

**Appendix C**  
**Data Dictionary of Database for ADRs**  
**(ER diagram was in Figure 3.5)**

Table name	Attribute	Type	Comment
canada_adr_drug	ad_index (PK)	varchar(7)	Index of canada_adr_drug table
	reaction_iden	varchar(10)	Reaction identifier
	aer_number	varchar(10)	Adverse Reaction Report (AER) number
	drug_iden	varchar(10)	Drug product identifier
	drugbank_id (FK)	varchar(7)	Drugbank ID
	role	text	Product role
	pt_code (FK)	varchar(8)	Code to identify the Preferred Term
center	record (PK)	varchar(10)	Record number
	pt_code (FK)	varchar(8)	Code to identify the Preferred Term
	hgnc_id (FK)	varchar(10)	Human protein ID from HUGO
	drugbank_id (FK)	varchar(7)	Drugbank ID
drugbank_mapping	drugbank_id (PK)	varchar(7)	Drug ID from DrugBank
	generic_name	text	Generic name of drug
	category	text	Categories of drug (separated by semi-colon)
	wiki_link	text	Wikipedia link of drug
	group	text	Group of drug (separated by semi-colon)
	brand	text	Trade name of drug (separated by semi-colon)
	atc_code	text	First level of ATC classification
	caco2	double	Caco-2 permeability coefficients
	dosage_route	text	How the drug is dispensed
	pka	double	The negative decadic logarithm of the ionization constant ( $K_a$ )
	half_life	double	Half life of drug in body (minute)
	logp	double	Water/octanol partition coefficient

Table name	Attribute	Type	Comment
	logs	double	Water solubility
	protein_binding	double	Percentage of the drug that is bound in plasma proteins
	elimination_route	text	Route by which the drug is eliminated
	toxicity	text	Description of side effects and toxic effects seen in humans
	hepat_score	double	Predictive score of drug to hepatobiliary disorders (mined by Weka)
drugbank_target	dp_index (PK)	varchar(6)	Index of drugbank_target table
	drugbank_id (FK)	varchar(7)	Drug ID from DrugBank
	hprd_id	varchar(5)	Human protein ID form HPRD as drug target
	hgnc_id (FK)	varchar(10)	Human protein ID from HUGO as drug target
meddra_llt	llt_code (PK)	varchar(8)	Code to identify the Lowest Level Term
	llt_name	text	Full name of the Lowest Level Term
	pt_code (FK)	varchar(8)	Code to identify the Preferred Term
meddra_pt	pt_code (PK)	varchar(8)	Code to identify the Preferred Term
	pt_name	text	Full name of the Preferred Term
	soc_code (FK)	varchar(8)	System Organ Class to which Preferred Term is linked
meddra_soc	soc_code (PK)	varchar(8)	Code to identify the System Organ Class
	soc_name	text	Full name of the System Organ Class
	soc_abbrev	varchar(8)	System Organ Class abbreviation
protein	hgnc_id (PK)	varchar(10)	Human protein ID from HUGO
	approved_symbol	varchar(15)	Gene symbol
	approved_name	text	Nomenclature of protein
	previous_symbol	text	Provious gene symbols (separeted by comma)
	aliase	text	Other aliases (separated by comma)
	chromosome	text	Location of gene coded protein

Table name	Attribute	Type	Comment
	enzyme_id	varchar(10)	Enzyme of protein
	entrez_id	varchar(10)	Entrez gene ID mapped by NCBI
	hprd_id	varchar(5)	Human protein ID from HPRD
	molec_term	text	Molecular function term and GO number (separated by semi-colon)
	bio_term	text	Biological process term and GO number (separated by semi-colon)
	cell_term	text	Cellular component term and GO number (separated by semi-colon)
taverna_adr_protein	ap_index (PK)	varchar(7)	Index of taverna_adr_protein table
	pt_code (FK)	varchar(8)	Code to identify the Preferred Term
	pt_name	text	Full name of the Preferred Term
	pmid	varchar(10)	PubMed ID as reference of ADR-protein association
	protein_name	text	Name of protein associated with ADR (searched and mapped by Taverna)
	uniprot_id	varchar(6)	Human UniProt ID which is searched and mapped by Taverna
	adr_protein_count	int(11)	Counting number of ADR and protein in PubMed
	adr_count	int(11)	Counting number of ADR in PubMed
	protein_count	int(11)	Counting number of protein in PubMed
	entrez_id	varchar(10)	Entrez gene ID mapped by NCBI
	ncbi_taxon	varchar(4)	NCBI taxonomy code (mapped by UniProt)
	hgnc_id (FK)	varchar(10)	Human protein ID from HuGO that found association with ADR using Taverna
	pmi	double	PMI to indicate the relevance of extracted information between ADR and protein which are observed together