# Converting the Clinical Word Document to a Sensely-format Web Document

## Requirements and Examples

The task is to take a formatted Word document and convert it to a web document that can be displayed and accessed within the Sensely conversational platform.

The original Word document for conversion contains a table of contents with section headings, text in paragraphs, a specially-formatted references section, internal document links to citations and links from references to external web pages, images with captions, and tables of data.

These elements need to be converted into HTML files with special formatting according to a CSS stylesheet you will be provided. Image files from the Word document will also need to be saved, resized, and included in the HTML document so that they display correctly.

### Word Files to Be Converted to HTML

- PBM Patient Module 3 Word Document
- PBM Business Module 4 Word Document

### Word to HTML Conversion Examples

#### Clinician Module 1 Example

The following is the original Word document version of Clinician Module 1:
https://sensely-conversation-content.s3.us-west-1.amazonaws.com/zuellig/20200821_MS51_MM_PBM_Clinician+Module+Draft_MM_v1_THS.DOCX

This is the complete Sensely-format web document for Clinician Module 1, which is one of the expected outputs for this task:
https://sensely-conversation-content.s3.us-west-1.amazonaws.com/zuellig/pbm-toolkit.html

**Duplicating features** of the Sensely-format web document that recreate features of the original Word document:

- Document content is represented in a specially formatted Table of Contents
- Statistics and mathematical notations are appropriately formatted

- Clickable citation links in parentheses in the main document bring the user directly to the corresponding citation in the references
- Clickable external links in the references section point to the appropriate linked web pages.
- Section headings are numbered and specially formatted to distinguish between main headings, sub-headings, sub-sub-headings, etc.

**Added features** of the Sensely-format web document that were not present in the original Word document:

- Table of content entries are clickable links that point directly to the appropriate section of the document, using an anchor tag.
- Each paragraph is followed by a link in parentheses that takes users back to the table of contents at the top of the document.
- Table of contents links to tables have been added for Table 1 and Table 2
- Tables in the original document, which are too wide to be converted into an HTML table comfortably for a 345px-width maximum frame, have been converted into images <345px in width, which link to external pages where the table is represented in a full readable format.
- Specially formatted small links after each citation in the references section will take the user back to the part of the document where that citation was used (note that the Clinician Module 1 uses a different format of links, but you should use the format of links in the Operational Module 2, and they should appear as below). For citations that are used multiple times, each link should be labeled with the numerical section of the document where the citation occurs, and with a letter if the citation is used multiple times in the same section of the document, as in the following example with the Abdullah et. al. citation, where it is used four times in Section 8, and each link points to an anchored point in the Section 8 where the in-text citation to Abdullah et. al. occurs:

  Abdullah, H. R., Ang, A. L., Froessler, B., Hofmann, A., Jang, J. H., Kim, Y. W., . . . Asia-Pacific, P. B. M. E. C. M. W. G. (2020). Getting patient blood management Pillar 1 right in the Asia-Pacific: a call for action. *Singapore Med J, 61*(6), 287-296. doi:10.11622/smedj.2019037 §8a, §8b, §8c, §8d

- External URLs, which can be very long and may exceed the maximum allowed page-width of 345px, are broken up with line breaks so that they do not exceed the boundaries of the display frame, as with the URL in the following BD citation, where a line breaks have been added to the URL

  BD. (2016). Retrieved from http://catalog.bd.com/nexus-ecat/categories?categoryID=R116,%20&%20&categoryName=VenousBloodCollectionTubes [Accessed March 31, 2020]. §3.3

**Missing features** in the Sensely-format web document that were present in the original Word document:

- Page numbers. The table of contents for the web document does not need page number references, and the HTML document does not need to be divided by page numbers. Instead, anchor links from the table of contents are used.

## Operational Module 2 Example

This is the original Word document version of Operational Module 2:
https://sensely-conversation-content.s3.us-west-1.amazonaws.com/zuellig/20210324_MS51_MM_AP_PBM_Operational+Module_Draft+3_THS.docx

This is the complete Sensely-format web document for Operational Module 2:
https://sensely-conversation-content.s3.us-west-1.amazonaws.com/zuellig/pbm-toolkit-operational.html

As with the Clinician Module, most features of the original Word document are duplicated in the HTML file. There are also added features that appear specifically in the Operational Module.

