 (5) (2009) 468-474.
[10] G. Hripcsak, J. J. Cimino, S. B. Johnson, P. D. Clayton, The columbiapresbyterian medical center decision-support system as a model for implementing the arden syntax., in: Proceedings of the Annual Symposium on Computer Application in Medical Care, American Medical Informatics Association, 1991, pp. 248-252.
[11] S. Perera, C. Henson, K. Thirunarayan, A. Sheth, S. Nair, Semantics driven approach for knowledge acquisition from emrs, IEEE Journal of Biomedical and Health Informatics 18 (2) (2014) 515-524.
[12] V. Gomoi, M. Vida, R. Robu, V. Stoicu-Tivadar, E. Bernad, O. Lupse, From data mining rules to medical logical modules and medical advices., Studies in health technology and informatics 192 (2012) 1094-1094.
[13] M. Toussi, J.-B. Lamy, P. Le Toumelin, A. Venot, Using data mining techniques to explore physicians' therapeutic decisions when clinical guidelines do not provide recommendations: methods and example for type 2 diabetes, BMC medical informatics and decision making 9 (1) (2009) 28.
[14] E. Shalom, Y. Shahar, M. Taieb-Maimon, G. Bar, A. Yarkoni, O. Young, S. B. Martins, L. Vaszar, M. K. Goldstein, Y. Liel, et al., A quantitative assessment of a methodology for collaborative speccation and evaluation of clinical guidelines, Journal of biomedical informatics 41 (6) (2008) 889{903.
[15] P. L. Miller, Domain-constrained generation of clinical condition sets to help test computer-based clinical guidelines, Journal of the American Medical Informatics Association 8 (2) (2001) 131-145.
[16] R. Bellazzi, B. Zupan, Predictive data mining in clinical medicine: current issues and guidelines, International journal of medical informatics 77 (2) (2008) 81-97.
[17] N. Pombo, P. A, J. Viana, Knowledge discovery in clinical decision support systems for pain management: A systematic review, Artficial intelligence in medicine 60 (1) (2014) 1-11.
[18] A. C. Lorena, L. F. Jacintho, M. F. Siqueira, R. D. Giovanni, L. G. Lohmann, A. C. De Carvalho, M. Yamamoto, Comparing machine learning classifiers in potential distribution modelling, Expert Systems with Applications 38 (5) (2011) 5268-5275.
[19] IBMSPSS User Guide, Ibm spss decision trees 21, http://www.sussex.ac.uk/its/pdfs/SPSS Decision Trees 21.pdf, (Accessed: 24 December, 2014) (2012).
[20] P. A. de Clercq, J. A. Blom, H. H. Korsten, A. Hasman, Approaches for creating computer-interpretable guidelines that facilitate decision support, Artificial intelligence in medicine 31 (1) (2004) 1-27.
[21] ArdenSyntax WG, Health level seven arden syntax for medical logic systems, Ann Arbor, MI:Health Level Seven, Inc, (Version 2.10) (2014).
[22] K. Kawamoto, G. Del Fiol, H. R. Strasberg, N. Hulse, C. Curtis, J. J. Cimino, B. H. Rocha, S. Maviglia, E. Fry, H. J. Scherpbier, et al., Multinational, multi-institutional analysis of clinical decision support data needs to inform development of the hl7 virtual medical record standard,in: AMIA Annual Symposium Proceedings, Vol. 2010, American Medical Informatics Association, 2010, pp. 377-381.
[23] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, I. H. Witten, The weka data mining software: an update, ACM SIGKDD explorations newsletter 11 (1) (2009) 10-18.
[24] D. M. Lopez, B. G. Blobel, A development framework for semantically interoperable health information systems, International journal of medical informatics 78 (2) (2009) 83{103.