**Added features** of the Sensely-format web document that were not present in the original Word document:

- Section numbers were added to the sections of the document detailed in the original Word document's table of contents. Main headings are given numbers like 1, 2, 3, and subheadings are given decimal headings like 3.1, 3.2, 3.3. These are added to the HTML document's table of contents, to the section headings within the document, and to the names of anchors within the HTML (see Naming Conventions).
- Tables and images that are too wide for the 345px limit width of the presentation frame have been reformatted as narrower images, which link to another page with an original-size version of the table in a new window. In the case of images and figures where resizing the image to a 345px width would have made text unreadable, the image has been recreated to be readable at a 345px width, often by allowing it to go longer in vertical dimension (there is no frame limit on the length of an image)

Figure 1: The original table from the Word document, which is too large for a 345px frame

Table 1: History of events related to PBM.

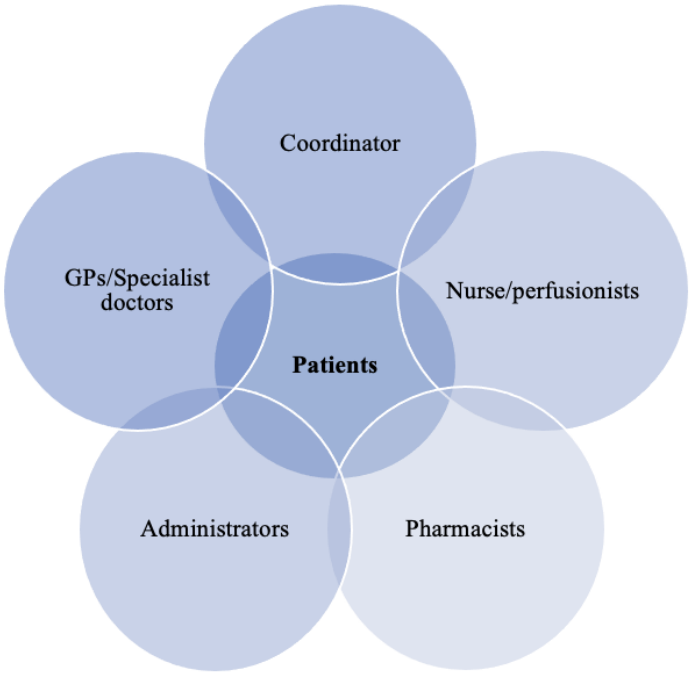
| Year/Period | Event |
|---|---|
| Early 1960s | Blood conservation methodology pioneered by Denton Cooley |
| 1988–2016 | Epidemics of hepatitis C and HIV |
| 1990 | Patient Self Determination Act increased awareness of the ability and need to increase communication between patients and caregiver |
| 1995 | Formation of The National Association for Bloodless Medicine and Surgery |
| Late 1990s | The NATA conceived |
| 1999 | Transfusion Requirements in Critical Care (TRICC) trial published |
| 2000 | SABM created |
| 1988–2016 | Numerous “Transfusion Guidelines” published |

Figure 2: The same table recreated as a much narrower image

**Table 1: History of events related to PBM.**

| Year/Period | Event |
|---|---|
| Early 1960s | Blood conservation methodology pioneered by Denton Cooley |
| 1988–2016 | Epidemics of hepatitis C and HIV |
| 1990 | Patient Self Determination Act increased awareness of the ability and need to increase communication between patients and caregiver |
| 1995 | Formation of The National Association for Bloodless Medicine and Surgery |
| Late 1990s | The NATA conceived |
| 1999 | Transfusion Requirements in Critical Care (TRICC) trial published |
| 2000 | SABM created |
| 1988–2016 | Numerous “Transfusion Guidelines” published |

Figure 3: Original image from the Word document that is too wide for the 345px frame

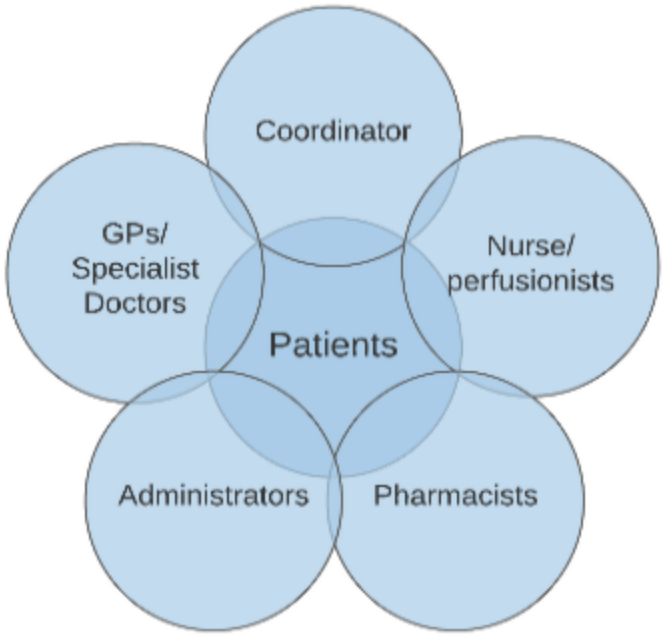



*ation*: GP, General practitioner

Figure 2: Primary stakeholders in PBM implementation

Figure 4: The same image, reformatted to be readable at 345px width

**Figure 2: Primary stakeholders in PBM implementation**




*Abbreviation*: GP, General practitioner

## Expected Outputs for this Task

- **HTML files** in .html file format and named according to a standard, consistent convention (see Naming Conventions).
    - **One main HTML file** containing all the content of the Word document. This HTML page will be the direct equivalent to the original Word file and will function identically to it, with images, tables, references, and links included.
    - **Several subordinate HTML files** that contain single sections of the whole Word document. These HTML files break up the original document into smaller pieces. All document-internal links in the subordinate HTML files can point to the larger comprehensive HTMl file described above—so, links will remain the same across all files.
    - **Several image files**, representing the images found in the original Word document, as well as images created to represent tables in the original Word document that are too large to comfortably fit in a 345px-width display frame.
- **Image files** in .png format, unless the original document image is in a different format and can be resized while remaining readable to the 345px width limit.
    - **All figures** should be represented by individual image files
    - **Any tables** that cannot be easily rendered in a 345px-width limit should be represented as <345px w small image files and larger, full-width image files, which will be paired through a link to a pop-up window.

# Steps for Conversion

1. Render the original text of the Word document into an HTML document with the same header and footer content as the examples (see Requirements and Examples), adjusted for the appropriate title for the module being encoded. It is fine to use an automated Word-to-HTML conversion tool for this task, but formatting will have to be corrected to match the Sensely-provided CSS classes and formatting standards. In our previous experience creating these HTML documents, we used an automated Word-To-HTML conversion tool, then used regular expressions to strip unneeded formatting tags and add CSS style references as needed, with some manual work required. The final HTML document should not have any extraneous formatting tags left over from any autoconversion, and should only use the Sensely-provided CSS classes for major text formatting.

2. Add anchors (and section numbers) to all section headers, so that they can be linked to within the document. Use the standardized format in Naming Conventions for the anchors.

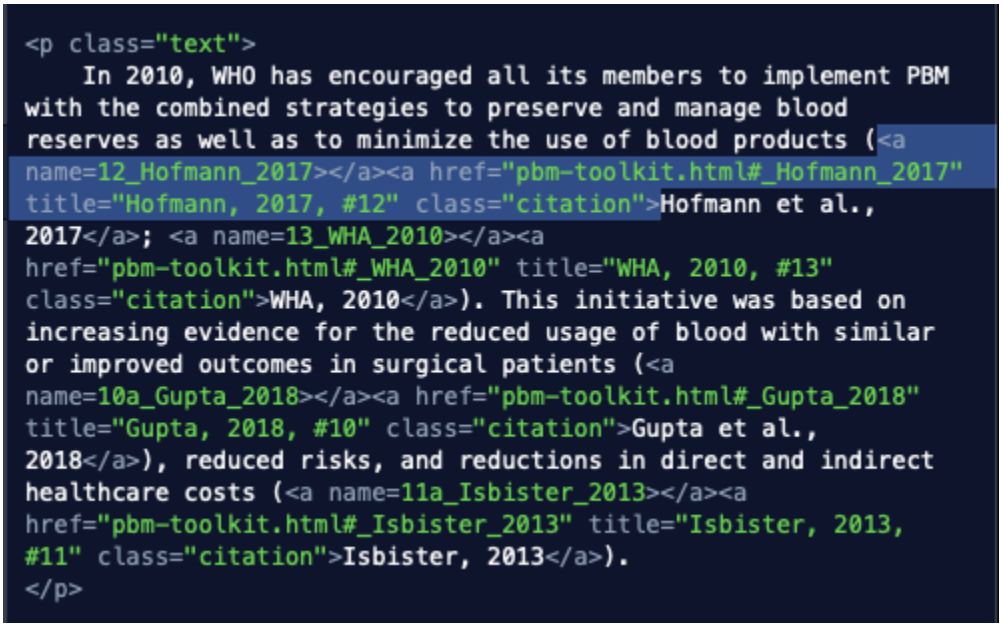
Figure 5: A formatted anchor tag added to a section heading in the HTML file

```
<div class="list">
    <div class="subhead">
            <a
name="2-1_evidence_of_successful_implementation_of_pbm"></a><h2>2.
1 Evidence of Successful Implementation of PBM</h2>
    </div>
</div>

<p class="text">
    In 2010, WHO has encouraged all its members to implement PBM
with the combined strategies to preserve and manage blood
reserves as well as to minimize the use of blood products (<a
name=12_Hofmann_2017></a><a href="pbm-toolkit.html#_Hofmann_2017"
title="Hofmann, 2017, #12" class="citation">Hofmann et al.,
2017</a>; <a name=13_WHA_2010></a><a
href="pbm-toolkit.html#_WHA_2010" title="WHA, 2010, #13"
class="citation">WHA, 2010</a>). This initiative was based on
increasing evidence for the reduced usage of blood with similar
or improved outcomes in surgical patients (<a
name=10a_Gupta_2018></a><a href="pbm-toolkit.html#_Gupta_2018"
title="Gupta, 2018, #10" class="citation">Gupta et al.,
2018</a>), reduced risks, and reductions in direct and indirect
healthcare costs (<a name=11a_Isbister_2013></a><a
href="pbm-toolkit.html#_Isbister_2013" title="Isbister, 2013,
#11" class="citation">Isbister, 2013</a>).
</p>
```

3. Link all headings in the table of contents to the appropriate sections of the HTML document, using the correct style for this category of link.
4. Link all in-text citations to their corresponding references in the references section (this may have been done for you if you used an autoconverting utility, in which case, make sure that the links are using the appropriate CSS style for citation links).
5. Place anchors on all in-text citations so that they can be linked back to from the references section. We used regular expressions to create anchors automatically from the content of the citations.

Figure 6: An in-text citation that links to an anchor before the appropriate reference (“ptm-toolkit.html#_Hofmann_2017”) and an anchor that allows the reference to link back to the original in-text citation (“12_Hofmann_2017”)



```
<p class="text">
    In 2010, WHO has encouraged all its members to implement PBM
with the combined strategies to preserve and manage blood
reserves as well as to minimize the use of blood products (<a
name=12_Hofmann_2017></a><a href="pbm-toolkit.html#_Hofmann_2017"
title="Hofmann, 2017, #12" class="citation">Hofmann et al.,
2017</a>; <a name=13_WHA_2010></a><a
href="pbm-toolkit.html#_WHA_2010" title="WHA, 2010, #13"
class="citation">WHA, 2010</a>). This initiative was based on
increasing evidence for the reduced usage of blood with similar
or improved outcomes in surgical patients (<a
name=10a_Gupta_2018></a><a href="pbm-toolkit.html#_Gupta_2018"
title="Gupta, 2018, #10" class="citation">Gupta et al.,
2018</a>), reduced risks, and reductions in direct and indirect
healthcare costs (<a name=11a_Isbister_2013></a><a
href="pbm-toolkit.html#_Isbister_2013" title="Isbister, 2013,
#11" class="citation">Isbister, 2013</a>).
</p>
```

6. Link all reference entries to an occurrence of the citation in the text, using a,b,c,d notation for references with more than one corresponding in-text citation.

Figure 7: Hoffman 2017 reference that links back to the original in-text citation

```
<p class="reference">
    <a name=_Hofmann_2017></a>Hofmann, A., N&oslash;rgaard, A.,
Kurz, J., Choorapoikayil, S., Meybohm, P., Zacharowski, K., &amp;
Kastner, P. G. H. (2017). Building national programmes of Patient
Blood Management (PBM) in the EU. <em>A Guide for Health
Authorities. Call for tender EAHC/2013/health/02
Contract</em>(20136106). <a class="referencelink"
href="pbm-toolkit.html#12_Hofmann_2017">&#8689;</a>
<p>
```

7. Make sure that all URLs in the references section that link to external pages are properly formatted and point to target="_blank" to open a new window. For all external URLs, use the following prefix text to properly format the link to be opened from the Sensely conversation window: "https://assets.sense.ly/webWidget/launchPage/launchPage.html?msg=Opening%20link&url=[EXTERNAL URL]". If the URL is long, break it up with line breaks so that the text does not exceed 345px.

Figure 8: An external link in the references section

```
<p class="reference">
    <a name=_WHA_2010></a>WHA. (2010). WHA 63.12. Sixty-Third
World Health Assembly,   Availability, safety, and quality
of blood products 2010. Retrieved from <a class="external"
href="https://assets.sense.ly/webWidget/launchPage/launchPage.html
?msg=Opening%20link&url=http://apps.who.int/gb/ebwha/pdf_files/WHA
63/A63_R12-en.pdf&urlLinkText=">http://apps.who.int/gb/ebwha/pdf_f
iles/<br />WHA63/A63_R12-en.pdf</a> [Accessed on February 05,
2020] <a class="referencelink"
href="pbm-toolkit.html#13_WHA_2010">&#8689;</a>
</p>
```

8. Convert all tables and images to a maximum 345px format. You can assume the images will be stored in the same folder as the HTML file and use bare file references. If a table cannot easily be converted to this width using HTML, capture the table as a small image and create a corresponding larger image. Then link the smaller image to the larger image using a target="blank" to open the image in a new window. For links to files opening in an external window, all files will be stored in the following location on sensely servers: https://sensely-conversation-content.s3-us-west-1.amazonaws.com/zuellig/

Figure 9: A table converted to a small image, which linked to a larger image version of the same table

```
<p class="tableheader"><a
name=table-2_suggested_approaches_to_pbm></a><strong>Table 1:
Patient blood management matrix</strong><p>

<p class="text"><em>Click table to open in a new tab</em></p>

<a
href="https://sensely-conversation-content.s3-us-west-1.amazonaws.
com/zuellig/pbm-toolkit-table_2.html" class="external"
target="_blank"><img
src="https://sensely-conversation-content.s3-us-west-1.amazonaws.c
om/zuellig/pbm-clinician-table2-image-small.png" /></a>
```

9. Review the text for any mathematical symbols that should be rendered with HTML entities, and ensure they display correctly when the HTML document is viewed from a browser.

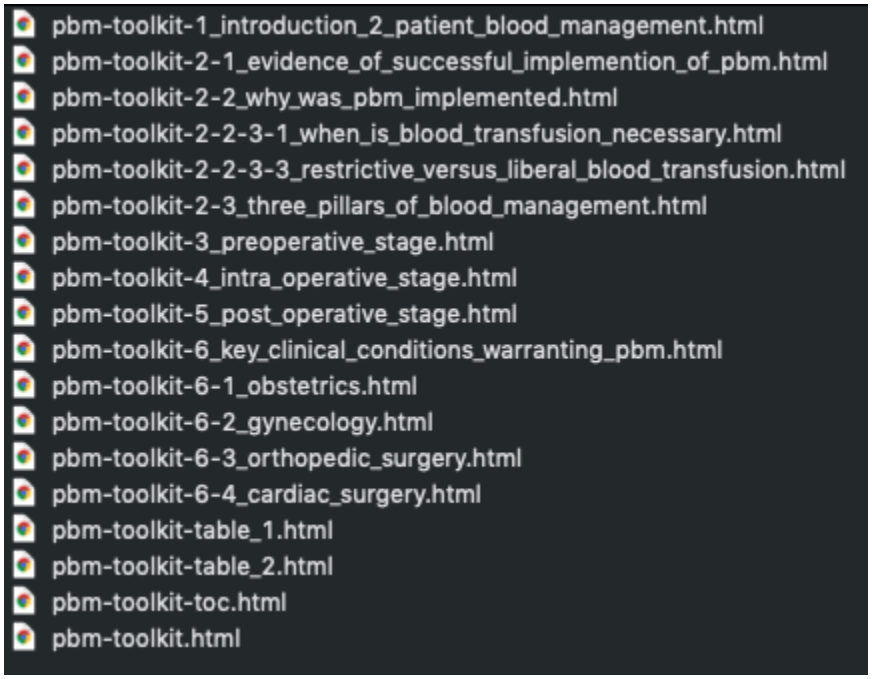
Figure 10: A paragraph containing > and < signs that have been converted to HTML entities.

```
<p class="text"> In an open-label, multicenter, randomized study,
FCM was shown to correct preoperative IDA (in patients with
menorrhagia) rapidly and achieve Hb &gt;10 g/dL in 2 weeks after
the first administration (78.8%) as effectively as iron sulfate
(IS) (72.3%) and in a shorter time period of 7.7 days vs. 10.5
days in IS group  (<a name=113c_Lee_2019></a><a
href="pbm-toolkit.html#_Lee_2019" title="Lee, 2019, #113"
class="citation">Lee et al., 2019</a>).   An RCT demonstrated
that a high dose (1000 mg over 15 minutes) FCM is more effective
in treating patients with anemia, ID, and heavy menstrual
bleeding as compared to 325 mg of ferrous sulfate. When compared
to the FS group, a higher number of patients in the FCM group
responded with 2.0 g/dL or more increase in Hb (82% vs. 62%;
<em>P</em>&lt;0.001); 3.0g/dL or more increase (53% vs. 36%,
P&lt;0.001). Additionally, a higher number of patients in the FCM
group demonstrated anemia correction (73% vs. 50%,
<em>P</em>&lt;0.001) as compared with the FS group (<a
```

10. Create additional HTML files containing subsections of the document. Subsection files should be created for just the table of contents, and for each major section. Links do not have to be changed (thus, a “Go back to table of contents” link can point to an anchor in

the main HTML document. It doesn't have to point to anywhere in the subsection HTML document).

Figure 11: A list of all the sub-section files for the clinician module HTML file, in addition to the original all-encompassing file "pbm-toolkit.html"




# CSS Classes for Word Document Content

Use the following CSS classes in pbm-stylesheet.css to classify all elements from the Word document in the HTML documents you generate.
p.doctitle: The class for the paragraph containing the document title and subtitle

- p.doctitle: The class for the paragraph containing the document title and subtitle
- span.title: The class for the title "Patient Blood Management Toolkit"
- span.subtitle: The class for the subtitle "[Patient|Business] Module"
- a.toc: The class for links that take the user back to the table of contents, at the end of each section
- div.toc: The container div for the table of contents
- .toclayer1: First level headings in the ToC (1 Introduction)
- .toclayer2: Second level headings in the ToC (1.1 Development of the Patient Blood Management Program)

- .toclayer3: Third level headings (3.3.1 Basic key performance indicators)
- .toclayer4: Fourth level headings (3.3.3.1 Anemia clinic)
- .toclayer5: Fifth level headings
- .tocspacer: Class to wrap ToC headings
- p.tocheader: Paragraph containing the headings "Table of Contents"
- p.text: Any paragraph of text in the document
- p.returntoc: Paragraph class that wraps all "Go back to table of contents" links after each section
- p.reference: Paragraph class that wraps all citations in the Works Cited section
- div.list: Wrapper for all section headings within the text of the document
- ol.decimalhead: Class for 'ol' tags that start off each first level heading (make sure to use start="x" for the number)
- div.subhead: Class for second level headings in the text
- div.subsubhead: Class for third level or lower headings in the text
- h1: Heading tag to wrap the text of first level headings
- h2: Second level heading text-wrapping tag
- h3: Third level
- h4: Fourth level
- h5: Fifth level (may not be needed)
- a.citation: Class for in-text citation links that will take the user down to the appropriate citation in the References section
- a.referencelink: A link after a citation in the References section that will take the user directly back to the place where that in-text citation was used
- a.toclink: A link in the Table of Contents

# Naming Conventions

**Sections**

Anchor tags for sections should always be identical in words, capitalization, and length to the actual section title in the document. Dashes should replace decimal points, and underscores replace spaces. For example, section “1.1 Development of the Patient Blood Management Program” should be given the anchor tag <a name="1-1_Development_of_the_Patient_Blood_Management_Program"></a>

**HTML Files**

The name of the overall HTML files for the Patient and Business modules should be “pbm-toolkit-patient.html” and “pbm-toolkit-business.html”. All HTML files that break up each document into sections should maintain the same name and added “-NUM-SECTIONNAME.html”. For example, the file broken off from “pbm-toolkit-operational.html” that contains just 1 Introduction section is named “pbm-toolkit-operational-1-Introduction.html

**Images**

Images can use a namescheme you like, as long as it is consistent about identifying all the images associated with one document and connecting smaller and larger images. For example, figures in the Operational Module 2 are labeled “Zuellig_operator_figure_06.png” for full-sized versions and “Zuellig_operator_figure_06_med.png” for smaller versions.

# Internal and External Links and Anchors

All elements of the document must have consistently namespaced anchors. All sections and subsections, all in-text citations, all figures and tables, and all references must have unique, identifiable anchors that can be directly linked to within or outside the document. Except for Sections, which must use the naming convention specified above, the structure of these anchor tags is up to your workflow, as long as they are consistent.

# Reference Files

**CSS Stylesheet**
pbm-stylesheet.css

**PBM Clinician Module 1**
PBM Clinician Module Original Word Document

pbm-toolkit.html (main converted HTML file, all files below are subsection files)

pbm-toolkit-1_introduction_2_patient_blood_management.html
pbm-toolkit-2-1_evidence_of_successful_implemention_of_pbm.html
pbm-toolkit-2-2_why_was_pbm_implemented.html
pbm-toolkit-2-2-3-1_when_is_blood_transfusion_necessary.html
pbm-toolkit-2-2-3-3_restrictive_versus_liberal_blood_transfusion.html
pbm-toolkit-2-3_three_pillars_of_blood_management.html
pbm-toolkit-3_preoperative_stage.html
pbm-toolkit-4_intra_operative_stage.html
pbm-toolkit-5_post_operative_stage.html
pbm-toolkit-6_key_clinical_conditions_warranting_pbm.html
pbm-toolkit-6-1_obstetrics.html
pbm-toolkit-6-2_gynecology.html
pbm-toolkit-6-3_orthopedic_surgery.html
pbm-toolkit-6-4_cardiac_surgery.html
pbm-toolkit-table_1.html
pbm-toolkit-table_2.html
pbm-toolkit-toc.html

**PBM Operational Module 2**

PBM Operational Module Original Word Document

HTML Files
pbm-toolkit-operational.html (main converted HTML file)
pbm-toolkit-operational-1-Introduction.html
pbm-toolkit-operational-2-Role-of-Stakeholders-in-the-Implementation-of-PBM.html
pbm-toolkit-operational-3-Developing-and-Sustaining-the-PBM-Program.html
pbm-toolkit-operational-5-Guidance-for-Hospitals-to-Establish-Preoperative-Anemia-Clinic.html
pbm-toolkit-operational-8-Successful-Cases-of-National-Health-Authorities-and-Hospitals-with-PBM-Operation.html
pbm-toolkit-operational-10-Recommendations-for-Implementation-of-PBM-Program.html
pbm-toolkit-operational-Table-4.html

Figure and Table Images
Zuellig_operator_figure_01_med.png
Zuellig_operator_figure_01.png
Zuellig_operator_figure_02_med.png
Zuellig_operator_figure_02.png
Zuellig_operator_figure_03_med.png
Zuellig_operator_figure_03.png
Zuellig_operator_figure_04_med.png
Zuellig_operator_figure_04.png
Zuellig_operator_figure_05_med.png
Zuellig_operator_figure_05.png
Zuellig_operator_figure_06_big.png
Zuellig_operator_figure_06_med.png
Zuellig_operator_figure_06.png
Zuellig_operator_table_01_med.png
Zuellig_operator_table_01.png
Zuellig_operator_table_02_med.png
Zuellig_operator_table_02.png
Zuellig_operator_table_03_med.png
Zuellig_operator_table_03.png
Zuellig_operator_table_04.png